Serviço Público Federal
UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

FICHA DE INSCRIÇÃO PARA SELEÇÃO 2011/1

Nível do Curso: **( ) Mestrado**
(X) Doutorado

1. DADOS PESSOAIS

Nome Completo: Dionisio Doering
Naturalidade: São Leopoldo Nacionalidade:Brasil UF: RS
Data de Nascimento:12/11/1977 Sexo: (X) Masculino ( ) Feminino
Estado civil: Casado
Filiação (nome do pai e da mãe): Paulo César Oliveira Doering e Rejane Luisa Doering

CPF n.°: ____________________ RG n.°: ________________________
Passaporte n.°: __________________________________ (para candidatos estrangeiros)
Endereço: Rua Santo Ângelo, 106 Cidade: São Leopoldo Estado: RS
País: Brasil CEP:93037-210
Telefone Comercial:510-486-5636 Telefone Pessoal:510-637-8395
E-mail: ddoering@lbl.gov

2. INFORMAÇÕES GERAIS

2.1) O PPGEE oferece as áreas de concentração e respectivas linhas de pesquisa abaixo. Assinale aquela à qual irá se candidatar e indique possíveis Orientadores:

1) Área de concentração: CONTROLE E AUTOMAÇÃO
(X) 1.A) Sistemas de Automação;
( ) 1.B) Sistemas de Controle;
( ) 1.C) Robótica;
( ) 1.D) Instrumentação.

2. Área de concentração: TECNOLOGIA DE INFORMAÇÃO E COMUNICAÇÕES
( ) 2.A) Processamento de Sinais;
( ) 2.B) Micro e Nano Eletrônica;
( ) 2.C) Sistemas de Comunicações.

3. Área de concentração: SISTEMAS DE ENERGIA
( ) 3.A) Sistemas de Potência;
( ) 3.B) Dispositivos Eletromagnéticos.

Orientador(es): _Prof. Dr. Carlos Eduardo Pereira________________________

2.2) Regime de Dedicação:
( ) Tempo Integral (X) Tempo Parcial

2.3) Solicita bolsa?
( ) Sim.
(X) Não. Por quê?__Aluno em tempor parcial__________________________________

____

2.4) Manterá vínculo empregatício durante o curso? (X) Sim ( ) Não
Empresa/Instituição:__Lawrence Berkeley National Lab.______________________
Cargo/Função:__Engenheiro de hardware______________________________
Telefone:__510-486-5636__________________________

3. FORMAÇÃO ACADÊMICA

3.1 – Graduação:
Título Universitário:Engenheiro Eletricista
Universidade:UFRGS
Nome do Curso:Engenharia Elétrica
Data de conclusão: Mês: Abril Ano: 2003
Foi bolsista de Iniciação Científica? ( ) Não (X) Sim – Qual agência de fomento/projeto:
_FAPERGS____
Data de início: ____/____/____ Data de término: ____/____/____

3.2 – Especialização:
Título Universitário:____
Universidade: ____
Nome do Curso/Programa de Pós-Graduação:___________________________
____
Data de conclusão: Mês:__________________________ Ano:______________________

3.3 – Mestrado:
Título Universitário:____
Universidade:UFRGS
Nome do Curso/Programa de Pós-Graduação: PPGEE
____
Data de conclusão: Mês:Janeiro Ano:2005

4. CONHECIMENTO DE IDIOMAS (Responder: excelente, bom, regular ou não se aplica)

Idioma	Leitura	Conversação	Redação	Compreensão
Inglês	Excelente	Excelente	Excelente	Excelente
Espanhol				
Francês				
Alemão				
Outro				

5. DOCUMENTOS EXIGIDOS

Para o julgamento de sua solicitação de inscrição, serão necessários os documentos abaixo relacionados, os quais deverão ser enviados à Coordenação do Curso, dentro do prazo pré-estabelecido:

- Ficha de inscrição preenchida (modelo PPGEE);
- Uma fotografia 3x4;

- Cópia do documento de identidade oficial e do CPF;
- Ao menos duas cartas de recomendação fornecidas por professores universitários ou profissionais de nível superior detentores do título de doutor, atuantes na área e preferencialmente com reconhecida experiência em pesquisa (modelo PPGEE);
- ***Curriculo Lattes*** (modelo CNPq - <http://www.cnpq.br/>);
- Anexar relação de disciplinas previamente cursadas em nível de pós-graduação nos últimos 06 (seis) anos, para as quais solicita equivalência e/ou aproveitamento – modelo PPGEE);
- Cópia autenticada do diploma de graduação (frente e verso), em curso reconhecido pelo MEC, ou atestado de que o aluno concluirá a graduação em 2011/1;
- Cópia autenticada do histórico escolar da graduação (Se for obtido eletronicamente, deverá ser carimbado e assinado pela Instituição de origem);
- Cópia autenticada frente e verso do Diploma de Mestrado (somente candidatos ao Doutorado – no caso de concluído);
- Cópia autenticada do histórico escolar de Mestrado (somente candidatos ao Doutorado – se for obtido eletronicamente, deverá ser carimbado e assinado pela Instituição de origem);
- Cópia das duas publicações mais relevantes (somente candidatos ao Doutorado);
- Descrição da área de interesse e proposta preliminar de trabalho de até quinze (15) páginas, com referências bibliográficas (somente candidatos ao Doutorado);
- Concordância de um Orientador devidamente credenciado pelo PPGEE (somente candidatos ao Doutorado – modelo PPGEE);
- O candidato em tempo parcial deve, também, providenciar documento junto à empresa em que trabalha, onde esta irá declarar que está ciente que o funcionário irá cursar o mestrado ou doutorado, que o mesmo será liberado nos horários da(s) disciplina(s) pretendida(s) e informar o número de horas semanais em que o funcionário será liberado;
- Candidatos do PQI: declaração do coordenador do PQI da IES a que é vinculado;
- Candidatos estrangeiros devem apresentar cópia autenticada do documento de passaporte (ou apresentar uma cópia simples e o documento original) e cópia do visto de estudante, se houver;
- CELPE-BRAS (Exame de Proficiência em Português para Estrangeiros) – esta é uma exigência do MEC e da CAPES para alunos estrangeiros. Para obter informações, consultar: <http://portal.mec.gov.br/index.php?option=com_content&view=article&id=12270&Itemid=518>
- Comprovante de pagamento da taxa de inscrição, no valor de R$ 100,00 (cem reais).

6. DECLARAÇÃO

Eu, DIONISIO DOERING, acima identificado declaro verdadeiras as informações prestadas, para fins de inscrição no processo de Seleção do Programa de Pós-Graduação em Engenharia Elétrica.

Local: BERKELEY, CA-USA Data: 09 / 12 / 2010

Assinatura do Candidato

7. ENDEREÇO PARA CORRESPONDÊNCIA

PPGEE - PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA
Universidade Federal do Rio Grande do Sul - UFRGS
Escola de Engenharia - Departamento de Engenharia Elétrica
Av. Osvaldo Aranha, 103 - Fone: (51) 3308-3129 Fax: (51)3308-3293
CEP: 90035-190 - Porto Alegre - RS
e-mail: <ppgee@.ufrgs.br> <http://www.ufrgs.br/ppgee>

República Federativa do Brasil
Ministério da Educação
UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

A Reitora da Universidade Federal do Rio Grande do Sul, no uso de suas atribuições e tendo em vista a conclusão do Curso de **Engenharia: Habilitação Engenharia Elétrica**, em 22 de março de 2003, e a colação de grau em 25 de abril de 2003, confere o título de

Engenheiro Eletricista

a

Dionísio Doering

nacionalidade brasileira, nascido a 12 de novembro de 1977, em São Leopoldo, Rio Grande do Sul, Cédula de Identidade nº 3015012465-I.I./RS, e outorga-lhe o presente Diploma a fim de que possa gozar de todos os direitos e prerrogativas legais.

Porto Alegre, 25 de abril de 2003.

Renato Machado de Brito
Diretor

Diplomado

Wrana Maria Panizzi
Reitora

CASA DA MOEDA DO BRASIL

Curso Reconhecido conforme Decreto nº 727 de 08/12/1900

MINISTÉRIO DA EDUCAÇÃO
Universidade Federal do Rio Grande do Sul
REITORIA

DIPLOMA registrado sob o nº 21 fl 04 do livro E-31 por competência atribuída pelo Art. 48, § 1º, da Lei nº 9.394, de 20 de dezembro de 1996.

Processo nº 23078.04915/03-49

Luis H. Godoy
Seção de Registro de Diplomas e Certificados

Em 03 de julho de 20 03

VISTO:
Ariane Wellis de Souza– Diretora do DECORDI
por delegação de competência da Pró-Reitoria Adjunta
de Graduação da UFRGS, conferida pela Portaria
nº 07 de 07-11-2000.

№ 016318

República Federativa do Brasil
Ministério da Educação
UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

O Reitor da Universidade Federal do Rio Grande do Sul, no uso de suas atribuições, tendo em vista a conclusão do curso de mestrado do Programa de Pós-Graduação em Engenharia Elétrica, da Escola de Engenharia, em 06 de janeiro de 2005, confere o título de

Mestre em Engenharia Elétrica
Área de Concentração: Automação e Instrumentação Eletro-Eletrônica

a

Dionísio Doering

nacionalidade brasileira, nascido a 12 de novembro de 1977, em São Leopoldo, Rio Grande do Sul, e outorga-lhe o presente diploma a fim de que possa gozar de todos os direitos e prerrogativas legais.

Porto Alegre, 25 de agosto de 2006.

Alberto Tamagna
Diretor

Diplomado

José Carlos Ferraz Hennemann
Reitor

CASA DA MOEDA DO BRASIL

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
PRÓ-REITORIA DE PÓS-GRADUAÇÃO

Curso reconhecido pela Portaria nº. 2.878, do Ministro de Estado da Educação, de 24/08/2005, publicado do D.O.U., 26/08/2005, nº. 165, Seção 1, p. 21. Diploma registrado sob o nº. 0711, fls. 119, do Livro PG 20, de acordo com artigo 48, parágrafo 1°, da Lei 9394, de 20 de dezembro de 1996.
Processo nº 23078.019594/06-93.

Porto Alegre, 25 de agosto de 2006.

Zaira Brasbiel de Azevedo
Diretora da Divisão de Diplomas e Certificados

Lia Teresinha Silva
Vice-Pró-Reitora de Pós-Graduação

№ 031724

Informações Acadêmicas do aluno DIONÍSIO DOERING - 96170

Histórico Escolar

Identificação e Endereço
Curso e Situação Acadêmica
Histórico Escolar Completo
Histórico do Curso Vigente
Possibilidade de Matrícula
Horário de Matrícula
Atividades de Ensino Corrente
Avaliação de desempenho
Pedido de Matrícula
SAIR

Lista das atividade de ensino cursadas pelo aluno na UFRGS.

Ano/Semestre	Atividade de Ensino	Conceito	Situação	Créditos
2002/2	ESTÁGIO SUPERVISIONADO III - ELE	A	Aprovado	0
2002/2	TÓPICOS ESPECIAIS EM SISTEMAS DIGITAIS I	A	Aprovado	4
2002/2	PROCESSAMENTO DIGITAL DE SINAIS	B	Aprovado	4
2002/2	INSTRUMENTAÇÃO A	B	Aprovado	4
2002/2	INSTALAÇÕES ELÉTRICAS	B	Aprovado	4
2002/2	COMUNICAÇÕES ÓTICAS	B	Aprovado	4
2002/2	APLICAÇÕES INDUSTRIAIS DA ELETRÔNICA	B	Aprovado	4
2002/2	SISTEMAS DE CONTROLE	B	Aprovado	6
2002/1	ELETRÔNICA APLICADA	A	Aprovado	4
2002/1	SISTEMAS DE TELEVISÃO	B	Aprovado	4
2002/1	PRINCÍPIOS DE COMUNICAÇÃO	B	Aprovado	4
2002/1	ANÁLISE DE SISTEMAS	B	Aprovado	4
2002/1	SISTEMAS DIGITAIS	A	Aprovado	5
2002/1	MICROPROCESSADORES I	A	Aprovado	5
2002/1	CONVERSÃO ELETROMECÂNICA DE ENERGIA II	B	Aprovado	6
2001/2	ENGENHARIA ECONÔMICA E AVALIAÇÕES	A	Aprovado	2
2001/2	TÓPICOS JURÍDICOS E SOCIAIS	B	Aprovado	2
2001/2	TEORIAS DE REDES	B	Aprovado	4
2001/2	TELEFONIA A	B	Aprovado	4
2001/2	TÉCNICAS DIGITAIS	A	Aprovado	6
2001/2	ELETRÔNICA FUNDAMENTAL II-A	A	Aprovado	6
2001/2	ONDAS ELETROMAGNÉTICAS	B	Aprovado	6
2001/2	CONVERSÃO ELETROMECÂNICA DE ENERGIA I	C	Aprovado	6
2001/1	FUNDAMENTOS DE ENGENHARIA DE SEGURANÇA DO TRABALHO - ELE	B	Aprovado	2
2001/1	FUNDAMENTO DE ECOLOGIA APLICADA - A	B	Aprovado	2
2001/1	LINGUAGEM C PARA ENGENHARIA	A	Aprovado	3
2001/1	HIDRÁULICA E HIDROLOGIA APLICADA II	B	Aprovado	4
2001/1	ECONOMIA A	B	Aprovado	4
2001/1	MATERIAIS ELÉTRICOS E MAGNÉTICOS	A	Aprovado	5
2001/1	TEORIA ELETROMAGNÉTICA APLICADA A	A	Aprovado	6
2001/1	ELETRÔNICA FUNDAMENTAL I-A	A	Aprovado	6
2001/1	CIRCUITOS ELÉTRICOS II A	B	Aprovado	6
2000/2	RESISTÊNCIA DOS MATERIAIS A	A	Aprovado	4
2000/2	MATEMÁTICA APLICADA II	A	Aprovado	6
2000/2	FÍSICA IV-C	A	Aprovado	6
2000/2	CÁLCULO NUMÉRICO	A	Aprovado	6
2000/2	CIRCUITOS ELÉTRICOS I-A	B	Aprovado	6
2000/1	ALGORITMOS E PROGRAMAÇÃO	A	Aprovado	4
2000/1	ADMINISTRAÇÃO E FINANÇAS	A	Aprovado	4
2000/1	MECÂNICA	B	Aprovado	4
2000/1	EQUAÇÕES DIFERENCIAIS II	A	Aprovado	6
2000/1	FÍSICA III-C	B	Aprovado	6
2000/1	ELETRICIDADE A	C	Aprovado	6
1999/2	INTRODUÇÃO À INFORMÁTICA	A	Aprovado	4
1999/2	PROBABILIDADE E ESTATÍSTICA	B	Aprovado	4
1999/2	DESENHO TÉCNICO II-A	B	Aprovado	4
1999/2	ÁLGEBRA LINEAR I - A	C	Aprovado	4
1999/2	FÍSICA II-C	A	Aprovado	6
1999/2	CÁLCULO E GEOMETRIA ANALÍTICA II - A	A	Aprovado	6
1999/2	QUÍMICA FUNDAMENTAL	B	Aprovado	6
1999/1	INTRODUÇÃO À ENGENHARIA ELÉTRICA	A	Aprovado	2
1999/1	GEOMETRIA DESCRITIVA II-A	B	Aprovado	2
1999/1	DESENHO TÉCNICO I-A	B	Aprovado	4
1999/1	FÍSICA I-C	A	Aprovado	6
1999/1	CÁLCULO E GEOMETRIA ANALÍTICA I - A	B	Aprovado	6

Histórico Escolar

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA **Área de Concentração:** PROCESS. SINAIS

DIONÍSIO DOERING 0988/99-5 Mestrado (x) Doutorado ( )

Disciplina							Professor Responsável	
Trimestre	**Código**	**Denominação**	**ob/el**	**cr**	**h/a**	**con**	**Professor**	**Título**
03/1	AD1	Atividade Didática I - Estágio-Docência na Graduação	E	1	15	A	Carlos Eduardo Pereira	Doutor
03/1	ELE00001	Instrumentação	E	3	45	A	Renato Machado de Brito	Doutor
03/1	ELE00017	Métodos Matemáticos para Engenharia Elétrica	E	3	45	B	Alexandre Sanfelice Bazanella	Doutor
03/1	ELE00019	Introdução aos Processos Estocásticos	E	3	45	A	Adalberto Schuck Júnior	Doutor
03/2	ELE00004	Processamento Avançado de Sinais	E	3	45	A	Adalberto Schuck Júnior	Doutor
03/2	ELE00014	Método de Elementos Finitos Aplicados a Campos	E	3	45	A	Ály Ferreira Flores Filho	Doutor
03/2	PROFIC	Proficiência em Língua Inglesa	E	0	0	APR	Departamento de Línguas Modernas	
03/3	ELE00005	Processamento de Imagens	E	3	45	A	Altamiro Amadeu Susin	Doutor
03/3	ELE00040	Trabalho Individual	E	3	45	A	Adalberto Schuck Júnior	Doutor
03/3	ELE00081	Tópicos Especiais em Processamentos de Sinais I	E	3	45	A	Adalberto Schuck Júnior	Doutor

Dissertação: Título -

Conclusão:

Professor Orientador: Adalberto Schuck Júnior

Data: 05/07/2004

Dionísio Doering

Possui graduação em Engenharia Elétrica pela Universidade Federal do Rio Grande do Sul (2003) e mestrado em Engenharia Elétrica pela Universidade Federal do Rio Grande do Sul (2005). Atualmente é hardware engineer - Lawrence Berkeley National Laboratory. Tem experiência na área de Engenharia Elétrica, com ênfase em Engenharia Elétrica, atuando principalmente nos seguintes temas: visão computacional, robótica móvel, espaço de escalas, células sangüíneas e digital signal processing.

(Texto gerado automaticamente pelo Sistema Lattes)

Última atualização em 25/11/2010

Endereço para acessar este CV:
http://lattes.cnpq.br/5394174946405180

Dados Pessoais

Nome	Dionísio Doering
Nome em citações bibliográficas	DOERING, Dionísio;Doering, D.;Doering, Dionisio;DOERING, D
Sexo	masculino
Filiação	Paulo César Oliveira Doering e Rejane Luiza Doering
Nascimento	12/11/1977 - São Leopoldo/RS - Brasil
Carteira de Identidade	3015012465 SSP - RS - 02/08/1990
CPF	91609089049
Endereço residencial	Rua Santo Ângelo,106 Santa Tereza - Sao Leopoldo 93037-210, RS - Brasil Telefone: 51 5723027 URL da home page: http://
Endereço profissional	Lawrence Berkeley National Laboratory One Cyclotron Rd. MS50A6134 - Berkeley 94720, - Estados Unidos Telefone: 510 4865636
Endereço eletrônico	e-mail para contato : ddoering@gmail.com e-mail alternativo : ddoering@yahoo.com

Formação Acadêmica/Titulação

2003 - 2005 Mestrado em Engenharia Elétrica.
Universidade Federal do Rio Grande do Sul, UFRGS, Porto Alegre, Brasil
Título: O uso do espaço de escala no conceito da transformada wavelet, Ano de obtenção: 2005
Orientador: Adalberto Schuck Junior
Bolsista do(a): Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
Palavras-chave: Espaço de Escalas, Image processing, CWT2D
Áreas do conhecimento : Engenharia Biomédica,Engenharia Elétrica

1999 - 2003 Graduação em Engenharia Elétrica.
Universidade Federal do Rio Grande do Sul, UFRGS, Porto Alegre, Brasil

Atuação profissional

1. Lawrence Berkeley National Laboratory - LBL

Vínculo institucional

2005 - Atual Vínculo: Term engineer , Enquadramento funcional: hardware engineer , Carga horária: 40, Regime: Integral

2. Grupo Gerdau S/A - GG*

Vínculo institucional

2005 - 2005 Vínculo: trainee , Enquadramento funcional: trainee , Carga horária: 40, Regime: Integral

Áreas de atuação

1. Engenharia Elétrica
2. Instrumentação Eletrônica
3. Hardware

Idiomas

Inglês Compreende Bem , Fala Bem, Escreve Bem, Lê Bem

Produção em C, T& A

Produção bibliográfica

Artigos completos publicados em periódicos

1. doi> Battaglia, Marco, Contarato, Devis, Denes, Peter, Doering, Dionisio, Duden, Thomas, Krieger, Brad, Giubilato, Piero, Gnani, Dario, Radmilovic, Velimir
Characterisation of a CMOS active pixel sensor for use in the TEAM microscope. Nuclear Instruments & Methods in Physics Research. Section A, Accelerators, Spectrometers, Detectors and Associated Equipment (Print). , v.622, p.669 - 677, 2010.
Referências adicionais : Inglês. Meio de divulgação: Meio digital, Home page: [doi:10.1016/j.nima.2010.07.066]

2. doi> Battaglia, Marco, Bisello, Dario, Contarato, Devis, Denes, Peter, Doering, Dionisio, Giubilato, Piero, Sung Kim, Tae, Mattiazzo, Serena, Radmilovic, Velimir, Zalusky, Sarah
Radiation hardness studies on CMOS monolithic pixel sensors. Nuclear Instruments & Methods in Physics Research. Section A, Accelerators, Spectrometers, Detectors and Associated Equipment (Print). , v.624, p.425 - 427, 2010.
Referências adicionais : Inglês. Meio de divulgação: Meio digital, Home page: [doi:10.1016/j.nima.2010.03.156]

3. ☆ doi> Denes, P., DOERING, Dionísio, Padmore, H. A., Walder, J.-P., Weizeorick, J.
A fast, direct x-ray detection charge-coupled device. Review of Scientific Instruments. , v.80, p.083302 - , 2009.
Referências adicionais : Inglês. Meio de divulgação: Meio digital, Home page: [doi:10.1063/1.3187222]

4. ☆ doi> BATTAGLIA, M, CONTARATO, D, DENES, P, DOERING, D, GIUBILATO, P, KIM, T, MATTIAZZO, S, RADMILOVIC, V, ZALUSKY, S
A rad-hard CMOS active pixel sensor for electron microscopy. Nuclear Instruments & Methods in Physics Research. Section A, Accelerators, Spectrometers, Detectors and Associated Equipment (Print). , v.598, p.642 - 649, 2009.
Referências adicionais : Inglês. Meio de divulgação: Meio digital

5. ☆ doi> Anderson, John, Brito, Renato, Doering, Dionisio, Hayden, Todd, Holmes, Bryan, Joseph, John, Yaver, Harold, Zimmermann, Sergio
. IEEE Transactions on Nuclear Science. , v.56, p.258 - 265, 2009.
Referências adicionais : Inglês. Meio de divulgação: Meio digital

6. doi> Battaglia, Marco, Contarato, Devis, Denes, Peter, Doering, Dionisio, Radmilovic, Velimir
CMOS pixel sensor response to low energy electrons in transmission electron microscopy. Nuclear Instruments & Methods in Physics Research. Section A, Accelerators, Spectrometers, Detectors and Associated Equipment (Print). , v.605, p.350 - 352, 2009.
Referências adicionais : Inglês. Meio de divulgação: Meio digital, Home page: [doi:10.1016/j.nima.2009.03.249]

7. doi> Lizarazo, J., Doering, Dionisio, Lawrence Doolittle, James Galvin, Caspi, S., Daniel R. Dietderich, Helene Felice, Ferracin, P., Arno Godeke, Joseph, John, Lietzke, A.F., Alessandro Ratti, Sabbi, G.L., Frederic Trillaud, X. R. Wang, Zimmermann, Sergio
Use of High Resolution DAQ System to Aid Diagnosis of HD2b, a High Performance Nb3Sn Dipole. IEEE Transactions on Applied Superconductivity. , v.19, p.2345 - 2349, 2009.
Referências adicionais : Inglês. Meio de divulgação: Meio digital, Home page: [doi:10.1109/tasc.2009.2019036]

8. ☆ DOERING, Dionísio, Schuck Jr., A.
A Novel Method for Generating Scale Space Kernels based on Wavelet Theory. Revista de Informática Teórica e Aplicada. , v.XV, p.121 - 137, 2008.
Referências adicionais : Inglês. Meio de divulgação: Vários

9. doi> Lizarazo, J., Lietzke, A.F., Sabbi, G.L., Ferracin, P., Caspi, S., Zimmermann, Sergio, Joseph, John, Doering, D.
Measurement of Fast Voltage Transients in High-Performance ${rm Nb}_{3}{rm Sn}$ Magnets. IEEE Transactions on Applied Superconductivity. , v.18, p.1581 - 1584, 2008.
Referências adicionais : Inglês. Meio de divulgação: Vários, Home page: [doi:10.1109/tasc.2008.921306]

10. ☆ doi> Roque, Waldir L., DOERING, Dionísio
Trajectory planning for lab robots based on global vision and Voronoi roadmaps. Robotica (Cambridge. Print). , v.23, p.467 - 477, 2005.
Palavras-chave: Robótica móvel, Visão computacional, Processamento de Imagem
Referências adicionais : Inglês. Meio de divulgação: Meio digital
O artigo está na fase de impressão pela editora (CUP, UK).

Trabalhos publicados em anais de eventos (completo)

1. doi> Zimmermann, Sergio, DOERING, D
EFFICIENT GAMMA-RAY SIGNAL DECOMPOSITION ANALYSIS BASED ON ORTHONORMAL TRANSFORMATION AND FIXED POLES In: Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on, 2010, Dallas, TX.
Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on. , 2010.
Palavras-chave: Digital signal processing
Referências adicionais : Estados Unidos/Inglês. Meio de divulgação: Meio digital

2. doi> Anderson, John, Brito, Renato, Doering, D., Hayden, Todd, Holmes, Bryan, Joseph, John, Yaver, Harold, Zimmermann, Sergio
Data acquisition and trigger system of the Gamma Ray Energy Tracking In-Beam Nuclear Array (GRETINA) In: IEEE Nuclear Science Symposium 2007, 2007, Honolulu/Hawaii.
Nuclear Science Symposium Conference Record. , 2007. v.3. p.1751 - 1756
Referências adicionais : Estados Unidos/Inglês. Meio de divulgação: Meio digital

3. DOERING, Dionísio, Schuck Jr., A., GUIMARÃES, Letícia Vieira
Contagem de Células Sangüíneas usando o Espaço de Escalas In: III Congresso Latino Americano de Eng. Biomédica e XIX Congresso Brasileiro de Engenharia Biomédica, 2004, João Pessoa.
IFMBE Proceedings. João Pessoa: Coral e IFMBE, 2004. v.5. p.1239 - 1242
Palavras-chave: Contagem de Células, Células Sangüíneas, Espaço de Escalas
Áreas do conhecimento : Engenharia Biomédica,Bioengenharia,Processamento de Sinais Biológicos
Referências adicionais : Brasil/Português. Meio de divulgação: Meio digital, Home page: [http://www.claeb2004.org.br]

4. DOERING, Dionísio, Schuck Jr., A.
Segmentação de cavidades em Imagens de Ecocardiografia em Espaço de Escalas In: III Congresso Latino Americano de Eng. Biomédica e XIX Congresso Brasileiro de Engenharia Biomédica, 2004, João Pessoa.
IFMBE Proceedings. João Pessoa: SBEB, CORAL & IFMBE, 2004. v.5. p.1323 - 1326
Palavras-chave: Eco-cardiografia, Segmentação, Espaço de Escalas
Áreas do conhecimento : Engenharia Biomédica,Processamento de Sinais Biológicos
Referências adicionais : Brasil/Português. Meio de divulgação: Meio digital, Home page: www.claeb2004.org.br

5. ROQUE, W. L., DOERING, Dionísio
Constructing Approximate Voronoi Diagrams from Digital Images of Generalized Polygons and Circular Objects In: The 11-th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision, 2003, Plzen.
Proceedings of WSCG 2003.. , 2003. v.1.
Palavras-chave: Voronoi Diagrams, Pattern Recognition, Computational Geometry, Image processing, Robotics
Áreas do conhecimento : Ciência da Computação
Referências adicionais : República Tcheca/Inglês. Meio de divulgação: Impresso

6. ROQUE, W. L., DOERING, Dionísio
Multiple Visiting Stations Trajectory (Re)Planning for Mobile Lab Robots based on Global Vision and Voronoi Roadmaps In: 33rd International Symposium on Robotics - ISR2002, 2002, Estocolmo.
Proceedings of ISR2002. , 2002. v.1.
Palavras-chave: Mobile robots, Path planning, Voronoi roadmaps, Computer vision
Referências adicionais : Suécia/Inglês. Meio de divulgação: Vários
A publicação do Proceedings do evento foi através de material impresso e através de CDROM. A referência do artigo no CDROM é: Service Robot Development I, paper 02.

7. DOERING, Dionísio, ROQUE, W. L.
Processamento de Imagem para Identificação da Estrutura Geométrica dos Obstáculos Presentes no Ambiente de Trabalho de Robôs Móveis In: Congresso Regional de Iniciação Científica e Tecnologia em Engenharia, 2002, Passo Fundo.
Anais do Cricte. , 2002. v.1.
Referências adicionais : Brasil/Português. Meio de divulgação: Meio digital

8. ROQUE, W. L., DOERING, Dionísio
Sistema de Planejamento de Trajetórias para Robôs Móveis com base em Visão Global In: II Congresso Brasileiro de Engenharia Mecânica - CONEM, 2002, João Pessoa.
Anais do II Congresso Nacional de Engenharia Mecânica. , 2002. v.01.
Palavras-chave: Robótica móvel, Planejamento de Trajetórias, Visão computacional
Referências adicionais : Brasil/Português. Meio de divulgação: Outro

9. DOERING, Dionísio
Instrumentação de pavimentos flexíveis In: XVI Congresso Regional de Iniciação Científica e Tecnologia em Engenharia, 2001, Ijuí.
Anais do Cricte 2001. , 2001. v.1.
Palavras-chave: Instrumentação, Pavimentação
Áreas do conhecimento : Medidas Elétricas,Infra-Estrutura de Transportes
Referências adicionais : Brasil/Português. Meio de divulgação: Meio digital

Trabalhos publicados em anais de eventos (resumo)

1. DOERING, Dionísio
Uso da Análise Espectral na Instrumentação In: XIII Salão de Iniciação Científica, 2002, Porto Alegre.
Anais do XIII Salão de Iniciação Científica. , 2002. v.1.
Áreas do conhecimento : Medidas Elétricas,Infra-Estrutura de Transportes
Referências adicionais : Brasil/Português. Meio de divulgação: Meio magnético

2. DOERING, Dionísio
Concepção de um sistema de aquisição de dados para instrumentação de pavimentos In: XII Salão de Iniciação Científica, 2001, Porto Alegre.
Anais do XII Salão de Iniciação Científica. , 2001. v.1.
Áreas do conhecimento : Medidas Elétricas,Infra-Estrutura de Transportes
Referências adicionais : Brasil/Português. Meio de divulgação: Meio digital

3. DOERING, Dionísio, PEIXOTO, Leonardo
CIT-Intelligent Traffic Controler In: 47th-Internaional Science and Engineering Fair, 1996, Tucson.
. , 1996.
Referências adicionais : Estados Unidos/Inglês. Meio de divulgação: Impresso

Eventos

Participação em eventos

1. Apresentação Oral no(a) **IEEE Nuclear Science Simposium**, 2007. (Congresso)
Data acquisition and trigger system of the Gamma Ray Energy Tracking In-Beam Nuclear Array (GRETINA).

2. Apresentação de Poster / Painel no(a) **7º SBPC Jovem**, 1999. (Congresso)
7º SBPC JovemLocal: PUC-RSKit Médico.
Áreas do conhecimento : Medidas Elétricas

Totais de produção

Produção bibliográfica

Artigos completos publicados em periódico	**10**
Trabalhos publicados em anais de eventos	**12**

Eventos

Participações em eventos (congresso)	**2**

Página gerada pelo Sistema Currículo Lattes em 25/11/2010 às 23:58:14.

Nuclear Instruments and Methods in Physics Research A 622 (2010) 669–677

Contents lists available at ScienceDirect

Nuclear Instruments and Methods in Physics Research A

journal homepage: www.elsevier.com/locate/nima

Characterisation of a CMOS active pixel sensor for use in the TEAM microscope

Marco Battaglia [a,b,*], Devis Contarato [a], Peter Denes [a], Dionisio Doering [a], Thomas Duden [a], Brad Krieger [a], Piero Giubilato [a,c], Dario Gnani [a], Velimir Radmilovic [a]

[a] *Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*
[b] *Santa Cruz Institute of Particle Physics, University of California at Santa Cruz, CA 95064, USA*
[c] *Dipartimento di Fisica, Università degli Studi, Padova I-35131, Italy*

ARTICLE INFO

Article history:
Received 10 June 2010
Accepted 22 July 2010
Available online 3 August 2010

Keywords:
Monolithic active pixel sensor
Transmission electron microscopy
Point spread function
Detection quantum efficiency

ABSTRACT

A 1M- and a 4M-pixel monolithic CMOS active pixel sensor with $9.5 \times 9.5\,\mu\text{m}^2$ pixels have been developed for direct imaging in transmission electron microscopy as part of the TEAM project. We present the design and a full characterisation of the detector. Data collected with electron beams at various energies of interest in electron microscopy are used to determine the detector response. Data are compared to predictions of simulation. The line spread function measured with 80 and 300 keV electrons is (12.1 ± 0.7) and $(7.4 \pm 0.6)\,\mu$m, respectively, in good agreement with our simulation. We measure the detection quantum efficiency to be 0.78 ± 0.04 at 80 keV and 0.74 ± 0.03 at 300 keV. Using a new imaging technique, based on single electron reconstruction, the line spread function for 80 and 300 keV electrons becomes (6.7 ± 0.3) and $(2.4 \pm 0.2)\,\mu$m, respectively. The radiation tolerance of the pixels has been tested up to 5 Mrad and the detector is still functional with a decrease of dynamic range by $\simeq 30\%$, corresponding to a reduction in full-well depth from ~ 39 to ~ 27 primary 300 keV electrons, due to leakage current increase, but identical line spread function performance.

1. Introduction

The TEAM (Transmission Electron Aberration-corrected Microscope) Project [1] has developed the most powerful electron microscope to date, with unprecedented sensitivity and spatial resolution, $\simeq 50$ pm [2] in both TEM (Transmission Electron Microscopy) and Scanning TEM (STEM). This was accomplished through advances in electron optics, particularly aberration correction, and the TEAM project included the development of new kinds of specimen stages and detectors.

One of the goals of the TEAM project is to be able to observe the dynamics of processes at the atomic scale [3,4], which requires advances over conventional TEM detectors. These include film, image plates and phosphors fiber-coupled to CCDs. Each of these techniques has fundamental limitations for high-speed in situ imaging. Film and image plates directly detect electrons with high spatial granularity but require relatively long exposures and are obviously not high-speed. Phosphors fiber-coupled to CCDs have a modest time granularity but are limited in their point spread function (PSF) and detection quantum efficiency (DQE) due to the physics limitations of the steps needed to convert the primary electrons into charge on the CCD cell. A high-speed detector able to directly detect electrons appears an ideal solution for TEM imaging. The key requirements for TEAM are a minimum of $1\text{k} \times 1\text{k}$ pixels, in order to obtain a field of view of $\simeq 200$ Å with a magnified pixel size of $\simeq 3$ pixels/0.5 Å (provided the PSF is ≤ 1 pixel) and a readout time of ≤ 10 ms, providing 100 or more frames/s. As TEAM is designed for material science applications, suitable radiation hardness is essential. Considering the standard way of operation of experiments in TEM, a radiation tolerance of ≥ 1 Mrad would enable its use for approximately one year, which appears to be a valid requirement.

There has been significant interest in the application of both CMOS [5–7] and hybrid [8–10] pixel sensors for direct imaging in TEM, to replace conventional CCDs optically coupled to phosphor plates [11]. In two recent papers [12,13] in this journal we discussed the design of a CMOS sensor prototype with a radiation-hard pixel cell and characterised its response in terms of energy deposition and line spread function. We found that by proper design of the pixel cell the sensor can be made enough radiation tolerant to operate properly up to several Mrad of ionising dose and the line spread function for 10 μm pixel varies from 12 to 8 μm for electrons in the energy range $80\,\text{keV} < E_e < 300\,\text{keV}$ of interest in TEM. These results motivated the development of a larger size pixel chip to perform tests of fast imaging at TEAM.

---

[*] Corresponding author at: Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. Tel.: +1 510 486 7029.
E-mail address: MBattaglia@lbl.gov (M. Battaglia).

0168-9002/$ - see front matter Published by Elsevier B.V.
doi:10.1016/j.nima.2010.07.066

The final TEAM detector contains 2k × 2k pixels with 100 frames/s readout. As an intermediate stage, 1k × 1k detectors were fabricated. In this paper we present the design of the TEAM1k pixel chip and discuss the results of its characterisation on a TEAM project test column.

2. The TEAM pixel sensors

The TEAM1k and TEAM2k detectors are fabricated in a 0.35 μm CMOS process. For the TEAM1k, four $1 \times 1\,\mathrm{cm}^2$ chips are placed within the $2 \times 2\,\mathrm{cm}^2$ reticle area and for the TEAM2k the entire reticle is used. All chips employ an identical $9.5 \times 9.5\,\mu\mathrm{m}^2$ pixel. In the TEAM1k reticle, two of the chips have a 1024 × 1024 pixel imaging area ("imaging" version), and the other two chips have the same imaging area, except that a central 500 μm circular area is replaced with a simple diode ("diffraction" version). One "imaging" and one "diffraction" chips are designed with analog and digital sections very similar to the design previously reported, as backup designs. The other two "imaging" and "diffraction" chips have several improvements over previous designs, and these new versions are the ones discussed in this paper. The TEAM1k chips are organised as 16 identical slices each with 1024 rows and 64 columns. Each slice has a novel analog output buffer, capable of driving a capacitive load of up to 20 pF. In addition, each slice has a variable bottom-of-column current load, which can "look ahead" a programmable number of columns. In a conventional active pixel sensors (APS), there is a fixed current bias at the bottom of the column. In order to reduce power, as the sensor has to be operated in vacuum, the standing current in this bias should be small. For speed, though, this current needs to be large enough for the signal to slew within the time needed to digitise a pixel. In this implementation, a larger bias current is switched on when the given column is selected ($N=0$), and m clock cycles before ($N=m$) and then maintained for the next clock cycle when the column is selected. Sending the output to digitising electronics outside of the microscope vacuum requires an amplifier capable of providing high currents as needed to drive the capacitive load. Again, with the desire to minimise power, the output stage consists of two parts: a conventional low-power operational amplifier, along with a higher-power slew rate enhancer (SRE). The small-signal amplifier has the necessary gain, bandwidth, and noise performance to settle to the required precision, while the SRE minimises the fraction of the sample period that is spent slewing a large voltage step. The SRE is activated when the error voltage at the small-signal amplifier input exceeds a designed value.

Since charge generation is confined primarily to the thin epitaxial layer, just $\simeq 14\,\mu$m thick, it is possible to remove most of the underlying detector bulk silicon using a back-thinning process. The charge collection, noise and charge-to-voltage calibration have been studied by characterising a batch of CMOS pixel sensors before and after back-thinning and no significant effects have been observed [14]. Back-thinning is performed by Aptek Industries [15] using a proprietary technique. The process has been extensively tested and yields in excess of 90% have been obtained for thicknesses down to 50 μm [14]. TEAM1k detectors are thinned to 50 μm. The use of thin sensors minimises the charge spread due to back-scattering in the bulk, thus maximising the contrast ratio in TEM imaging.

3. Response simulation

The energy deposition in the sensor active layer and the lateral charge spread are simulated with the `Geant4` program [16], using the low-energy electromagnetic physics models [17]. The CMOS pixel sensor is modelled according to the detailed geometric structure of oxide, metal interconnect and silicon layers, as specified by the foundry. Electrons are incident perpendicular to the detector plane and tracked through the sensor. For each interaction within the epitaxial layer, the energy released and the position are recorded.

Charge collection is simulated with `PixelSim`, a dedicated digitisation module [18], developed in the `Marlin` C++ reconstruction framework [19] as discussed in Ref. [12]. The simulation has a single free parameter, the diffusion constant σ_{diff}, used to determine the width of the charge carrier cloud. Its value is extracted from data by a χ^2 fit to the pixel multiplicity in the clusters obtained for 300 keV electrons. We find $\sigma_{\mathrm{diff}} = (14.5^{+2.0}_{-1.0})\,\mu$m, which is compatible with both the value obtained for 1.5 GeV e^- s with a prototype having $20 \times 20\,\mu\mathrm{m}^2$ pixels produced in the same CMOS process [13] and that inferred from the diffusion length (estimated from the doping in the epitaxial layer) and the measured charge collection time [12].

4. Sensor tests

The sensor response to ionising radiation is studied with 5.9 keV X-rays and electrons of energy ranging from 80 to 300 keV. The microscope tests are carried out in the FEI Titan test column at the National Center for Electron Microscopy (NCEM). The prototype chip is mounted on a proximity board which is held on the film insertion plate of the microscope. The insertion plate is built on a removable mechanical assembly which allows easy and quick replacement of the sensor. The plate can be inserted in and retracted from the column by means of a pneumatic actuator.

The temperature is monitored during operation by measuring the resistance of a temperature-sensitive resistor included in the chip. In absence of active cooling, the thick sensor operates at a temperature of $\simeq 90\,^\circ$C, while the thin sensor between 40 and 50 °C, depending on running conditions. Data are taken for 6.25 and 25 MHz clock frequency and different beam intensities. For response characterisation, four contiguous sectors of the chip are tested at the same time, the four analog outputs being read out in parallel. Two flex cables, exiting from the column through a vacuum feed-through, provide the interconnection with the data acquisition and monitoring electronics located outside the microscope. An intermediate board provides the chip with the biases and driving clocks, performs signal amplification and deploys differential stages for adapting the analog output signals to the data acquisition (DAQ) system.

4.1. *Data acquisition and analysis*

For this study a custom DAQ system [20] is used. Data from four of the chip analog outputs are digitised by four 100 ms/s, 14-bit ADCs on a custom-designed board interconnected with a commercial development board [21], implementing a Xilinx Virtex-5 FPGA device, 64 Mb of on-board DDRAM memory and communication devices. Clock pattern generation and chip control is performed via high-speed LVDS lines by the Virtex-5, which also controls and synchronises the acquisition with the ADCs reference clock.

The DAQ system is connected to a control PC via a USB 2.0 bus, for acquisition control and DAQ steering through dedicated registers and for data retrieval. The bandwidth for data transfer is ~40 Mbytes/s. A graphical user interface based on a series of dedicated C/C++ classes interfaced to the ROOT [22] framework classes provides an intuitive and easy-to-use system for acquisition monitoring, fast data display and on-line processing [20]. Raw data are stored in ROOT format, and further converted in the LCIO [23] format for offline analysis. This is performed with a set of dedicated processors, developed in the `Marlin` reconstruction framework.

4.2. Noise, uniformity and calibration

The sensor noise and uniformity among the different analog outputs are measured in the lab in dark. The pedestal level for each pixel is computed from the average pulse height over 100 consecutive acquisitions. A pixel-to-pixel dispersion of less than 8 mV is obtained. The pedestal dispersion varies within 2% among the different analog outputs of the same sensor. An average pixel leakage current below 10 fA is inferred from the results of measurements performed at different clock frequencies ranging from 25 MHz down to 6.25 MHz, corresponding to integration times of 2.6 and 10.5 ms, respectively. The sensor charge-to-voltage conversion gain is measured from the response to 5.9 keV X-rays from a ^{55}Fe source, corresponding to a charge generation of 1640 electrons. Fig. 1 shows the spectrum obtained on one analog output for a clock frequency of 6.25 MHz, at room temperature. A calibration of 14.9 e$^-$/ADC count is obtained. The gain of the different analog outputs of the same sensor is found to vary within 3%. The charge-to-voltage conversion gain is expected to depend on the voltage applied to the pixel diode junction. For small signals, as those induced by one or few electrons, the capacitance of the junction can be considered constant and the conversion linear. For larger signals obtained when many electrons hit the same pixel in a single frame the capacitance of the diode junctions is expected to increase by several percent, up to ~20% for signals corresponding to the full sensor dynamic range. Since most of the analyses reported in this paper deal with relatively small signals, up to few tens of electrons or less, we assume a constant calibration factor in the analysis. Typical average pixel noise figures of (32 ± 2)e$^-$ and (28 ± 2)e$^-$ of equivalent noise charge (ENC) are measured under operating condition in the microscope set-up for clock frequencies of 25 and 6.25 MHz, respectively.

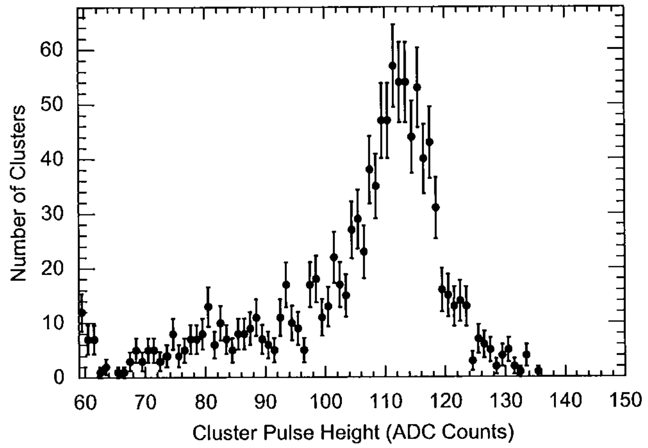


Fig. 1. ^{55}Fe spectrum obtained on one analog output at room temperature and for 6.25 MHz clock frequency.

4.3. Electron response

The response to single electrons is characterised in terms of energy deposition and cluster size. We study the energy deposition in the sensor pixels, varying the number of electrons recorded per pixel and per frame. First, we operate with a flux of $\simeq 50\,e^-\,mm^{-2}\,frame^{-1}$, which allows us to resolve individual electrons. Under these conditions, electron hits are reconstructed as pixel clusters. For these events the sensor response is characterised in terms of the number of pixels associated to a cluster and the pulse height measured in a 5 × 5 pixel matrix centred around a seed pixel. The response measured on data is compared to the `Geant4`+`PixelSim` simulation. Fig. 2 shows the reconstructed energy deposited by 300 keV electrons for data and simulation. A good agreement is observed.

The uniformity across a detector sector has been studied for bright field illumination at a flux corresponding to an energy deposition of $\simeq 200\,keV\,pixel^{-1}\,frame^{-1}$. The distribution of the energy recorded in each pixel in a single frame is shown in the right panel of Fig. 2. The dispersion of signal on the pixels fully exposed to the electron beam is obtained by a Gaussian fit to the recorded energy distribution. The relative r.m.s. dispersion at 300, 200, 120 and 80 keV is 0.063, 0.058, 0.080 and 0.11, respectively. The increase at lower energies is interpreted as an effect of the larger energy loss fluctuations, since measurements are performed at constant energy per pixel instead of constant number of electrons per pixel.

The study of the cluster size allows us to investigate the charge spread around the point of impact of each electron onto the detector. This is used to validate the simulation and to assess the

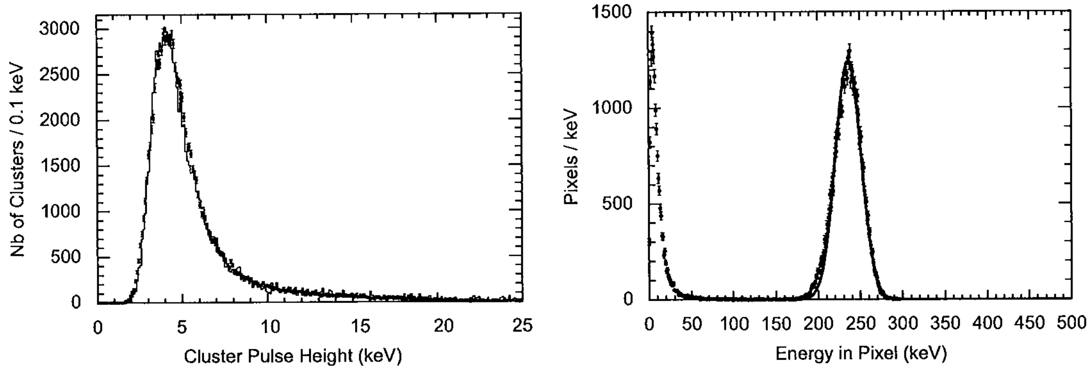


Fig. 2. Reconstructed deposited energy: (left) energy recorded in the 5 × 5 pixel matrix for single 300 keV electrons. The points with error bars show the data and the histogram the result of the `Geant4`+`PixelSim` simulation. (right) Energy recorded per pixel for multiple electrons in bright field illumination. The points show the data recorded in a single frame: pixels fully exposed to the electron beam exhibit a rather uniform response described by a Gaussian distribution. Pixels with low recorded energy correspond to the areas of the detector screened with an Au wire and a metal plate knife edge (see text). Only pixels with at least 1 keV of recorded energy are shown.

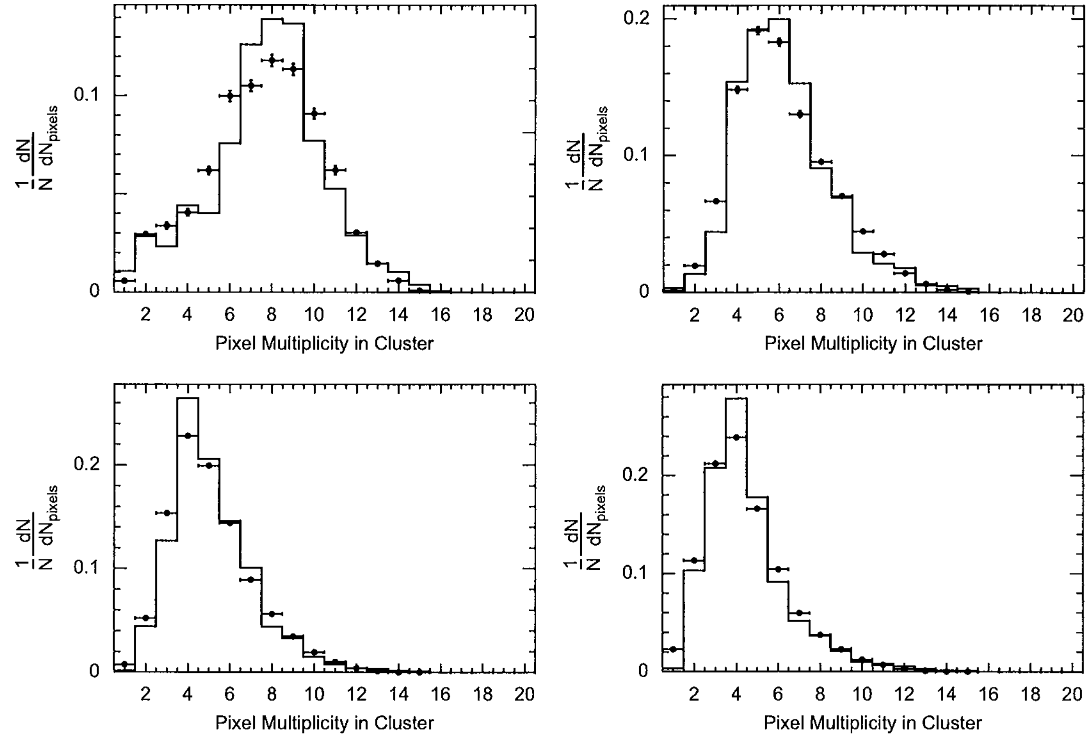


Fig. 3. Number of pixels in single electron clusters at various electron energies: 80 keV (upper left panel), 150 keV (upper right panel), 200 keV (lower left panel) and 300 keV (lower right panel). Points with error bars represent the data and the line the result of the `Geant4+PixelSim` simulation.

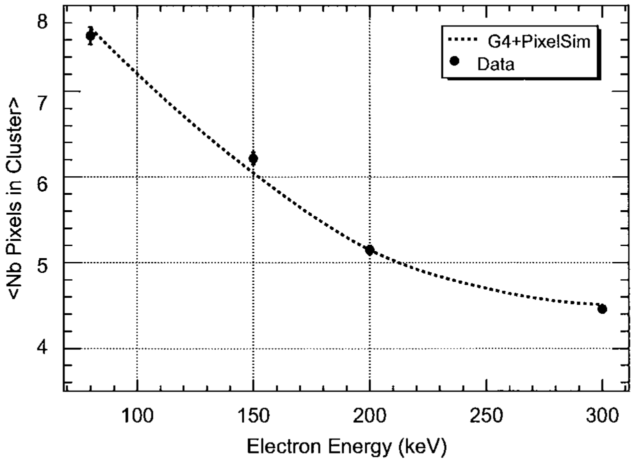


Fig. 4. Average pixel multiplicity vs. electron energy. The points with error bars show the data and the line the result of the `Geant4+PixelSim` simulation.

variation of the charge spread as a function of the electron energy. A clustering algorithm with two thresholds is used to determine the cluster size [12]. First the detector is scanned for "seed" pixels with pulse height values over a S/N threshold set to 3.5. Seeds are sorted according to their pulse heights and the surrounding neighbouring pixels are added to the cluster if their S/N exceeds 2.5. In simulation, the σ_{diff} parameter is tuned to minimise the χ^2 of the pixel multiplicity distribution in the clusters for simulation vs. data at 300 keV, where multiple scattering is lower, as discussed above. The pixel multiplicities in electron clusters at various energies are shown in Fig. 3 and the evolution of the average pixel multiplicity vs. electron energy in Fig. 4. The agreement of the tuned `PixelSim` simulation with data is good and we observe an increase of the pixel multiplicity due to the increased multiple scattering and energy deposition at lower energies.

5. Imaging characterisation

The pixel imaging performance is characterised in terms of three observables: the line spread function (LSF), the modulation transfer function (MTF) and the detection quantum efficiency (DQE). We study these observables for two different imaging regimes, bright field illumination, where each pixel typically receives several electrons per frame, and for cluster imaging, a recently proposed alternative imaging technique [24], where the electron flux is kept low enough that the position of impact of each individual electron is reconstructed by interpolating the charge collected on the pixels of a signal cluster. This allows us to achieve spatial sampling with frequencies much larger than the Nyqvist frequency.

The point spread function is one of the key features for an imaging sensor in transmission electron microscopy. It depends on several parameters of which the most important are pixel size, electron multiple scattering in the active layer and in its vicinity and charge carrier diffusion. For these tests the detector is mounted on a proximity board which is cut below its active area to eliminate back-scattering from the board material. Pixel sensors of both 300 μm- and 50 μm–thickness are tested. A 60 μm–diameter Au wire is mounted parallel to the pixel rows on top of each sensor at a distance of ~2 mm above its surface.

Another important feature is the image contrast, defined by the ratio between the signal on the pixels directly exposed to the beam to the response of those which are covered. The contrast depends on charge leakage, scattering and noise. We measure it using both the wire and a metal foil covering the upper portion of the chip. We expect the contrast to improve for thin sensors, where the contribution of electrons back-scattered in the bulk Si and depositing energy in the shadowed area is smaller.

5.1. *Bright field illumination*

In simulation, a monochromatic, point-like beam of electrons is sent onto the surface of the detector. The line spread function (LSF) is determined as the r.m.s. of the predicted distribution of the detected charge on the pixels. On the data the LSF for bright field illumination is determined using the image projected onto the sensor by the thin Au wire, using the same technique as in Ref. [12]. Since the gold wire has well-defined edges, the profile of the deposited energy in the pixels, measured across the wire, allows us to study the charge spread due to scattering and diffusion along the projected image of the wire edge and compare to simulation. We study the change in the recorded signal, by scanning along pixel rows across the gold wire. We extract the LSF from images such as those in Fig. 5, which show the pulse heights measured on the pixels along a set of rows. In Ref. [12], we parametrised the measured pulse height on pixel rows across the wire image with a box function smeared by a LSF Gaussian term and extracted the LSF by a 1-parameter χ^2 fit with the Gaussian width as free parameter. Here we adopt an extension of this method which uses the sum of several sigmoid functions to describe the edge as proposed in [25]

$$PH = a_0 + \sum_i a_i Erf\left(\frac{x_0 - x}{\sqrt{2}\sigma_i}\right) \tag{1}$$

where the error function is defined as $Erf(x) = (2/\sqrt{\pi})\int_0^x dte^{-t^2}$. We find that one or two sigmoid functions are sufficient to describe the data. We use a single sigmoid function to obtain LSF values of (7.6 ± 0.6) μm at 300 keV and (12.6 ± 0.7) μm at 80 keV, where the quoted uncertainty is statistical. These results are consistent with those obtained from simulation and with the data in our previous study. Since it is customary to quote the imaging resolution in terms of the modulation transfer function (MTF), which is the Fourier transform of the line spread function, we repeat the fit using the sum of two sigmoid functions in Eq. (1). The width of the first, σ_1, is fixed to its value obtained in the previous fit, using a single function, while the width of the second function, σ_2, as well as the two normalisation coefficients a_1 and a_2 are left free. We observe that the χ^2 at the minimum improves only marginally by performing the fit with two functions compared to a single function, as used to extract the LSF. Fig. 6 shows the MTF obtained with this method on data, for 80 and 300 keV electrons.

Next we determine the detection quantum efficiency, a quantity widely used to characterise electron detection systems, introduced in Ref. [26]. We compute the DQE accounting for the charge migration between neighbouring pixels due to charge carrier diffusion following the method adopted in [27,28]

$$DQE \equiv \frac{(S/N)^2_{out}}{(S/N)^2_{in}} = \frac{(S/N)^2_{out}}{m \times N} \tag{2}$$

where N is the number of electrons pixel^{-1} frame^{-1} and m is the mixing factor, given by $m = 1/\sum f_{ij}^2$, with f_{ij} being the fraction of the

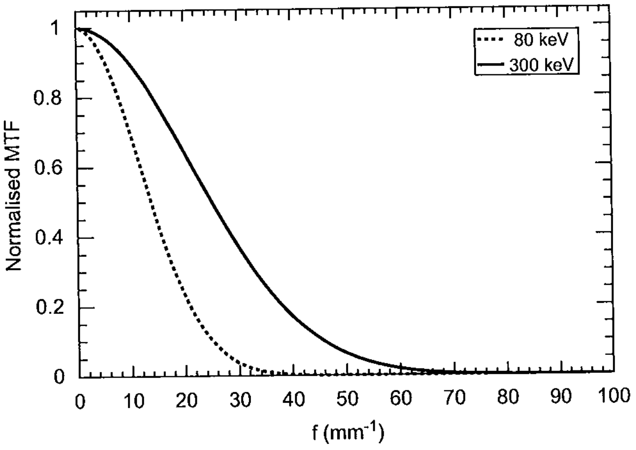


Fig. 6. Normalised modulation transfer function obtained on data for bright field illumination at 80 keV (dotted line) and 300 keV (continuous line).

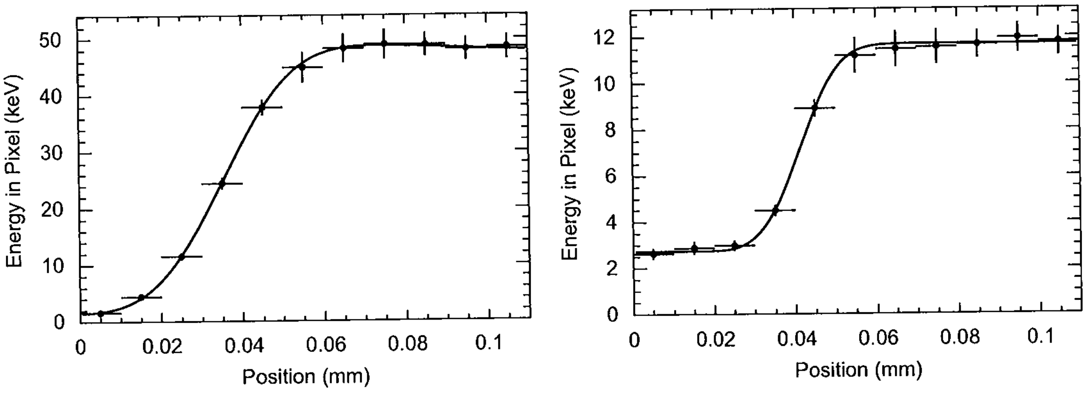


Fig. 5. Pulse heights measured on pixels along a row across the Au wire stretched above the pixels for 80 keV (left panel) and 300 keV (right panel) beam. Data are shown as points with error bars while the continuous line shows the sigmoid function corresponding to the best fit.

charge collected at pixel j for an electron hitting the detector on pixel i. We extract the f_{ij} coefficients by studying the charge distribution in the signal clusters for single electrons, as discussed in Section 4.3. We perform this measurement varying the electron flux and measuring the current on the calibrated phosphor screen of the Titan microscope. We take data at different fluxes in the range 1.3–21 e^- pixel^{-1} frame^{-1} at 300 keV and 0.3–3.1 e^- pixel^{-1} frame^{-1} at 80 keV. First, we determine the number of electrons per pixel and per frame from the pixel single electron response in terms of pulse height and charge spread, discussed in Section 4.3. The left panel of Fig. 7 shows the number of electrons per pixel and frame measured from the pixel response as a function of that from the beam current on the phosphor screen, which exhibits a linearity within $\simeq 10\%$. We measure the DQE at different fluxes using Eq. (2). Results are consistent within the statistical uncertainties (see the right panel of Fig. 7). By averaging these results, we obtain DQE $= 0.74 \pm 0.03$ at 300 keV and 0.78 ± 0.04 at 80 keV, where the quoted uncertainties are statistical.

5.2. Cluster imaging

In bright field illumination the point spread function has a contribution from the lateral charge spread due to charge carrier diffusion in the active detector volume. At high rate, the signal recorded on each individual pixel is the superposition of the charge directly deposited by a particle below the pixel area with that collected from nearby pixels through diffusion, multiple scattering and back-scattering from the bulk Si. If the electron rate is kept low enough so that individual electron clusters can be reconstructed, the position of passage of each electron can be obtained from the centre of gravity of the observed signal charge. The leakage of charge on neighbouring pixels is taken into account through the centre of gravity calculation and the precision depends on the signal-to-noise ratio. This technique, widely adopted in tracking applications for accelerator particle physics, provides us with a significant gain in spatial resolution. In a recent letter to this journal [24] we proposed to adopt the same technique for imaging, with low electron fluxes, i.e. $\leq 10^2\, e^-\, mm^{-2}$ frame^{-1}. We named this technique "cluster imaging". In this regime, the image is formed by adding many subsequent frames. We showed that for cluster imaging the LSF depends only on the detector pixel size and cluster S/N (determining the single point resolution) and on the multiple scattering. A significant improvement in the LSF can be obtained. Here, we repeat the analysis by extracting the LSF from the wire edge images built by the superposition of multiple frames taken with $\simeq 10$–$40\, e^-\, mm^{-2}$ frame^{-1}, using the same sigmoid fit adopted above. First we use a single sigmoid to extract the value of the LSF (see Fig. 8). Results are summarised in Table 1 and Fig. 9, where the values for bright field illumination and cluster imaging measured at different energies are compared. Then, we fix the width of the first function to this fitted value and we introduce a second sigmoid with free width and fit it together with the normalisation coefficients, as discussed above. The LSF function, parametrised according to Eq. (1), is used to extract the MTF. Results are shown in Fig. 10 for 300 and 80 keV electrons.

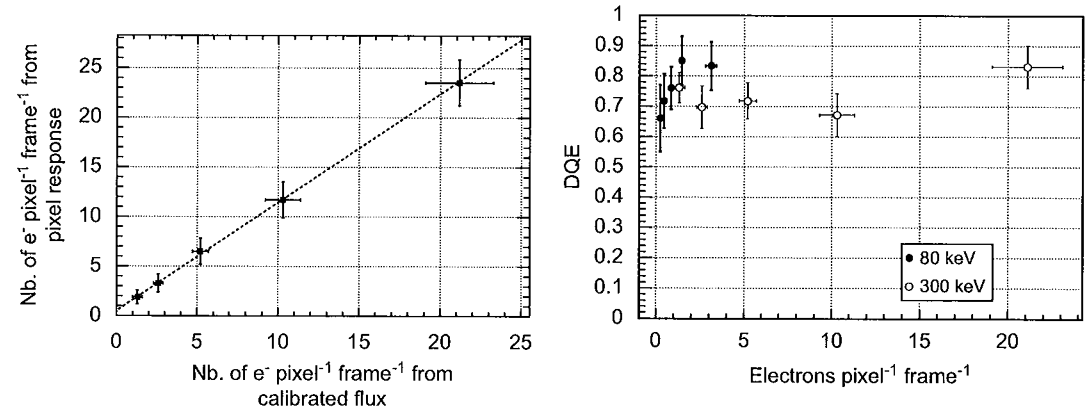


Fig. 7. Reconstructed electron flux from pixel response and DQE: (left) number of electrons per pixel and frame obtained from the measured energy in the pixels after unfolding the contribution of charge spread as a function of the result from the electron flux obtained from the beam current read on the microscope calibrated screen at 300 keV. The fitted line has a slope of 1.09 ± 0.11. (right) Measured DQE as a function of the electron flux at 80 and 300 keV.

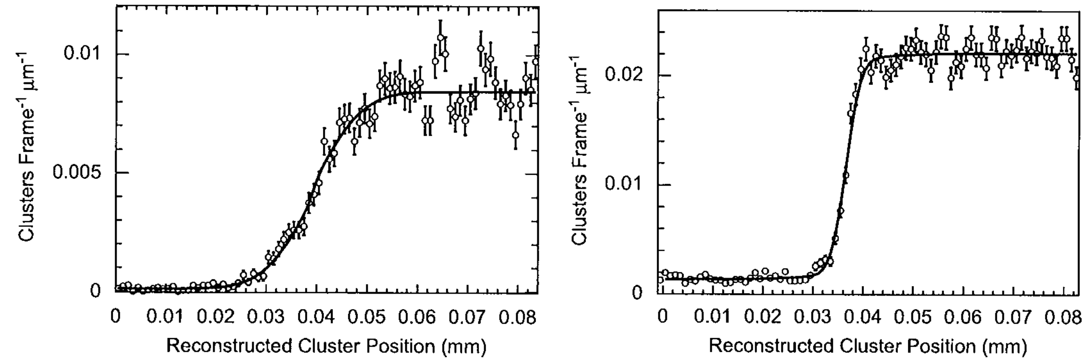


Fig. 8. Pulse heights measured on pixels along a row across the Au wire stretched above the pixels for 80 keV (left panel) 300 keV (right panel) beam. Data are shown as points with error bars while the continuous line shows the sigmoid function corresponding to the best fit.

Table 1
Line spread function values measured with bright field illumination and cluster imaging at various electron energies.

Electron energy (keV)	Bright field LSF (μm)	Cluster imaging LSF (μm)
80	12.1 ± 0.7	6.68 ± 0.34
150	11.2 ± 0.6	4.64 ± 0.27
200	9.4 ± 0.6	3.53 ± 0.16
300	7.4 ± 0.6	2.35 ± 0.15

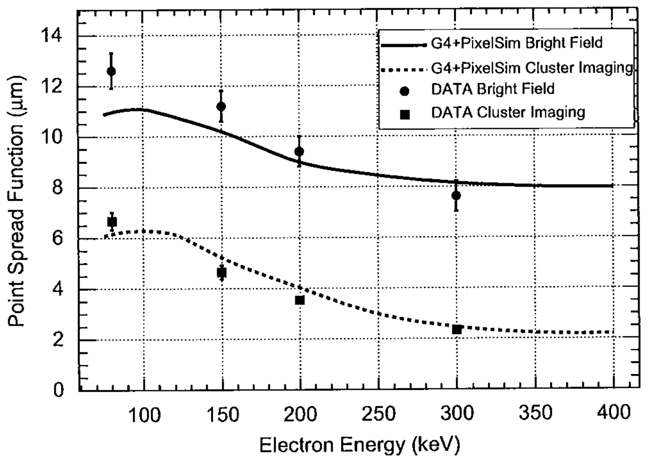


Fig. 9. Fitted line spread function as a function of beam energy. Cluster imaging (squares with error bars) and bright field illumination (circles with error bars) data are compared to `Geant4`+`PixelSim` simulation.

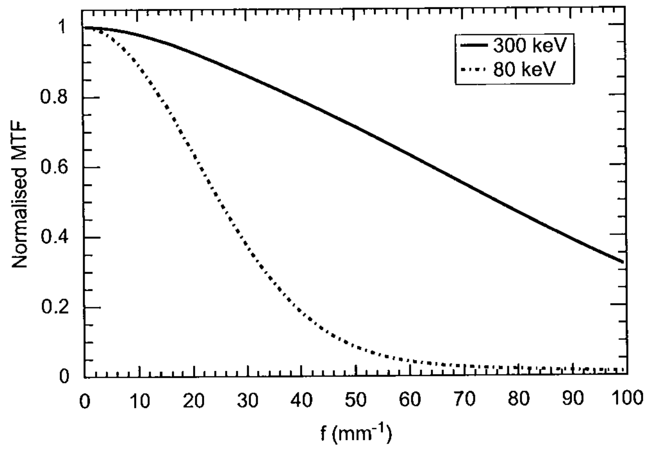


Fig. 10. Normalised modulation transfer function obtained on data for cluster imaging illumination at 80 keV (dotted line) and 300 keV (continuous line).

5.3. *Image contrast and imaging tests*

We study the contrast ratio obtained for a 50 μm–thick sensor and a sensor which has the original 300 μm Si thickness, by comparing the signal observed on the pixel area covered by the wire to that obtained on the pixels directly exposed to the beam. We measure a ratio of 2.95 for the 300 μm–thick and 4.49 for the 50 μm–thick sensor, using bright field illumination. The contrast is affected by multiple scattering effects in the detector and the thinned sensor performs better due to the reduced back-scattering in the bulk Si underneath the sensitive epitaxial layer. The contrast decreases at lower energies, we measure 3.9 on the thin sensor at 150 keV, until the range of the electrons becomes short enough that the scattering contribution to the energy leaking under the shadow of the wire becomes small and a contrast ratio in excess of 20 is measured at 80 keV. Then, we study the pixel response at the edge of a metal plate on the thick and the thinned sensor. The plate covers an area where pixels do not receive direct electron hits. Fig. 11 shows the pulse height measured on pixels along rows across the plate edge. Beyond the sharp edge we observe that the recorded signal in the thick sensor falls less rapidly compared to that in the thin sensor, which we interpret as an effect of charge deposited by back-scattered electrons. We also measure the LSF by fitting a single sigmoid function to the pixel response. We obtain (10.4 ± 0.5) and (7.7 ± 0.4) μm for the 300 μm- and the 50 μm–thick sensor, respectively. The latter result is compatible to that obtained with the thin Au wire. Finally, we perform an imaging test using a Si sample on the TEAM1 microscope with 300 keV electrons. Fig. 12 gives a qualitative demonstration of the performance of the TEAM1k detector. The image shows a single 2.6 ms exposure revealing the Si dumbbells close to the ⟨1 1 0⟩ zone axis, with a spacing of 1.36 Å.

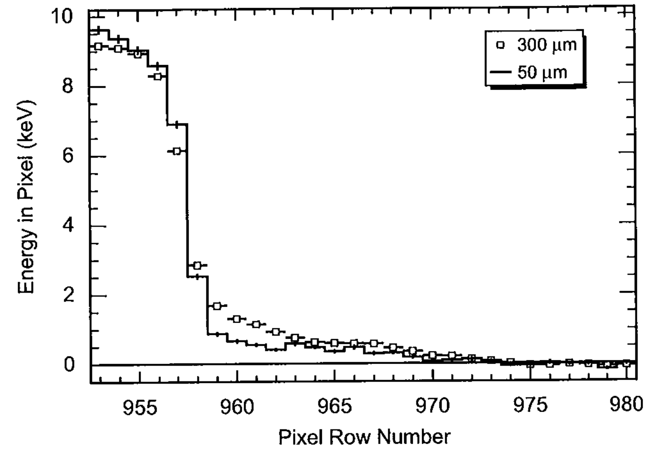


Fig. 11. Pulse heights measured on pixels along rows across a metal knife edge for 300 keV beam with a 300 μm (squares with error bars) and a 50 μm (line with error bars) thick TEAM1k sensor. The enhanced leakage of charge in the pixels below the metal plate due to electron scattering in the thick sensor is evident.

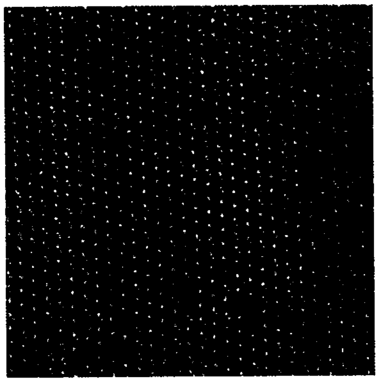


Fig. 12. Raw image of a single 2.6 ms exposure of a Si sample in ⟨1 1 0⟩ orientation taken with the TEAM1k chip and a 300 keV beam at the TEAM1 microscope. The image has no corrections applied, which gives an impression of the overall uniformity across the detector.

6. Sensor irradiation

The radiation tolerance of the chip has been assessed by irradiating a TEAM1k sensor with 300 keV electrons, up to a dose of 5 Mrad. The damage mechanism is an increase of leakage current, which not only increases noise, but also decreases the dynamic range. CMOS active pixel sensors are capable of good S/N because the charge collection node capacitance is small, resulting in a high conversion gain (V/e$^-$) and thus high sensitivity to leakage current. The irradiation has been performed in subsequent steps with a flux of 875 e$^-$ s^{-1} μm^{-2}, corresponding to a dose rate of 250 rad s^{-1}, from the measured beam current on a known beam spot area onto the Titan calibrated phosphor screen. In-between consecutive irradiation steps, several dark frames are acquired without beam, in order to follow the evolution of the pixel leakage current with the dose, and with very low intensity beam, to monitor the pixel response to single electrons and the gain calibration. A small portion of the sensor active surface is covered with a gold wire, as described in Section 5. Bright field images of this wire are acquired throughout the irradiation in order to monitor also the sensor imaging performance. All tests are performed with the detector cooled at +5 °C. The sensor is glued with thermally conductive epoxy to an AlN substrate board, and wire-bonded to a flex circuit glued on top of the AlN board. Heat is removed from the sensor by a copper finger that contacts the AlN board and is cooled by a double-Peltier system.

Fig. 13 shows the evolution of the pixel leakage current with dose. The leakage current is determined by comparing the pixel base level at two different clock frequencies, 25 and 6.25 MHz, corresponding to integration times of 2.6 and 10.5 ms, respectively. A sub-linear increase of the current with dose is observed, reaching about 0.4 pA after 5 Mrad. The leakage current increase is due to ionising damage in the field oxide that leads to trapping of positive charge. The inversion of the Si interface results in an increased current in the charge collecting diode. Three hundred keV electrons are not expected to significantly affect the silicon bulk through displacement damage.

Results of the irradiation test of an earlier prototype chip with a comparable pixel cell manufactured in the same process gave a leakage current of 0.2 pA after a dose of 1.11 Mrad with 200 keV electrons on 20 × 20 μm^2 pixels. The test was performed at room temperature and for an integration time of 737 μs. Considering that the integration time of the TEAM1k chip is 14 times longer, the beneficial effect of cooling on the sensor radiation tolerance is evident.

From a device operation point of view, the increase of the leakage current results in an increase of the pixel base level. This is removed by subtraction of the pedestal level. However, the leakage current increase causes a decrease of the dynamic range and an increase of the single pixel noise. At 25 MHz clock frequency, after 5 Mrad of dose we observe a decrease of the pixel dynamic range by about 30%, which still allows proper operation of the chip. Table 2 summarises the pixel charge-to-voltage gain calibration and noise as a function of the dose, for a pixel clock frequency of 25 MHz. The gain calibration is estimated from the position of the most probable value of the Landau distribution for single electrons measured at low flux after each irradiation step. A slight decrease of the pixel gain is observed with the increasing dose, while the noise increases by about 25%. The single electron cluster pulse height is not significantly affected after irradiation (see Fig. 14) and the average cluster signal-to-noise ratio (S/N) measured for single 300 keV electron detection changes from 12.3 before irradiation to 10.9 after 5 Mrad of dose. Finally, we check the imaging properties of the detector by determining the LSF after irradiation. We repeat the fit to the measured pulse height on pixel rows across an image of the stretched gold wire. By performing a single sigmoid fit, as discussed in Section 5.1, we obtain a LSF value of (7.2 ± 0.6) μm at 300 keV, which is consistent with the value of (7.4 ± 0.6) μm obtained before irradiation.

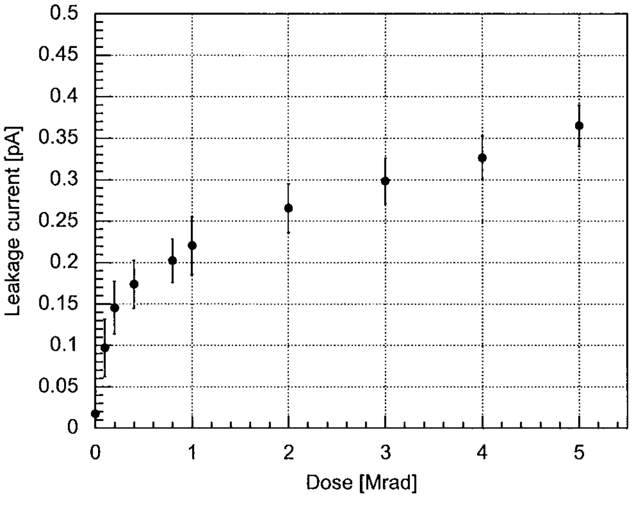


Fig. 13. Leakage current as a function of the 300 keV electron dose.

Table 2
Charge-to-voltage conversion gain and pixel noise as a function of the delivered 300 keV electron dose.

Dose (Mrad)	**Calibration** (e$^-$/ADC)	**Noise** (e$^-$ ENC)
0	11.2	56 ± 10
1	8.0	60 ± 7
2	8.2	75 ± 7
3	7.3	70 ± 6
4	6.7	67 ± 6
5	7.3	69 ± 6

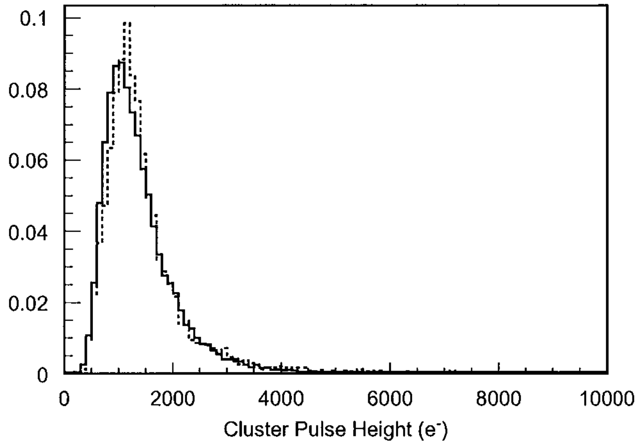


Fig. 14. Single electron cluster signal (in e$^-$ units) for 300 keV electrons before (continuous line) and after 5 Mrad of dose (dashed line) and recalibration. The comparison of the width of the two distribution shows that the cluster noise after irradiation is comparable to that of the unirradiated detector.

7. Conclusions

A direct detection detector which meets the requirements for TEAM has been developed and characterised using 80–300 keV energy electrons. We measure a line spread function of (12.1 ± 0.7) to (7.4 ± 0.6) μm for $80 \leq E_e \leq 300$ keV with bright field illumination and (6.7 ± 0.3) to (2.4 ± 0.2) μm with cluster imaging and a DQE = 0.78 ± 0.04 and 0.74 ± 0.03 at the two ends of the energy range explored. The imaging performances of the detector are identical after irradiation with 300 keV electrons up to a dose of 5 Mrad, while the dynamic range is reduced by $\simeq 30\%$ due to leakage current increase. The good agreement obtained in pixel response and line spread function between measurements and simulation demonstrates that the tools to understand the detailed performance of the detector are in hand. R&D on new detectors with enhanced performances is currently under way.

Acknowledgements

We wish to thank N. Andresen and R. Erni who contributed to the detector installation on the electron microscope. We are grateful to the INFN Group of Padova, Italy, for their support with the DAQ system, in particular to D. Bisello, D. Pantano and M. Tessaro. This work was supported by the Director, Office of Science, of the U.S. Department of Energy under Contract no. DE-AC02-05CH11231. The TEAM Project is supported by the Department of Energy, Office of Science, Basic Energy Sciences.

References

[1] U. Dahmen, Microsc. Microanal. 13 (2) (2007).
[2] C. Kisielowski, R. Erni, B. Freitag, Microsc. Microanal. 14 (2) (2008) 78.
[3] V.E. Cosslett, et al., Nature 281 (1979) 49.
[4] C.O. Girit, et al., Science 323 (2009) 1705.
[5] A.-C. Milazzo, et al., Ultramicroscopy 104 (2005) 152.
[6] G. Deptuch, et al., Ultramicroscopy 107 (2007) 674.
[7] P. Denes, J.-M. Bussat, Z. Lee, V. Radmilovic, Nucl. Instr. and Meth. A 579 (2007) 891.
[8] G.Y. Fan, et al., Ultramicroscopy 70 (1998) 113.
[9] A.R. Faruqi, D.M. Cattermole, C. Reburn, Nucl. Instr. and Meth. A 513 (2003) 317.
[10] A.R. Faruqi, et al., Ultramicroscopy 94 (2003) 263.
[11] A.R. Faruqi, R. Henderson, Curr. Opin. Struct. Biol. 17 (2007) 549.
[12] M. Battaglia, et al., Nucl. Instr. and Meth. A 598 (2009) 642 [arXiv:0811.2833 [physics.ins-det]].
[13] M. Battaglia, et al., Nucl. Instr. and Meth. A 605 (2009) 350 [arXiv:0904.0552[physics.ins-det]].
[14] M. Battaglia, et al., Nucl. Instr. and Meth. A 579 (2007) 675 [arXiv:physics/0611081].
[15] Aptek Industries, San Jose, CA 95111, USA.
[16] S. Agostinelli, et al., Nucl. Instr. and Meth. A 506 (2003) 250.
[17] S. Chauvie, et al., Prepared for CHEP'01: Computing in High-Energy Physics and Nuclear, Beijing, China, 3–7 September 2001.
[18] M. Battaglia, Nucl. Instr. and Meth. A 572 (2007) 274.
[19] F. Gaede, Nucl. Instr. and Meth. A 559 (2006) 177.
[20] M. Battaglia, et al., Nucl. Instr. and Meth. A 611 (2009) 105.
[21] Manufactured by Avnet Inc., Phoenix, AZ 85034 USA.
[22] R. Brun, F. Rademakers, Nucl. Instr. and Meth. A 389 (1997) 81.
[23] F. Gaede, et al., in: Proceedings of Interlaken 2004, Computing in High Energy Physics and Nuclear Physics, Report CERN 2005-002, 471.
[24] M. Battaglia, D. Contarato, P. Denes, P. Giubilato, Nucl. Instr. and Meth. A 608 (2009) 363 [arXiv:0907.3809[physics.ins-det]].
[25] A.L. Weickenmeier, W. Nüchter, J. Mayer, Optik 99 (4) (1995) 147.
[26] K.-H. Herrmann, D. Krahl, Adv. Opt. Electr. Microsc. 9 (1984) 1.
[27] K. Ishizuka, Ultramicroscopy 52 (1993) 7.
[28] M. Horacek, Rev. Sci. Instr. 76 (2005) 093704.

Nuclear Instruments and Methods in Physics Research A 624 (2010) 425–427

Contents lists available at ScienceDirect

Nuclear Instruments and Methods in Physics Research A

journal homepage: www.elsevier.com/locate/nima

Radiation hardness studies on CMOS monolithic pixel sensors

Marco Battaglia [a,b], Dario Bisello [c], Devis Contarato [b,*], Peter Denes [b], Dionisio Doering [b], Piero Giubilato [b,c], Tae Sung Kim [b], Serena Mattiazzo [c], Velimir Radmilovic [b], Sarah Zalusky [a,b]

[a] *Department of Physics, University of California at Berkeley, CA 94720, USA*
[b] *Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA*
[c] *Dipartimento di Fisica, Università di Padova and INFN, Sezione di Padova, I-35131 Padova, Italy*

ARTICLE INFO

Available online 2 April 2010

Keywords:
Monolithic active pixel sensors
Silicon pixel detectors
Radiation tolerance

ABSTRACT

This paper presents irradiation studies performed on a CMOS monolithic pixel sensor prototype implementing different optimizations of the pixel cell aimed at a superior radiation tolerance. Irradiations with 200 keV electrons up to a total dose of 1.1 Mrad have been performed in view of the utilization of such a design in Transmission Electron Microscopy (TEM) applications. Comparative irradiations were performed with 29 MeV protons up to a 2 Mrad total dose and with 1–14 MeV neutrons up to fluences in excess of $10^{13}\,n_{eq}\,cm^{-2}$. Experimental results show an improved performance of pixels designed with Enclosed Layout Transistor (ELT) rules and an optimized layout of the charge collecting diodes.

1. Introduction

CMOS monolithic pixel sensors have been shown in the past decade to be an attractive option for applications ranging from high resolution particle tracking [1] to fast imaging in Transmission Electron Microscopy (TEM) [2,3] and beam monitoring. In order to withstand the large radiation fluxes received by the sensors in these applications, an optimized design is needed with respect to standard devices.

In High-Energy Physics (HEP), CMOS monolithic pixels are a favored option for future vertex trackers at e^+e^- colliders due to their higher spatial resolution and much reduced material budget compared to hybrid pixel sensors. The ionising dose at the position of the innermost vertex detector layer is expected to be $\sim 50\,krad\,yr^{-1}$ at a high energy linear collider and $\sim 0.5-1\,Mrad\,yr^{-1}$ at a high luminosity B-factory, while the anticipated yearly non-ionising fluences are in the range 10^{11}–10^{12} n_{eq} $cm^{-2}yr^{-1}$. In TEM applications, CMOS pixels achieve single electron sensitivity through direct detection, while the thin collection layer, limited to the typically $10-15\,\mu m$ thick epitaxial layer, ensures an excellent point spread function. Moreover, a high readout speed can be achieved, thus favoring fast dynamic imaging. Considering the typical yearly usage of such a detector in TEM, a tolerance to ionizing doses in excess of 1 Mrad is desirable. Furthermore, radiation hardness to doses up to several Mrad may be required for beam imaging applications, such as real-time monitoring of beam position and profile in hadrontherapy.

A CMOS monolithic pixel prototype, the LDRD2-RH chip described in Section 2, has been designed exploring several layout optimizations of the pixel cell aimed at a tolerance to a few Mrad ionising dose and to non-ionising fluences of order $10^{13}\,n_{eq}$ cm^{-2}. Its design is discussed in detail in Ref. [4]. The sensor was originally intended as the prototype for a large scale CMOS imager to be deployed in TEM applications. The response of the LDRD2-RH chip to electrons in the energy range of interest for TEM (80–300 keV) has been extensively characterized in terms of energy deposition, charge spread and point spread function in Refs. [4,5]. This paper will report on comparative irradiation studies performed with electrons, protons and neutrons in order to evaluate the performance of the different layout options (Section 3).

2. Sensor design and experimental setup

The LDRD2-RH chip (Fig. 1) was manufactured in AMS $0.35\,\mu m$ CMOS-OPTO technology with an epitaxial layer of $14\,\mu m$ nominal thickness, and features an array of 96×96 pixels on a $20\,\mu m$ pitch. The sensor, designed as a variation of a previous non-radiation-hard design [6], is subdivided in several sectors with different layout optimizations of the charge collecting diode. Pixels implement a simple 3-transistor (3T) analog architecture, designed both with and without Enclosed Layout Transistor (ELT) rules for comparison. The chip is read out in a rolling-shutter mode, ensuring a constant integration time across the pixel

---

* Corresponding author.
E-mail address: DContarato@lbl.gov (D. Contarato).

0168-9002/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2010.03.156

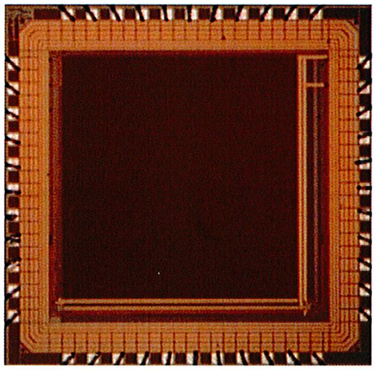


Fig. 1. Picture of the LDRD2-RH chip.

matrix. Pixels are clocked up to a maximum frequency of 25 MHz, corresponding to an integration time of 184 μs, and are read out through a custom FPGA-driven acquisition board equipped with a set of 14 bit, 40 MSample s^{-1} ADCs. The DAQ is connected to a National Instrument digital acquisition board installed on a control PC, and data are acquired and monitored by a LabView-based custom software.

3. Irradiation studies

The radiation hardness of the LDRD2-RH chip has been assessed by comparing the sensor response before and after irradiation with 200 keV electrons, 29 MeV protons and $\simeq 1-14$ MeV neutrons. Pixel noise, leakage current and charge collection properties have been studied for the different irradiations. All tests have been performed at room temperature.

In CMOS sensors, ionising radiation leads to trapping of positive charge in the field oxide, resulting in inversion of the silicon at the interface and thus in an increased leakage current in the charge collecting diode. Non-ionising energy loss (NIEL) can displace silicon bulk atoms and create defects in the detector sensitive volume; this is true especially in the vicinity of the diode where the silicon is depleted, and can lead to a bulk current that adds up to the diode leakage current. The threshold energy for electrons to cause displacement damage in silicon is 260 keV [7], and 200 keV electrons are thus expected to create only ionising damage and to not affect the silicon bulk. On the other hand, 29 MeV protons damage the sensor via both ionisation and NIEL, while neutrons are expected to affect the silicon only via bulk damage.

The irradiation with 200 keV electrons has been performed at the LBNL National Center for Electron Microscopy (NCEM). The sensor has been irradiated in steps up to a total dose of 1.11 Mrad, and the pixel leakage current has been monitored after each step as reported in Ref. [4]. The irradiation with 29 MeV protons was performed at the BASE facility of the LBNL 88-in. Cyclotron [8]. A sensor was irradiated in steps up to a total integrated fluence of 8.5×10^{12} p cm^{-2}, corresponding to a total dose of 1.98 Mrad. A flux of $\sim 3 \times 10^{8}$ p cm^{-2} s^{-1} was employed throughout the irradiation. Similarly to the electron irradiation, the pixel dark level was recorded at the end of each irradiation step, in order to monitor the variation of leakage current and pixel noise as a function of the delivered dose.

Fig. 2 compares the results from the 200 keV electron and the 29 MeV proton irradiation for pixels designed with ELT rules and three layout options (labelled as A–C) for the charge collecting diodes, implemented in different sectors of the chip. All three layouts implement a p^+ guard-ring around the charge-collecting diode. Layout A has thick field oxide between the diode implant and the guard-ring. Layouts B and C have thin oxide between the diode implant and the guard-ring, with layout B having a shallower n^+ diode implant [4]. In both irradiations, the best performance among all ELT designs is obtained in pixels designed with layout option C, which shows only a moderate increase of the leakage current; a thinner oxide limits the build-up of positive charge trapped in the oxide and is thus beneficial in limiting the pixel leakage. At equal doses, a larger leakage current is associated with proton irradiation, hinting at a possible contribution from displacement damage.

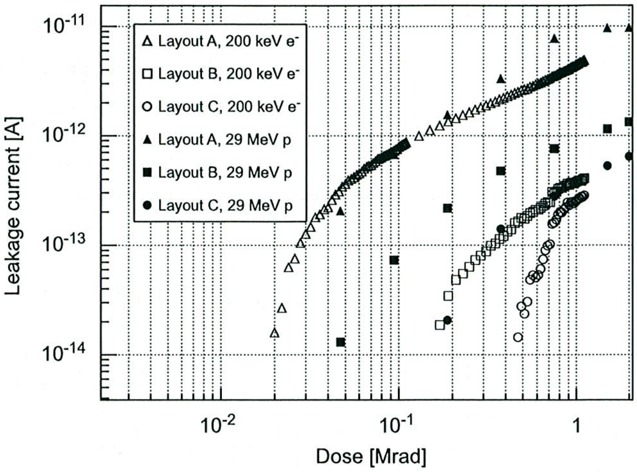


Fig. 2. Pixel leakage current as a function of the 200 keV electron and 29 MeV proton dose for pixels with ELT transistor design and different layouts of the charge collecting diode. The data referring to electron irradiation is taken from [4].

Neutron irradiation has been performed at the LBNL 88-in. Cyclotron by exposing one LDRD2-RH sensor to neutrons produced from 20 MeV deuteron breakup on a thin Be target [9]. The detector was located 8 cm downstream from the target and activation foils were placed just behind it to monitor the delivered fluence. Deuteron breakup produces neutrons on a continuous spectrum from ~ 1 MeV up to ~ 14 MeV. A beam current of 800 nA was used, corresponding to an estimated flux at the chip position of $\simeq 4 \times 10^{8}$ n cm^{-2} s^{-1}. The total fluence was measured from the γ activity of activation foils to be in excess of 1.2×10^{13} n cm^{-2}. After irradiation, no significant change in the sensor leakage current or noise could be observed.

The detector charge-to-voltage conversion gain and its noise in equivalent noise charge (ENC) have been determined by exposing the sensor to a ^{55}Fe source and by reconstructing the position of the 5.9 keV X-ray peak before and after irradiation. Results obtained on the pixels designed with layout C after the three irradiation experiments described above are summarised in Table 1. We here note that the pre-irradiation noise values for the electron and proton irradiations are significantly higher compared to the neutron case because of the longer interconnection cables between the detector board and the DAQ system required in the first two experiments. No significant change in pixel noise is observed, while a 35% decrease in gain is observed after both electron and proton irradiation. A similar decrease in gain of $\simeq 1.35$ was observed on the electron irradiated sensor from the test of the pixel response to charged particles (200 keV and 1.5 GeV electrons). As reported in Ref. [4], after correcting for the gain change, the noise and energy deposition distributions were in good agreement with the pre-irradiation ones, indicating that the pixel charge collection properties have not been affected, and that the change in gain might be due to

Table 1
Charge-to-voltage conversion gain and pixel noise measured before and after the three irradiations in pixels implementing diodes designed with layout option C and ELT transistors.

	Before irradiation	After irradiation
e^- irradiation		
Pixel noise	(130 ± 6) e^- ENC	(122 ± 6) e^- ENC
Calibration	(26.7 ± 0.6) e^-/ADC	(36.3 ± 0.8) e^-/ADC
p Irradiation		
Pixel noise	(128 ± 5) e^- ENC	(126 ± 5) e^- ENC
Calibration	(25.9 ± 0.5) e^-/ADC	(34.9 ± 0.9) e^-/ADC
n Irradiation		
Pixel noise	(69 ± 3) e^- ENC	(64 ± 5) e^- ENC
Calibration	(25.2 ± 0.3) e^-/ADC	(24.4 ± 0.4) e^-/ADC

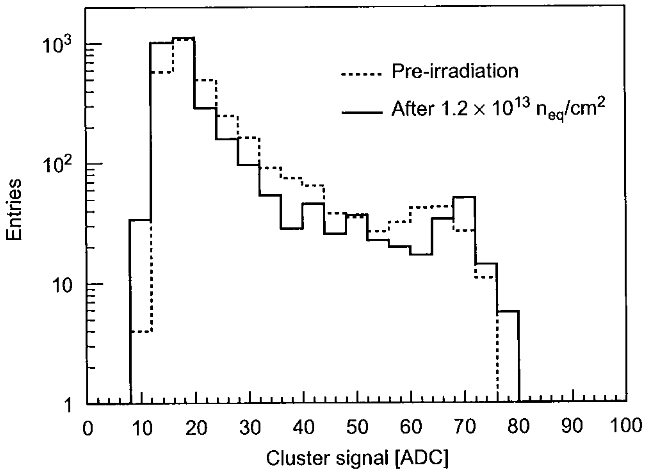


Fig. 3. Single pixel spectrum obtained with ^{55}Fe on the pixels designed with layout C and ELT transistors before and after neutron irradiation. The position of the 5.9 keV peak is visible at 65–70 ADC counts.

radiation effects on the detector output electronics, which is designed with a standard layout and is thus more sensitive to (ionizing) radiation than the transistors in the ELT pixel cells.

Fig. 3 shows the ^{55}Fe spectrum in the pixel sector equipped with layout C diodes and ELT transistors before and after neutron irradiation, while Fig. 4 shows the corresponding single pixel noise distributions. Neutron irradiation leaves the sensor gain basically unchanged, suggesting that the effect of bulk damage is still negligible at the fluences attained, and that the increase of leakage current observed after proton irradiation is mainly due to ionisation damage. However, we note here that the hardness factor of 29 MeV protons with respect to 1 MeV neutrons is 2.346 [10], yielding a maximum equivalent fluence for protons of 2×10^{13} $n_{eq}\,cm^{-2}$, almost a factor two larger than the total fluence achieved with the neutron irradiation.

4. Conclusions

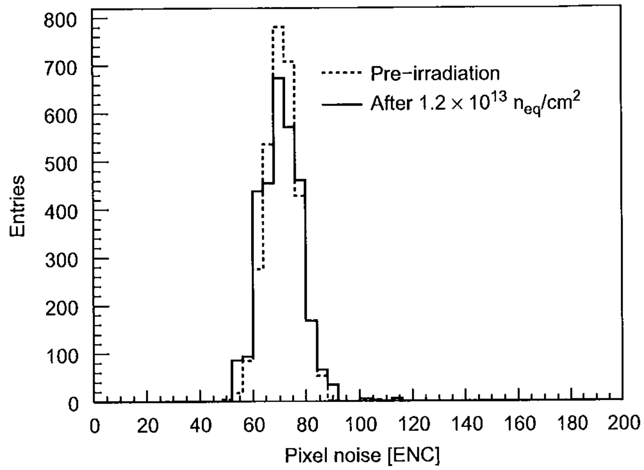


Fig. 4. Pixel noise distribution for the pixels designed with layout C and ELT transistors before and after neutron irradiation.

A CMOS monolithic pixel sensor with an optimized design of the pixel cell for improved radiation tolerance has been designed and tested. The sensor radiation tolerance has been assessed against 200 keV electrons up to a total dose in excess of 1 Mrad, 29 MeV protons up to a total dose of 2 Mrad and with 1–14 MeV neutrons up to equivalent fluences in excess of 1×10^{13} $n_{eq}\,cm^{-2}$. After electron and proton irradiation, only a moderate increase in the pixel leakage current was measured for pixels implementing transistors designed with ELT rules and an adequate layout of the charge collecting diode. At equal ionizing doses, a larger leakage current was found associated with proton irradiation. No significant change in leakage current or noise could be observed after neutron irradiation, suggesting that the effect of bulk damage is still negligible at the fluences considered.

The best performing pixel layout has been used in the design of a large scale prototype, the TEAM1K detector, featuring 1024×1024 pixels of 10 μm pitch and multiple parallel outputs for high frame rate. The chip, also manufactured in the AMS 0.35 μm CMOS-OPTO process, is currently under test and will be deployed as the demonstrator of a high-resolution, fast and radiation-hard CMOS pixel imager for the TEAM electron microscope at the LBNL National Center for Electron Microscopy.

Acknowledgements

This work was supported by the Director, Office of Science of the US Department of Energy under Contract no. DE-AC02-05CH11231. The authors would like to thank the staff of the LBNL National Center for Electron Microscopy and 88-in. Cyclotron for the assistance in the irradiation experiments.

References

[1] R. Turchetta, et al., Nucl. Instr. and Meth. A 458 (2001) 677.
[2] G. Deptuch, et al., Ultramicroscopy 107 (2007) 674.
[3] P. Denes, et al., Nucl. Instr. and Meth. A 579 (2007) 891.
[4] M. Battaglia, et al., Nucl. Instr. and Meth. A 598 (2009) 642.
[5] M. Battaglia, et al., Nucl. Instr. and Meth. A 605 (2009) 350.
[6] M. Battaglia, et al., IEEE Trans. Nucl. Sci. NS-55 (2008) 3746.
[7] G. Lindström, Nucl. Instr. and Meth. A 512 (2003) 30.
[8] M.A. McMahan, Nucl. Instr. and Meth. B 241 (2005) 409.
[9] D.L. Bleuel, et al., Nucl. Instr. and Meth. B 261 (2007) 974.
[10] A. Vasilescu, G. Lindström, Displacement damage in silicon, online compilation available at 〈http://sesam.desy.de/members/gunnar/Si-dfuncs.html〉.

REVIEW OF SCIENTIFIC INSTRUMENTS **80**, 083302 (2009)

A fast, direct x-ray detection charge-coupled device

P. Denes,[1] D. Doering,[1] H. A. Padmore,[2] J.-P. Walder,[1] and J. Weizeorick[3]
[1]*Engineering Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Rd., Berkeley, California 94720, USA*
[2]*Advanced Light Source Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Rd., Berkeley, California 94720, USA*
[3]*X-ray Science Division, Advanced Photon Source, Argonne National Laboratory, 9700 South Cass Ave., Argonne, Illinois 60439, USA*

(Received 14 June 2009; accepted 7 July 2009; published online 6 August 2009)

A charge-coupled device (CCD) capable of 200 Mpixels/s readout has been designed and fabricated on thick, high-resistivity silicon. The CCDs, up to 600 μm thick, are fully depleted, ensuring good infrared to x-ray detection efficiency, together with a small point spread function. High readout speed, with good analog performance, is obtained by the use of a large number of parallel output ports. A set of companion 16-channel custom readout integrated circuits, capable of 15 bits of dynamic range, is used to read out the CCD. A gate array-controlled back end data acquisition system frames and transfers images, as well as provides the CCD clocks. *© 2009 American Institute of Physics.* [DOI: 10.1063/1.3187222]

I. INTRODUCTION

Over the past 20 years, the brightness of synchrotron radiation sources has increased by six orders of magnitude, whereas the detectors used in experiments are in many cases unchanged. One of the ubiquitous detectors in synchrotron radiation research is a fiber-coupled phosphor screen, read out by a charge-coupled device (CCD). CCDs have been the scientific imager of choice for more than 3 decades due to their noiseless, efficient charge transfer which allows linear recording of signals over a wide dynamic range. The primary drawback of CCDs is the serial nature of their charge transfer mechanism that tends to be slow.

Our goal was to improve this kind of detector by increasing the CCD readout speed by a factor of 100 while maintaining all of the other excellent characteristics of CCDs. At the same time, fabricating the CCD on thick, high-resistivity silicon opens up the possibility of direct x-ray detection (the CCD is thick enough to absorb essentially all x rays up to 10 keV) as well as improve the point spread function (PSF) (as the CCD is fully depleted, charge is collected solely by drift, without diffusion).

We describe below the design and fabrication of a 96 port CCD on thick, high-resistivity silicon, as well as a companion 16-channel custom readout integrated circuit and the data acquisition system. First results from tests with x rays are presented.

II. CCD DESIGN

Conventional CCDs achieve higher readout speed with increased clock rates and higher digitization frequency. The speed increase results in a corresponding increase in bandwidth, so that the readout noise increases as $f^{1/2}$. With fixed well depth, this results in an $f^{1/2}$ decrease in dynamic range. An alternate way to increase speed is to increase the number of readout ports. In a single port CCD, the image matrix is surrounded by the parallel clock distribution interconnect on two sides and by the output (serial) shift register with a readout port on one side (see Fig. 1). This is commonly extended to two ports by splitting the serial shift register, adding another output port to the other side, and reading half the CCD out of one port and the other half out of the other port. Duplicating this arrangement on the top of the CCD and splitting the parallel clocking directions results in a four port device. One way to further increase the number of readout ports would be to read out every column [column-parallel CCD (Ref. 1)]. When the pixel pitch is narrower than the width of the output stage, this introduces interconnection problems which are insurmountable in most CCD processes. Alternately, one could have many short serial shift registers [almost column-parallel CCD (Ref. 2)], but this requires finding a way to intersperse the output stages into the area occupied by the serial shift register.

The approach we take is the latter: an output stage for every ten columns. In order to fit the output stage in the area of the serial shift register, each ten-column serial shift register is compressed by a short (6 pixels) pitch-adapting taper (see Fig. 2). The pixels in the taper are designed to have the same capacitance (and therefore the same properties) as pixels in the imaging matrix. For reasons described below, the pixel pitch is 30 μm, so that the output pitch is 300 μm. The output stages share a common drain connection for every four outputs, so that together with a V_{DD} pad, the resulting 240 μm pad pitch allows simple wire bonding. In conventional scientific CCDs, reset and output V_{DD} pads are not shared in order to minimize cross-talk, whereas for the high density of this CCD, individual power pads would lead to a prohibitive number of power supplies.

The CCD was fabricated in a process originally developed at LBNL,[3] where a conventional CCD structure is grown on a high-resistivity silicon wafer. The high resistivity allows the complete volume of the imaging matrix to be

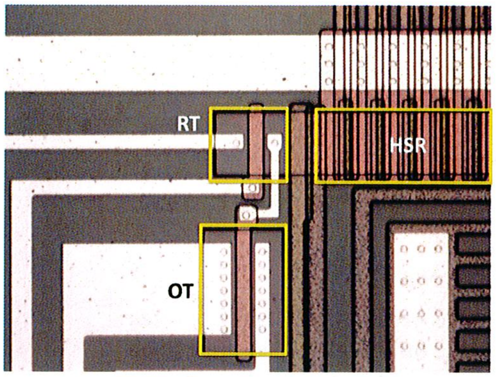


FIG. 1. (Color online) Microphotograph of the output stage of a typical three-phase CCD. The serial shift register is labeled HSR, and the reset and output transistors are labeled RT and OT, respectively.

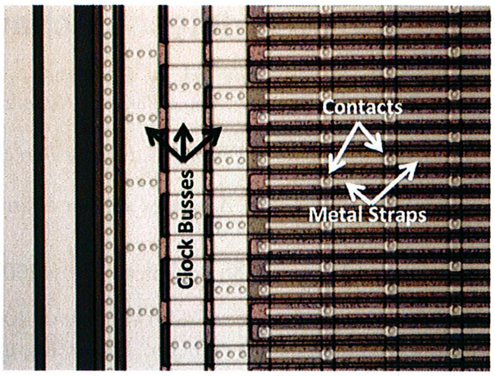


FIG. 3. (Color online) Microphotograph of metal-strapped CCD gates. Each polysilicon gate is covered with metal and contacted (over the channel stop) every other pixel.

depleted so that all of the charge is collected and diffusion is minimized. In addition, the entrance window (backside) is specially designed[4] to present a minimum of dead material while maintaining good conductivity. This allows the same device to simultaneously be optically sensitive (with excellent blue and red quantum efficiencies), sensitive to x rays from the VUV to ~10 keV, and sensitive to low energy electrons.

The time needed to read out a CCD with N_XN_Y pixels having m output ports on the top and bottom each is $T = (N_Y/2)[T_V + (N_X/m)T_H]$, where T_V is the parallel clock time and T_H is the time to serially shift one column and digitize the output. In a three-phase CCD (like this one) three parallel clocks must be toggled during time T_V, so that T_V should be a small fraction of $(N_X/m)T_H$ in order to not introduce significant dead time. Unlike the serial (mini)shift registers, the parallel clock gates traverse the width of the CCD. The polysilicon gate resistivity is three orders of magnitude higher than that of aluminum. For normal (slow) CCD readout, this is not a concern, but at high speeds the RC time constant (R from the polysilicon gate resistivity and C from the combination of gate overlap and channel capacitance) can make T_V prohibitively long. For this reason, the gates have been metal strapped, whereby (as shown in Fig. 3) each polysilicon gate is covered with metal, and contacted every second channel stop. The lithographic tolerances needed for the contact and metal etching steps result in a minimum metal-strapped pixel pitch of about 25 μm.

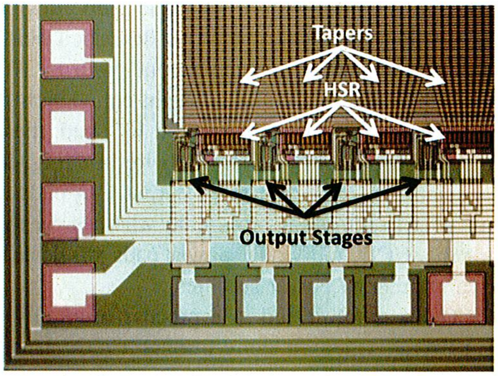


FIG. 2. (Color online) Microphotograph of the output stages of the almost column-parallel CCD. The constant-area tapers are visible, as are the 10 pixel miniserial shift registers labeled HSR. Output stages are nestled between the shift registers.

The CCD lot was fabricated on high-resistivity 6 in. wafers, 675 μm thick by Dalsa Semiconductor. Dalsa provides the front-end processing (implants, polysilicon gates), after which thinning, backside processing, and metallization are performed at the LBNL Microsystems Laboratory. The backside processing involves high temperatures so the aluminum metallization must be performed last. A few control wafers were fully processed by Dalsa (i.e., including metallization). Those wafers can only be frontside illuminated and are 675 μm thick. CCDs processed at LBNL can be front side or back side illuminated and for this lot were thinned to 200 μm.

III. CCD SIGNAL PROCESSING INTEGRATED CIRCUIT

The high density of analog outputs from the CCD makes discrete readout impractical. We have therefore developed a custom integrated circuit, fCRIC, to acquire and digitize the CCD signals. The fCRIC is based upon a floating-point architecture developed at LBNL (Ref. 5) and is shown in Fig. 4. The CCD output is ac coupled to the input stage, where the signal is amplified and converted from single ended to differential. The differential output voltage signal is then integrated by a differential multislope integrator which functions as follows: Initially, all switches are open, so that the

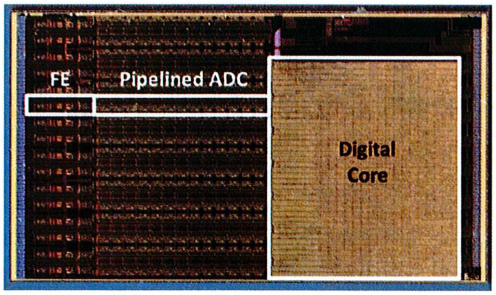


FIG. 4. (Color online) Photograph of fCRIC readout integrated circuit. For one of the 16 identical channels, the analog front end, labeled FE, is shown, along with its pipelined ADC. The digital core, common to the chip, is also shown.

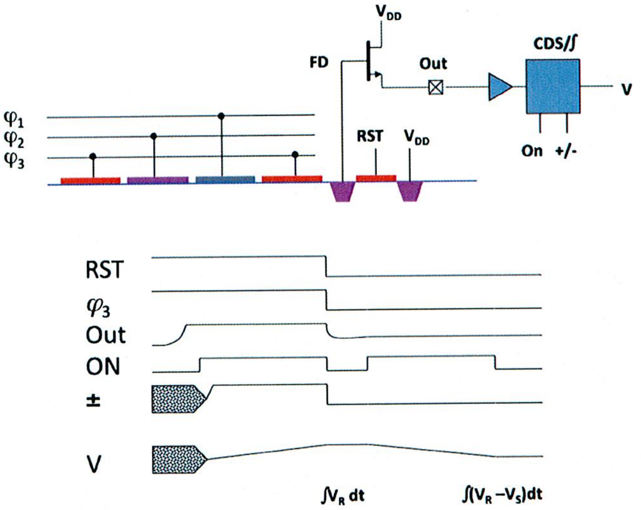


FIG. 5. (Color online) Schematic of the end of the serial shift register, output stage, and fCRIC along with the timing diagram. While RST and ON are high, the reset level is integrated. The signal charge to be integrated is then clocked onto the floating diffusion, the sign of integration is changed, and the signal level is integrated for the same amount of time as the reset level. The resulting voltage to be digitized is thus $\propto \int (V_R - V_S)dt$.

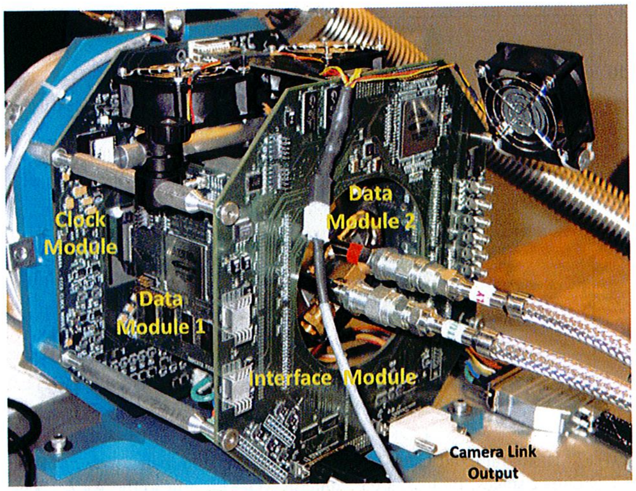


FIG. 6. (Color online) Back end readout showing the clock, data, and interface modules.

feedback capacitance is C. If the integrator output exceeds a preset threshold, then capacitors $3C$ are switched in. The feedback capacitance now goes from C to $4C$, which means that the gain has been reduced by a factor of 4. Similarly, if the integrator output again exceeds the preset threshold, then capacitors $4C$ are switched on. With a feedback capacitance of $8C$ defined as unity gain, the three effective gains are 8, 2, and 1. As shown in the timing diagram in Fig. 5, first the reset level of the CCD is integrated. The sign of the integration is now changed, and the signal level is integrated. This subtraction of the reset level from the signal level is known as correlated double sampling and is used in CCDs to reduce low frequency noise.

The analog result, signal reset, is then digitized by a 12 bit pipelined analog to digital converter (ADC). The digital result consists of the ADC mantissa and 2 bits representing the gain (1, 2, or 8). The overall gain has been set so that for a nominal 400 ns signal integration time, 0.5 V at the input of the fCRIC corresponds to the full scale of the ADC. With a typical conversion gain of 3.5 $\mu V/e^-$, full scale corresponds to the $2^{17}e^-$ on the gain 1 scale, $2^{16}e^-$ on the gain 2 scale, and $2^{14}e^-$ on the gain 8 scale. With 12 ADC bits, one analog to digital unit (ADU) thus corresponds to $32e^-$, $16e^-$, or $4e^-$.

The fCRIC contains 16 identical channels sharing a common digital back end. The digital circuitry includes a command decoder and four, serial data output lines, circuitry to assemble the pipelined ADC output into data words, circuitry to self-calibrate the ADCs, and a digital timing generator. The timing generator consists of a number of 8 bit counters, incrementing on a master clock of up to 250 MHz. Each counter controls the transition of the internal timing signals used in the analog signal acquisition and digitization. The 16 channels also share common analog services, including a bandgap voltage reference. The fCRIC was designed and fabricated in commercial 0.25 μm complementary metal oxide semiconductor and measures 4.8×8.3 mm^2.

IV. DIGITAL READOUT AND TIMING CONTROL

The back end electronics, shown in Fig. 6, control the digital readout and the timing of the detector. It consists of an interface module, two data modules, and a clock module. The data from the six fCRICs flow into the two data modules where they are buffered and organized. From the data modules the data are sent to the interface module, which buffers them again and then sends them out a camera link port. The camera link data are received by a Dalsa X64-CL full camera link frame grabber, which is plugged into a PCI 64 slot in a computer. The Dalsa frame grabber is capable of acquisition rates up to 680 Mbits/s and is controlled through a user interface that makes calls to Dalsa's application-programming interface.

The functions of the interface module are to generate all of the CCD clock signals, to provide a method of synchronizing external equipment to the detector, to receive data from the two data modules, and to send the CCD data out to a camera link interface. The logic in the interface module, controlling the CCD clock signals, allows developers to modify many parameters without having to reprogram the field programmable gate array (FPGA). For example, one can change the serial cycle time, the number of pixels read out, the shape of the waveforms, and the clock voltages all from a graphical user interface (GUI) without reprogramming the FPGA.

The functions of the data modules are to receive 12 low voltage differential signaling (LVDS) data streams from three fCRICs, provide a serial bus to program the fCRICs, provide the digital power to the fCRICs, and allow some real time data manipulations. One real time data manipulation is the rearrangement of the data from three fCRICs to match the CCD geometry before sending them to the interface module. The logic implements this descrambling with a ping-

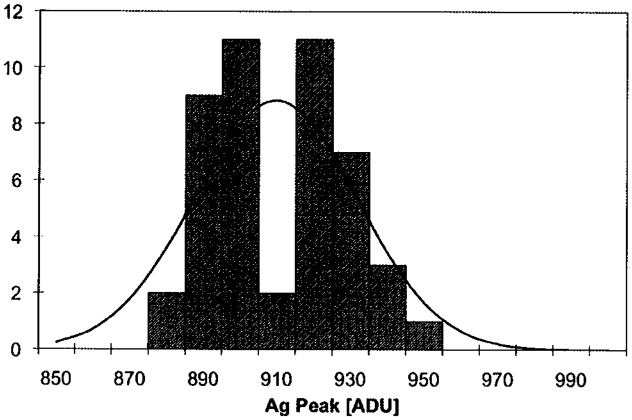


FIG. 7. Histogram and Gaussian fit to the conversion gain for each of the CCD output stages. Plotted is the value in ADU for a fit to the Ag K_α peak for each individual output. The average of all outputs is 909.5 ADU.

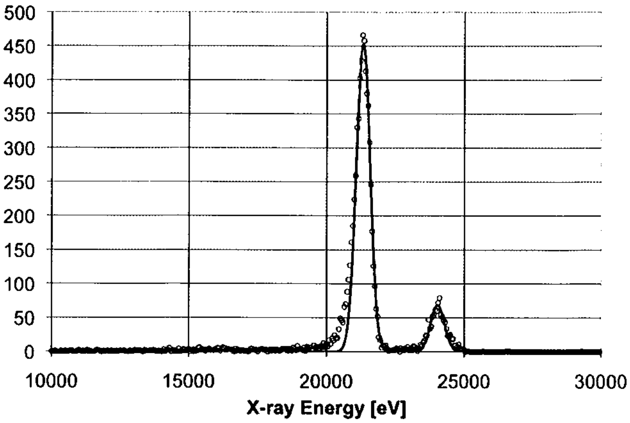


FIG. 8. Histogram and fit for Ag fluorescence photons over the surface of the CCD. The K_α peak has a Gaussian fit width of 250 eV.

pong line buffer: one buffer receives data from the fCRICs while the other buffer is transferring data to the interface module.

The functions of the clock module are to convert the CCD clocks to programmable voltage levels, provide enough current to the clock signals to drive the CCD, provide programmable bias voltages to the CCD, and provide analog power to the fCRICs. The CCD clock signals come from the interface module and are optically isolated from the digital power supply. All of the programmable voltages are set through a serial bus from the interface module and can be modified through the GUI.

V. RESULTS

A prototype CCD with 480×480, 30 μm pixels was fabricated, along with various test devices, including a four port version (20×480 pixels) for use with conventional CCD readout systems. The prototype CCD implements several new design features compared to previous LBNL CCDs: a large number of output stages—48 on each side; use of the constant-area taper, metal strapping, and large (30 μm) pixels. A front-illuminated four port version was characterized first with slow (100 kHz digitization rate) readout in order to assess the performance of the large, metal-strapped pixels and to compare the behavior to previous LBNL CCDs. As the devices are 675 μm thick, there is considerable bulk thermal leakage current, so that measurements must be performed at low temperatures. With ^{55}Fe, the conversion gain was measured at 140 K to be 3.3 $\mu V/e^-$, consistent with expectations. At that temperature, long exposures gave a leakage current of $5e^-/\mathrm{h/pixel}$, or 2.5×10^{-17} A/cm^2.

Backside illuminated CCDs thinned to 200 μm have also been characterized on ALS Beamline 5.3.1 (Ref. 6) with fluorescence x rays from thin foils. For all of the subsequent measurements, the CCD was read out at 5 ms/frame. Figure 7 shows a histogram of the gain for each output stage for Ag fluorescence photons (22 keV). At 500 mV full scale on the ×1 gain range, the average peak value of 910 ADU on the ×8 gain range corresponds to a conversion gain of about 2.3 $\mu V/e^-$. This conversion gain is about 1/3 lower than that measured for low-speed operation due to the settling time of the output stage. The single-stage output source follower is biased at 1 mA and has a settling time approaching 1 μs. At low speed, the voltage signal to be integrated is fully settled while integration takes place, but at high speed, it is not, resulting in a lower conversion gain. The output gain uniformity is quite good, as seen in Fig. 7, with a standard deviation of ~3% (which includes the tolerance of the source follower bias resistors, and hence the gain of the output transistor). This demonstrates good matching across the CCD chip.

The spectrum for Ag fluorescence photons, corrected for the output stage gains, is shown in Fig. 8. The measured resolution, 250 eV, is currently read noise limited and will be improved in the next implementation of the readout circuit board. Although the Gaussian PSF for 200 μm fully depleted silicon is around 5 μm, which seems negligible compared to a 30 μm pixel, it is easy to show that for uniform illumination, a 30 μm pixel contains only 75% of the total charge for a 5 μm PSF. This can be seen in Fig. 9, which shows the energy-ordered sum, $S(n)=\Sigma_{i=1}^{n}P_i$, where P_i is the energy recorded in the ith pixel and $P_i > P_{i+1}$. The progression seen in Fig. 9 is consistent with a 5 μm Gaussian PSF. Lastly, different depths of conversion lead to subtle differ-

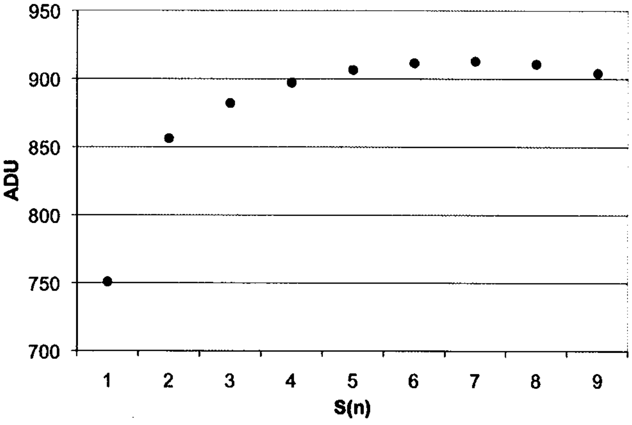


FIG. 9. Charge collected for Ag fluorescence photons for an energy-ordered sum; $S(i)$ is the sum of the ith most energetic pixels.

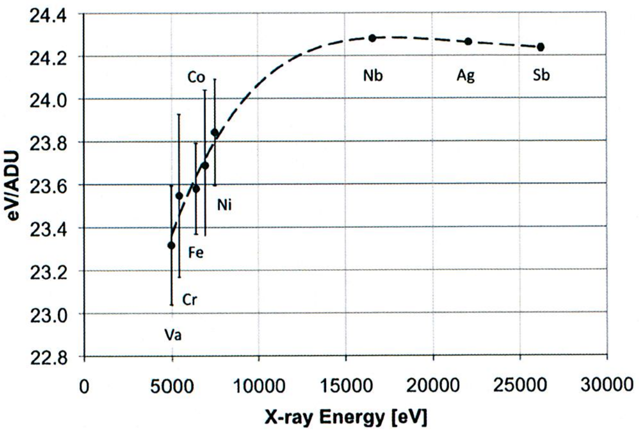


FIG. 10. Calibration constants (eV/ADU) for different x-ray energies.

ences in conversion gain. Figure 10 shows the measured conversion gains as a function of energy, and therefore different average conversion depths.

VI. CONCLUSIONS

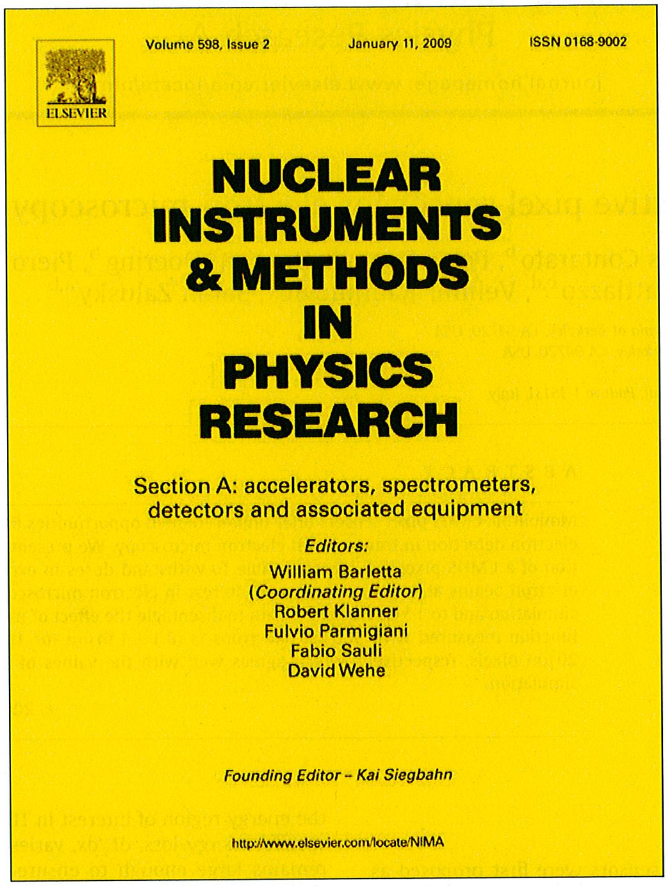
A custom CCD, capable of high-speed readout due to multiple output ports, has been fabricated on thick, high-resistivity silicon. This CCD, in addition to high-speed readout for optical photons, has excellent efficiency for direct x-ray detection. To take full advantage of the high performance, a 16-channel custom readout integrated circuit was developed in parallel with the CCD. The complete system includes a digital back end control, clocking, and data acquisition system, capable of acquiring images at 200 frames/s.

The CCD described here was originally developed for optical imaging, to read out, for example, a fiber-coupled phosphor x-ray detector. It is clear, however, that there are several applications of this CCD in direct x-ray detection, such as an imaging energy-measuring (spectroscopic) detector. We will therefore optimize the readout system for direct x-ray detection in a future iteration: the well depth for x rays is significantly less than for optical photons, so a future version will emphasize higher readout speed at the expense of reduced ADC resolution. For applications requiring larger areas, along with an electronic shutter, a 1000×2000 version of this CCD will be fabricated.

ACKNOWLEDGMENTS

We gratefully acknowledge the contributions from Mathew Church for the design of the camera mechanics, Rossana Cambie for the design and processing of the silicon substrate holding the CCD and fCRICs, and Jacque Wycoff, Rhonda Whitharm, and John Eames for assembling the substrates. Tim Madden and Antonino Miceli helped in the design of the clock module and detector characterization. Robert Abiad, Geroge Chao, Dario Gnani, and Brad Krieger were part of the fCRIC design team. We were aided in laboratory testing by Bill Kolbe and Armin Karcher and at ALS Beamline 5.3.1 by Rich Celestre. Steve Holland's profound knowledge and expertise in CCDs was essential for the design of the CCD presented here, and Chris Bebek provided both guidance and a fruitful link to the astronomical CCD community.

[1] C. J. S. Damerell, Nucl. Instrum. Methods Phys. Res. A **541**, 178 (2005).
[2] R. Bredthauer, Proc. SPIE **1155**, 89 (1989).
[3] R. J. Stover, M. Wei, D. K. Gilmore, S. E. Holland, D. E. Groom, W. W. Moses, S. Perlmutter, G. Goldhaber, C. R. Pennypacker, N. W. Wang, and N. Palaio, Proc. SPIE **3505**, 13 (1998).
[4] S. E. Holland, D. E. Groom, N. P. Palaio, R. J. Stover, and M. Wei, IEEE Trans. Electron Devices **50**, 225 (2003).
[5] J. P. Walder, G. Chao, J. F. Genat, A. Karcher, B. Krieger, S. Kurz, S. J. Steckert, and H. von der Lippe, IEEE Trans. Nucl. Sci. **51**, 1936 (2004).
[6] See: http://www.als.lbl.gov/als/techspecs/bl5.3.1.html which lists the characteristics of ALS beamline 5.3.1 where testing of the CCD was performed.



Volume 598, Issue 2 January 11, 2009 ISSN 0168-9002

ELSEVIER

NUCLEAR
INSTRUMENTS
& METHODS
IN
PHYSICS
RESEARCH

Section A: accelerators, spectrometers,
detectors and associated equipment

http://www.elsevier.com/locate/NIMA

Nuclear Instruments and Methods in Physics Research A 598 (2009) 642–649

Contents lists available at ScienceDirect

Nuclear Instruments and Methods in Physics Research A

journal homepage: www.elsevier.com/locate/nima

A rad-hard CMOS active pixel sensor for electron microscopy

Marco Battaglia [a,b,*], Devis Contarato [b], Peter Denes [b], Dionisio Doering [b], Piero Giubilato [b,c,d], Tae Sung Kim [b], Serena Mattiazzo [c,d], Velimir Radmilovic [b], Sarah Zalusky [a,b]

[a] *Department of Physics, University of California at Berkeley, CA 94720, USA*
[b] *Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*
[c] *INFN, Sezione di Padova, I-35131, Italy*
[d] *Dipartimento di Fisica, Università degli Studi, Padova I-35131, Italy*

ARTICLE INFO

Article history:
Received 9 September 2008
Accepted 22 September 2008
Available online 11 October 2008

Keywords:
Monolithic active pixel sensor
Transmission electron microscopy

ABSTRACT

Monolithic CMOS pixel sensors offer unprecedented opportunities for fast nano-imaging through direct electron detection in transmission electron microscopy. We present the design and a full characterisation of a CMOS pixel test structure able to withstand doses in excess of 1 Mrad. Data collected with electron beams at various energies of interest in electron microscopy are compared to predictions of simulation and to 1.5 GeV electron data to disentagle the effect of multiple scattering. The point spread function measured with 300 keV electrons is $(8.1 \pm 1.6)\,\mu m$ for $10\,\mu m$ pixel and $(10.9 \pm 2.3)\,\mu m$ for $20\,\mu m$ pixels, respectively, which agrees well with the values of 8.4 and $10.5\,\mu m$ predicted by our simulation.

1. Introduction

CMOS monolithic active pixel sensors were first proposed as photo-detectors 40 years ago. About a decade ago their use as charged particle detectors started to be explored, recognising that they represent an appealing technology offering small pixel size and the possibility of high speed readout [1,2]. In addition to their use in charged particle tracking, they have been proposed and investigated for direct electron detection in electron microscopy, in particular in Transmission Electron Microscopy (TEM) as electron imagers [3–5] together with hybrid pixels [6–8], as an advantageous alternative to CCDs optically coupled to phosphor plates (for a review of different detection technologies, see Ref. [9]).

For single particle tracking, it has been demonstrated that CMOS pixels of order $10 \times 10\,\mu m^2$ can reconstruct the position of impact of the particle with an accuracy of $\mathcal{O}(1\,\mu m)$, by using the centre of gravity of the charge distribution [10]. Owing to their thin sensitive region, confined to the $\simeq 10\,\mu m$ lightly doped epitaxial layer, these sensors can be back-thinned down to $\leqslant 50\,\mu m$, without loss in performance [11]. The thin charge-collection region and total thickness of CMOS pixel sensors is ideal in TEM applications, where multiple scattering effects are significant, for obtaining a good Point Spread Function (PSF). In the energy region of interest in TEM, which is $\simeq 60-400$ keV, the specific energy loss, dE/dx, varies by an order of magnitude but remains large enough to ensure single electron detection with large signal-to-noise ratio (S/N) values. For TEM, the detector should record the position of a collection of electrons that represents a magnified image. Whereas diffusion is beneficial for tracking applications, as it improves the determination of the charge centre of gravity, and thus of the particle impact point, to an accuracy significantly better than the pixel pitch, for TEM diffusion is harmful, as it increases the PSF and needs to be carefully evaluated.

While significant progress has been made in the last few years to develop CMOS pixel sensors for tracking and vertexing applications in particle collider experiments, several issues, relevant to their application for fast nano-imaging at TEM, remain to be addressed. In particular, their radiation hardness is a key challenge in the design of a detector for TEM. Unlike HEP, where the total dose requirement can be estimated, based on integrated luminosity, TEM requirements depend on the mode of operation. Whereas the radiation hardness for the specimen can be easily specified (in electrons per square angstrom), as the magnification is variable, this does not translate into a requirement on the detector. In addition, large numbers of electrons may be contained in a limited number of Bragg spots, when operating in diffraction mode, causing large doses to be received by a small number of pixels. Given that a very intense bright field image could deposit order of $10\,\mathrm{rad\,s^{-1}\,pixel^{-1}}$, a target radiation tolerance of $\geqslant 1$ Mrad would enable its use for approximately one year, which appears to be a valid requirement.

---

[*] Corresponding author at: Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. Tel.: +1 510 486 7029.
E-mail address: MBattaglia@lbl.gov (M. Battaglia).

0168-9002/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2008.09.029

In this paper we present a prototype CMOS active pixel sensor with radiation tolerant pixel cell design for use in TEM. We evaluate its performance applying high energy physics (HEP) modelling and experimental techniques to their characterisation as TEM detectors. We study the particle energy loss, charge spread, PSF and response after irradiation with 200 keV electrons and 29 MeV protons for doses in excess to 1 Mrad. Data are compared to results of detailed simulation of electron interactions, energy loss and signal generation.

2. Sensor design and readout

We developed a new circuit, named LDRD2-RH, optimised for radiation tolerance, as a variation on a previous circuit, developed for HEP applications [12]. The LDRD2-RH chip retains the original addressing and output circuitry. The sensor is divided into several sectors, each consisting of 24×48 pixels arrayed on a $20\,\mu m$ pitch, in order to test different pixel layouts. The principle radiation damage mechanism is radiation-induced leakage current in the charge-collecting diode. This leakage current is primarily caused by the positive potential generated by trapped (positive) charge in the oxide above the p-type silicon, which results in inversion of the silicon at the Si/SiO_2 interface. In each $20 \times 20\,\mu m^2$ pixel, two $10 \times 20\,\mu m^2$ pixels are drawn: each with the same design of the charge-collecting diode, and with the classical 3T-readout. For both sub-pixels, the transistor sizes (W/L) are the same, but in one sub-pixel the transistor has a conventional linear layout, and in the other, an enclosed gate layout. Fig. 1 shows some of the different diode layouts used. The detector has been fabricated in the AMS $0.35\,\mu m$ 4-metal, 2-poly CMOS-OPTO process, which has an epitaxial layer with a nominal thickness of $14\,\mu m$.

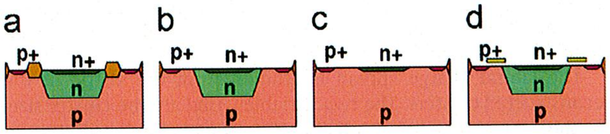


Fig. 1. Cross-section layouts of different diodes used in the test chip. The diode in (a), GR, is an n-well diode with an enclosing $p+$ guard ring; (b) NW draws thin oxide over the diode; (c) is like (b) but without the well implant; and (d) PO is like (b) but with a polysilicon ring covering the potentially inverted region.

Pixels are read out in rolling shutter mode, which ensures a constant integration time across the pixel matrix. Pixels are clocked at 6.25 and 25 MHz, corresponding to an integration time of 737 and 184 μs, respectively. The detector is read out through a custom FPGA-driven acquisition board. A set of 14 bits, 40 MSample s^{-1} ADCs reads the chip outputs, while an array of digital buffers drives all the required clocks and synchronisation signals. The FPGA has been programmed to generate the clock pattern and collect the data sampled by the ADCs. A 32 bit-wide bus connects the FPGA to a National Instruments PCI 6533 digital acquisition board installed on the PCI bus of a control PC. Data is processed on-line by a LabView-based program, which performs the computation of the pixel noise and pedestal levels. Results are given for the PO pixel design, which was found to be best performing in terms of radiation tolerance, read out at 6.25 MHz, unless otherwise noted.

3. Response simulation

The energy deposition in the sensor active layer and the lateral charge spread is simulated using two independent programs.

The simulation of electron scattering in matter for applications in electron microscopy has been described in Ref. [13]. The first simulation program is inspired by this formalism, but with the relativistic corrections added, as required for accuracy at our energies of interest. The second simulation is based on the `Geant-4` program [14] and uses the low energy electromagnetic physics models [15].

In both simulations, the CMOS sensor is modelled according to the detailed geometric structure of oxide, metal interconnect and silicon layers, as specified by the foundry. Electrons are incident perpendicular to the detector plane and tracked through the sensor. For each interaction within the epitaxial layer, the energy released and the position are recorded. Fig. 2 shows the sensor layout adopted in the simulation with the simulated trajectories of 200 keV electrons.

The two simulation programs give a good description of the measured energy deposition as shown in Fig. 2. `Geant-4` provides with a better description of the high energy part of the Landau distribution which tends to be systematically over-estimated by the results of the first simulation program. This effect becomes more significant with increasing electron energies.

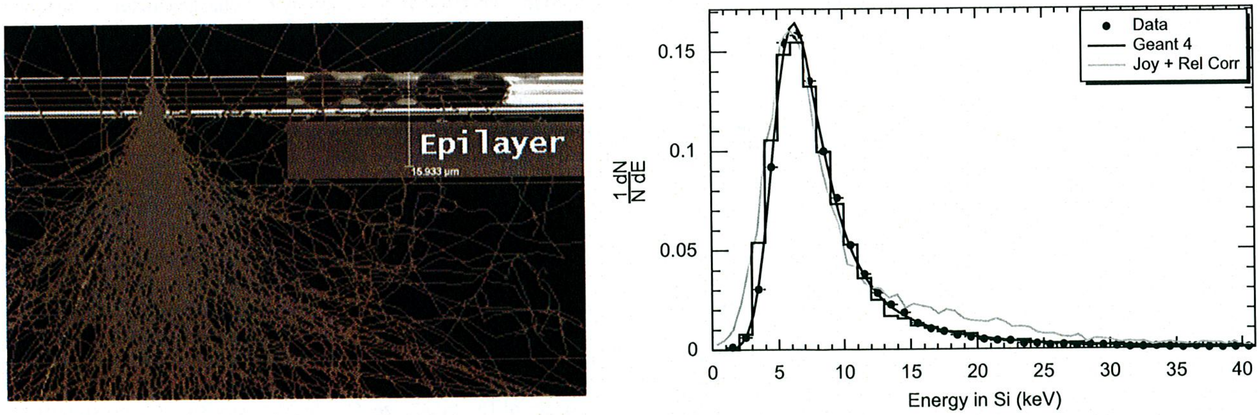


Fig. 2. Response simulation of the LDRD2-RH chip. (Left) Simulated trajectories of 200 keV electrons in the pixel sensor. The metal layers on top and the sensitive epitaxial layer below are highlighted. (Right) Reconstructed deposited energy for 200 keV electrons. The points with error bars show the data, the grey smoothed line the prediction of the simulation according to Ref. [13] with relativistic corrections added and the histogram the result of the `Geant-4` simulation. The continuous line is the fit to the data points with a Landau function convoluted with a Gaussian noise term.

Table 1
Lateral charge spread due to multiple scattering predicted by `Geant-4` for various electron energies

E_e	Lateral spread (μm)
120 keV	15.6
160 keV	11.5
200 keV	9.3
300 keV	6.8
1.5 GeV	1.3

The lateral charge spread due to multiple scattering in the sensor is evaluated in simulation by computing the lateral position of the electron interaction points in the epitaxial layer, weighted by the deposited energy. Results from the `Geant-4` simulation are given in Table 1.

Charge collection is simulated with `PixelSim`, a dedicated digitisation module [16], developed in the `Marlin` C++ reconstruction framework [17]. This processor starts from the ionisation points generated along the particle trajectory by `Geant-4` and models the diffusion of charge carriers from the epitaxial layer to the collection diode. `PixelSim` provides us with full simulation of the response of all the individual pixels in the detector matrix, including electronics noise and efficiency effects, which can be processed through the same analysis chain as the data. The simulation has a free parameter, the diffusion parameter σ_{diff}, used to determine the width of the charge carrier cloud, which is tuned to reproduce the pixel multiplicity in the cluster measured for 1.5 GeV electrons in data, as discussed in the next section. We find the best agreement between the simulated and measured pixel multiplicity in the cluster for $\sigma_{\mathrm{diff}} = 16.5\,\mu\mathrm{m}$. This can be compared with an estimate of the charge diffusion length $L_n = \sqrt{D_n \tau_n}$, obtained from the diffusion coefficient, $D_n = kT/e\mu_n$, and the charge-collection time, τ_n. The diffusion coefficient is computed for an estimated doping of $10^{14}-10^{15}\,\mathrm{cm}^{-3}$ of the epitaxial layer. The charge-collection time has been measured on data. We focus a 1060 nm laser to a $\simeq$ 10 μm spot onto a single pixel and pulse the laser for 2 ns. The pixel analog output is recorded on a digital oscilloscope and we observe that the pixel analog level reaches a plateau 150 ns after the arrival of the laser pulse. From these data we estimate a diffusion length of 14–19 μm, which agrees well with the simulation result.

4. Sensor tests

4.1. Energy deposition

The detector calibration is obtained by recording the position of the 5.9 keV full energy peak of a collimated 2.2 mCi ^{55}Fe source. We find a conversion factor of 0.98 keV/ADC count or 26.7 e^-/ADC count at 6.25 MHz readout frequency.

Electrons in the energy range from 120 to 200 keV from the 200CX electron microscope at the National Center for Electron Microscopy (NCEM) are used to characterise the detector response to low energy particles. The response to high momentum particles is studied with the 1.5 GeV electron beam extracted from the LBNL Advanced Light Source (ALS) booster. Data are converted into the `lcio` format [18]. Data analysis is performed offline by a set of dedicated processors developed in `Marlin` and proceeds as follows. Events are first scanned for noisy pixels. The noise and pedestal values computed on-line are updated, using the algorithm in Ref. [19], to follow possible variations in the course of data taking. The measured average pixel noise is $(130 \pm 6)\,e^-$ and $(71 \pm 4)\,e^-$ of equivalent noise charge (ENC) for the electron microscope and the ALS data, respectively, at 6.25 MHz and $(110 \pm 8)\,e^-$ and $(67 \pm 5)\,e^-$ at 25 MHz. The noise is partly due to the readout electronics, and is larger in the electron microscope set-up due to the longer cable needed to route the analog signals out of the vacuum enclosure at the bottom of the microscope column, with a minor contribution from leakage current from operating the detector at $\simeq +25\,^{\circ}\mathrm{C}$.

Electron hits are then reconstructed from the recorded pixel pulse heights. The detector is scanned for pixels with pulse height values over a given S/N threshold. These are designated as cluster 'seeds'. Seeds are sorted according to their pulse heights and the surrounding neighbouring pixels are tested for addition to the cluster. The neighbour search is performed in a 5×5 matrix around the seed. Pixel thresholds at 3.5 and 2.0 units of noise have been used for seed and additional pixels, respectively. Clusters are not allowed to overlap, i.e. pixels already associated with one cluster are not considered for populating another cluster around a different seed. Finally, we require that clusters are not discontinuous, i.e. pixels associated to a cluster cannot be interleaved by any pixel below the neighbour threshold. Reconstructed hits are characterised in terms of the energy recorded in a 3×3 pixel matrix centred around the seed pixel, the pixel multiplicity of the reconstructed cluster and the fraction of the total charge collected by the pixels in the matrix, sorted in order of decreasing pulse height.

Results for the reconstructed deposited energy are shown in Fig. 3, where data and simulation for 120, 160, 200 keV and 1.5 GeV electrons are compared. The simulation reproduces well the measured energy deposition, over this range of particle energies. The observed average values of energy deposition exceeds those predicted by the thin straggling model [20] for 14 μm of Si, due to the effect of electron interactions in the SiO_2 and metal layers on top of the sensitive volume.

4.2. Charge spread

The lateral charge spread is studied from the shape of the reconstructed clusters. The two variables used are the cluster size, i.e. the average pixel multiplicity $\langle N_{\mathrm{pixels}} \rangle$, and the distribution of the fraction of the total charge collected in the pixels with the highest pulse height in a matrix around the seed pixel. Table 2 compares the cluster size for electrons of different energies measured in data with the simulation prediction. The cluster size for 1.5 GeV electrons is governed by the charge carrier diffusion in the epitaxial layer, since the multiple scattering is negligible compared to the pixel size, at this energy. Both data and simulation show an increase of the multiplicity for lower energies consistent with the combined effect of the larger collected charge, which pushes the pulse height on more pixels above the S/N threshold for additional pixels, and of the multiple scattering in the detector, which displaces the locations of energy deposition from the point of entrance of the electron. The two effects can be disentangled by analysing the fraction of the cluster charge collected in the pixels with the highest pulse height in a 3×3 matrix around the cluster seed. Fig. 4 shows the results obtained for data at 120 and 200 keV compared to 1.5 GeV electrons. We indeed observe an increase of the area over which the charge spreads with 120 keV electrons, consistent with the effect of lateral charge spread due to multiple scattering which becomes of the order of the pixel size. On the contrary there is no significant increase on the charge distribution for 200 keV electrons, where we also observe only a very modest increase in cluster size. These results are well reproduced by simulation as shown in Fig. 4. This first validation of simulation in the description of charge spread is

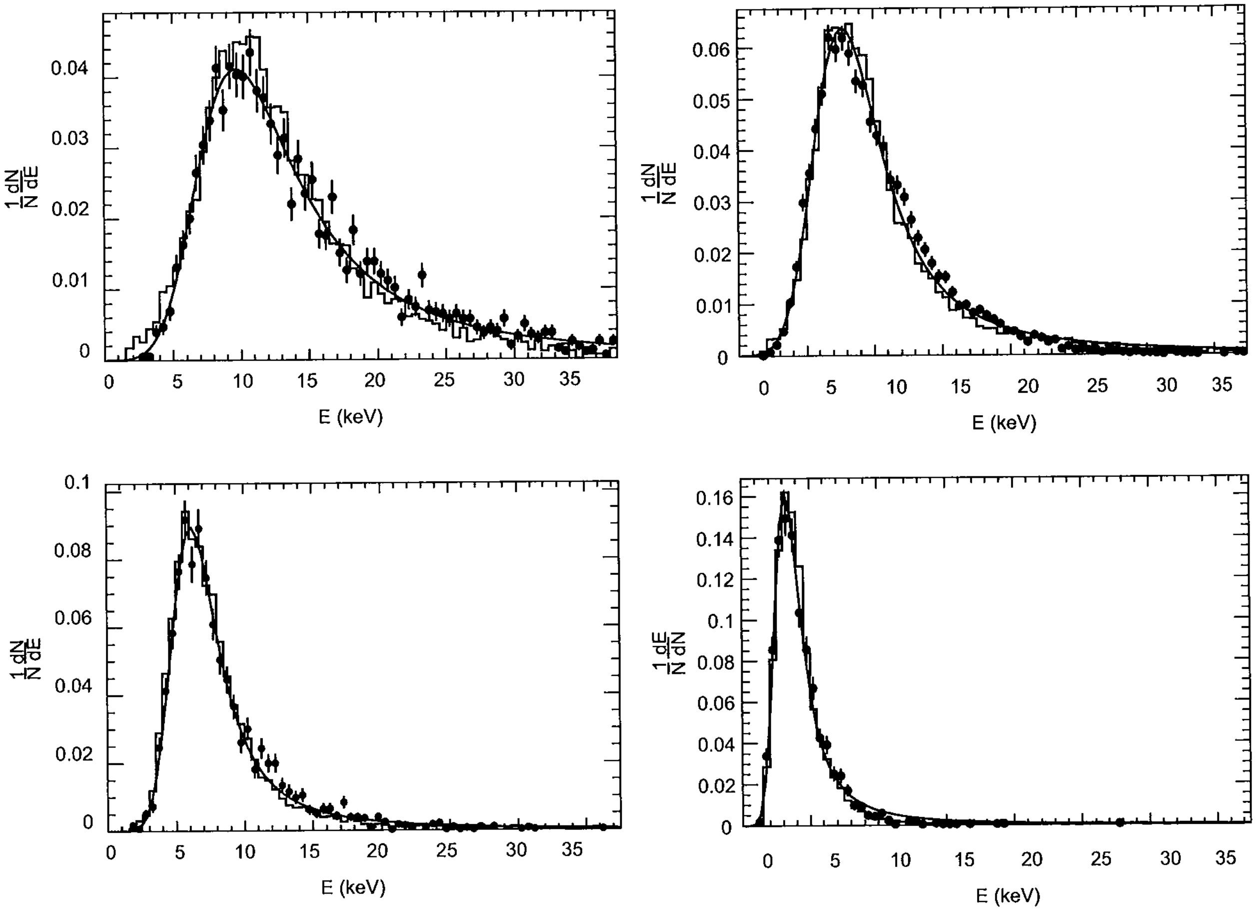


Fig. 3. Reconstructed deposited energy in the 3 × 3 pixel matrix for 120 keV electrons (upper left), 160 keV electrons (upper right), 200 keV electrons (lower left) and 1.5 GeV electrons (lower right). The points with error bars show the data and the histogram the result of the `Geant-4` simulation. The continuous line shows a Landau function convoluted with a Gaussian noise term fit to the data.

Table 2
Cluster size for electrons of different energies in data and simulation

E_e	$\langle N_{\text{pixels}} \rangle$ data	$\langle N_{\text{pixels}} \rangle$ simulation
120 keV	3.35 ± 0.06	3.01 ± 0.02
160 keV	2.82 ± 0.04	2.73 ± 0.03
200 keV	2.25 ± 0.04	2.33 ± 0.03
1.5 GeV	2.20 ± 0.02	2.21 ± 0.02

important for its use in the PSF estimation, which is discussed next.

4.3. PSF and imaging

The detector PSF originates from charge spread due to charge carrier diffusion and ionising particle scattering, as well as by the finite spatial sampling frequency, which depends on the pixel pitch. We estimate the PSF using `Geant-4` + `PixelSim` simulation and test the simulation results with data. In simulation, a monochromatic, point-like beam of electrons is sent onto the surface of the detector. The point spread function is determined as the r.m.s. of the predicted distribution of the detected charge on the pixels. Results are summarised in Table 3 for different electron energies and 10 and 20 μm pixel pitch. Simulation predicts a PSF better than 10 μm for 300 keV electrons imaged with a 10 μm pixel pitch.

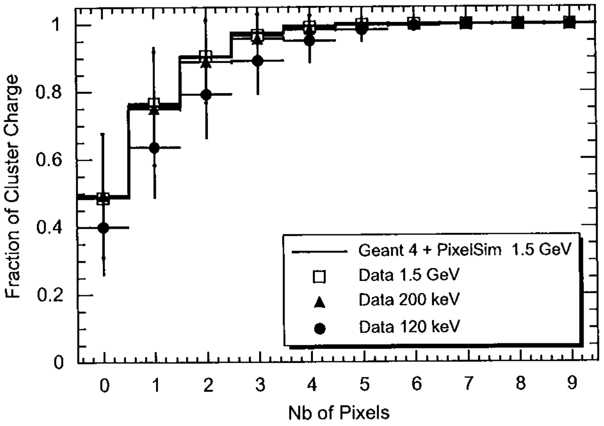


Fig. 4. Fraction of the charge collected in a 3 × 3 pixel matrix as a function of the number of pixels added, for 120 keV (dots), 200 keV (triangles), 1.5 GeV (squares) data, and 1.5 GeV simulation (line). The error bars show the r.m.s. of the data distribution. Pixels in the matrix are sorted in decreasing pulse height order.

We validate these results using data taken at different energies, from 100 up to 300 keV, with 10 and 20 μm pixels at the TITAN test column at NCEM. We use the LDRD2-RH as well as an earlier sensor, the LDRD1 chip, featuring simple 3T pixels with 10, 20 and 40 μm pitch and fabricated in the same AMS 0.35-OPTO process [21]. A gold wire was mounted, parallel to the pixel columns, at a distance of $\simeq$ 3 mm from the detector surface. The wire diameter is measured to be (59.6 ± 0.7) μm, using a high resolution optical survey system. Since the gold wire has well-defined edges, the profile of the deposited energy in the pixels, measured across the wire allows us to study the charge spread due to scattering and diffusion along the projected image of the wire edge and compare to simulation.

The set-up has been simulated in detail in `Geant-4` and the pixel response extracted from `PixelSim`. We study the change in the recorded signal, by scanning along pixel rows across the gold wire. Fig. 5 shows the pulse heights measured on the pixels along a set of rows, comparing data at 300 and 120 keV with simulation, for the 20 μm pixel pitch. The good agreement observed validates the estimation of the point spread function obtained from simulation.

Further, we extract the point spread function directly from the data and compare the results to the simulation predictions. We parametrise the measured pulse height on pixel rows across the image projected by the wire with a box function smeared by a Gaussian term, which describes the point spread function. The maximum and minimum pulse height levels, for pixels away from the wire region and for the pixel exactly below the wire centre, respectively, are fixed to those observed in data and we perform a simple 1-parameter χ^2 fit to extract the Gaussian width term, which gives the PSF. At 300 keV, we measure a PSF value of (8.1 ± 1.6) μm for 10 μm pixels on the LDRD1 chip (see Fig. 6) and of (10.9 ± 2.3) μm for 20 μm pixels on the LDRD2-RH chip, to be compared to simulation which predicts 8.4 and 10.5 μm, respectively. The PSF scaling with electron energy and pixel size is shown in Fig. 6. Again data and simulation are in good agreement. It is interesting to observe that for increasing electron energies the point spread function improves, due to the reduction of multiple

Table 3
Point spread function predicted by Geant 4 + PixelSim for different electron energies and 10 and 20 μm pixel pitch

Energy (keV)	10 μm pixels	20 μm pixels
120	11.4	13.1
200	9.1	11.2
300	8.4	10.5

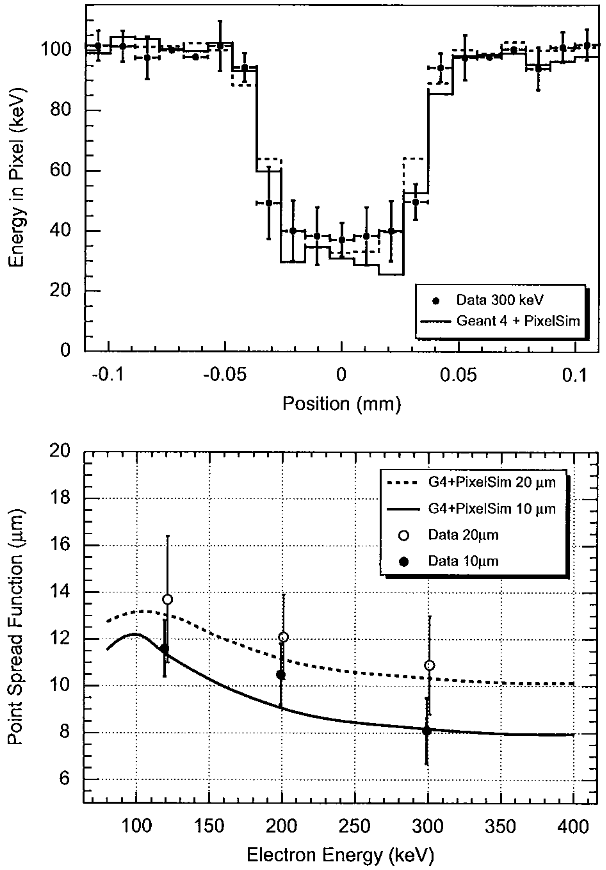


Fig. 6. Point spread function determination. Top: Pulse heights measured on pixels with 10 μm pitch, along a row across the Au wire stretched above the 10 μm pixels of the LDRD1 sensor for 300 keV beam. Data (points with error bars) are compared with the prediction of the `Geant-4` + `PixelSim` (solid line) simulation and the box+Gaussian function fit (dashed line). Bottom: PSF as a function of beam energy. Data (points with error bars) are compared to simulation (lines) for 10 and 20 μm pixels.

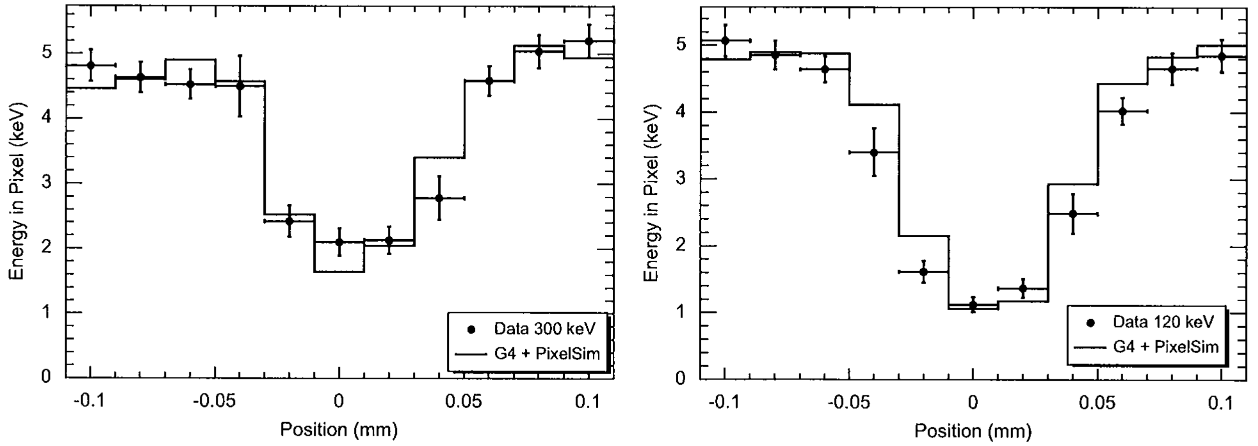


Fig. 5. Pulse heights measured on pixels with 20 μm pitch, along a row across the Au wire stretched above the LDRD2-RH sensor. Data at 300 keV (left) and 120 keV (right) are compared with the prediction of the `Geant-4` + `PixelSim` simulation.

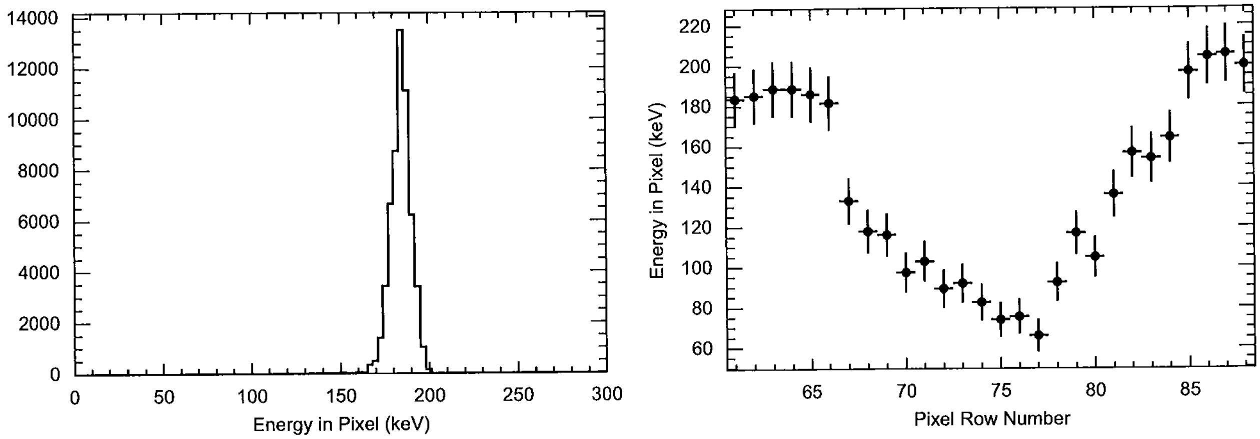


Fig. 7. (Left) Energy recorded on individual pixels with 200 keV electrons flat field illumination, (right) energy recorded on a pixel column across the projected image of a Pb–Sn alloy nanoparticle in an Al matrix.

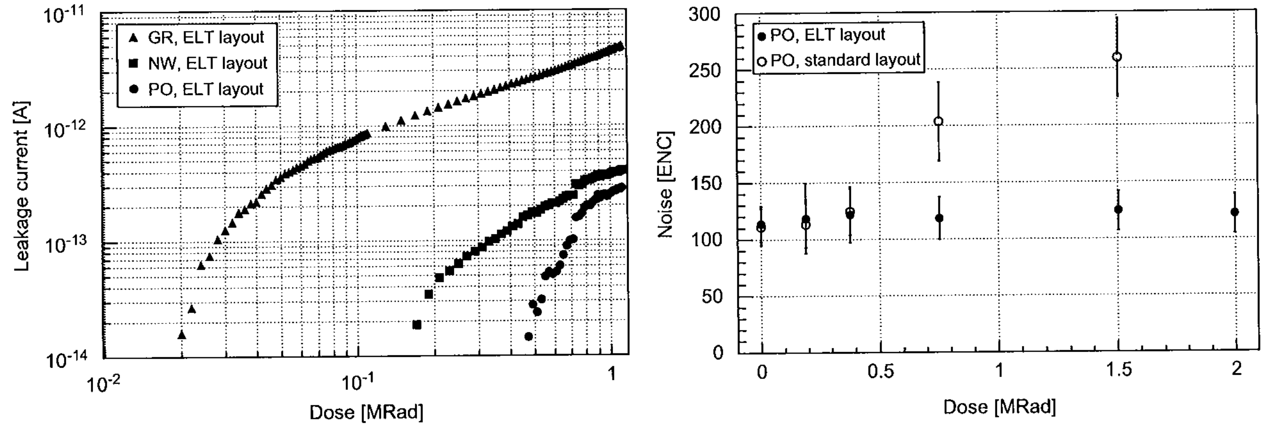


Fig. 8. Results of the LDRD2-RH sensor irradiation. (Left) Sensor leakage current as a function of dose for 200 keV electrons. (Right) Average pixel noise as a function of dose for 29 MeV proton in cells designed with ELT and standard layouts. The measurement at the highest dose for the standard layout is not reported since the sensor output signal saturated the ADC dynamic range due to the increased pixel leakage current. The error bars represent the r.m.s. of the pixel noise distributions.

scattering, up to $\simeq 300$ keV, where it approaches an asymptotic value for 10 μm pixels, dominated by the effect of charge spread and spatial sampling frequency. On the opposite end of the energy range, the point spread function grows to a maximum around 100 keV and simulation predicts it to fall for lower energies, due to the short range in Si of these soft electrons.

Finally, we have performed a realistic imaging test with the LDRD2-RH sensor, by acquiring images of nanoparticles of a 91.8% Pb + 8.2% Sn alloy embedded in a solid Al matrix [22] at the 200CX electron microscope. The images are obtained with a 200 keV electron flat field illumination at 6.25 MHz readout frequency and 50 frames s^{-1} are written to disk. We estimate the flux to be $\simeq 35\,e^-$ $pixel^{-1}$ $event^{-1}$, where one event corresponds to a 737 μs integration time. Fig. 7 shows the pulse height recorded on the pixels across the projected image of a nanoparticle.

5. Sensor irradiation

The radiation tolerance of the LDRD2-RH chip for use in TEM has been assessed by comparing the sensor response before and after irradiation with 200 keV electrons up to a dose in excess of 1 Mrad. Results have been integrated by an irradiation with 29 MeV protons at the LBNL 88-in. cyclotron. We study the pixel noise, leakage current and charge collection. All tests are performed at room temperature.

For the electron irradiation, the sensor active surface has been covered with a same gold mesh, having 50 μm wide bars and 204 μm wide holes. This allows to compare the response of irradiated and non-irradiated pixels on the same chip. The sensor has been irradiated with a flux of $\simeq 2300\,e^-\,s^{-1}\,\mu m^{-2}$, in multiple steps, up to a total estimated dose of 1.11 Mrad. In between consecutive irradiation steps, 100 events are acquired without beam and the pixel pedestals and noise computed, in order to monitor the evolution of the pixel leakage current with dose. Fig. 8 shows the increase of the pixel pedestal levels, which measure the leakage current, as a function of the integrated dose.

After irradiation the sensor response is tested with 200 keV and 1.5 GeV electrons. The deposited energy in the 3×3 pixel matrix is compared to that obtained before irradiation. We fit the energy spectrum with a Landau function convoluted with a Gaussian distribution to represent the noise contribution. The fit function has three free parameters, the Landau peak position (Landau m.p.v.), the Gaussian width (Gaussian noise) and an overall normalisation. Due to the large correlation between the Gaussian width and the Landau width, the latter is fixed in the fit

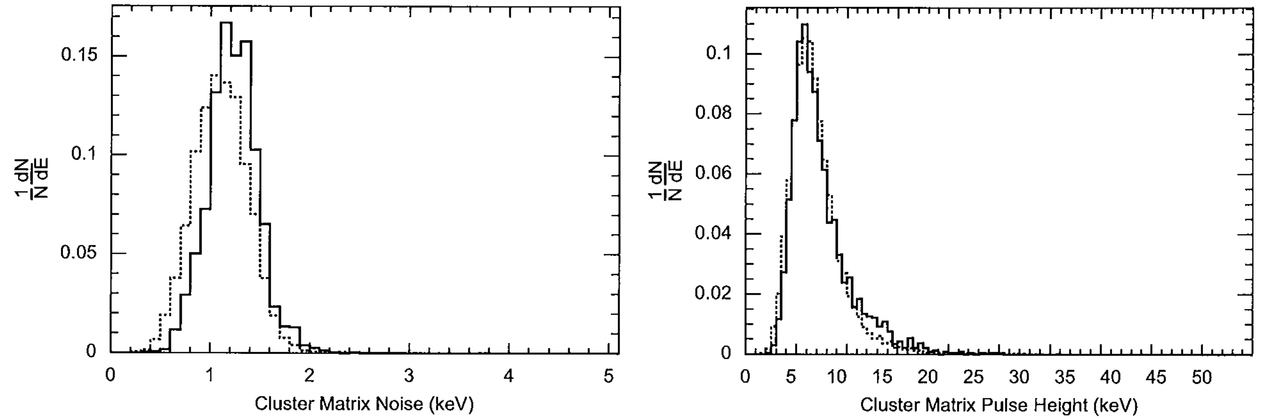


Fig. 9. Measured noise (left) and deposited energy (right) in a 3 × 3 pixel matrix for 200 KeV electrons before (continuous line) and after (dashed line) irradiation, obtained after applying the gain correction described in the text.

Table 4
Results of the fits to the distributions of the deposited energy in a 3 × 3 pixel matrix for 200 keV and 1.5 GeV electrons before and after the 1.1 Mrad electron irradiation

	Before irradiation (keV)	After irradiation (keV)
200 keV e^-		
Landau m.p.v.	(5.57 ± 0.04)	(5.54 ± 0.04)
Gaussian noise	(1.19 ± 0.04)	(0.98 ± 0.03)
3 × 3 matrix noise	(1.22 ± 0.20)	(1.11 ± 0.20)
1.5 GeV e^-		
Landau m.p.v.	(3.05 ± 0.03)	(3.37 ± 0.07)
Gaussian noise	(0.60 ± 0.06)	(0.81 ± 0.07)
3 × 3 matrix noise	(0.74 ± 0.01)	(0.88 ± 0.01)

The noise of the 3 × 3 pixel matrix is also given for comparison. Data after the irradiation is re-calibrated to account for the observed gain shift.

to the value obtained on simulation. We observe a change of the gain by ≃ 1.35, confirmed by calibration with ^{55}Fe. After correcting for this gain shift, the noise and energy deposition distributions before and after irradiation are in good agreement. In the ALS data, taken with shorter connections between the detector and the digitiser board, the electronics noise is lower and an hint of a small increase of the detector noise after irradiation may be observed. The deposited energy spectrum is unchanged after irradiation, confirming that the charge-collection properties of the pixel cell are not affected. Results are summarised in Fig. 9 and Table 4.

Since the threshold energy for electrons to cause displacement damage in silicon is 260 keV [23], we expect 200 keV to create only ionising damage. We investigated the effect of 29 MeV protons, which damage the sensor via both ionisation and non-ionising energy loss. The irradiation has been performed at the BASE Facility of the LBNL 88-in. cyclotron [24]. Results are reported in details in Ref. [25]. One prototype sensor was irradiated with 29 MeV protons up to a total integrated fluence of 8.5×10^{12} pixel cm^{-2}, corresponding to a total dose of ≃ 2 Mrad. At equal doses, a larger increase of the leakage current, compared to the electron irradiation, hints at a probable contribution from displacement damage. Fig. 8 (right) shows the average pixel noise in the PO sectors for ELT and linear transistors; it can be seen that while the noise of the ELT layout cells remains basically unchanged over the dose range and only slightly increases at the highest doses, the noise of the standard layout cells is doubled after ~1 Mrad. The increased leakage current can be controlled by cooling the chip. The PO design with ELT layout appears to be the most radiation tolerant of the designs tested in the LDRD2-RH chip.

6. Conclusions

Thin monolithic CMOS pixel sensors with small, radiation tolerant pixels offer a very attractive solution for fast nano-imaging in transmission electron microscopy. The thin sensitive layer ensures a direct detection with small multiple scattering contribution to the point spread function. Fast readout and single electron sensitivity in small pixel cells result in high resolution dynamic imaging. We have developed a prototype CMOS pixel sensor implementing radiation tolerant cells. The sensor response in terms of energy deposition, charge spread and point spread function has been extensively tested for electrons in the energy range of interest for TEM. A detailed simulation based on `Geant-4` and a dedicated charge generation and collection simulation package has been validated on the data recorded.

The point spread function measured with 300 keV electrons is (8.1 ± 1.6) μm for 10 μm pixel and (10.9 ± 2.3) μm for 20 μm pixels, respectively, which agrees well with the values of 8.4 and 10.5 μm predicted by our simulation.

The radiation tolerance of pixels with enclosed transistors and specific diode design has been verified with 200 keV electrons and 29 MeV protons. The pixel cells withstand doses in excess to 1 Mrad, without significant degradation in noise and charge collection efficiency, making them well suited for a sensor for direct detection in TEM.

Acknowledgements

We wish to thank Thomas Duden, Rolf Erni, Michael Johnson, Zhongoon Lee, Peggy McMahan, Marta Rossell Abrodos and the staff of the ALS and the 88-in. cyclotron for assistance and for the excellent performance of the accelerators. This work was supported by the Director, Office of Science of the US Department of Energy under Contract no. DE-AC02-05CH11231.

References

[1] E.R. Fossum, IEEE Trans. Electron. Devices ED-44 (1997) 1689.
[2] R. Turchetta, et al., Nucl. Instr. and Meth. A 458 (2001) 677.
[3] A.-C. Milazzo, et al., Ultramicroscopy 104 (2005) 152.
[4] G. Deptuch, et al., Ultramicroscopy 107 (2007) 674.

[5] P. Denes, J.-M. Bussat, Z. Lee, V. Radmillovic, Nucl. Instr. and Meth. A 579 (2007) 891.
[6] G.Y. Fan, et al., Ultramicroscopy 70 (1998) 113.
[7] A.R. Faruqi, D.M. Cattermole, C. Reburn, Nucl. Instr. and Meth. A 513 (2003) 317.
[8] A.R. Faruqi, et al., Ultramicroscopy 94 (2003) 263.
[9] A.R. Faruqi, R. Henderson, Curr. Opinion Struct. Biol. 17 (2007) 549.
[10] Yu. Gornushkin, et al., Nucl. Instr. and Meth. A 513 (2003) 291.
[11] M. Battaglia, D. Contarato, P. Giubilato, L. Greiner, L. Glesener, B. Hooberman, Nucl. Instr. and Meth. A 579 (2007) 675.
[12] M. Battaglia, et al., IEEE Trans. Nucl. Sci. 55 (6) (2008), doi:10.1109/TNS.2008.2007484.
[13] D. Joy, in: Monte Carlo Modeling for Electron Microscopy and Microanalysis, Oxford University Press, New York, 2005.
[14] S. Agostinelli, et al., Nucl. Instr. and Meth. A 506 (2003) 250.
[15] S. Chauvie, G. Depaola, V. Ivanchenko, F. Longo, P. Nieminen, M.G. Pia, Prepared for CHEP'01: Computing in High-Energy Physics and Nuclear, Beijing, China, 3–7 September 2001.
[16] M. Battaglia, Nucl. Instr. and Meth. A 572 (2007) 274.
[17] F. Gaede, Nucl. Instr. and Meth. A 559 (2006) 177.
[18] F. Gaede, T. Behnke, N. Graf, T. Johnson, in: Proceedings of the 2003 Conference for Computing in High-Energy and Nuclear Physics (CHEP 03), La Jolla, CA, 24–28 March 2003, TUKT001 [arXiv:physics/0306114].
[19] V. Chabaud, et al., Nucl. Instr. and Meth. A 292 (1990) 75.
[20] H. Bichsel, Nucl. Instr. and Meth. B 52 (1990) 136.
[21] M. Battaglia, et al., in: Proceedings of the International Symposium on Detector Development for Particle, Astroparticle and Synchrotron Radiation Experiments (SNIC 2006), Menlo Park, CA, 3–6 April 2006, ECONF-C0604032, p. 0108.
[22] H. Gabrisch, L. Kjeldgaard, E. Johnson, U. Dahmen, Acta Mater. 49 (2001) 4259.
[23] G. Lindström, Nucl. Instr. and Meth. A 512 (2003) 30.
[24] M.A. McMahan, Nucl. Instr. and Meth. B 241 (2005) 409.
[25] M. Battaglia, et al., IEEE Trans. Nucl. Sci. (2008), submitted for publication.

Data Acquisition and Trigger System of the Gamma Ray Energy Tracking In-Beam Nuclear Array (GRETINA)

John Anderson, Renato Brito, Dionisio Doering, Todd Hayden, Bryan Holmes, John Joseph, Harold Yaver, and Sergio Zimmermann, *Senior Member, IEEE*

***Abstract*—The Gamma Ray Energy Tracking In-Beam Nuclear Array (GRETINA), capable of determining the energy and position (within 2 mm) of each gamma-ray interaction point and tracking multiple gamma-ray interactions, has been designed. GRETINA will be composed of seven detector modules, each with four highly pure germanium crystals. Each crystal has 36 segments and one central contact instrumented by charge sensitive amplifiers. Two custom designed modules, the Digitizer/Digital Signal Processing (DSP) and the Trigger Timing and Control, compose the electronics of this system. The Digitizer/DSP converts the analog information with 14-bit analog to digital converters (operating at 100 MS/s, and digitally processes the data to determine the energy and timing information of the gamma interactions with the crystal. Each Digitizer/DSP is controlled by and sends trigger information to the Trigger Timing & Control system through a bidirectional Gbit link. Presently four different trigger algorithms are planned for the trigger system and can be selected for trigger decision. In this paper the details of the electronics and algorithms of the GRETINA data acquisition and trigger system are presented and the performance is reviewed.**

***Index Terms*—Data acquisition systems, digital signal processing, gamma-ray spectrometer, trigger systems.**

I. INTRODUCTION

We have designed a data acquisition and trigger system for the Gamma Ray Energy Tracking In-Beam Nuclear Array (GRETINA). GRETINA is based on germanium detectors and it will be capable of determining the energy and position (within 2 mm) of the points of interaction of the gamma-rays with the germanium crystal and of tracking multiple gamma-ray interactions [1], [2]. GRETINA is composed of seven detector modules, each with four high purity germanium crystals (see Fig. 1), comprising a quarter or 1-π of a sphere. The detector module components include charge sensitive amplifiers [3] assembled inside the detector enclosure to instrument each of the 36 segments and the central contact. The gamma ray interaction with the germanium crystal induces charge on the segments and central contact. The amplifiers integrate this charge and drive an analog voltage to the GRETINA front-end electronics.

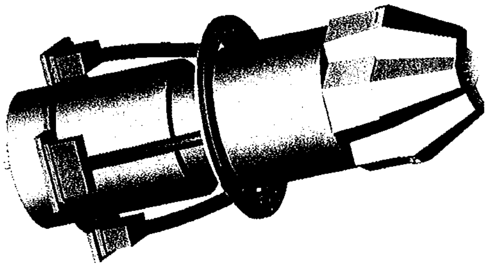


Fig. 1. GRETINA detector module.

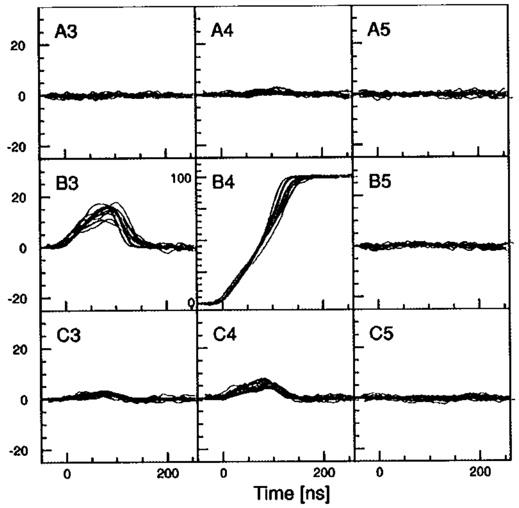


Fig. 2. Waveforms for interactions below segment B4.

For illustration, Fig. 2 shows examples of typical waveforms on the output of the amplifiers. The gamma-rays from a collimated ^{137}Cs source are interacting with the detector at a specific location below the B4 segment. The segments shown are the one that collects the charge (B4) and its eight nearest neighbor segments. The neighbor segments detect induced charge. A total of 16 measured shapes are plotted in gray and calculated signals at this given position are plotted in black; the agreement between calculation and measurement is very good. Observe that the measured signals include noise.

The detector modules are supported by a mechanical structure composed of two quasi-hemispheres shells that surround

Manuscript received January 22, 2008; revised August 14, 2008. Current version published February 11, 2009. This work was supported by the Director, Office of Science, Office of Nuclear Physics, of the U.S. Department of Energy under Contract DE-AC02-05CH11231.

J. Anderson and T. Hayden are with Argonne National Laboratory, Argonne, IL 60439 USA.

R. Brito was with the Lawrence Berkeley National Laboratory, Berkeley, CA 94720 USA. He is now with the Department of Electrical Engineering, Federal University of Rio Grande do Sul, Porto Alegre, RS 90035, Brazil.

D. Doering, B. Holmes, J. Joseph, H. Yaver, and S. Zimmermann are with the Lawrence Berkeley National Laboratory, Berkeley, CA 94720 USA (e-mail: szimmermann@lbl.gov).

Digital Object Identifier 10.1109/TNS.2008.2009444

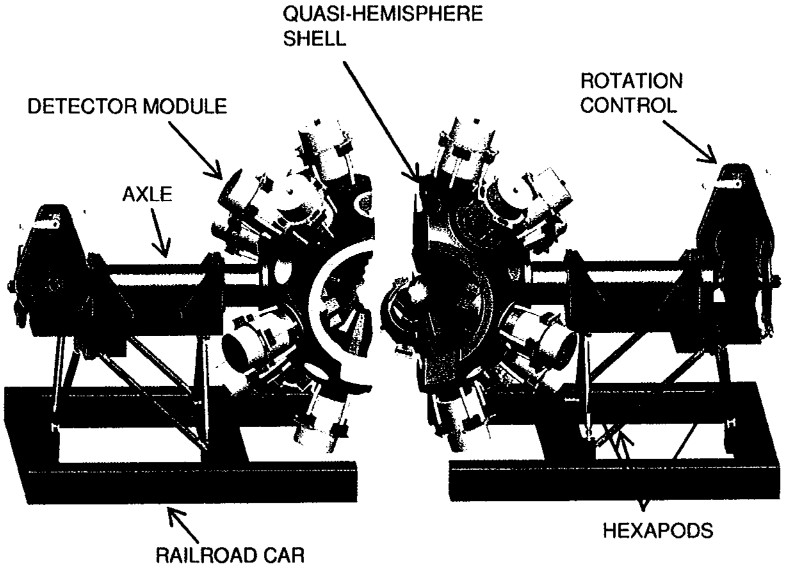


Fig. 3. GRETINA mechanical support structure.

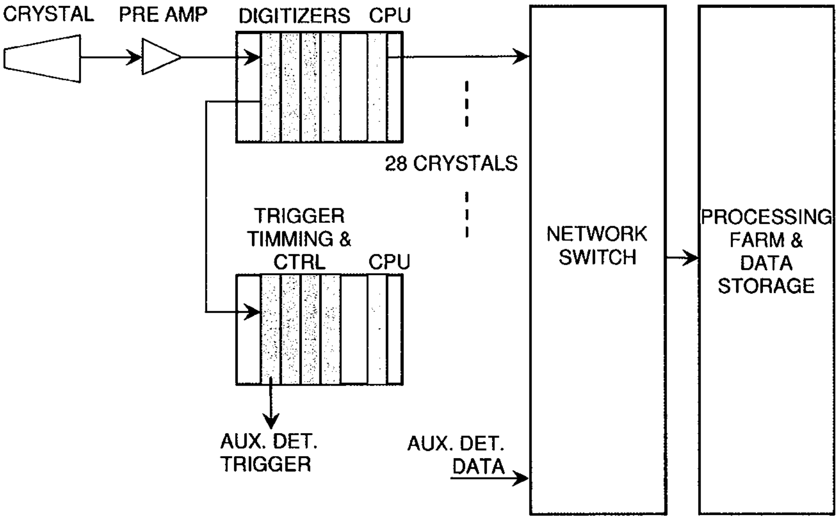


Fig. 4. GRETINA electronics and computing systems.

the target chamber. This shell structure is capable of supporting 21 detector modules. The structure allows rotation for detector mounting (through a gear box in the end of the axles) and translation to access the target chamber (through railroad cars). Hexapods connect the structure to the railroad cars. Fig. 3 shows a sketch of the support structure, with all possible positions instrumented with a detector module. There are plans to construct the full 4-π detector module array, GRETA [4], and the electronics system foresees this possible scenario.

In this paper the details of the electronics and algorithms of the GRETINA data acquisition and trigger system will be presented and the performance will be reviewed. In addition, grounding and filtering techniques used to achieve the 14-bit analog to digital conversion (ADC) performance will be discussed, as well as transmission line techniques for the very low bit error rate of the gigabit links.

II. System Architecture

Fig. 4 shows a block diagram of the GRETINA Electronics and Computing Systems. The oblong shape on the left represents the detector modules and its crystals. Charge sensitive amplifiers instrument the segments and central contact. Fifteen meters of shielded twisted pair cable connect the pre-amplifier outputs to the digitizer modules. Two custom designed modules, the Digitizer/Digital Signal Processing (Digitizer/DSP) and the Trigger Timing & Control (TTC), compose the electronics of this system. Four Digitizer/DSP modules instrument one crystal: each master Digitizer/DSP interfaces with the TTC system and controls three slave Digitizer/DSP modules. A digital bus in the front panel allows the digitizers serving a crystal to synchronize among themselves for clock and trigger information. A very simple communication protocol based on a single master controlling the operations is used.

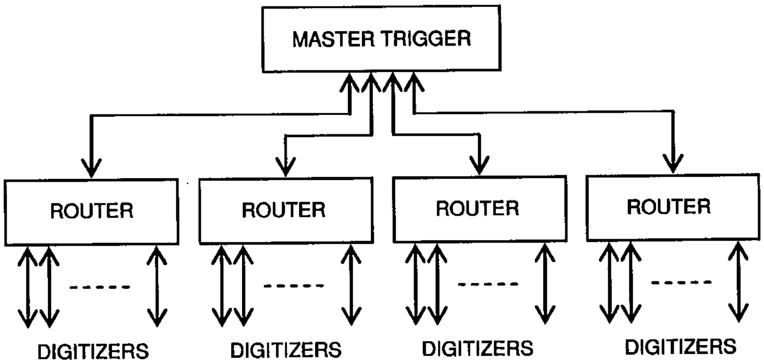


Fig. 5. TTC router and master connection.

The master digitizer monitors the crystal central contact. If a gamma ray deposits charge in the crystal above the programmable threshold of the leading edge (LE) discriminator, the master Digitizer/DSP recognizes the event and reads the segment hit pattern using the front panel bus. In parallel the Master digitizer estimates the energy of the central contact signal using a fast algorithm. It then assembles the trigger information (time stamp (TS) of the LE discriminator, central contact energy and segment hit pattern) and sends it to the TTC over one pair of a bidirectional 1 Gb/s serial link. A buffer records the timestamps of all recent discriminator decisions.

The serial link connecting the TTC and Digitizer/DSP modules is implemented using the National Semiconductor DS92LV18 Serializer/Deserializer (SerDes). It transmits 20 bits per word at a 50 MHz rate, where 18 bits are available for trigger and control information. The link meets the stability requirement of the system and has a predictable latency. The predictable latency allows the same twisted pair to transmit control information and provide the 50 MHz master clock to each Digitizer/DSP. The trigger information exchange between the master trigger and Digitizer/DSP employs a synchronous protocol and the master trigger module determines the synchronism. A synchronous implementation is easier to implement and maintain. The protocol consists of an endlessly repeating series of 20 command frames transmitted every 2 μs that allows the TTC system to regularly synchronize and control activities at the Digitizer/DSPs. Also, all trigger information from the Digitizer/DSP is transmitted to the TTC within this 2 μs window. The reason for a 2 μs period is to accommodate instantaneous fluctuations on the maximum average of one gamma interaction per crystal every 20 μs.

The whole TTC system is assembled around the same hardware module. The firmware is configured for two different functions: the Router module and the Master TTC. The Router routes all information between the Master TTC and the Digitizer/DSP modules and assists in fast multiplicity trigger decisions. The Master TTC collects all messages from the Routers plus additional information from auxiliary trigger modules and uses it to make a global trigger decision (refer to Fig. 5). Once a global trigger decision is made, the TTC system sends a trigger decision command to the Routers for distribution to all master Digitizer/DSP modules and auxiliary detectors. Each master Digitizer/DSP identifies a match between the timestamp embedded within the trigger decision message and the saved LE discriminator timestamps, and requests all slave Digitizer/DSP to transfer the data from the circular buffer into its own readout FIFO. Later, the VME readout CPUs in the digitizer crates read the event data from each Digitizer/DSP FIFO, assemble the crystal event, and send the data to the network switch. The switch routes the events to the computing farm where they are processed. The processing uses the segment information to estimate the position (r, θ, z) and energy of the interaction points. Additional processing establishes the tracks by connecting the individual interaction points.

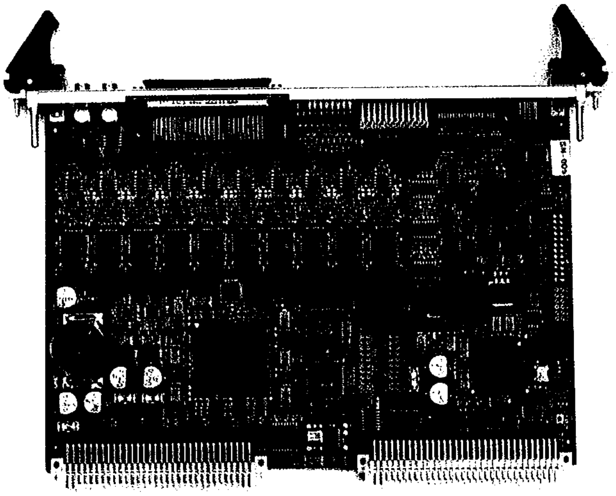


Fig. 6. Digitizer/DSP module.

We have extensively tested the performance of the DS92LV18 SerDes for bit error rate (BER). We have run 12 SerDes links in parallel for more than 10 consecutive days with no errors, which corresponds to a BER better than 10^{-16} or less than one error per day for GRETINA (four per day for GRETA). This component is a very simple SerDes that adds minimum protocol (an important characteristic for the constant latency). In order to guarantee the DC balancing of the communication we encoded the data stream in the FPGA. This encoding operates in the following way. First, before sending a word, the FPGA sums the number of zero and one bits in the word. Then it determines the DC balance of the line due to previous transmissions (i.e., if it has sent more zero bits or more one bits). Finally, to keep

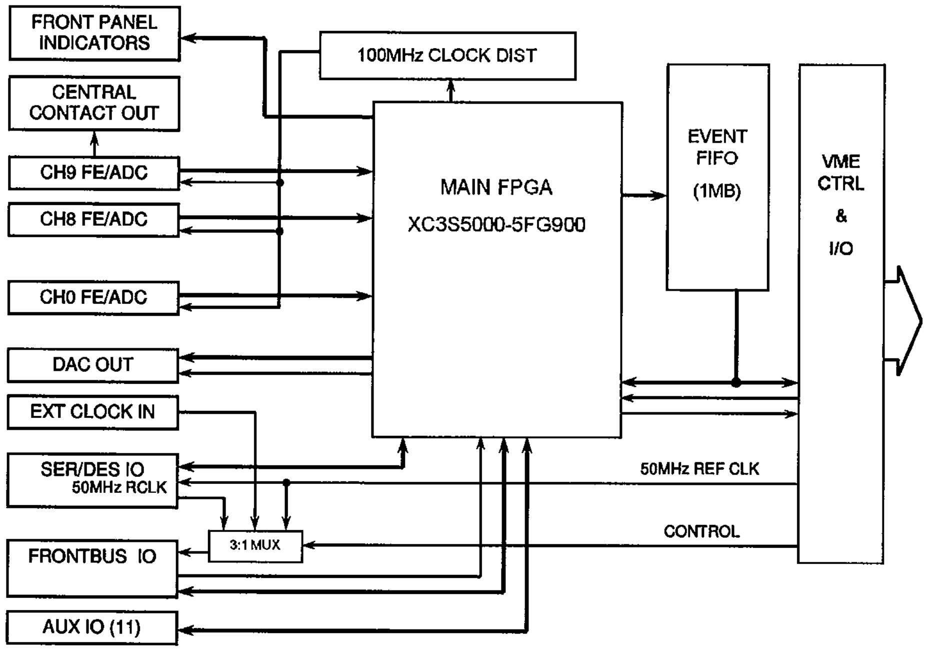


Fig. 7. Digitizer/DSP block diagram.

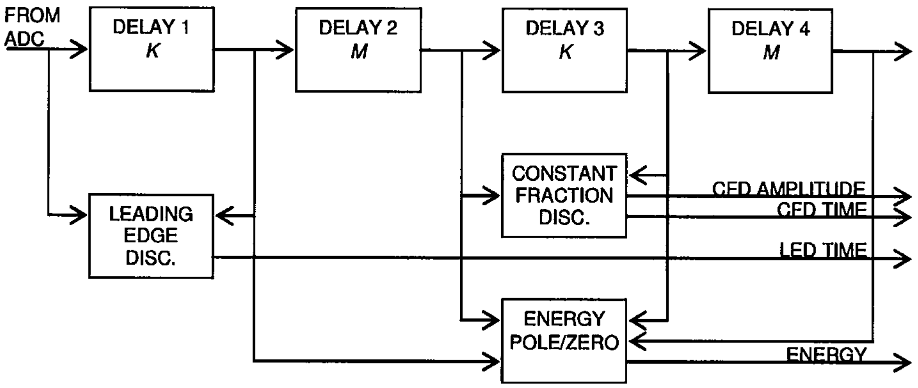


Fig. 8. Digital signal processing.

the line balanced, it then transmits the word itself or its complement. A dedicated bit in the word is set or reset to indicate to the receiver how to decode it. Also, we observed that replacing the cable from unshielded CAT5 to a good quality shielded CAT5 decreased the BER to the level reported above. LVDS drivers (DS90LV004) with programmable pre-emphasis are used in all modules to compensate for cable losses.

Each Digitizer/DSP and TTC module is implemented using two field programmable gate arrays (FPGA): one is a smaller FPGA, which controls the VME interface and has a steady configuration, and the other is a larger FPGA which executes the module specific algorithms and is easily re-configured through VME.

III. DIGITIZER/DSP

The Digitizer/DSP module samples the crystal segment and central contact analog information using 14-bit ADC operating at 100 MS/s. The ADC used is the AD6645 from Analog Devices. Fig. 6 shows a picture of the module and Fig. 7 shows its block diagram. The Digitizer/DSP has a total of 10 analog inputs. The ADCs are connected to an FPGA, which digitally processes the conversion and executes the following operations: leading and constant fraction discrimination, trapezoidal filtering, energy determination, and pole/zero cancellation.

Fig. 8 shows a block diagram of the digital processing algorithms and how they interconnect and Table I describes the algorithms. The processing is optimized to be implemented using the internal FPGA resources and occurs at a clock rate of 100 MHz, which is generated inside the FPGA by multiplying the 50 MHz master clock distributed from the TTC. This results in approximately 20 giga-operations/s. While the processing occurs, the raw data is stored in 40 μsec circular buffers (designed around the FPGA block RAMs). For further details about these algorithms refer to [5].

The present noise performance of the ADC inputs at full digital processing rate is approximately 1.5 RMS counts for all inputs (refer to Fig. 9). Achieving this performance in a board with digital power supplies and heavy digital activity required special

TABLE I
DIGITAL SIGNAL PROCESSING

Leading Edge Discrimination
$y(n)=x(n)-x(n\text{-}k)$ (differentiation)
$y(n)=(x(n)+x(n\text{-}2))+x(n\text{-}1)<<1$ (×4, Gaussian filtering)
Threshold comparison → LE discriminator time
Constant Fraction Discrimination
$y(n)=x(n)\text{ - }x(n\text{-}k)$ (differentiation)
$y(n)=(x(n)+x(n\text{-}2))+x(n\text{-}1)<<1$ (×2, Gaussian filtering)
$y(n)=x(n\text{-}k)<<ab\text{ - }x(n)$ (constant fraction)
Zero crossing comparison → CFD time
Trapezoidal filter and energy determination [6]
$y(n)=y(n\text{-}1)+((x(n)+x(n\text{-}2m\text{-}k))-(x(n\text{-}m)+x(n\text{-}m\text{-}k))$
Maximum tracking → energy
Pole-Zero cancellation
$I(n)=I(n\text{-}1)+x(n)$
$y(n)=x(n)+I(n)/t$ (where t is the pre-amp time constant)

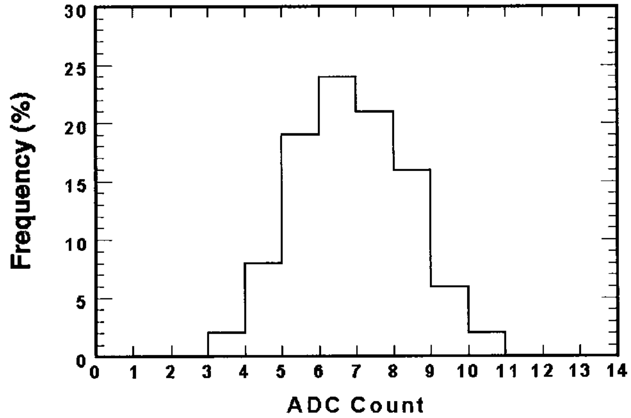


Fig. 9. Noise distribution: $\sigma = 1.5$ counts.

attention. We have filtered the +5 V with a π-filter (observe the inductor on the left portion of Fig. 6). Also, the analog inputs are all implemented using balanced differential techniques.

During prototyping we found two sources of noise due to board layout: one associated with a stray capacitance and the other with the analog and digital ground separation. For the first source we measured about 15 RMS counts and we observed that the ADC noise spectrum was not constant: it was mostly flat until ~ 15 MHz; then it increased by 6 dB/oct until ~ 30 MHz where it became mostly flat again. The increase at 15 MHz was traced to a stray capacitance between the input of the differential amplifier and the ground plane. This capacitance limited the amplifier feedback and, therefore, increasing the gain for higher frequencies (i.e., larger than ~ 15 MHz). At ~ 30 MHz the amplifier reached the maximum frequency response and the gain stabilized again. To reduce this stray capacitance we modified the layout and removed the ground plane right below the differential amplifier inputs. For the second source of noise we observed that the RMS noise performance was not similar for all ADC channels: the ADCs on the left portion of the board (i.e., close to the π-filter, refer to Fig. 6) had approximately 1.5 RMS counts while the ones on the right had approximately 3 RMS counts. We then changed the layout and cut the ground plane around the analog section to create an analog ground mostly free of digital switching. We connected the analog ground together with the digital ground under the ADCs (as suggested on the datasheet) and also close to the π-filter. With these two layout changes we now obtain approximately the same noise performance to all channels.

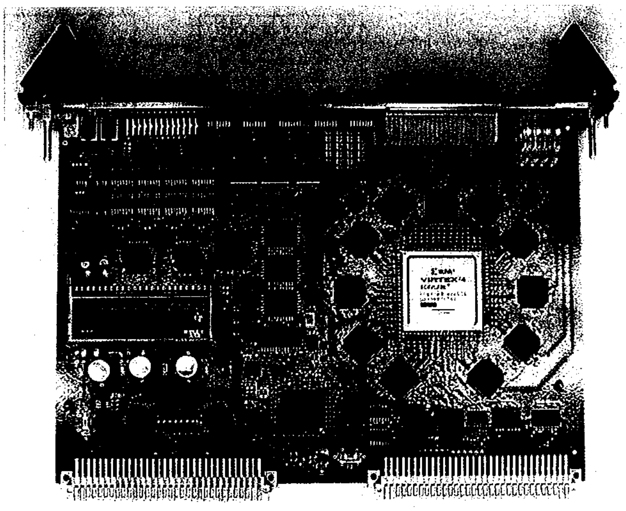


Fig. 10. Trigger, timing and control module.

The Digitizer/DSP module also has programmable digital inputs and outputs for interfacing with other units and front panel LE discriminators for status. All the processing described above is done in a XC3S5000 FPGA from Xilinx. Presently about 50% of the FPGA is utilized, allowing for further increase in algorithm complexity.

IV. TRIGGER, TIMING AND CONTROL

Fig. 10 and 11 show the picture and block diagram of the TTC module, respectively. The TTC is presently configured for multiplicity algorithm, and we intend to add three more in the next few months. When completed, the trigger algorithms will be:

a) *Multiplicity*: Uses the LE discriminator detection of the crystals central contacts. The TTC generates a trigger when it detected that the sum of LE discriminators crosses some threshold within a time window.
b) *Energy*: uses the energy of a central contact or the sum of energies of central contacts. When this energy falls within some programmable window the TCC generates a trigger.
c) *Pattern distribution*: It also uses central contact energy as described above. The TTC triggers when it detects coincidence of gamma-rays energy above threshold in any two pre-programmed crystals.
d) *Auxiliary detector trigger*: The TTC receives an external trigger within the 40 μsec pipeline depth of the Digitizer/DSP.

We call these algorithms prompt triggers, because a trigger is generated as soon as the proper condition is detected. A delayed coincidence trigger can further refine the meaning of these prompt triggers. When this option is enabled, the TTC generates a trigger command when it detects two prompt trigger conditions (e.g., two multiplicity triggers) in two time windows within the allowed overall trigger window. Finally, observe that all the trigger parameters are configurable through VME.

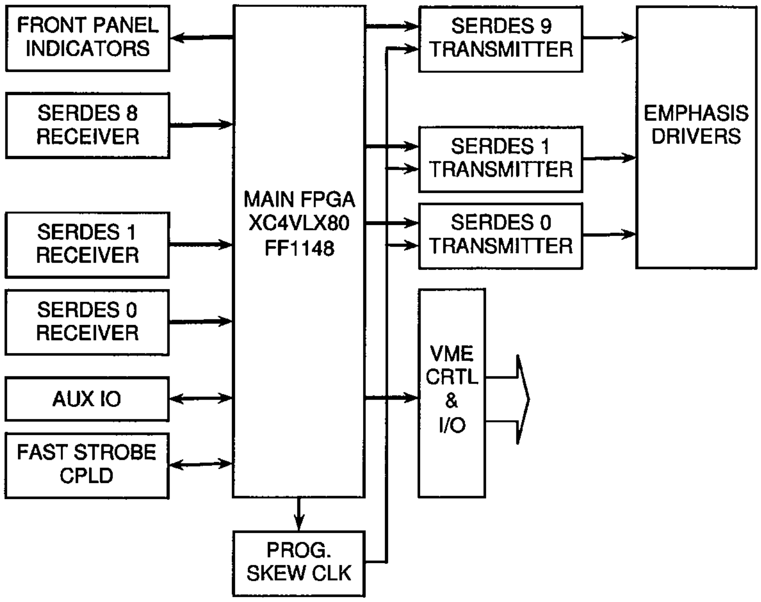


Fig. 11. Trigger, timing and control block diagram.

The main FPGA in the TTC module is the Virtex 4 series XC4VLX80 FPGA. This part executes the trigger algorithms and interfaces with eight SerDes. As already described, the SerDes components transfer information between the Digitizers/DSP. The base multiplicity algorithm presently uses only 15% of the resources inside this FPGA, which allows for further development of the trigger algorithms.

The high-speed switching of the SerDes has edge transition times in the order of 100 ps range. These high-speed signals require special attention during layout, and are more critical on the TTC (when compared with the Digitizer/DSP module), since it has eight SerDes and not all can be mounted close to the connector. Carefully matched, impedance controlled differential traces are used. Additionally, the number of vias was minimized: the component side transmits trigger information, and it does not have any vias, while the solder side is used to receive trigger information, and it has just one via per trace of the differential signal. To further improve signal fidelity, a mixed stackup is used where the dielectric under the component and solder layers of the board is Rogers 4350, which has lower losses than FR4.

The electronics requirements [7] state that the Digitizers/ DSPs shall sample all crystals within a 2 ns clock phase window. To synchronize all these modules, all CAT5 cables between the TTC and the Digitizer/DSP are approximately of the same length. However, small cable and component delay differences may generate delays that exceed this limit. To compensate for small delay variations and allow field tuning without recompiling the FPGA firmware, the TTC uses programmable skew clock buffers (Cypress Semiconductor CY7B992) to adjust the delay of the transmit clock (TCLK) of each SerDes. This, in turn, changes the delay of the sampling clock of the Digitizer/DSP at the other end of the cable. To adjust the clock phases we intend to use an oscilloscope monitoring multiple clock outputs on the front bus of master digitizers followed by manual reconfiguration of the VME registers controlling the programmable skew clock buffers. Also, the system phase adjustment can be validated before, during and after an experiment by measuring coincidences between any pair of detectors using a radioactive ^{60}Co source that provides two gamma-rays in prompt coincidence.

GRETINA may provide trigger to auxiliary detectors. Some of these detectors use old technology and they do not have digital pipelines; the trigger decision has to be done when the signals are actually traveling on cables, and the sampling has to start when they arrived at the auxiliary electronics. Only the multiplicity algorithm will be used when interfacing with this type of auxiliary detectors. We have implemented a parallel path that by-passes the SerDes to meet the more stringent timing requirements. The SerDes add a delay of approximately 70 ns to serialize and de-serialize the data. This new path uses a dedicated available pair of the CAT5 cable to transmit from the Digitizer/DSP to the TTC system the status of the central contact LE discriminator detection. To determine multiplicity, the TTC makes the sum of the LE discriminators using front panel connectors and CPLDs. The partial sums travel from the Routers to the Master Trigger module, which executes the final sum and determines the multiplicity. We estimate that this processing will take less than 250 ns. Also, it is interesting to highlight that this requirement also imposed one of the criteria to select the ADC: the latency of the AD6645 is just three clock cycles, which allows for faster LE discriminator determination.

The TTC module also has programmable digital inputs and outputs for interfacing with other units (e.g., auxiliary detectors & NIM electronics) plus front panel LED for status.

V. CABLE AND CONNECTOR SELECTION

We have dedicated substantial effort in selecting the proper cabling for this electronics. One example already discussed is

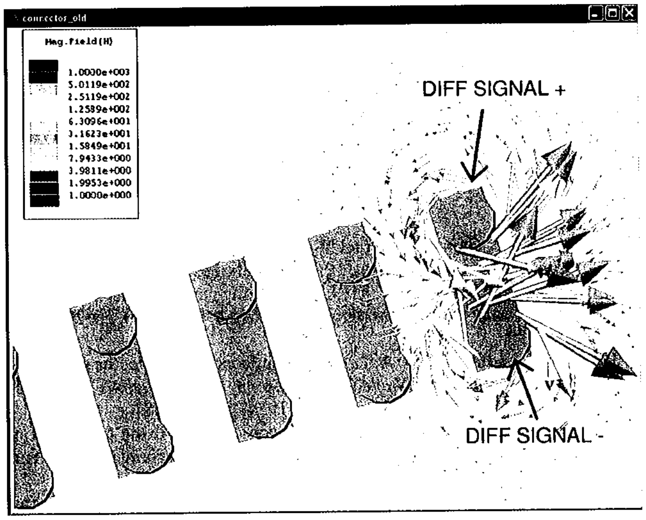


Fig. 12. Magnetic coupling in 0.1" spacing connectors.

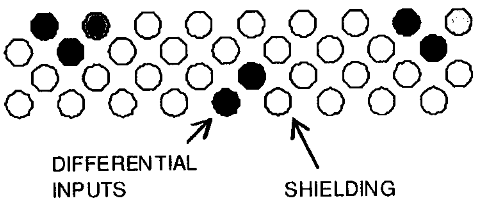


Fig. 13. IIT-Cannon connector pin assignment.

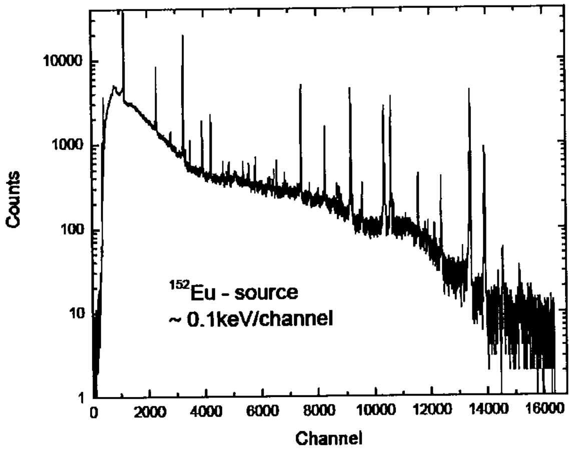


Fig. 14. Energy spectrum for a ^{152}Eu source.

the CAT5 cable for the SerDes. We will now describe the selection of another very important cable, the one that connects the detector pre-amplifiers to the Digitizers/DSPs, which carries the crystal interaction information. Reducing crosstalk between the signals in this interconnection is very important for estimating the position (r, θ, z) of the interactions points: excessive crosstalk decreases the position resolution of the gamma ray interaction. The requirement of total crosstalk in the detector module is less than 0.2% and, therefore, the rest of the electronics crosstalk has to be substantially less than this (required of $< 0.04\%$). To avoid ground loops we decided that the input of the Digitizers/DSPs would be differential and that we would employ twisted pairs. We tested several cables. We obtained the best performance with twisted individually shielded pairs. The crosstalk between adjacent pairs was $< 0.04\%$ for a 15 m cable span. For comparison, a similar cable, but non-twisted (used for LVDS transmission), had a crosstalk of $\sim 2.5\%$. This crosstalk is explained by the magnetic field created in one pair passing through the shielding and magnetically coupling to an adjacent pair. The shielding between pairs is a thin aluminum layer, and it does not attenuate the magnetic field significantly for these speeds. These tests considered a constant rise time (10%–90%) of 30 nsec, well within the expected rise times in GRETINA.

For the connector between the cable and the digitizer we have also tested several connector types, and we also observed crosstalk between signals. We traced the crosstalk, again, to the magnetic field of one signal coupling into the next signal. This can be visualized in Fig. 12. This figure shows the results of Maxwell simulations [from Ansoft] where the arrows represent the direction and the magnitude of the magnetic field. These simulations were done considering that the positive input of the differential signals are connected on the top row and the negative on the bottom row, forming these differential connections. The offending signal uses the differential connection on the right. One can observe that the magnetic field propagates well into the adjacent differential connections, causing crosstalk. The connector simulated has 0.1" pin spacing. Ways to mitigate this effect are threefold: (a) increasing the distance between differential connections, (b) selecting the direction of the pairs (ideally in an angle of 90°), and (c) short circuiting unused connector pairs (to create an opposing magnetic field that decreasing the offending field). For example, the simulations and subsequent tests have demonstrated that skipping two differential connections (on the right of the offending differential connection) and using the third for the next signal decreased the crosstalk by $\sim 10\times$, and that short circuiting these two unused set of pins reduces the crosstalk on the third pair by a further $\sim 2.5\times$. These tests were run with a 100 MHz sinusoidal signal. To meet the crosstalk mitigation strategies outlined above we used a 100 pin Double Density, Subminiature D type connector manufactured by ITT-Cannon. Fig. 13 shows a sketch of how we connected the differential pairs (represented by back circles) and its individual shielding (represented by gray circles). The figure shows the sketch of just a portion of the connector. The unused pins were short circuited in pairs. With this arrangement the measured crosstalk was below what we can measure using 14-bits ADC of the Digitizer/DSP (i.e., crosstalk $< 0.025\%$). The signal used had a constant rise time (10–90%) of 30 nsec.

VI. Test Results and Conclusions

We have tested the performance of the DAQ and trigger system connected to a detector module and we will now describe a few results we have obtained. For these tests we used the production cables and floating low- and high-voltage (bias) supplies. Fig. 14 shows a ^{152}Eu source energy spectrum obtained with the central contact of one of the crystals.

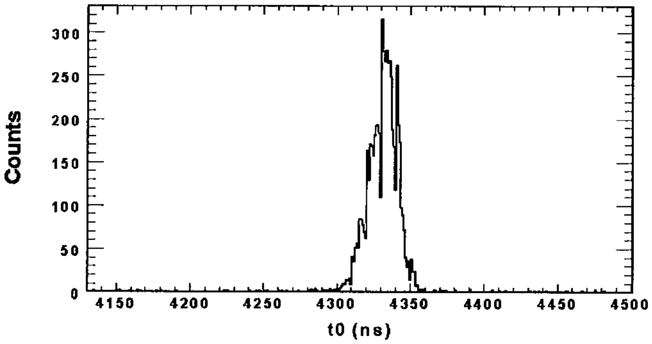


Fig. 15. Constant Fraction discriminator performance.

Fig. 2 shows waveforms that are similar to the waveforms the DSP/Digitizer module records. Fig. 15 shows a time spectrum used to measure the performance of the constant fraction discriminator in the digital processing. The T_O, derived from the FPGA algorithm is referenced to a very fast external trigger provided by a CsF scintillator with essentially no contribution to the width. The standard deviation of the Gaussian fit is 9 ns.

All tests with the prototype system indicate that the DAQ and Trigger systems for GRETINA will meet the requirements. We are presently preparing the production of the system to enter in operation in 2011.

ACKNOWLEDGMENT

The authors would like to thank the GRETINA Advisory Committee and M. Cromaz, S. Gros, I-Y. Lee, A. Macchiavelli and D. Radford for their valuable suggestions.

REFERENCES

[1] I. Y. Lee, "Gamma-ray tracking detectors," *Nucl. Instrum. Methods Phys. Res. A*, vol. A422, pp. 195–200, Feb. 1999.

[2] M. A. Deleplanque *et al.*, "GRETA: utilizing new concepts in γ-ray detection," *Nucl. Instrum. Methods Phys. Res. A*, vol. A430, pp. 292–310, Jul. 1999.

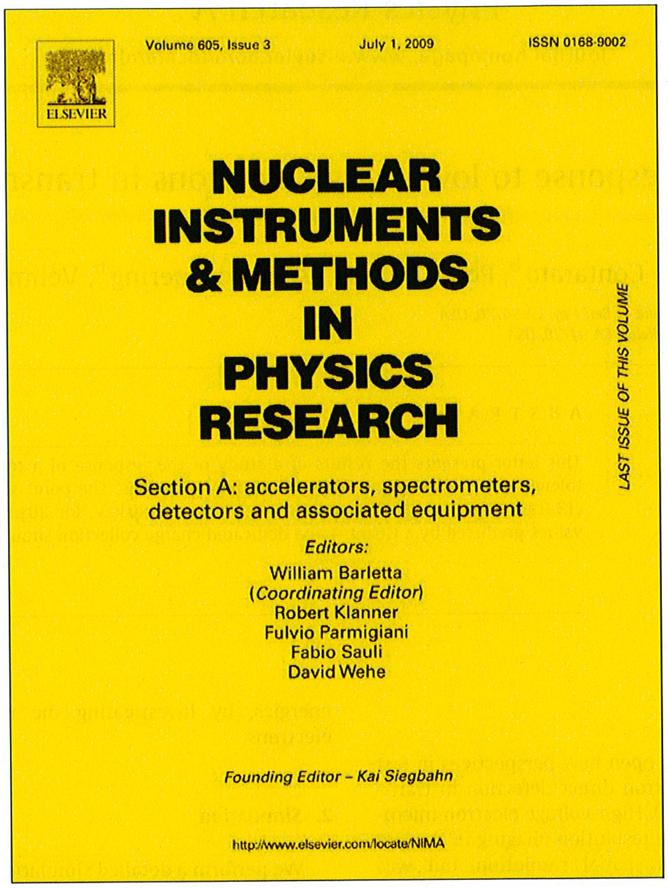
[3] F. S. Goulding and D. A. Landis, "Semiconductor detector spectrometer electronics," in *Nuclear Spectroscopy and Reactions Part A*, J. Cerny, Ed. New York: Academic, 1974, pp. 413–481.

[4] The Future of Gamma-Ray Spectroscopy: GRETA, the Gamma-Ray Energy Tracking, 2006 [Online]. Available: http://fsunuc.physics.fsu.edu/~gretina/GRETA_WP_Jan07_4.pdf

[5] M. Cromaz *et al.*, "A digital signal processing module for gamma-ray tracking detectors," *Nucl. Instrum. Methods Phys. Res. A*, vol. A597, pp. 233–237, Dec. 2008.

[6] V. T. Jordanov and G. F. Knoll, "Digital synthesis of pulse shapes in real time for high resolution radiation spectroscopy," *Nucl. Instrum. Methods Phys. Res. A*, vol. A345, pp. 337–345, Jun. 1994.

[7] F. Bieser, I.-Y. Lee, A. Macchiavelli, J. Weizeorick, and S. Zimmermann, GRETINA Electronics Requirement Document LBNL, Berkeley, 2005, Doc. GRT-3-041115.



Volume 605, Issue 3 July 1, 2009 ISSN 0168-9002

ELSEVIER

NUCLEAR
INSTRUMENTS
& METHODS
IN
PHYSICS
RESEARCH

Section A: accelerators, spectrometers,
detectors and associated equipment

LAST ISSUE OF THIS VOLUME

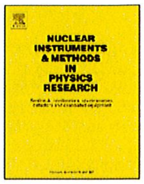
http://www.elsevier.com/locate/NIMA

Nuclear Instruments and Methods in Physics Research A 605 (2009) 350–352

Contents lists available at ScienceDirect

Nuclear Instruments and Methods in Physics Research A

journal homepage: www.elsevier.com/locate/nima



Letter to the Editor

CMOS pixel sensor response to low energy electrons in transmission electron microscopy

Marco Battaglia [a,b,*], Devis Contarato [b], Peter Denes [b], Dionisio Doering [b], Velimir Radmilovic [b]

[a] *Department of Physics, University of California at Berkeley, CA 94720, USA*
[b] *Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*

ARTICLE INFO

Article history:
Received 26 March 2009
Accepted 27 March 2009
Available online 10 April 2009

Keywords:
Monolithic active pixel sensor
Transmission electron microscopy

ABSTRACT

This letter presents the results of a study of the response of a test CMOS sensor with a radiation tolerant pixel cell design to 80 and 100 keV electrons. The point spread function is measured to be (13.0 ± 1.7) μm at 100 keV and (12.1 ± 1.6) μm at 80 keV, for 20 μm pixels. Results agree well with values predicted by a Geant-4 and dedicated charge collection simulation.

Published by Elsevier B.V.

1. Introduction

Monolithic CMOS pixel sensors open new perspectives in fast nano-imaging through single electron direct detection in transmission electron microscopy (TEM). High-voltage electron microscopy, developed and used for high resolution imaging in the late 1970s [1] produced advances in spatial resolution, but was abandoned due to the severe displacement damage of the sample. As the displacement damage threshold is proportional to $\sqrt{E}$, there is now much interest in TEM of organic samples with energies of 80–100 keV, where recent advances in electron optics ensure deep sub-angstrom spatial resolution [2]. For example, the maximum energy transferred by an 80 keV electron to a carbon atom is 15.6 eV, which is below the threshold for knock-on damage. This makes low energy TEM necessary for atomic resolution studies of samples such as single atomic layers of carbon in graphene or carbon nanotubes [3] and in biology. There are two main issues to be considered for imaging with low energy electrons. The first is the large fluctuations in the energy deposition. The second is the degradation of the point spread function (PSF) due to the $1/E$ increase of the electron multiple scattering in the detector.

In an earlier paper [4], we presented the design of a radiation tolerant CMOS pixel cell and investigated the response of 10 and 20 μm pixels to electrons in the energy range 120 keV up to 300 keV for TEM. In this letter we extend that study to lower energies, by investigating the response to 80 and 100 keV electrons.

2. Simulation

We perform a detailed simulation of the charge deposition and signal formation in the CMOS pixel sensor based on the `Geant-4` program [5] using the low energy electromagnetic physics models [6]. The CMOS sensor is modelled according to the detailed geometric structure of oxide, metal interconnect and silicon layers, as specified by the foundry. Electrons are incident perpendicular to the detector plane and tracked through the sensor. For each interaction within the epitaxial layer, the ionisation point position and the amount of energy released are recorded.

Charge collection in the sensor is simulated with `PixelSim`, a dedicated digitisation module [7], developed in the `Marlin` C++ reconstruction framework [8], originally deployed for the International Linear Collider particle physics project. The processor starts from the ionisation points generated along the particle trajectory by `Geant-4` and stored in `lcio` format [9]. The `PixelSim` simulation models diffusion of charge carriers from their production point in the epitaxial layer to the collection diode. This provides us with full simulation of the response of each individual pixels in the detector matrix, including sensor geometry and electronics noise effects, which can be processed through the same analysis chain as the experimental data. The simulation has a single free parameter, the diffusion parameter σ_{diff}, used to determine the width of the charge carrier cloud. Its value is extracted from data by a χ^2 fit to the pixel multiplicity in the clusters of 1.5 GeV electrons since, at this energy, the multiple

---

[*] Corresponding author at: Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. Tel.: +1 510 486 7029.
E-mail address: MBattaglia@lbl.gov (M. Battaglia).

0168-9002/$ - see front matter Published by Elsevier B.V.
doi:10.1016/j.nima.2009.03.249

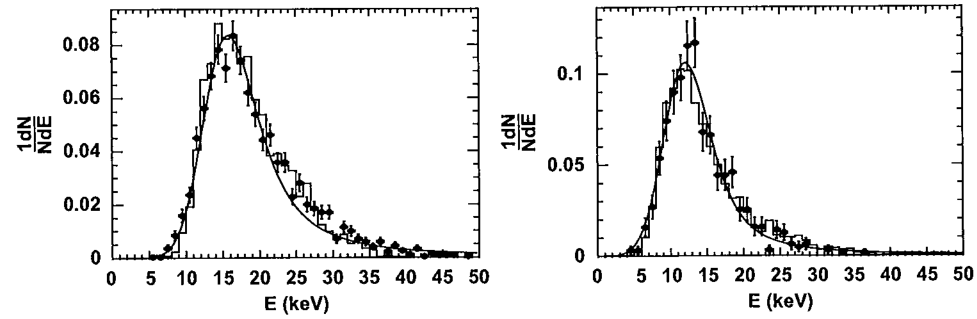


Fig. 1. Reconstructed deposited energy in a 3 × 3 pixel matrix for 80 keV (left) and 100 keV electrons (right). The points with error bars show the data and the histogram as the result of the `Geant-4` simulation. The continuous line shows a Landau function convoluted with a Gaussian noise term fit to the data.

scattering contribution to the charge distribution is negligible. We find $\sigma_{\mathrm{diff}} = (16.3 \pm 1.4)\,\mu$m, which agrees well with the diffusion length estimated from the doping in the epitaxial layer and the charge collection time [4].

3. Measurement

The detector charge-to-voltage conversion is 0.98 keV/ADC count or 26.7e$^-$/ADC count at 6.25 MHz readout frequency, obtained by recording the position of the 5.9 keV full energy peak of a collimated 2.2 mCi ^{55}Fe source.

We use the TITAN test column at the National Center for Electron Microscopy (NCEM) to characterise the detector response to 80 and 100 keV electrons and validate the simulation. The signal pulse height in a 3 × 3 matrix around each seed pixel having a signal-to-noise in excess of 4.5 is shown in Fig. 1 for data and simulation. The broadening of the energy distribution compared to electrons of higher energy is evident, however, the ratio of the Landau width to the Landau most probable value does not increase significantly compared to that for electrons of higher energy. We estimate the uncertainty on the number of electrons per pixel that can be reconstructed from the measured pulse height in a single pixel. We simulate a flat field illumination by generating multiple electrons hitting each pixel and reconstruct the pixel pulse height. This accounts for cross-feed between neighbouring pixels due to charge diffusion and multiple scattering. We determine the number of electrons on each pixel by dividing the simulated pixel pulse height by the average pulse height induced by a single electron and study the distribution of the reconstructed number of electrons as a function of that simulated. We find that the relative uncertainties on this number scale from 0.17 (0.14) for 10e$^-$/pixel to 0.12 (0.10) for 20e$^-$/pixel and to 0.08 (0.06) for 50e$^-$/pixel at 80 keV (100 keV), respectively. These results are comparable to relative uncertainties of 0.15, 0.11, 0.07 obtained for 200 keV electrons.

Finally, we determine the point spread function following the same method discussed in Ref. [4]. We reconstruct the image of a gold wire with a diameter measured to be $(59.6 \pm 0.7)\,\mu$m and mounted parallel to the pixel columns, at a distance of $\simeq 3$ mm from the detector surface. The profile of the deposited energy in the pixels, measured across the wire allows us to determine the charge spread due to electron multiple scattering and charge carrier diffusion. We describe the measured pulse height on the pixel rows across the image projected by the wire with a box function having the same width as the measured wire diameter smeared by a Gaussian term, which describes the point spread function. The contrast factor, i.e. the ratio of maximum to minimum pulse height levels, for the pixels away from the wire shadow and for those exactly below the wire centre, respectively, are set to those observed in data and we perform a simple 1-parameter χ^2 fit to extract the Gaussian width term, which gives the estimation of the PSF. Results are given in Table 1 and Fig. 2. A good agreement is found between the measurement and the prediction from simulation. These results are compared with those obtained at higher energies, presented in Ref. [4]. It is interesting to observe how the degradation of the PSF at decreasing energies, caused by multiple scattering, reaches a plateau around 120 keV. This is due to the decrease of the electron range with the particle energy, which limits the distance over which charge can be spread. In fact, we measure a point spread function value for 80 and 100 keV electrons which is compatible with that measured at higher energies, as predicted by the simulation.This result is quite encouraging for extending the application of CMOS pixel sensors to fast TEM imaging of organic and biological samples with low energy electrons.

Table 1
Point spread function predicted by `Geant 4+PixelSim` and measured with data for 20 μm pixel pitch.

Energy (keV)	`Geant-4+PixelSim` (μm)	Data 20 μm pixels (μm)
80	12.7 ± 0.5	12.1 ± 1.6
100	13.2 ± 0.5	13.0 ± 1.7

The uncertainty quoted for simulation is the systematics from σ_{diff} that for data accounts for statistical and systematics from pixel response equalisation.

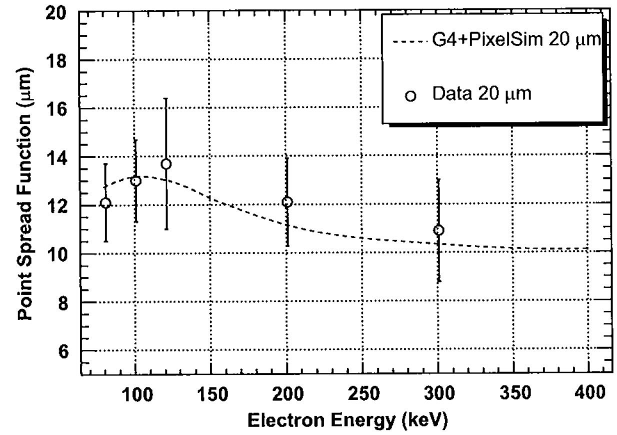


Fig. 2. Point spread function vs. electron energy with data (points with error bars) compared to simulation (lines) for 20 μm pixels. Data points from 120 to 300 keV are from Ref. [4]. The new measurements at 80 and 100 keV show the saturation of the point spread function contribution from multiple scattering in the sensor due to the reduced range of electrons as predicted by the simulation, shown by the line.

Acknowledgements

We wish to thank Thomas Duden, Rolf Erni and Zhongoon Lee. This work was supported by the Director, Office of Science, of the U.S. Department of Energy under Contract no. DE-AC02-05CH11231.

References

[1] V.E. Cosslett, et al., Nature 281 (1979) 49.
[2] C.O. Girit, et al., Science 323 (2009) 1705.
[3] J.C. Meyer, C. Kisielowski, R. Erni, M.D. Rossell, M.F. Crommie, A. Zettl, Nano Lett. 8 (2008) 3582.
[4] M. Battaglia, et al., Nucl. Instr. and Meth. A 598 (2009) 642 (arXiv:0811.2833 [physics.ins-det]).
[5] S. Agostinelli, et al., Nucl. Instr. and Meth. A 506 (2003) 250.
[6] S. Chauvie, G. Depaola, V. Ivanchenko, F. Longo, P. Nieminen, M.G. Pia, Prepared for CHEP'01: Computing in High-Energy Physics and Nuclear, Beijing, China, 3–7 September 2001.
[7] M. Battaglia, Nucl. Instr. and Meth. A 572 (2007) 274.
[8] F. Gaede, Nucl. Instr. and Meth. A 559 (2006) 177.
[9] F. Gaede, T. Behnke, N. Graf, T. Johnson, in: Proceedings of the 2003 Conference for Computing in High-Energy and Nuclear Physics (CHEP 03), La Jolla, California, 24–28 March 2003, pp. TUKT001 (arXiv:physics/0306114).

Use of High Resolution DAQ System to Aid Diagnosis of HD2b, a High Performance Nb_3Sn Dipole

Juan Lizarazo, Dionisio Doering, Lawrence Doolittle, James Galvin, Shlomo Caspi, Daniel R. Dietderich,
Helene Felice, Paolo Ferracin, Arno Godeke, J. M. Joseph, A. F. Lietzke, Alessandro Ratti, GianLuca Sabbi,
Frederic Trillaud, X. R. Wang, and Sergio Zimmerman

***Abstract*—A novel voltage monitoring system to record voltage transients in superconducting magnets is being developed at LBNL [1]. This system has 160 monitoring channels capable of measuring differential voltages of up to 1.5 kV with 100 kHz bandwidth and 500 kS/s digitizing rate. This paper presents analysis results from data taken with a 16 channel prototype system. From that analysis we were able to diagnose a change in the current-temperature margin of the superconducting cable by analysing Flux-Jump data collected after a magnet energy extraction failure during testing of a high field Nb_3Sn dipole.**

***Index Terms*—HD2, Nb_3Sn, superconducting magnets.**

I. INTRODUCTION

SEVERAL improvements to magnet test and diagnostics have been made over the past year by the Superconducting Magnet program at Lawrence Berkeley National Laboratory (LBNL). The new and under development systems include: magnet voltage monitoring, magnetic measurements, magnet fault trigger, strain gauge measurements, power supply control, etc. The end purpose of all these upgrades is to better understand the magnets being developed, and in this way provide better feedback to the design and manufacturing process. This paper focus on the use of a novel magnet data acquisition system to obtain an insight on a change in performance after an energy extraction failure during the training of a high field dipole magnet called HD2b. After that, this paper will discuss on the steps being taken on studying the possibility of expanding our data acquisition system by developing a multichannel Flux-Jump antenna with high spatial resolution.

II. HD2b TEST INCIDENT

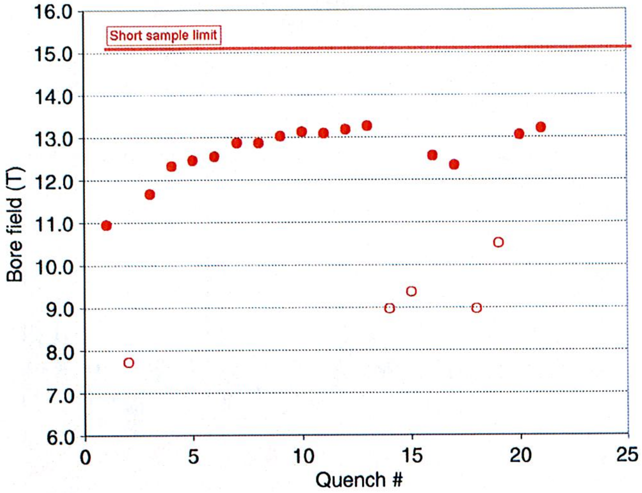


Fig. 1. HD2b training curve. Solid-dots correspond to training quenches, empty circles are quenches due to excessive ramp rate.

HD2a, HD2b and HD2c are part of an ongoing series of high field Nb_3Sn race track dipoles being design, assemble and tested at LBNL [2], [3]. During testing of HD2b on March 30 2008, after performing 13 successful training quenches, an SCR in the energy extraction rack broke during extraction of the magnet at the end of quench 13 (Q13). After this failure, a change in the magnet performance was observed in subsequent training ramps.

Fig. 1 shows the training curve for HD2b. The solid dots correspond to training quenches and the empty circles correspond to quenches due to excessive ramp rate.

Q13 occurred at 15.127 kA, after that, the quenching current went down to 9.928 kA for Q14 and 10.393 kA for Q15. The power supply ramping rate was then changed from 50 A/s up to 11 kA and 20 A/s to quenching, to a ramping rate of 50 A/s up to 5 kA, then 20 A/s up to 11 kA and 10 A/s up to quenching. Small variations of this current ramping profile were used in subsequent ramps, reaching the same level of performance than Q13 at Q21 with a quenching current of 15.059 kA.

When a magnet quenches from superconducting state the extraction rack must shut down the current fast enough to avoid excess heat, but slow enough to keep the magnet from arching. Fig. 2 shows the derivative of the voltage in each of the two coils in HD2b during a normal magnet extraction. Time 0 corresponds to the trigger point; at time equal -1 ms the magnet started quenching; at time $+5$ ms the protection heaters were

Manuscript received August 27, 2008. First published June 30, 2009; current version published July 15, 2009. This work was supported by the Director, Office of Energy Research, Office of High Energy and Nuclear Physics, High Energy Physics Division, U.S. Department of Energy under Contract DE-AC02-05CH11231.

The authors are with Lawrence Berkeley National Laboratory, Berkeley, CA 94720 USA (e-mail: jlizarazo@lbl.gov).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TASC.2009.2019036

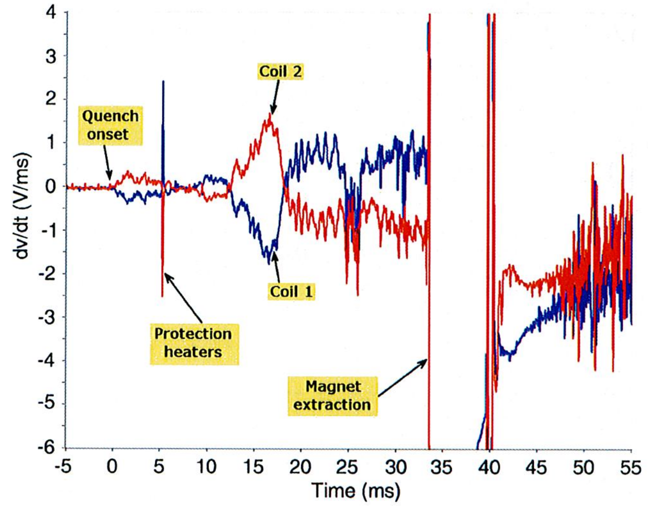


Fig. 2. Coil voltages during a normal magnet extraction. A negative dV/dt after extraction indicates decay in magnet current.

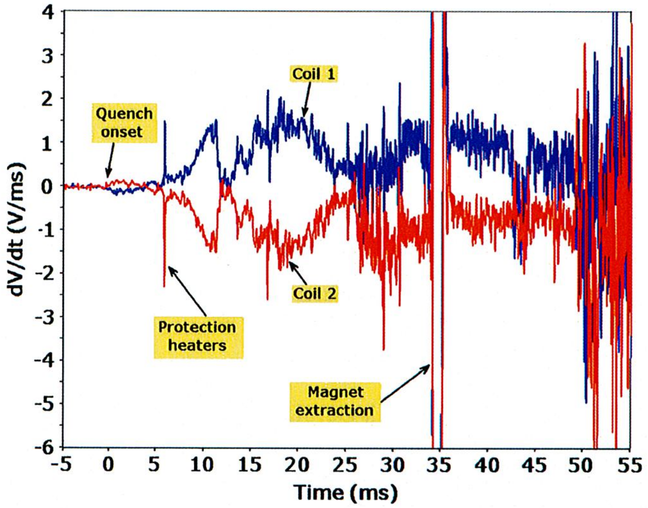


Fig. 3. Coil voltages during HD2b magnet extraction at the end of quench 13. There is no indication of current decay after extraction.

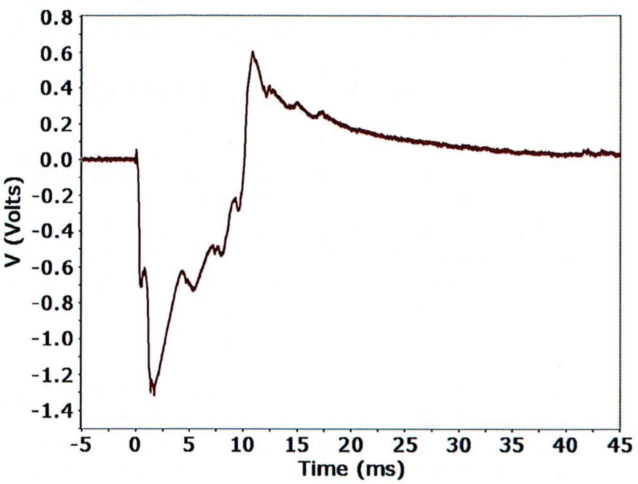


Fig. 4. Voltage unbalance induced between the two coils in HD2b by a typical Flux-Jump.

fired; and at time +34 ms the extraction system was triggered. A normal extraction time is around 6 ms and afterwards the value of dV/dt becomes negative for both coils, indicating that the current in the magnet is decaying.

Fig. 3 corresponds to the derivative of the voltage in each coil during energy extraction at the end of Q13. It shows that the magnet extraction time was reduced to less than 1 ms, and the value of dV/dt in coil1 remained positive and practically unchanged, suggesting persistence of the magnet current. Afterwards, the SCR in the extraction rack that failed and caused this abnormal behavior was replaced and operations resumed the next day.

III. HD2b Energy Extraction Failure Diagnostic Approach

After the change in performance shown in Fig. 1, a basic question arose on whether or not the energy extraction system failure at the end of Q13 had any effect on the Jc distribution in the magnet. With the purpose of answering this question, data from a prototype high resolution Magnet Voltage Monitoring System (MVMS) [1] developed by the Superconducting Magnet Program at LBNL was used.

MVMS on its current developmental stage consists of 16 channels capable of measuring differential voltages of up to 1500 V with a common mode of 1500 V. Each channel has a bandwidth of 100 kHz and is being used to collect data at a sampling rate of 400 kS/s during a window of 100 ms around a trigger point. This system was used to monitor the voltage along 8 sectors on each coil of HD2b, and it allowed us to record with unprecedented resolution the fast voltage transients that occur during the ramping of the magnet current. Two kinds of events were collected, Flux-Jumps (FJ) and Slip-Sticks (SS), named after their hypothesized process of origin.

Fig. 4 shows a typical FJ collected during HD2b training. FJs are caused by loss of diamagnetism in a volume of superconductor, producing a change in B field and inducing a voltage signal across the terminals of the magnet [4]–[6].

Fig. 5 shows a signal from a typical SS collected during HD2b training, SS are mechanical vibrations caused when dynamic competing internal forces cause sections of the magnet to detach and release energy in the form of dumped oscillations [4].

FJs are seen as an undesired effect due to the substantial amount of energy they release and its potential to quench the cable from superconducting state. However, they can also be seen as a way to diagnose changes in Jc distribution. In particular, the flux change induced by a FJ is an indication of the current-temperature margin of the volume of superconductor where the FJ took place.

To use FJs as a diagnostic tool, the distribution of integrated flux unbalance per FJ for two training ramps before and two after Q13 were compared. For that purpose a software analysis tool was developed, which allowed us to process hundreds of events in a semi-automatic way. This program runs over all events collected during a training ramp and decides whether an event is a FJ or a SS by using a Fourier filter and then setting a threshold of

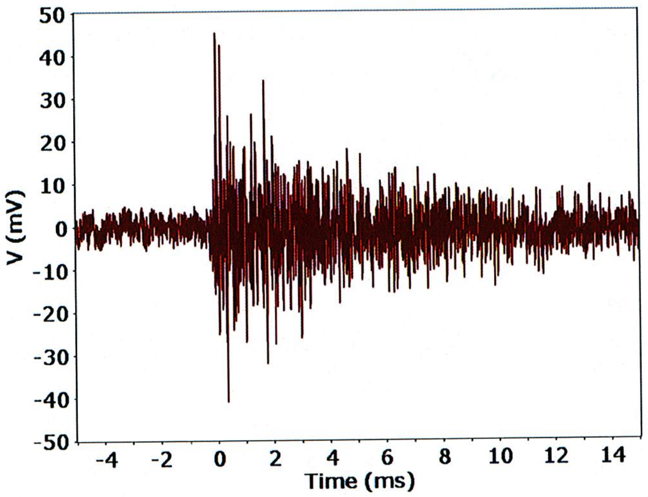


Fig. 5. Voltage unbalance induced between the two coils in HD2b by a typical Slip-Stick.

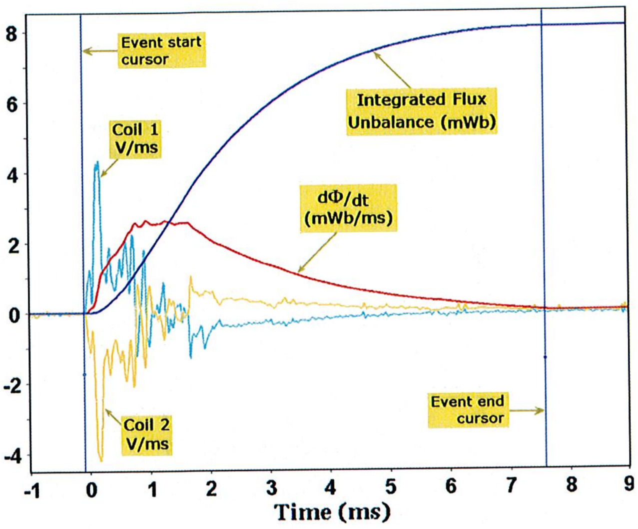


Fig. 6. Flux-Jump signal after processing by the analysis software.

acceptance on the output. The program then identifies the beginning and end of the event within the 100 ms window collected by the MVMS, and places a cursor at each boundary. At this point the user is given the option of rejecting the event, accepting the recognition done by the software, or changing the type of event and setting new event boundaries before going into further processing.

Fig. 6 shows a typical event after processing. The traces labeled as"Coil1" and "Coil2" represent the derivative of the voltage in coils 1 and 2 during this particular event in units of V/ms. The trace labeled as "$d\Phi/dt$" represents the voltage induced in coil 1 minus the voltage induced in coil 2 in Volts, which corresponds to the time derivative of the integrated flux unbalance. The integrated flux unbalance is the integrated flux crossing coil 1 minus the integrated flux crossing coil 2. A coil integrated flux refers to the sum over all its turns of the flux crossing each turn in the coil, which is only a fraction of the actual flux change produced by the FJ.

The trace labeled as "Integrated Flux Unbalance" in Fig. 6 is the time integral of $d\Phi/dt$ in miliWebers. Multiple shapes of this parameter had been collected, reflecting the complex way flux lines go between coils and affect the instantaneous Jc margin through the cable, potentially triggering similar events in other parts of the magnet. The value of the integrated flux unbalance at the end of the event is the net integrated flux unbalance left by the event. In this manner this parameter refers to integration over area and time of the difference in flux enclosed by the coils. The cursors shown in Fig. 6 correspond to the start and end boundaries set for this event.

IV. HD2b FLUX-JUMP OBSERVATIONS

All events collected during training ramps Q11, Q12, Q14 and Q15 up to 9.7 kA were analysed using the analysis tool described above. Q11 and Q12 were collected before the energy extraction failure at the end of Q13, while Q14 and Q15 were collected after the failure. These ramps were chosen because they share the same ramp rate and trigger threshold. Fig. 7 shows two histograms corresponding to the distribution of net integrated flux unbalance per FJ for the two ramps before and for the two ramps after Q13.

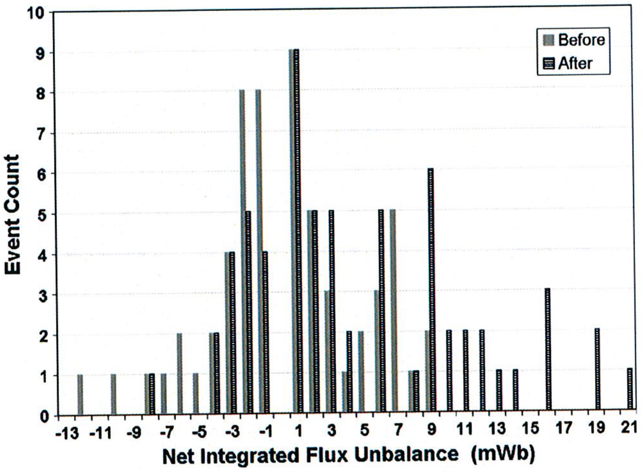


Fig. 7. Net integrated flux unbalance histogram before and after failure.

TABLE I
FJ EVENT COUNT FOR COILS 1 AND 2 BEFORE AND AFTER Q13

	Coil 1	Coil 2
Num Events Before	31	29
Num Events After	47	16

Each point making up these histograms corresponds to the net integrated flux unbalance for a particular event. A positive (*negative*) value means that the event mostly or entirely occurred in coil 1 (*coil 2*). Thus, the histogram made up of events before Q13 suggests a symmetrical FJ activity between the two coils, while the histogram with events after Q13 suggests a FJ activity bias towards coil 1. The event count for each coil before and after Q13 is summarized in Table I.

Fig. 8 shows the net integrated flux unbalance before and after Q13 as a function of event count using only data from coil 1.

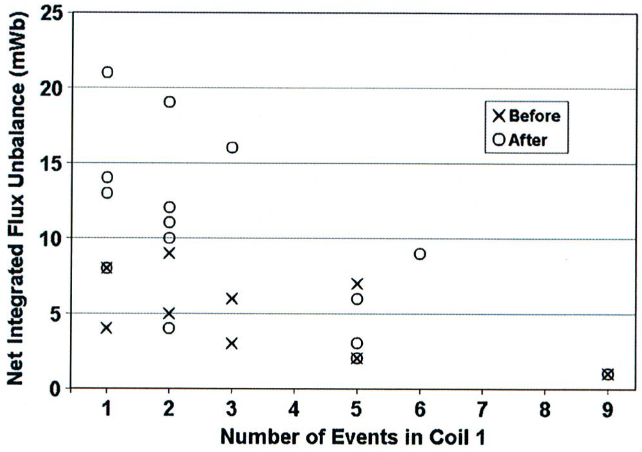


Fig. 8. Net integrated flux unbalance for events in coil 1 before (exes) and after (circles) energy extraction failure at the end of Q13.

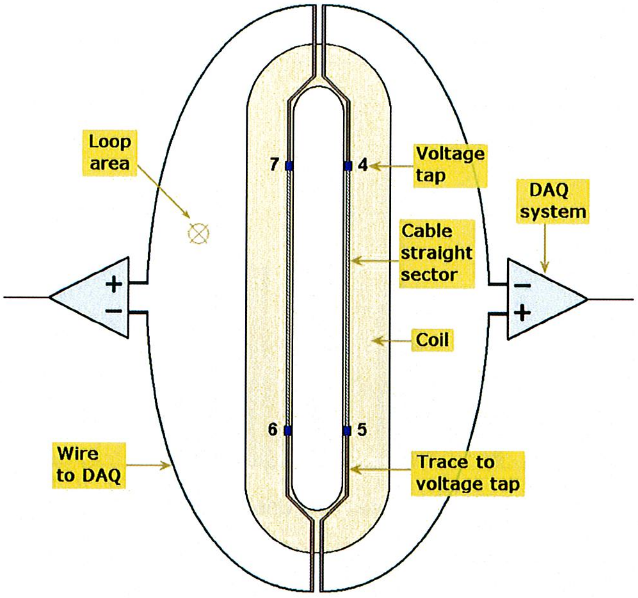


Fig. 9. Schematic for the two 1-loop antennas placed on the inner layer of coil 1 monitoring opposite sides of the coil.

This plot shows a significant increase on the size of FJs after the extraction failure.

Due to the correlation between the cable's current-temperature margin and the magnet's FJ activity, the change in FJ patterns observed before and after Q13 suggests that the energy extraction failure that occurred at the end of Q13 indeed caused a change in the current-temperature margin of HD2b.

V. Flux-Jump Antenna Prototyping

Given the potential that FJ studies have on aiding magnet diagnostics, the Superconducting Magnet Program at LBNL is studying the possibility of implementing a multi-loop multi-channel FJ antenna to spatially locate FJs inside a magnet and measure their properties more precisely. For this purpose, the ability to detect a FJ using a 1-loop antenna was tested during the training of HD2b.

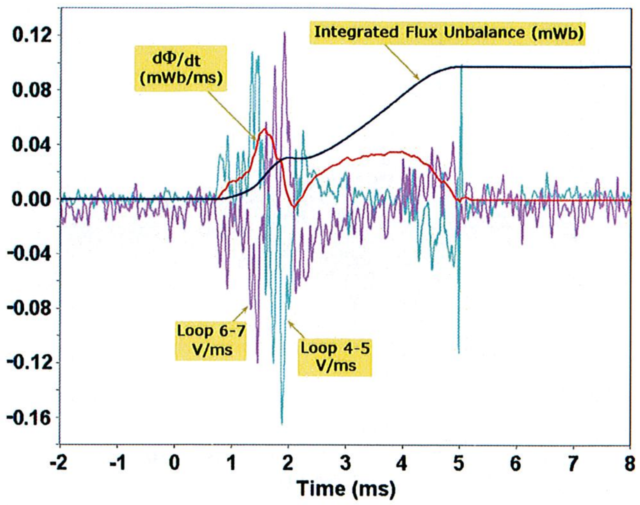


Fig. 10. Signals from two 1-loop antennas placed on each side of coil1, software reconstructed $d\Phi/dt$ and integrated flux unbalance.

Fig. 9 shows a schematic for the two 1-loop antennas placed on the surface of the inner face of coil 1. Voltage taps 4, 5, 6 and 7 on the inner layer monitor the voltage on the two straight sectors against the pole, however, voltage taps 4 and 7 were routed to the DAQ system through the top of the magnet, while voltage taps 6 and 5 were routed through the bottom of the magnet as depicted by Fig. 9.

If the straight sectors are not quenching, the resistive component of the voltage at the loop terminals is zero and only the voltage induced by the flux enclosed by the loop will be collected by the data acquisition system.

To further look into the effect of the energy extraction failure after Q13, the 1-loop antennas were used to locate the side of coil 1 that the 6 FJs with the highest net integrated flux unbalance collected on Q14 and Q15 were originated. These events correspond to the events making the three circular data points above 15 mWb shown in Fig. 8. Fig. 10 shows the raw signals collected from these two loops for event 142 in Q15 and the calculated integrated flux unbalance. This event corresponds to the event making the highest circular data point with net integrated flux unbalance of 21 mWb.

The 1-loop antennas indicate that 5 out of the top 6 events were located in the half portion of the coil covered by sector 4–5. Understanding how this correlates to the location of quench 13 is under way. This is an important result in that it suggests that the excess FJs are grouping around one side of coil 1. This result also tells us that a 1-loop antenna can pick up the signal induced by a FJ and provide information about its location within the magnet.

VI. Conclusions

- A comparison of FJs count before and after the magnet energy extraction failure shows an increase in the magnet's FJ activity.
- A comparison of the magnitude of FJs originated in coil1 before and after the energy extraction failure shows that the net integrated flux unbalance per FJ increased.

- A change in FJ behavior suggests a change in the current-temperature margin distribution of the cable.
- Data from the 1-loop antennas placed on coil 1 suggests that the excess FJs tend to cluster in one side of the coil.
- The 1-loop antenna placed in coil 1 proved the feasibility of a FJ antenna with multiple loops and higher spatial resolution. This approach will be further studied during the following year.

REFERENCES

[1] J. Lizarazo *et al.*, "Measurement of fast voltage transients in high-performance Nb_3Sn magnets," *IEEE Trans. Appl. Supercond.*, vol. 18, no. 2, p. 1581, June 2008.

[2] A. F. Lietzke *et al.*, "Test results for HD1: A 16 T Nb_3Sn dipole magnet," *IEEE Trans. Appl. Supercond.*, vol. 14, no. 2, p. 345, June 2004.

[3] A. F. Lietzke *et al.*, "Test results for RD3c, a Nb_3Sn, racetrack dipole," *IEEE Trans. Appl. Supercond.*, pp. 1292–1296, 2002, Houston.

[4] A. F. Lietzke *et al.*, "Differentiation of performance-limiting voltage transients during Nb_3Sn magnet testing," in *AIP Conference Proceedings.*, 2006, vol. 824B, pp. 550–557.

[5] S. Feher, B. Bordini, R. Carcagno, A. Makulski, D. F. Orris, Y. M. Pischalnikov, C. Sylvester, and M. Tartaglia, "Sudden flux change studies in high field superconducting accelerator magnets," *IEEE Trans. Appl. Supercond.*, vol. 15, no. 2, p. 1591, June 2005.

[6] S. Feher *et al.*, "Cable testing for Fermilab's high field magnets using small racetrack coils," *IEEE Trans. Appl. Supercond.*, vol. 15, no. 2, p. 1550, June 2005.

[7] A. F. Lietzke *et al.*, "Race-track coil technology validation (RT-1) test results," in *VLHC Workshop*, May 24, 2000, FNAL.

A Novel Method for Generating Scale Space Kernels Based on Wavelet Theory

Dionísio Doering1

Adalberto Schuck Júnior2

Abstract: The linear scale-space kernel is a Gaussian or Poisson function. These functions were chosen based on several axioms. This representation creates a good base for visualization when there is no information (in advanced) about which scales are more important. These kernels have some deficiencies, as an example, its support region goes from minus to plus infinite. In order to solve these issues several others scale-space kernels have been proposed. In this paper we present a novel method to create scale-space kernels from one-dimensional wavelet functions. In order to do so, we show the scale-space and wavelet fundamental equations and then the relationship between them. We also describe three different methods to generate two-dimensional functions from one-dimensional functions. Then we show results got from scale-space blob detector using the original and two new scale-space bases (Haar and Bi-ortogonal 4.4), and a comparison between the edges detected using the Gaussian kernel and Haar kernel for a noisy image. Finally we show a comparison between the scale space Haar edge detector and the Canny edge detector for an image with one known square in it, for that case we show the Mean Square Error (MSE) of the edges detected with both algorithms.

Index Terms—Kernel, Scale space, Wavelets.

---

1 Lawrence Berkeley National Laboratory. One Cyclotron Road, MS50A6134, Berkeley, CA, USA. {ddoering@lbl.gov}
2 DELET – EE – UFRGS. Av. Oswaldo Aranha, 103, sala 206. Centro Porto Alegre, RS. Brasil. {schuck@ufrgs.br}

1 Introduction

There are several different multi-scale signal processing techniques such as quad-tree [1], pyramid [1], wavelets [2][3] and scale-space [4]. The linear scale-space, for example, defined through its axioms has the Gaussian (and Poisson) as its kernel function, although several authors have been discussing if the Gaussian function is the best one for the scale-space. In order to answer this question it is necessary to compare the results from the Gaussian scale-space with other scale-spaces.

The scale-space theory was proposed by Witking [4] and expanded to the bi-dimensional case by Koenderingk [5]. This theory has been largely used to extract information from images. Some examples are edge detection [6], ridge detection [7], structure extraction [8, 9] and movement detection [10]. The scale space created from the convolution of an image with a Gaussian or Poisson function have the following properties [6][7][11]: (i) causality, (ii) isotropy, (iii) homogeneity, (iv) non creation of local extrema, (v) non enhancement of local extrema, (vi) semi-group structure and (vii) infinitesimal generator. Florack et al. [12] say that these are good properties for image representation where there is no previous knowledge about the image because they work equally in any direction of the image. But in applications where one wants to emphasize a specific direction the Gaussian (or Poisson) function may not be the best choice. Another issue of the Gaussian scale-space, as mentioned by Remaki and Cheriet [13], is the length of its kernel that goes from minus to plus infinite and this implies the necessity of the use of large masks. Based on these limitations several authors have been searching for new scale-space kernels.

Another multi scale theory widely used is the Wavelet Transform [14]. In this work the wavelet transform of a signal is defined as being the internal product between the signal and the mother wavelet function in a specific scale and shifted by some factor. In the definition of the wavelet transform it is not defined a mother wavelet function allowing this transform to be very flexible. The most important conditions for these functions (wavelets) are admissibility and regularity. Due to the former condition the frequency component of the function at frequency zero must be zero. And due to the latter condition the wavelet function must be local in both time and frequency domains. These two conditions together imply that a wavelet function must be a band pass filter. Due to the enormous flexibility of the wavelet transform it became a general case for several others transforms, such as the Canny edge detector [15] and the scale-space theory [4]. The scale-space is a particular case of the wavelet transform when the derivatives of the Gaussian functions are used to extract information [16].

The relationship between the wavelets and the scale-space that were demonstrated by Mallat [16] haven't been used to solve digital image processing tasks up to now. One of the challenges dealt by the scale-space community is the definition of new kernels that are able to overcome the Gaussian limitations. In this paper (i) we will use the scale-space theory, (ii) the Wavelet theory and (iii) the relationship between them in order to create a novel method to generate scale-space kernels from the wavelet mother functions. With this we want to enlarge the discussions about the scale space kernels.

2 Theoretical Fundamentals

2.1 The Scale Space

The scale-space representation of an image is a set of images that represent the original image in different scales. Mathematically, given a bi-dimensional function $f(x,y) \in \Re^2 \rightarrow \Re$, its scale space representation $L(x,y,t) \in \Re^2 \times \Re_+ \rightarrow \Re$ is:

$$L(x,y,t) = f(x,y) * g(x,y,t) \tag{1}$$

where t is the scale, $*$ is the convolution operator and $g(x,y,t) \in \Re^2 \times \Re_+ \rightarrow \Re$ is the Gaussian kernel defined as:

$$g(x,y,t) = \frac{1}{(2\pi t)} e^{-\left(x^2+y^2\right)/2t} \tag{2}$$

where the scale (t) is the variance (σ^2) of the Gaussian function.

2.1.1 Scale-space blob detector

When a gray scale image is represented as a tri-dimensional surface its objects will become blobs [8]. There are two types of blob detectors, (1) one that has clear object in a dark background and (2) one that has dark objects in a clear background. The blob detector [17, 18] is defined using the image Laplacian:

$$\nabla^2 L = L_{xx} + L_{yy} \tag{3}$$

where $\nabla^2 L$ is the image Laplacian, L_{xx} and L_{yy} are the second derivative of L in the x and y directions respectively. Images with clear regions and dark background, for example, have their blob detector operator defined as follows:

$$\begin{array}{ll} I(x,y) = 0, & \nabla^2 L > 0 \\ I(x,y) = 1, & \nabla^2 L \leq 0 \end{array} \tag{4}$$

where I is a binary image, where 0 = black= background and 1 = white = blob.

2.1.2 Scale space edge detector

Lindeberg [19] proposed the detection of edges based on the second and third derivatives of the scale space image using the (u,v) coordinate system. In this system v is the direction parallel to the gradient direction and the u direction is perpendicular to the gradient direction. Then, a pixel will be an edge pixel if the next two conditions are met:

$$L_{vv} = L_x^2 L_{xx} + 2L_x L_y L_{xy} + L_y^2 L_{yy} = 0 \tag{5}$$

$$L_{vvv} = L_x^3 L_{xxx} + 3L_x^2 L_y L_{xxy} + 3L_x L_y^2 L_{xyy} L_y^3 L_{yyy} < 0 \tag{6}$$

where L_{vv} and L_{vvv} are the second and third derivatives of *L* in the *v* direction.

2.2 Wavelet Theory

The wavelet transform is an efficient tool for local analysis of non-stationary and transient signals [20]. The wavelet functions are created by scaling and shifting the mother wavelet. For a function to be a wavelet function it must have the admissibility and regularity conditions. This means that these functions must be local in the time (or space) and frequency (or scale) domains.

The wavelet transform can be implemented in two different ways, the Continuous Wavelet Transform (CWT) or the Discrete Wavelet Transform (DWT). In this paper we are going to use the CWT definitions because the scale space is also continuous and then it is possible to relate these two theories.

2.2.1 Bi-dimensional Continuous Wavelet Transform (CWT2D)

The CWT can be extended to two or more dimensions and keep the same properties of the 1D case [21]. Given a bi-dimensional signal (2D), for example an image *f(x,y),* its CWT2D for a wavelet function ψ is:

$$W2D(s,\theta,b_x,b_y) = \frac{1}{s}\left\langle \psi\left(\frac{r_\theta(x-b_x, y-b_y)}{s}\right), f(x,y)\right\rangle = \frac{1}{s}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\psi\left(\frac{r_\theta(x-b_x, y-b_y)}{s}\right)^{*} f(x,y)dxdy \tag{7}$$

where $\langle\ \rangle$ is the internal product, ψ is a bi-dimensional wavelet, shifted by b_x in the x direction and b_y in the y direction, scaled by s ($s > 0$) and rotated by θ ($0 \leq \theta < 2\pi$) where r_θ is a rotation matrix. For (7) to hold it is necessary that the image *f(x,y)* needs to be squared integrable and defined in the plane $\Re^2$, or $f(x,y) \in L^2$ (or in other words, have finite energy). For ψ be a mother wavelet, it must satisfy the admissibility and regularity conditions as mentioned before. The CWT2D defined by (7) has four parameters (s, θ, b_x, b_y) [22]. In order to be possible to visualize this transform it is necessary to lock some of them. There are several different combinations but the combinations where (b_x, b_y) or (s, θ) are kept locked have been more used because it's physical interpretation.

2.2.2 The mother wavelet families

The mother wavelets can be defined in different ways. One option is using its equation in the time (space) domain. Another option is to use the function properties in the frequency domain. As examples of mother wavelets, there are those that are defined in the

Matlab version 5.3: Morlet; Mexican hat; Meyer; Haar; Daubechies; Symlets; Coiflets; Splines; bi-orthogonal.

2.3 Relationships between the Scale Space and the Continuous Wavelet Transform

It was shown by Mallat [16] that the scale space is a special case of the wavelet transform. In that work, besides proving that these two transforms are related, the author also defined a method to generate new wavelets from a smoothing function. This method is described as follows:

A smoothing function is defined as:

$$\theta_s(x) = \frac{1}{s}\theta\left(\frac{x}{s}\right) \tag{14}$$

The wavelet functions generated from the smoothing function are defined as:

$$\psi^1(x) = \frac{d\theta(x)}{dx} \tag{15}$$

$$\psi^2(x) = \frac{d^2\theta(x)}{dx^2} \tag{16}$$

In that work Mallat used the bi-dimensional Gaussian function as the smoothing function and he had generated two wavelets. In [23] Sheng claims that any derivative of order 'n' of the Gaussian function can be a wavelet. This allows us to expand the Malat's method to the 'n' order (for n > 0) and then the wavelets are defined as:

$$\psi^n(x) = \frac{d^n\theta(x)}{dx^n} \tag{17}$$

The scale space is a special case of the wavelet theory when the function $\theta(s)$ is the Gaussian function.

3 The Method Proposed

In this section the method proposed to generate new kernels for the scale space from wavelet functions is defined, as well the results obtained. The functions needed to realize this task were developed using the software Matlab 5.3. The image processing was done in a Athlon XP 1.2 GHz with 256 Mb of RAM memory.

3.1 Method to generalize the scale space from different wavelet families

One of the objectives of this work is to generate the scale space and its operators from different wavelet families. The scale space operators are defined using the Gaussian derivatives, and this function works as a smoothing filter. The wavelet functions are band pass filters. So it is necessary to associate each wavelet function to a smoothing function and from this smoothing function it is possible to apply the scale space operator definition. In order to associate a smoothing function to each wavelet function it will be used the Mallat's method described in section 2.3 backwards. This means that we are going to start with the wavelet function ($\psi^n(x)$) looking for a smoothing function ($\theta_s(x)$), as will be shown later.

It is assumed that the wavelet function under test is a derivative of some unknown smoothing function, but it is unknown which derivative order is the wavelet from this smoothing function. Based on wavelet theory, it is known that the wavelet function is a band pass filter and the smoothing function is a low pass filter. Using these characteristics it is going to be determined which is the derivative order of the wavelet function (if exist such a scaling or smoothing function). To do that the following algorithm is used:

1. The wavelet is considered the n-order derivative of some smoothing function.
2. The wavelet is integrated.
3. The frequency response of the new function is obtained by means of FFT.
4. If the frequency response shows that the new function is a band bass function then steps 2 and 3 are repeated for the new function.
5. If the new function is a low pass filter then it is defined as the smoothing function that is associated to the wavelet function under test.
6. The derivative order is defined as being equal to the number of times that steps 2 and 3 were executed, or the number of times that the function was integrated.

To exemplify the use of this algorithm we will apply it to the Haar wavelet function. Figure 1(a) shows the Haar wavelet and Figure 1(b) shows its frequency response. From the frequency response one can see that this function is a band pass filter as expected because this is a wavelet function. Following the algorithm, on step 1 this function is considered the n-order of a smoothing function. On step 2 we integrate this function and the resulting function is shown in Figure 1(c). On step 3 we generate its frequency response and show it in Figure 1(d). Step 4 we evaluate the frequency response and conclude that this is not a band pass filter and then we proceed to step 5. In 5 we define this new function as being the smoothing function of the wavelet Haar because its frequency response is a low pass filter. Step 6 determines the order of the wavelet function as being 1 since the wavelet function was integrated once to get to the smoothing function.

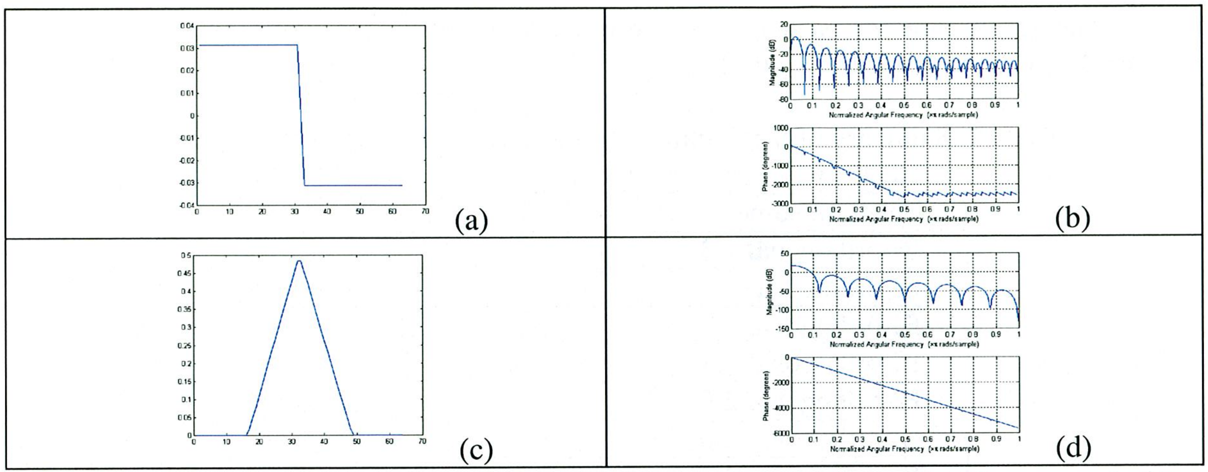


Figure 1- (a) wavelet Haar on scale 32 and (b) first integral of the wavelet Haar on scale 32, (c) magnitude and phase of the frequency response of the function wavelet Haar and (d) first integral of the wavelet Haar.

Once the smoothing functions are known and which is the derivative order of the wavelet function, it is possible then to apply scale space operator definition to them. As seen in the previous section the scale space operators are defined from the derivatives (L_x, L_{xx}, ...) of some smoothing function.

3.2 Choosing the mother wavelet function

The choice of the mother wavelet is of fundamental importance to the analyses of a given signal when this transformation is used. This is independent of the signal dimension (1D, 2D, …). Antoine [21] defines some options to solve this task. The first of them is to choose the mother wavelet that gives the best result for a given problem. The second option is to create a new mother wavelet for the application. The third option is to use some wavelet families that have their behavior and definition known, at least for the 1D case. This work uses the third option to choose the mother wavelet added of some requirements necessary to allow the comparison between the wavelet and the scale space. These conditions are postulated as follows: (1) the existence of a smoothing function created by integrating the wavelet function *n* times; (2) the smoothing function is real and symmetric; (3) the smoothing functions derivatives must follow the properties of the wavelet transform (admissibility and the regularity conditions).

It is important to remark that this work does not have the intention to prove the necessity or sufficiency of the postulates stated above because it is out of its scope. The wavelet functions that don't have such characteristics will not be used in this work. The list of functions that fulfill the postulated above are: the Gaussian derivatives (Mexican Hat); Haar; Bi-ortogonal 1.3; Bi-ortogonal 1.5; Bi-ortogonal 2.2; Bi-ortogonal 2.4; Bi-ortogonal 2.6; Bi-ortogonal 2.8; Bi-ortogonal 4.4; Bi-ortogonal 5.5; Bi-ortogonal 6.8.

The results got for each one of these functions using the algorithm explained in section 3.1 are summarized in Table 1.

Table 1: Summary of the features that the wavelet functions have when they were tested with the algorithm defined in section 4.3.

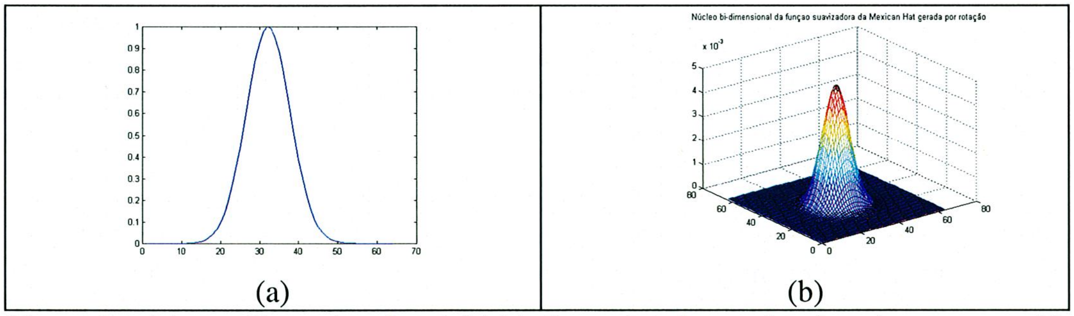
Function name	Function Order*
Bi-orthogonal 1.3	1
Bi-orthogonal 1.5	1
Bi-orthogonal 2.2	2
Bi-orthogonal 2.4	2
Bi-orthogonal 2.6	2
Bi-orthogonal 2.8	2
Bi-orthogonal 4.4	4
Bi-orthogonal 5.5	6
Bi-orthogonal 6.8	6
Haar	1
Mexican Hat	2

***number of times that the function was integrated in order to get a smoothing function.**

3.3 Generation of bi-dimensional functions from one-dimensional functions

Some of the functions mentioned on the last section don't have a mathematical definition in the time (space) domain for the bi-dimensional case. It was used in this work three different ways to solve this task.

The first method to generate the bi-dimensional functions uses a given function $y = f(x)$, symmetric and the bi-dimensional function is generated by rotating the uni-dimensional function. This rotation is done numerically and the function is rotated in relation to the z axis. Figure 2 shows the one-dimensional and bi-dimensional Gaussian function generated using this method.



(a) (b)

Figure 2 – (a) one-dimensional and (b) bi-dimensional gaussian function generated by rotation of (a).

The second method used to generate a bi-dimensional function from a one-dimensional function receives one function $f(x)$ and one function $f(y)$, both symmetric to some point and then generate a bi-dimensional function using tensor product. The tensor product between these two functions $f(x)$ and $f(y)$ is defined as:

$$f(x,y) = f(x) \, . \, f(y) \tag{18}$$

The third method used to generate a bi-dimensional function from a one-dimensional function receives a given function $y = f(x)$, symmetric to some point and generates a bi-dimensional function by convolution. The convolution between these two functions is defined as follows:

$$F(x,y) = f(x) * f(y) \tag{19}$$

where * is the convolution operator.

All the three methods shown before generate a bi-dimensional Gaussian function identical to the function generated by the mathematical bi-dimensional function definition. This demonstrates that all of them could be used. Although, when these methods are used for other functions other than the Gaussian the bi-dimensional function that results may vary for each method. In this work it is going to be shown only the results got using the bi-dimensional functions generated by rotation because the operators that result from this method are symmetric and this makes them rotational invariant. This is a feature that exists in the Gaussian scale space and it will be preserved in this work.

3.4 Examples of the new scale-space kernels use.

The scale space representation of an image can be done for the different scale spaces using different smoothing function as well as for different methods of generating bi-dimensional functions. One example is the Lena image representation. Figure 3 shows a scale space representation using the Haar's smoothing function for scales 2, 8 and 16. Figure 4 shows the blobs of the Lena image using the Haar's smoothing function for scales 2, 8 and 16. Figure 5 shows the edges of the Lena image using the scale space edge detector and the Haar's smoothing function for the scales 2, 8, and 16. These images provide good examples to show the different scale space features in a qualitative way.

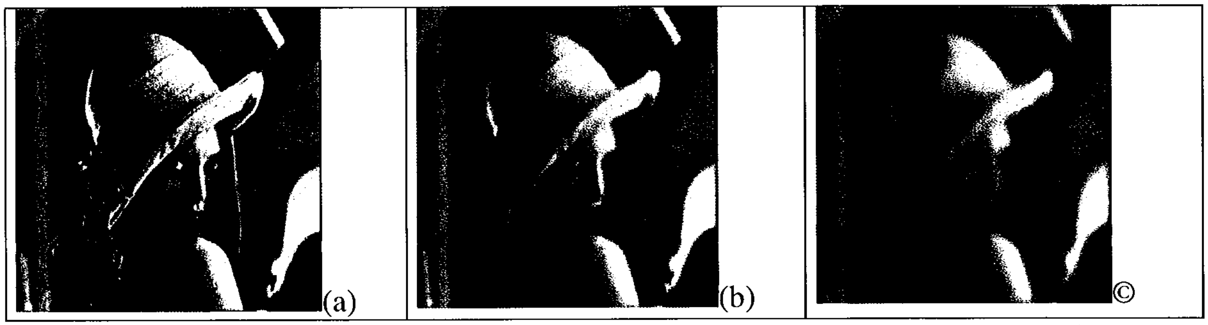


Figure 3 – Scale space representation of the Lena image using the Haar smoothing function, created by convolution on scales 2 (a), 8 (b) and 16 (c).

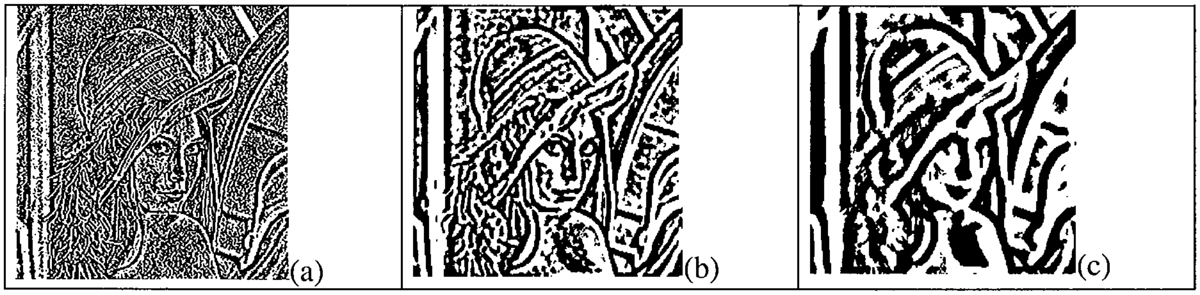


Figure 4 – Blobs from the Lena image, when using the scale space blob detector and the Haar smoothing function, created by rotation, on scales 2 (a), 8 (b) and 16 (c).

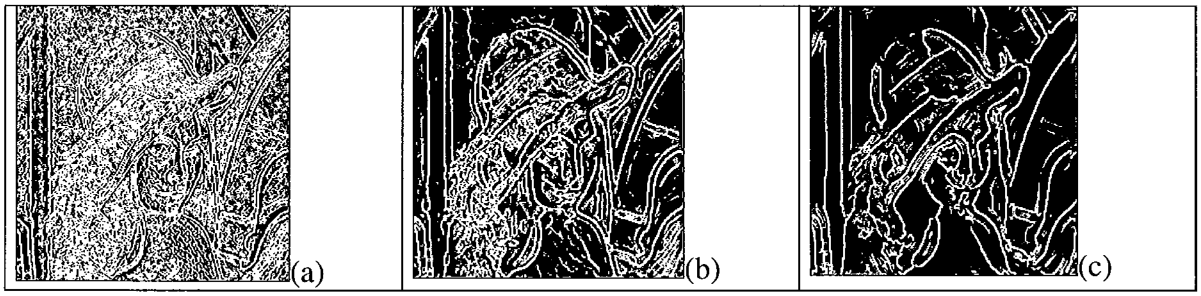


Figure 5 – Edges from the Lena image, when using the scale space blob detector and the Haar smoothing function, created by rotation, on scales 2 (a), 8 (b) and 16 (c).

In order to analyze some of the new scale spaces properties independently it was used an artificial image called phantom1 (Figure 6a). This image has objects with different shapes and sizes but the same contrast. This allows us to observe the behavior of the scale space operator in different scales. To understand the behavior of the different kernels with respect to noise a second phantom image was created (Figure 6b). This phantom besides objects of different intensities has noise added to it. The noisy image is defined as:

$$\hat{I}(x,y) = f(x,y) + c.n(x,y) \tag{20}$$

where $\hat{I}(x,y)$ is the noisy image, $f(x,y)$ is the original image, n(x,y) is the normal noise with standard deviation set to one and c is a constant that multiplies the noise. This constant was arbitrarily chosen to 8. In a future work, it will be presented the robustness of detection algorithms with noise level in images using the technique here presented.

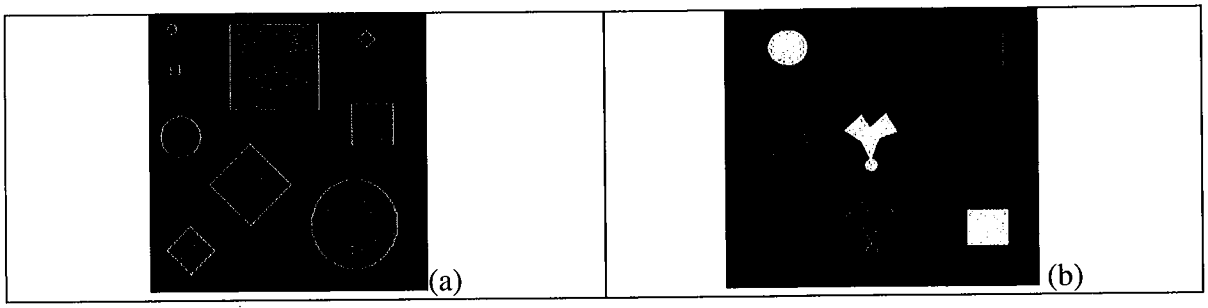


Figure 6 – Artificial images called (a) *phantom1* and (b) *phantom2* (b).

Figure 7 shows the phantom1 image's scale space blobs using the Haar smoothing function for scales 2, 54 and 80. Figure 8 shows the edges of the phantom1 image using the scale space edge detector for the Haar smoothing function for scales 2, 8 and 16. Figure 9 shows the blobs of the phantom1 image using the smoothing function of the wavelet Bi-orthogonal 4.4 for scales 2, 26 and 72. Figure 10 shows the edges of the phantom1 image using the scale space edge detector and the smoothing function of the wavelet Bi-orthogonal for scales 1, 4 and 16.

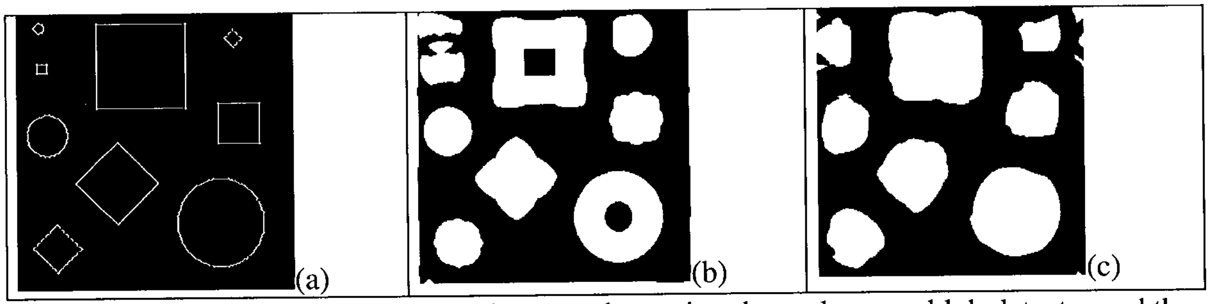


Figure 7 – Blobs from the *phantom1* image, when using the scale space blob detector and the Haar smoothing function, created by rotation, on scales 2 (a), 54 (b) and 80 (c).

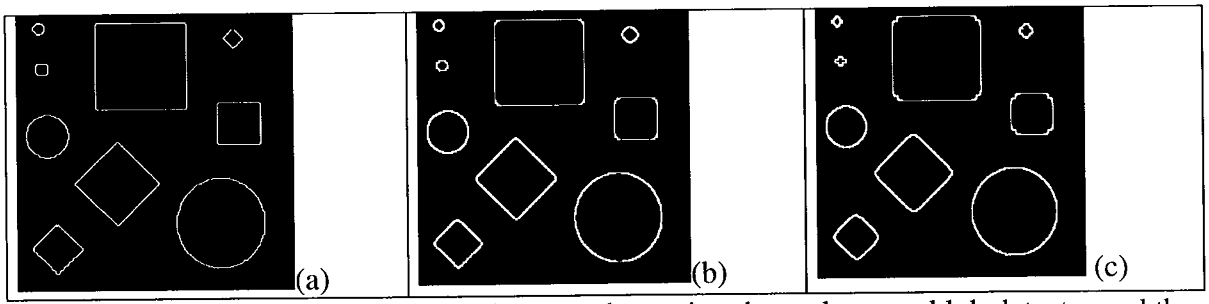


Figure 8 – Edges from the *phantom1* image, when using the scale space blob detector and the Haar smoothing function, created by rotation, on scales 2 (a), 8 (b), and 16 (c).

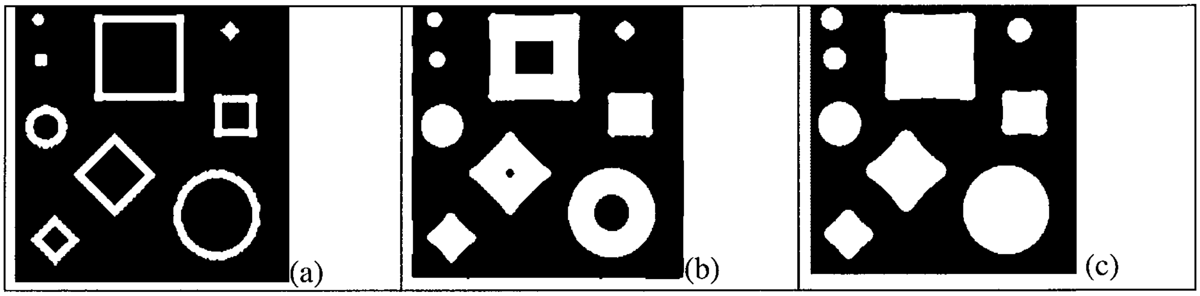


Figure 9 – Blobs from the *phantom1* image, when using the scale space blob detector and the Bi-ortogonal 4.4 smoothing functions, created by rotation, on scales 2 (a), 26 (b) and 72 (c).

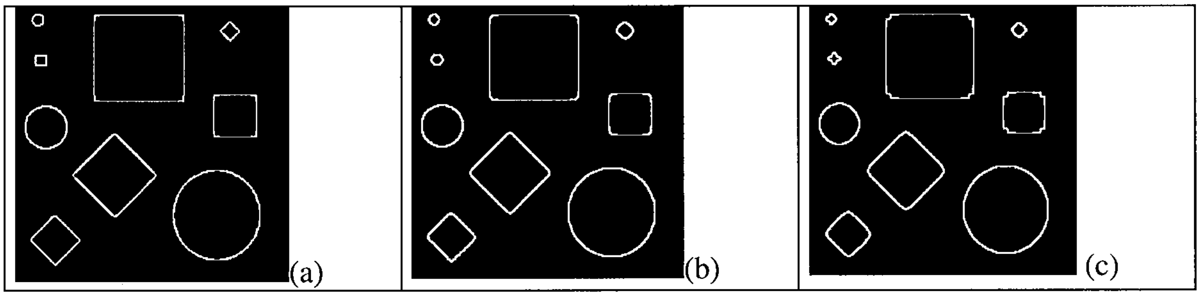


Figure 10 – Edges from the *phantom1* image, when using the scale space blob detector and the Bi-ortogonal 4.4 smoothing function, created by rotation, on scales 1 (a), 4 (b) and 16 (c).

Figure 11 shows the detected edges of the noisy phantom2 image (Figure 6(a)) using the Gaussian smoothing function, for scales 2 (a), 16 (b) and 32 (c). Figure 12 shows the detected edges of the noisy phantom2 image using the smoothing function of the Haar wavelet, for scales 2 (a), 16 (b) and 32 (c).

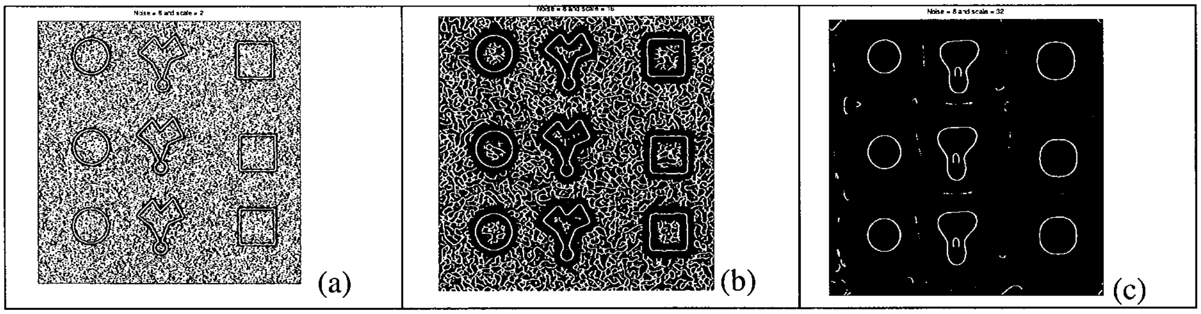


Figure 11 – Edges detected from the *phantom2* image when using the Gaussian scale space on scales 2 (a) 16 (b), e 32 (c).

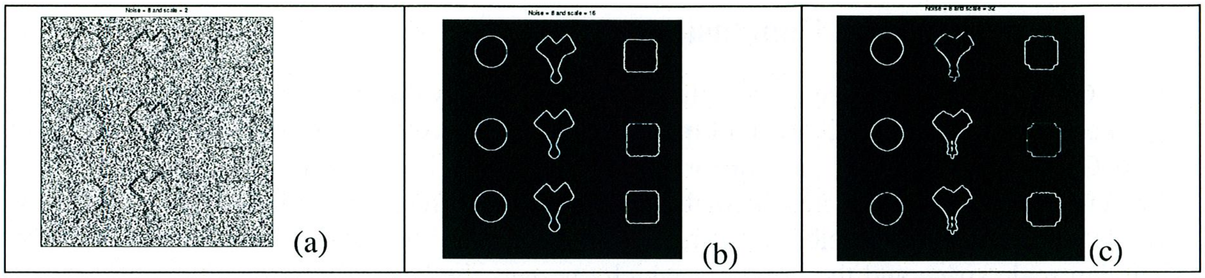


Figure 12 – Edges detected from the *phantom2* image when using the Haar scale space on scales 2 (a), 16 (b), e 32 (c).

In order to create a quantitative comparison between the scale space edge detector using the Haar kernel and the Canny edge detector we create a phantom with a white box in a black background (phantom 3). The edges of this phantom are known and this allowed us to calculate the mean square error (MSE) for the edges detected with both algorithms while adding different random noise levels. Figure 13 (a) shows the MSE for the Canny edge detector and Scale Space Haar edge detector (at scale set to 12) for the phantom 3 using different noise levels. Figure 13 (b) shows the number of edge points detected using the Canny edge detector and Scale Space Haar edge detector (at scale set to 12) for a phantom3 using different noise levels. These examples show that the scale space algorithm has a superior performance over the Canny edge detector for noisy images.

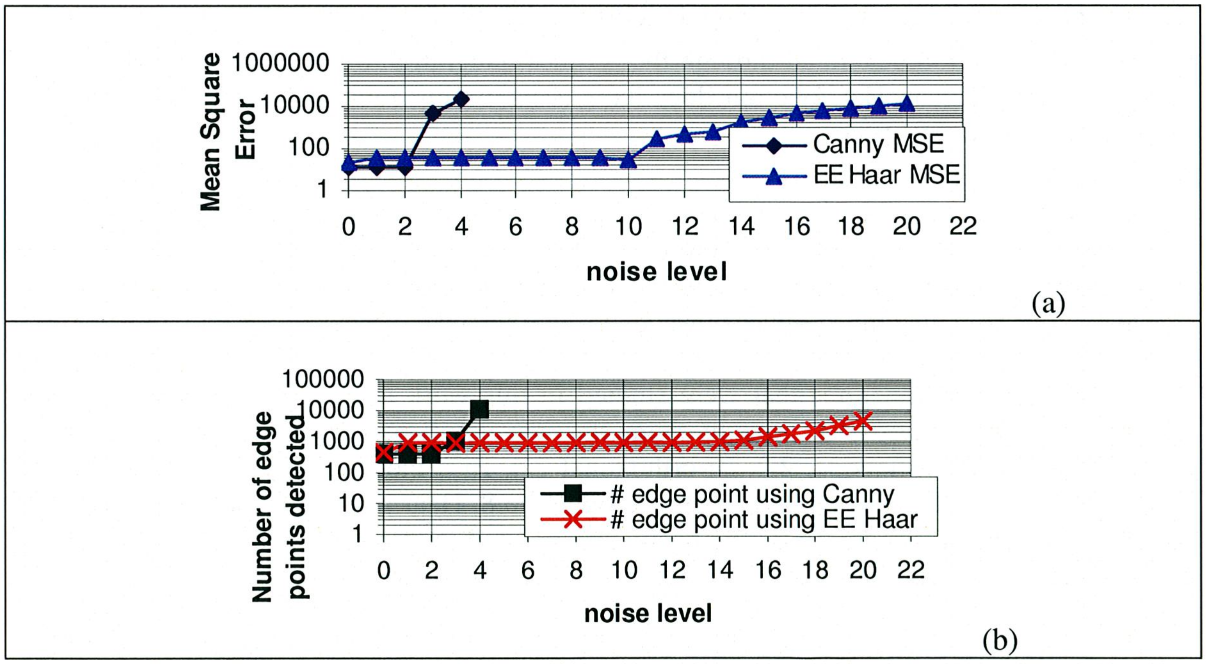


Figure 13 – Mean square error for the Canny edge detector and Scale Space Haar edge detector (at scale set to 12) for a phantom versus noise level (a). Number of points detected using the Canny edge detector and Scale Space Haar edge detector (at scale set to 12) for a phantom versus noise level (b).

4 Discussion and Conclusions

One of the scale space aspects that have been widely discussed by several researchers is its kernel (originally the Gaussian function). The objective of these discussions is to know if the Gaussian is the best function or in which cases it is the best one. Besides that the researches are developing new kernels with new or different features. Once these new kernels are known it is possible to compare the scale space generated with them against the Gaussian scale space and then find out which one gets the best results for different cases of study. In this work, it was used (1) the scale space theory, (2) the wavelet theory and (3) the relationships between them to develop a new method to create new scale space kernels from wavelet functions.

In the proposed method, we start from a known wavelet function. Then we find the smoothing function that is associated to this wavelet through numerical integration and analyses of its frequency domain response. Once the smoothing function is available it is used the original definition of the scale space operators with this new function. This work shows the results from the scale space blob and edge detectors. The only thing that was changed was the kernel.

According with the wavelet transform theory the wavelet function is band bass function (ψ). Some of these functions have a low pass function associated (ϕ). It is important to note that the smoothing function is NOT this low pass function (ϕ) and it actually is a function generated through the $\underline{n}$ times integration of the wavelet function. For example, the Mexican Hat function does have a smoothing function (the Gaussian) but it does not have a ϕ function.

The proposed method limits the set of wavelet functions that are used with it based on some postulates. These postulates were defined in such a way that once we use a wavelet function with the algorithm it must be possible to get a new function that has a low pass filter behavior. Besides this only symmetric wavelet functions were used.

The wavelet functions that were studied in this work are the functions that already exist in Matlab version 5.3. Among them the wavelet functions Bi-orthogonal 1.3, Bi-orthogonal 1.5, Bi-orthogonal 2.2, Bi-orthogonal 2.4, Bi-orthogonal 2.6, Bi-orthogonal 2.8, Bi-orthogonal 4.4, Bi-orthogonal 5.5, Bi-orthogonal 6.8, Haar and Mexican Hat are the ones that have the features required by the postulated defined in this work. Those functions were used with the proposed method to generate new scale space kernels.

Some of the wavelets that were used in this work don't have definition for the 2D case. Due to that it was used some methods to create 2D wavelets from 1D wavelets. Three different methods to solve this problem were used and they are rotation, tensor product and convolution. When these methods were used with the 1D Gaussian function the results match the 2D Gaussian function defined analytically. This indicates that any one of them could be used. But the rotation method generates 2D functions that are always symmetric and because of it this was the chosen method.

After we have the wavelet function, a method to create scale space kernels and 2D functions, it was possible to generate the scale space associated with each function. All kernels produced are local and rotational invariant.

The Haar wavelet function used in this work was defined from minus 0.5 to plus 0.5 because this makes it symmetric in relation to zero. As this function is defined inside a limited range it is correct to say that it has compact support. The Gaussian function, for example, is defined from minus to plus infinite. In order to apply the latter filter in a signal it is necessary to truncate it at some point. Values that are based on the standard deviation, such as one, two or three standard deviations are usually chosen. Because this filter is truncate, the result of filtering an image with it is an approximation of the continuous case. As bigger the point where the function is truncated the better the approximation is. On the other hand the bigger this value is then the bigger is the processing cost to solve this task. An important advantage for the Haar kernel in relation to the Gaussian is that it is completely defined in a given range therefore it doesn't need to be truncated. Besides that the total computational cost to use this kernel will also be lower than the Gaussian. Kernels with compact support have been developed as shown in [16].

The Haar wavelet function is usually used to detect discontinuities. For images these discontinuities are probably edges of the objects. When the smoothing function of the Haar function was used with the scale space edge detector this behavior was preserved, even for large scale as shown in Figure 5. For this case the biggest distortions had happened at the object's corners. It was also shown the edges detected with the scale space edge detector using the smoothing function of the wavelet Haar and Gaussian. The noise degradation of the edges was far less when using the smoothing function of the Haar wavelet than when using the Gaussian smoothing function. These results are examples that each kernel has its own features, and then the discussion about which is the best kernel for each application is extremely valid.

The wavelet transform has several kernels and it is necessary to choose which kernel will be used in a given application. The scale space also has several kernels and it is also necessary to choose which kernel to use in a given application. Antoine [21] defines some options to solve this task, as described in section 3.2. The authors suggest this method to be used for the scale space theory too. The main contribution of this work for the task of choosing the best scale space kernel is the method to generate new kernels and the ten new kernels described earlier. The Mexican hat wavelet generated the Gaussian smoothing function and it was used to validate the algorithm that generates new scale space kernels from wavelet functions and it is not considered a new kernel. With that there is now a bigger set of kernels that can be used to find out which one is the best for each application.

We also compared the results got using the Canny edge detector with the scale space Haar edge detector. The result of the MSE showed us that for low noise images the Canny algorithm has a better performance. But when we added higher noise levels the scale space Haar edge detector outperform the Canny algorithm. This shows that for such noisy images the scale space Haar edge detector is a better option. The number of edge points detected also show us that the Canny edge detector has a better performance when the noise level on the image is low while the scale space edge detector has a better performance for noisy images.

In this work it was proposed a new method to create new scale space kernels from wavelet functions. Some of the challenges in this field were solved, but there are several others that are still open and they are suggested for future work. Some suggestions are:

- Implement ridge detector;
- Implement scale-spaces with non-symmetric bases;
- Verify for the new kernels which axioms the satisfy;
- Test for detection robustness of noise images.

6 Acknowledgment

Dionísio Doering thanks CAPES for his scholarship.

7 References

[1] Lindeberg, T., "Linear scale-space: Basic theory", Kluwer Academic Publishers, 1994.

[2] Daubechies, I., "Ten Lectures on Wavelets", Society for Industrial and Applied Mathemathics (SIAM), 1992.

[3] A Wavelet Tour of Signal Processing, Second Edition: Wavelet Analysis & Its Applications) (Hardcover);Academic Press; 2nd edition September 15, 1999.

[4] Witkin, A. P., Scale-space Filtering. In Proc. 8th Int. Joint. Conf. Art. Intel. Pages 1019-1022. Karlsruhe, West Germany, 1983.

[5] Koendering, J.J., "The Structure of Images", Biol. Cybern., 1984.

[6] Lindeberg,T., "Scale-Space: A Framework for handling image structures at multiple scales", Proc. CERN School of Computingm Netherlands, September, 1996.

[7] Cunha, A. M., "Espaço de Escala e Detecção de Arestas", outubro de 2000.

[8] Lindeberg,T., Eklundh, J. "Scale-space primal sketch: construction and experiments", Image and Vision Computing, Volume 10, Issue 1, January-February 1992, Pages 3-18.

[9] ZUCCOLOTTO, M., OLABARRIAGA, S., SCHUCK JR., A., "Segmentação de cavidades em imagens cardiográficas utilizando Espaço de escalas", IICLAEB, 2002.

[10] Laptev, I., Lindeberg, T., "Velocity adaptation of spatio-temporal receptive fields for direct recognition of activities: an experimental study", Image and Vision Computing, 22, pp.:105-116, 2004.

[11] Bosworth, J. H., Acton, S. T., "Morphological scale-space in image processing", Digital Image processing, 2003, 338-367.

[12] Florack, L M J, Salden, A H, ter Haar Romeny, B M, Koenderink, J J, and Viergever, M A, "Nonlinear scale-space", Image and Vision Computing Volume 13 Number 4 May 1995.

[13] Remaki, L., Cheriet, M., "Visual Data Extraction From Bi-Level Document Images Using a Generalized Kernel Family with Compact Support, in Scale-Space", Proceedings of the 5th International Conference on Document Analysis and Recognition – IEEE Conference. pp. 609-612, 1999.

[14] Grossman, A, Morlet, J. , Decomposition of Hard Functions into square integrable wavelets of constant shape, SIAM, 1984.

[15] Canny, J., "A Computational approach to edge detection", IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 8, pp. 679-698, 1986.
[16] Mallat, S., Hwang, W. L., "Singularity Detection and Processing with Wavelets", IEEE Transactions on Information Teory, Vol.38, N°2, 1992.
[17] Lindeberg: "Detecting salient blob-like image structures and their scales with a scale-space primal sketch: A method for focus-of-attention", International Journal of Computer Vision, vol. 11, pp. 283--318, Dec. 1993.
[18] Gonzales, R. C., Woods, R. E., Processamento de Imagens digitais, 1992.
[19] Lindeberg, T., "Edge Detection and Ridge Detection with Automatic Scale Selection", IEEE, 1996
[20] Zhang, L., Lenders, P., Locating the Head Boundary with 2D Continuous Wavelet, International Simposium on Intelligent Multimedia, 2001.
[21] Antoine, J. P., "Image analysis with two-dimensional continuous wavelet transform", Signal Processing, 1993.
[22] Antoine, J. P., "Image analysis with two-dimensional continuous wavelet transform", Proc. of Int. Conf., France, 1989.
[23] Sheng, Y., "Wavelet Transform", The Transform and Applications Handbook, IEEE Press, 1996.

IEEE TRANSACTIONS ON APPLIED SUPERCONDUCTIVITY, VOL. 18, NO. 2, JUNE 2008

Measurement of Fast Voltage Transients in High-Performance $\mathrm{Nb_3Sn}$ Magnets

J. Lizarazo, A. F. Lietzke, G. L. Sabbi, P. Ferracin, S. Caspi, S. Zimmerman, J. Joseph, and D. Doering

***Abstract*—The Superconducting Magnet group at Lawrence Berkeley National Laboratory has been developing $\mathrm{Nb_3Sn}$ high-field accelerator magnet technology for the last fifteen years. In order to support the magnet R&D effort, we are developing a diagnostic system that can help identify the causes of performance limiting quenches by recording small flux-changes within the magnet prior to quench-onset. These analysis techniques were applied to the test results from recent $\mathrm{Nb_3Sn}$ magnets. This paper will examine various types of events and their distinguishing characteristics. The present measurement techniques are discussed along with the design of a new data acquisition system that will substantially improve the quality of the recorded signals.**

***Index Terms*—LARP, $\mathrm{Nb_3Sn}$, superconducting magnets.**

I. INTRODUCTION

A NOVEL Magnet Voltage Monitoring System (MVMS) is being developed at LBNL with the purpose of probing the voltage along multiple sectors of a coil cable during magnet testing. It has been observed that small transient imbalances occur inside high-performance $\mathrm{Nb_3Sn}$ magnets during training [1], [2]. These imbalances have different amplitudes and time profiles depending on the process that originated them.

For the most part the effect of a transient imbalance fades out and its induced voltage spike decays. However, as the current through the magnet increases, the effect of these transient imbalances becomes more relevant, especially when the magnet is approaching its short sample current and the superconducting margin is low.

The purpose of the new MVMS is to record the voltage spikes generated by magnet imbalances and the magnet quenching. This system will allow Supercon (Lawrence Berkeley Superconducting Magnet Group) to understand the origins of quenches and diagnose what is taking place inside a magnet. This diagnostic process is crucial to provide feedback for the manufacturing process.

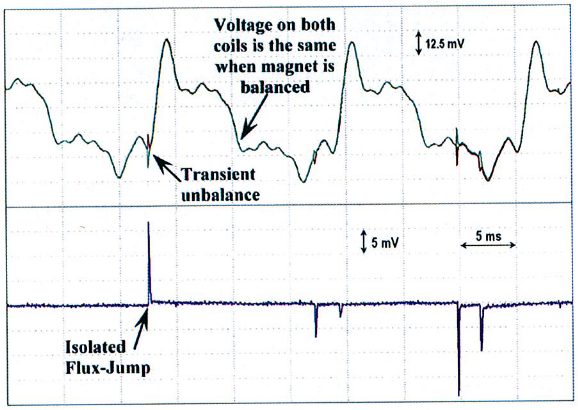


Fig. 1. The two traces at the top correspond to the raw voltage across each of the two coils in a dipole magnet (SD01b) for a Flux-Jump event. The signal at the bottom is obtained after subtraction of power supply noise (differential voltage between coils), uncovering 5 single Flux-Jumps.

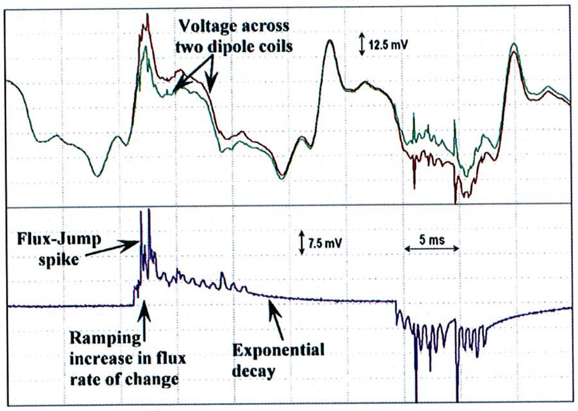


Fig. 2. Double Flux-Jump event at high field. As the current in the magnet is increased, the shape of Flux-Jump events becomes more complex. Top traces are the coil's raw voltage, bottom trace is their subtraction.

II. CHARACTERIZATION OF FAST VOLTAGE TRANSIENTS

Two kinds of fast transient events have been identified in $\mathrm{Nb_3Sn}$ magnets [2]–[4] and named after the suspected mechanism that originates them. The first kind is called Flux-Jump and it is usually the first kind of event observed during current ramping. It is believed that Flux-Jumps are a transient local collapse of diamagnetism in a superconductor, which generates an electric signal due to change in magnetic flux.

The second kind of fast voltage transient that has been observed is called Slip-Stick. It is believed that portions of a magnet can rearrange due to changes in magnet parameters like Lorentz force or temperature. When such rearrangements occur, a new equilibrium point is achieved with frictional forces holding the magnet at a more stable point. At the moment that a slippage occurs, energy is released and parts of the magnet experience a dumped vibration.

A. Flux-Jumps

Figs. 1 through 3 show a sequence of Flux-Jumps captured during training of the SD01b dipole. The two traces at the top

---

Manuscript received August 28, 2007. This work was supported by the Director, Office of Energy Research, Office of High Energy and Nuclear Physics, High Energy Physics Division, U. S. Department of Energy, under Contract DE-AC02-05CH11231.

The authors are with the Lawrence Berkeley National Laboratory, Berkeley, CA 94720 USA (e-mail: jlizarazo@lbl.gov).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TASC.2008.921306

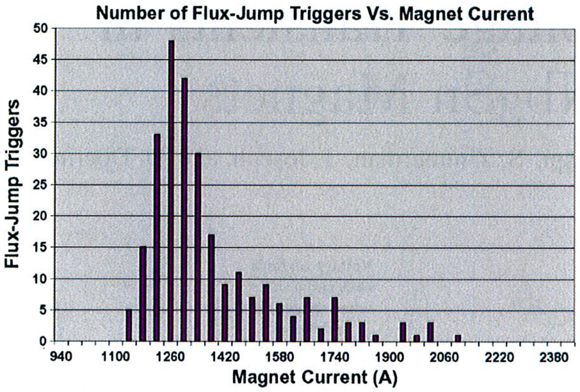


Fig. 3. Flux-Jump triggers distribution as a function of magnet current for SD01b superconducting magnet. Flux-Jump event rate peaks at 1300 A, with 70% of the Flux-Jumps happening within 15 seconds, between 1200 and 1400 A.

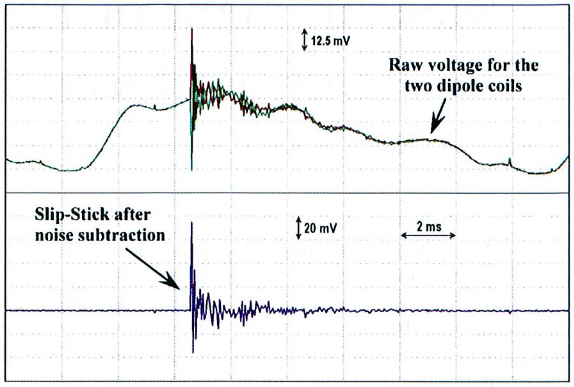


Fig. 4. The two traces at the top correspond to the raw voltage across each of the two coils in a dipole magnet (SD01b) for a Slip-Stick event. The signal at the bottom is obtained after subtraction of power supply noise (differential voltage between coils), showing a ringing associated with flux change due to mechanical vibration.

of figure 1 correspond to the voltage induced in each coil of the dipole. The signal amplitude is proportional to the change of flux through (enclosed by) the coil.

The plot at the bottom of Fig. 1 corresponds to the pure Flux-Jump event after common mode subtraction, showing 5 isolated Flux-Jumps at low field. These single Flux-Jumps are caused by the loss of diamagnetism at one single location, which did not trigger such same effect in its neighborhood.

As the magnetic field increases, the shape of Flux-Jump events becomes more complex, as shown in Fig. 2.

Three main characteristics can be seen in the high-field Flux-Jump events: 1) A series of spikes corresponding to a train of individual Flux-Jumps; Due to their proximity in time, it is reasonable to assume that they are correlated. 2) An integrated effect showing the net magnetic flux continuously increasing at a faster rate. This could be associated to the heat released by the Flux-Jump transferring into neighboring regions reducing their superconducting margin and creating further collapse of diamagnetism. 3) An exponential decay on the magnetic flux increase rate, with a decay time of a few ms.

Fig. 3 shows the Flux-Jump distribution during the first 10 training ramps in SD01b as a function of magnet current. Although the magnet current was varied from 0 to ~1800 A for all ramps, most Flux-Jumps occur at around one specific value of current (~1300 A) or strength of magnetic field. The width of the distribution is only ~15 seconds, ramping at 14 A/sec.

Since the concept of transient magnet imbalances was first introduced in FY2000 [5] a great deal of interest aroused in trying to connect this phenomenon with a theoretical model. For this purpose accurate measurements have to be made, which is one of the goals that we hope to achieve with the new MVMS being setup at LBNL.

B. Slip-Sticks

Fig. 4 shows a typical signal generated by a Slip-Stick in SD01b during magnet training. A characteristic common to all Slip-Sticks studied is the sudden start of an oscillation with frequency components in the order of several tens of kHz, and a damping factor with a time constant ~1.5 to 2 ms.]

Slip-Sticks start to occur at a higher field than Flux-Jumps when the Lorentz force on the coils is much higher, and becomes comparable to the force exerted on the coils by the magnet's mechanical structure [6].

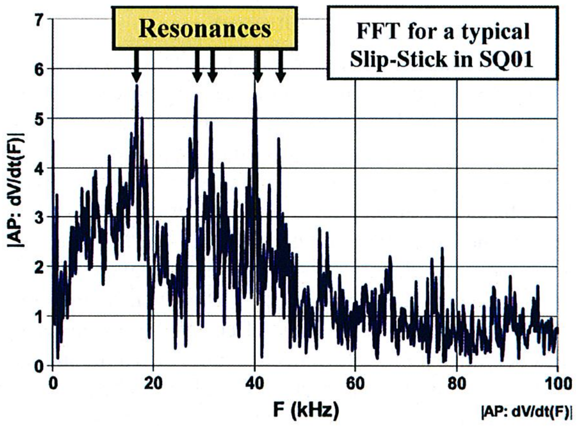


Fig. 5. Fast Fourier transform for a slip-stick collected during testing of a large-bore quadrupole magnet (SQ01) showing a collection of dominant frequencies [7].

While the total number of Flux-Jumps as a function of ramp number stays fairly constant, the number of Slip-Sticks rapidly decreases with ramp number. For the dipole discussed above (SD01b) there were 26 slips during the first training ramp, 7 in the second ramp, 1 in ramp 3, 6 in ramp 4, and only 1 or 2 slips in each of the subsequent 17 ramps. This data is telling us that most of the magnet's internal re-arrangement happens during the first few training ramps.

Fig. 5 shows a Fourier transform for a Slip-Stick event in a large-bore quadrupole called SQ01. This distribution shows a small set of resonant frequencies that could be related to the magnet's normal modes of mechanical vibration.

The new MVMS will allow us to make more precise measurements and find correlations, if any, between Slip-Stick parameters like frequency components and decay time, to magnet parameters that determine its normal modes of vibration like stress, material Young modulus, mechanical structure, etc.

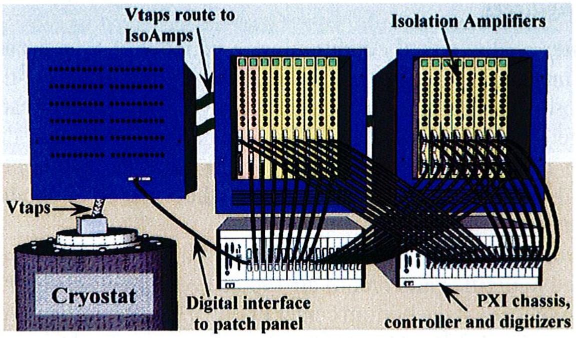


Fig. 6. 3D model for Supercon's new MVMS. An electronic patch panel connects to a cryostat and routes the out-coming signals to two chassis housing a set of isolation amplifiers (IsoAmps). The signals from the isolation amplifiers are digitized by a bank of commercial digitizers.

III. MVMS FUNCTIONALITY

Supercon's new MVMS is a custom made system to address our unique magnet testing needs. The following is a list of functionalities that the new MVMS will provide.

- Voltage monitoring of up to 136 magnet sectors including whole, half and quarter magnet. The system architecture provides easy scalability to much higher channel count.
- 1500 VDC can be applied between channels or channel-to-ground, in a safe and reliable manner.
- 120 kHz analog front end bandwidth.
- 500 kS/s sampling rate.
- 16-bit digitizing resolution.
- Three modes of Vtap (probing voltage tap) connectivity: 1) Consecutive channels monitor consecutive magnet sectors. 2) Re-routing of sector reference to bypass open Vtaps. 3) Debugging and calibration.
- User can select between derivative or linear amplification.
- Variable gain over two and a half orders of magnitude.
- Automatic creation of a test setup record. A log entry is made every time a change in the system setup is made.
- Independent acquisition of Quenches and voltage spikes.
- DC voltage measurement capability. This feature allows the user to check magnet connections and find open Vtaps before and after cool-down. This is an important feature to create confidence on the magnet test.
- BNC outputs for oscilloscope monitoring and debugging.

IV. MVMS ARCHITECTURE

Fig. 6 shows the front view of a 3D model for the new MVMS for Supercon's magnet test facility.

On the upper left corner we have an electronic patch panel that routes signals into two chassis housing high-voltage isolation amplifiers (top center and top right).

At the bottom center and bottom right there are two National Instruments PXI chassis housing a PC controller and a set of general purpose digitizer modules.

The upper section of the cryostat cooling the magnet is shown in the lower left corner of this same figure.

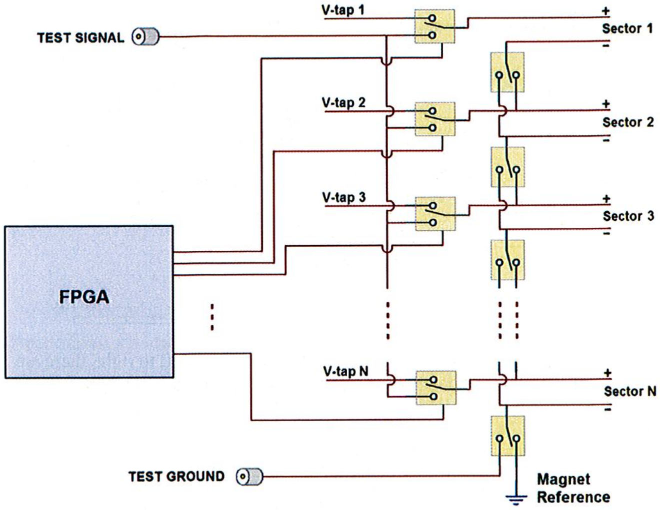


Fig. 7. Block diagram for the relays inside the electronic patch panel routing probing signals from the magnet to the isolation amplifiers. An FPGA communicates to the control PC and dynamically sets the relays state.

A. Electronic Patch Panel

This component receives all Vtaps coming from the magnet and uses a set of electromechanical relays to route them into a bank of amplifiers according to a user defined setup file.

Fig. 7 shows a schematic diagram illustrating how the relays are placed to bypass any number of consecutive V-taps and re-route reference signals to feed multiple consecutive channels. The state of each relay is controlled by an FPGA (Field Programmable Gate Array) inside the patch panel, which communicates to a PC using a commercial general purpose I/O module inside a PXI chassis.

The relays are rated for 3000 V coil-contact and 2500 V between contacts, this being the limiting factor for the patch panel voltage rating.

All relays can also be set to simultaneously connect all differential pair outputs to a test bus, which allows for easy debugging and calibration of the isolation amplifiers.

Every time a change in the connectivity settings of the voltage taps is made, a test setup record is created in software. This record serves as documentation for the experimental setup, which builds confidence in the test and avoids manual log entries and the errors that come with them.

B. Isolation Amplifiers

The bank of isolation amplifiers is split between 2 chassis capable of housing up to 10 8-channel amplifier cards each, providing a maximum of 160 monitoring channels.

Fig. 8 shows a simplified schematic of one isolation amplifier channel. A novel architecture is used to completely eliminate common mode rejection issues, without adjusting or matching any components. Our approach is to make each channel electrically floating, independent of the magnet or test facility grounds. Thus, each channel receives a pair of consecutive Vtaps and uses one of them as isolated reference.

The first stage of the channel is shown in Fig. 8 around the first Op-Amp from the input. This section brings the differen-

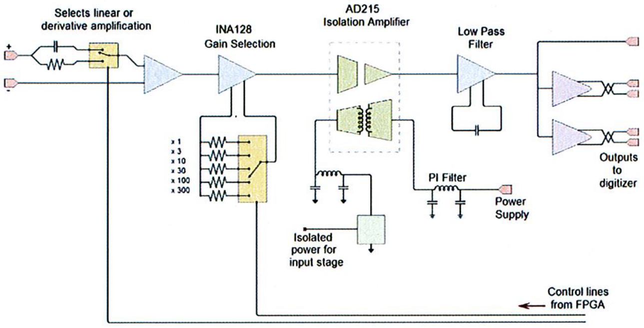


Fig. 8. Single channel isolation amplifier diagram. From left to right, there are 5 stages of signal conditioning: linear/derivative selection, gain set, high voltage isolation, low-pass filter, line drivers.

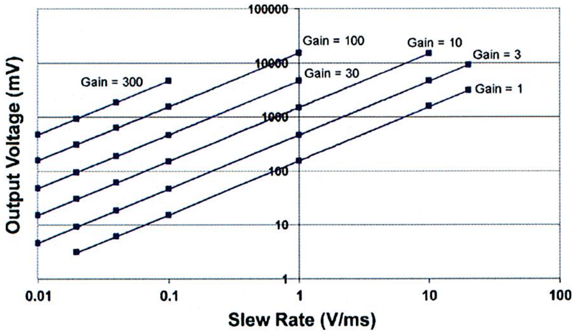


Fig. 9. Measured response for a single-channel isolation amplifier working in differential mode at several values of gain and input slew rates. The amplifier shows very good linearity over 4 orders of magnitude.

tial voltage of the Vtap pair to a range acceptable for commercial low-voltage devices. In addition, it allows the user to select between linear and derivative amplification. The main purpose of the derivative amplification is to provide a large signal range by measuring slew rates rather than the raw voltage from the magnet. On the other hand, the main purpose of the linear mode is to record small transient imbalances.

The second stage around the second Op-Amp from the input provides 5 gain settings for different digital resolutions. In linear mode the resolution goes from 20 mV at 8 bits, to 2000 V at 16 bits. In derivative mode the gain can be set to resolve from a slew rate of 10 mV/ms at 8 bits, to 25 V/ms at 16 bits.

The next device at the center of the schematic in Fig. 8, labeled as AD215, changes the reference of its input voltage. In our case, this device takes the voltage of the positive Vtap after the first two stages of conditioning referenced to the negative Vtap, and outputs the same differential voltage but referenced to the test facility ground. It also provides isolated power for the first two stages of signal conditioning.

The next stage in the channel's signal conditioning path is a low pass filter with frequency cutoff at ~200 kHz, increasing rejection to 60 dB per decade, with the purpose of providing the right Nyquist frequency to the digitizer. It also fans out to the differential line drivers and the general purpose BNC.

Fig. 9 shows the measured response of the prototype channel for several values of input slew rates and channel gains. The lines on top of the data are linear fits showing excellent level of linearity over 4 orders of magnitude.

C. *Digitizers*

After analog signal conditioning, a window of time around each magnet imbalance has to be digitized and recorded. For that purpose, the signals from the isolation amplifiers are duplicated and fed into two commercial digitizer modules. The two digitizers are triggered independently, such that one module only records Flux-Jumps and Slip-Sticks, while the other module only records the magnet quench.

Two PXI chassis filled each with 17 digitizer modules are needed to digitize and record all 136 analog channels. The PXI chassis shown at the bottom center in Fig. 6 houses the digitizers used to record voltage transients, while the chassis at the bottom right houses the digitizers that record the magnet quench. This architecture allows us to solve any issues related to digitizer dead-time after trigger.

For every trigger, each channel is digitized at a rate of 500 kS/s with 16 bit resolution during a window of time defined by the user. The default window being recorded around each event is 100 ms long, where 30 ms are used on pre-trigger samples and 70 ms on post-trigger samples. However, the system can be set to record up to 8 seconds long events.

The digital I/Os for each digitizer module are used to talk to the FPGA in each of the isolation amplifiers, and in this way read and write amplifier settings and status.

An embedded PC-controller running Windows Vista controls settings and communication with the isolation amplifiers and the electronic patch panel. This embedded PC also allows for remote-desktop connection to control the system remotely from the test facility control room.

V. CONCLUSIONS

Analysis of signals collected from recent superconducting magnets at LBNL shows that the fast transient voltages that occur before magnet quenching have the potential of providing useful information in understanding magnet behavior. For the purpose of collecting these events, a custom system is being setup and an incomplete version of it will be used for the testing of the HD2 magnet at the end of December 2007 at LBNL.

Several tests have been performed to build confidence on the data to be collected with the new MVMS. The data collected with HD2 will be analysed and used to provide feedback for the system's final design and implementation.

REFERENCES

[1] A. F. Lietzke, "Differentiation of performance-limiting voltage transients during Nb_3Sn magnet testing," in *ICMC-2005 Conference Proceedings*.

[2] S. Feher, B. Bordini, R. Carcagno, A. Makulski, D. F. Orris, Y. M. Pischalnikov, C. Sylvester, and M. Tartaglia, "Sudden flux change studies in high field superconducting accelerator magnets," *IEEE Trans. Appl. Supercond.*, vol. 15, no. 2, p. 1591, Jun. 2005.

[3] A. F. Lietzke *et al.*, "Test results for HD1: A 16 T Nb_3Sn dipole magnet," *IEEE Trans. Appl. Supercond.*, vol. 14, no. 2, p. 345, Jun. 2004.

[4] A. F. Lietzke *et al.*, "Test results for RD3c, a Nb_3Sn, racetrack dipole," *IEEE Trans Appl. Supercond.*, pp. 1292–1296, 2002, Houston.

[5] A. F. Lietzke *et al.*, "Race-track coil technology validation (RT-1) test results," in *VLHC Workshop*, May 24, 2000, FNAL.

[6] A. F. Lietzke *et al.*, "Test results for HD1: A 16 T Nb_3Sn dipole magnet," *IEEE Trans. Appl. Supercond.*, vol. 14, no. 2, p. 345, Jun. 2004.

[7] A. F. Lietzke *et al.*, Voltage-Spike Diagnostics of Under-Performing Nb_3Sn Magnets Internal LBNL communication.

Robotica (2005) volume 23, pp. 467–477. © 2005 Cambridge University Press
doi:10.1017/S0263574704001183 Printed in the United Kingdom

Trajectory planning for lab robots based on global vision and Voronoi roadmaps

Waldir L. Roque* and Dionísio Doering

Instituto de Matemática, Universidade Federal do Rio Grande do Sul, 91509-900 Porto Alegre, RS (Brazil)

(Received in Final Form: October 11, 2004)

SUMMARY

This paper discusses the techniques and their applications in the development of a path planning system composed of three modules, namely: global vision (GVM), trajectory planning (TPM) and navigation control (NCM). The GVM captures and processes the workspace image to identify the obstacle and the robot configurations. These configurations are used by the TPM to generate the Voronoi roadmap, to compute the maximal clearance shortest feasible path and the visibility pathway between two configurations. The NCM controls the robot functionalities and navigation. To validate the path planning system, three sets of experiments have been conducted using the Lab robot Khepera, which have shown very good results.

KEYWORDS: Mobile robots; Robot navigation; Path planning; Voronoi roadmap; Global vision

1. INTRODUCTION

Research in autonomous mobile robots (AMR) has shown great progress in the last decades. The development of AMR is getting closer to the final user as can now be seen, for instance, in domestic appliances and in the growing number of companies producing AMR for laboratory research.[1] The potential of applicability of mobile robots is high and their demand will be raised in the near future when both the technology and the cost benefit relation become much more favorable than today. Despite all this progress there are still many problems that have to be solved and better understood in mobile robotics, some of them are very much domain specific. One of the main concern for autonomous navigation in either unstructured or structured environments is collision-free path planning, which requires a precise navigation system.

Several techniques have been proposed in the literature[2] to deal with the path planning problem, among them the generalized Voronoi diagram (GVD) has been addressed by several authors as it exhibit some nice properties, like the generation of a roadmap with maximum clearance from the obstacles and its corresponding graph is fully connected for bounded workspaces, which provides connectivity and safe navigation among the obstacles. The application of the GVD has somewhat a drawback as to generate the roadmap the geometry of the workspace should be known *a priori*. In other words, it is important to know the obstacles and the robot configurations $\mathcal{C}_o$ in the workspace to obtain the corresponding free configuration space $\mathcal{C}_f$. This drawback may be circumvented if the robot can get to know the geometry, globally. This may be achieved by the use of sensorial data of the workspace, which may apply different sensors to construct the geometry.

Some results of the local GVD generation based on sonar sensor readings have been presented by Choset[3] and based on laser sensor readings by Blanco[4] and by Markovic.[5] However, in these approaches only local sensorial information of the workspace geometry is captured and only the very local roadmap can be constructed. An alternative approach to get the whole workspace geometry is collecting sensorial data from a global vision device and then letting the GVD roadmap to be generated based on this sensorial data.

The simple generation of the Voronoi diagram for a set of planar object shapes based on digital images was proposed in papers.[6–8] An advantage of the local sensing is that the sensors can normally be set on board of the robot, while for global vision it is somewhat more difficult. However, with the current wireless communication the sensorial data can be transferred to the robot or to a host computer in real-time, providing fully autonomy.

In this paper we present the development of a path planning system, called RoTA, based on the generation of GVD roadmap retrieved from the geometrical structure extracted from the obstacles image captured from the workspace. The path planning system has an architecture composed of three main modules, namely: i) the global vision module (GVM), ii) the trajectory planning module (TMP) and iii) the navigation control module (NCM). In section 2, we describe each module, the techniques involved and their integration into the path planning system. Some experimental results realized with an actual laboratory robot is discussed in section 3 and in section 4 some comments, final remarks and conclusions are addressed.

2. THE SYSTEM MODULES AND TECHNIQUES

In what follows we give a short description of each module describing the mathematical and computational techniques involved in their developments. Figure 1 illustrates the overall system architecture.

2.1. Global vision module

The global vision module is responsible for capturing and processing the image of the actual workspace. This module

---

* Corresponding author. roque@mat.ufrgs.br

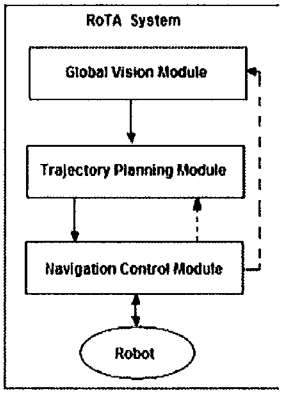


Fig. 1. The block diagram of the RoTA system architecture.

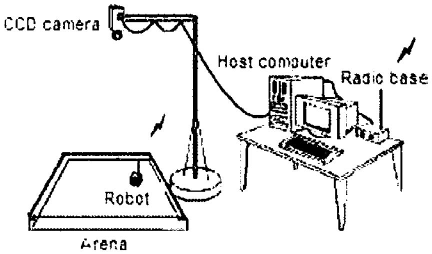


Fig. 2. Illustration of the GVM physical device.

is currently composed of a single CCD digital camera (NTSC, 640 × 480 resolution, RGB color model) that is placed 1.8 m high on the top center of the workspace arena (see Fig. 2). The arena is rectangular with dimensions 1.00 × 1.10 m. The workspace illumination is carefully set and the images are captured at a rate of 30 frames per second.

The obstacle shapes considered are generalized polygons and circles. Generalized polygons are polygons whose edges can be both straight line and arc segments. The actual obstacles are made of polystyrene with 3 cm hight, surrounded by an aluminium paper and with a top cover in black to contrast with the arena background that is set in white. To obtain the geometrical structure of the workspace scenario, a single image is captured, and standard image processing techniques for segmentation and binarization by thresholding are applied. The connected components of the obstacles are labelled and identified to provide their geometrical data structure.

2.1.1. Generalized polygon and circle representations.

Any simple polygon can be characterized by its ordered sequence of vertices[9] and a circle by its center and radius. A generalized polygon can be also represented by a set of vertices, where the arc edges can be approximated by the set of vertices that form its polygonal line approximation. Ultimately, a set $V = \{v_1, v_2, \ldots, v_n\}$, where v_i are the vertices, and $C = (x_0, y_0, r)$, where (x_0, y_0) and r are the coordinates of the center and r the radius of a circle, are the minimum geometrical data structure sufficient to represent the obstacles.

To be able to uniquely identify the geometrical data structure of polygonal obstacles, we first apply an algorithm for edge detection, based on the Laplacian technique,[10] which provides the ordered sequence of pixels forming the obstacles' border. Although the sequence of pixels of the border defines a 8-connected digital generalized polygon, the number of border pixels are very much higher then compared to its number of vertices. Therefore, to minimally characterize the obstacles, the polygon vertices should be identified. For that, a corner detection algorithm should be applied to get the corner pixels on the border sequence.

2.1.2. Vertex detection.

There are several algorithms proposed in the literature for corner detection.[11–14] The corner detection algorithms are generally classified into two major categories: i) the boundary-based techniques, that are based on the fact that the boundary pixels of the object are known, and ii) the gray-level techniques, that take into account morphological properties of the image. Recently, an algorithm was proposed by Tsai et al.,[15] based on the former approach, which is robust, easy to implement and quite stable to noise. It can also be able to detect both convex and nonconvex vertices.

In brief, Tsai's algorithm relies on the analysis of the eigenvalues of the covariance matrices of a digital curve segment. For that, it takes a sequence $P = \{p_1, p_2, \ldots, p_n\}$ of n pixels that form the border of an object, where $p_i = \{(x_i, y_i), i = 1, 2, \ldots, n\}$, corresponds to the Cartesian coordinates of the pixel and the pixel p_{i+1} is the neighbor of pixel p_i modulo n. A region of support is defined around a pixel p_i as $S_k(p_i) = \{p_j;\ j = i-k, i-k+1, \ldots, i+k-1, i+k\}$, where k is an integer number that defines the amplitude of the supporting region. The covariance matrix is computed by

$$C = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}, \tag{1}$$

where

$$c_{11} = \left(\frac{1}{2k+1} \sum_{j=i-k}^{i+k} x_j^2 \right) - c_x^2, \tag{2}$$

$$c_{22} = \left(\frac{1}{2k+1} \sum_{j=i-k}^{i+k} y_j^2 \right) - c_y^2, \tag{3}$$

$$c_{12} = c_{21} = \left(\frac{1}{2k+1} \sum_{j=i-k}^{i+k} (x_j\, y_j) \right) - c_x\, c_y, \tag{4}$$

$$c_x = \frac{1}{2k+1} \sum_{j=i-k}^{i+k} x_j, \quad c_y = \frac{1}{2k+1} \sum_{j=i-k}^{i+k} y_j. \tag{5}$$

As the covariance matrix is Hermitian, there are two real eigenvalues, which are given by:

$$\lambda_L = \frac{1}{2}\left(c_{11} + c_{22} + \sqrt{(c_{11} - c_{22})^2 + 4c_{12}^2}\right), \tag{6}$$

$$\lambda_S = \frac{1}{2}\left(c_{11} + c_{22} - \sqrt{(c_{11} - c_{22})^2 + 4c_{12}^2}\right). \tag{7}$$

According to this technique, a corner is detected when the eigenvalues λ_L or λ_S of the covariance matrices are respectively smaller than or greater than a predefined threshold value. The corners are separated by at least a set of k pixels. In fact, just one of the eigenvalues can be used to identify the corners. For the case of λ_S, it has been experimentally observed that pixels on a straight line have their λ_S eigenvalues very close to zero whereas eigenvalues much greater correspond to a vertex. Therefore, in our specific application, after some experimental measurements, good results were obtained by setting the region of support parameter as $k = 7$ and the threshold value for detecting a vertex as $\varepsilon = 10^{-1}$.

Unfortunately, Tsai's algorithm has shown to be very poor to identify digital circles, which according to the covariance matrix analysis, would happen when the eigenvalues are such that $\lambda_S \approx \lambda_L$. Therefore, an alternative approach has to be applied to identify circular obstacles.

2.1.3. Digital circles. To recognize digital circles an algorithm was proposed by Sauer[16] with linear time complexity $\mathcal{O}(n)$, where n is the number of pixels. In short, the algorithm is as follows: Let $C(r, x_m, y_m)$ be a digital closed curve where (x_m, y_m) is the centroid coordinate and r is the radius mean value, then for a digital curve[17] to be a digital circle the following relation has to be fulfilled:

$$p_y - 0.5 - y_m \leq \pm\sqrt{r^2 - (p_x - x_m)^2} \leq p_y + 0.5 - y_m, \quad (8)$$

where (p_x, p_y) are the pixel coordinates. The implementation of this algorithm was done to decide if the sequence of pixels on the border of an obstacle form a digital circle. Essentially, the identification of a circle translates to the computation of the centroid $C = (x_0, y_0)$ and the radius r under a threshold value.

The identification of vertices and edges allows a minimal characterization of the geometrical structure of an object and from that its full reconstruction. For generalized polygons, the arc edges can be approximated by polygonal lines, where their vertices can be obtained by the analysis of the λ_S eigenvalues. The geometrical data structure of the obstacles are the input data for the TPM, as will be seen in the section 2.2.

2.1.4. The robot configuration and geometric data structure. For the GVM to be able to identify the robot configuration (position and orientation) in the workspace, the robot provides two special landmarks. One is a small circle and the other is a small triangle, both are in black and attached on the top of the robot, with known areas and being far apart by a known displacement d, measured from their centroids. These landmarks are sufficient to specify the robot position and orientation on the scene, as the robot position is provided by the small circle centroid, which was chosen to coincide with the robot's center* and the orientation is measured with the vector pointing out from the circle's centroid to the triangle's one. With this approach the localization problem of the robot in the workspace is solved.

The geometrical data structure of the obstacles provided by the GVM are written in the following format:

E v h G
R r
T (x, y, θ)
P (x, y) r
S (x_0, y_0) (x_1, y_1) r
C (x, y) R r
A r N $(x_1, y_1)(x_2, y_2) \ldots (x_N, y_N)$
L r N $(x_1, y_1)(x_2, y_2) \ldots (x_N, y_N)$

In the above structure, E specifies the parameters for the vertical (v) and horizontal (h) workspace scales and G is the grid size. The R specifies the robot with radius r as we are considering a disc-like robot; T specifies the robot configuration with (x, y) being its position in Cartesian coordinates, where (0, 0) is measured at the top left corner of the workspace, and θ its orientation given in degrees with respect to the positive x direction; P specifies the Cartesian coordinates (x, y) of a single pixel; S corresponds to a straight line segment starting at pixel (x_0, y_0) and ending at pixel (x_1, y_1); C gives the pixel coordinates (x, y) of the center of a circle of radius R; A gives the ordered sequence of N vertices starting at pixel coordinate (x_1, y_1) and ending at (x_N, y_N) corresponding to the polygonal lines that form an arc; finally, L gives the ordered sequence of N vertices that form a generalized polygon.

As the robot model considered is holonomic disc-like, thus the technique of isotropic expansion[18] of the obstacles can be done to find the collision-free configuration space $\mathcal{C}_f$ preserving the Voronoi roadmap structure. In other words, the obstacles configuration space $\mathcal{C}_o$ is transformed in such way that the extended robot will be considered a point-like one and the obstacles are isotropically expanded by the robot radius.

2.2. *Trajectory planning module*

The trajectory planning module is responsible for the generation of a roadmap and from that to identify the shortest feasible path among the free configuration space $\mathcal{C}_f$. The roadmap will be constructed based on the retraction technique of the Generalized Voronoi Diagrams.

2.2.1. The generalized Voronoi diagram. The Voronoi diagram retraction technique has been of much interest to the robot path planning problem[19] as it has the desired property of generating maximal clearance roadmaps on the obstacles free space, where the mobile robot shall navigate. In what follows, we briefly introduce this geometrical construction.

The planar ordinary Voronoi diagram (OVD)[20] is defined as a partition of the plane into regions according to the principle of the *nearest neighbour*. More precisely, consider the Euclidean distance $d(p, p_i)$ from a point p to a set of non-collinear points p_i in the plane. A Voronoi region $R(p_i) \equiv R_i$ generated by a point p_i is defined as

$$R_i = \{p; d(p, p_i) \leq d(p, p_j), \forall p_i \neq p_j\}. \quad (9)$$

The Voronoi diagram $V(P)$ for a set of points $P = \{p_1, p_2, \ldots, p_n\}$, is defined as the union of all Voronoi regions $V(P) = \bigcup_{i=1}^{n} R(p_i)$.

---

[*] Recall that the robot is disc-like.

The points p_i are called Voronoi generators, the edges between two Voronoi regions are called Voronoi edges and the vertices where 3 or more Voronoi edges meet are called Voronoi vertices. We say that a Voronoi generator p_i is adjacent to p_j when their Voronoi regions share a common edge. According to its definition, the Voronoi diagram is such that any point on the edge of two neighboring regions is equidistant from the corresponding Voronoi generators. Therefore, points on the edges of two common Voronoi regions are as far apart from one generator as from the other. In other words, if one thinks of the Voronoi generators as point-like obstacles, the edges are the *maximal clearance paths* (MCP) among the obstacles.

The concept of planar Voronoi diagram can be extended to more complex sets. Therefore, we may construct the generalized Voronoi diagram (GVD) for a set of objects like lines, arcs, circles and polygons.[19]

Let $L = \{L_1, L_2, L_3, \ldots, L_n\} \subseteq \mathcal{R}^2$ be a set, where L_i is a line segment or an arc, such that they do not intersect each other, $L_i \cap L_j = \emptyset$, for $i \neq j$. Let's define the distance from a point p to L_i by the shortest distance between p and a point p_i on L_i by

$$d_s(p, L_i) = \min_{x_i}\{\|\mathbf{x} - \mathbf{x}_i\|\};\ \mathbf{x}_i \in L_i, \tag{10}$$

where $\mathbf{x}$ and $\mathbf{x}_i$ are the position vectors of p and p_i, respectively. The Voronoi region for $R(L_i)$ is given by

$$R(L_i) = \{p; d_s(p, L_i) \leq d_s(p, L_j),\ j \neq i,\ j \in I_n\}. \tag{11}$$

The union set of all Voronoi regions given by $V(L) = \bigcup_{i=1}^{n} R(L_i)$ generates the line Voronoi diagram for the set L.

Concerning our interest in the path planning problem, these objects may simulate the presence of one-dimensional obstacles like walls on the workspace through their orthogonal projections on the plane. The edges form the MCP among these obstacles.

A natural extension of the above concepts for the case of polygonal objects can be done by considering the generation of the Voronoi diagram for areas. The idea is to consider a set of areas $A = \{A_1, A_2, \ldots, A_n\}$ in $\mathcal{R}^2$. Assuming that the areas are connected closed sets with unity Euler number and that they do not intersect each other, we define a distance from a point p to an area A_i, as the shortest distance from p to p_i on A_i as follows:

$$d_s(p, A_i) = \min_{x_i}\{\|\mathbf{x} - \mathbf{x}_i\|\};\ \mathbf{x}_i \in A_i, \tag{12}$$

where $\mathbf{x}$ and $\mathbf{x}_i$ are the position vectors of p and p_i, respectively. According to this distance, we may define the Voronoi regions associated to each area $R(A_i)$ as

$$R(A_i) = \{p; d_s(p, A_i) \leq d_s(p, A_j),\ j \neq i,\ j \in I_n\}. \tag{13}$$

The area Voronoi diagram is the set $V(A) = \bigcup_{i=1}^{n} R(A_i)$.

It is important to note that the area Voronoi diagram subsumes the line and the ordinary Voronoi diagrams. The area Voronoi diagram can be seen as the diagram for a set of generalized polygonal obstacles, where the generalized polygon is identified by its area contour. Thus, in a general scenario an obstacle can be simulated by its corresponding orthogonal projection on the plane.

2.2.2. Computational generation of roadmaps. Several algorithms have been proposed to generate the Voronoi diagram for a set of objects.[21] In Sugihara[22] an incremental algorithm for OVD with average running time complexity of $\mathcal{O}(n)$ was proposed, where n is the number of generators. This algorithm has an additional advantage as it is robust to numerical errors. The construction of the OVD starts with a trivial diagram for 3 generators and then adds up the new generators one by one at a time. The principle is very simple: To generate the new Voronoi region $R(p_l)$ for a generator p_l we first identify the generator p_i, whose region contains the new generator p_l. Then we draw the perpendicular bisector between p_l and p_i until it intersects the edges of $R(p_i)$. The bisectors intersect the edges of $R(p_i)$ in two points. Let us call q one of them. The point q is at the boundary of the two Voronoi regions $R(p_i)$ and $R(p_j)$, respectively. Now take the perpendicular bisector of p_l and p_j, starting from q till it intersects another edge of $R(p_j)$. This procedure should be followed for all Voronoi generators surrounding p_l until we get back to region $R(p_i)$. Removing the edges enclosed by the sequence of bisectors the Voronoi region for the new generator p_l is found.

Figure 3 illustrates the generation of a new Voronoi region when a Voronoi generator P15 is introduced.

Departing from the OVD approach, an extension of the algorithm has been done to construct an *approximate* GVD (AGVD) for generalized polygons and circular obstacles. For that, let consider a straight line or arc segment as a sequence of n points with a constant small displacement δ between them. This provides a polygonal line approximation for the actual line or arc segment. The incremental type algorithm above can be applied to generate the AGVD for area objects if the Voronoi edges that cross the generalized polygon edges are invalidated and omitted, remaining only the Voronoi edges for adjacent objects (see Fig. 4).

In addition, to allow a more realistic obstacle model, we have considered obstacles with non-positive Euler numbers, that means obstacles with holes. As the regions inside the obstacles are not feasible free configuration space for the robot navigation, a point-in-polygon algorithm was implemented to invalidate these regions and let the AGVD be computed for obstacles with holes (see Fig. 5).

The resolution displacement parameter δ can be arbitrarily adjusted at the expense of a higher computational cost. As we are considering a disc-like robot with radius r, and taking into account that the Voronoi edges are the maximal clearance paths for the robot collision-free navigation, this will provide a natural threshold for the resolution displacement parameter as $\delta < 2r$. The Voronoi edges that are invalidated and omitted in the AGVD would, otherwise, lead to the robot collision with the obstacles. This algorithm has been implemented to generate the approximate roadmap for extended obstacles, based on the minimal geometrical structure that are extracted from the obstacles digital images.[23]

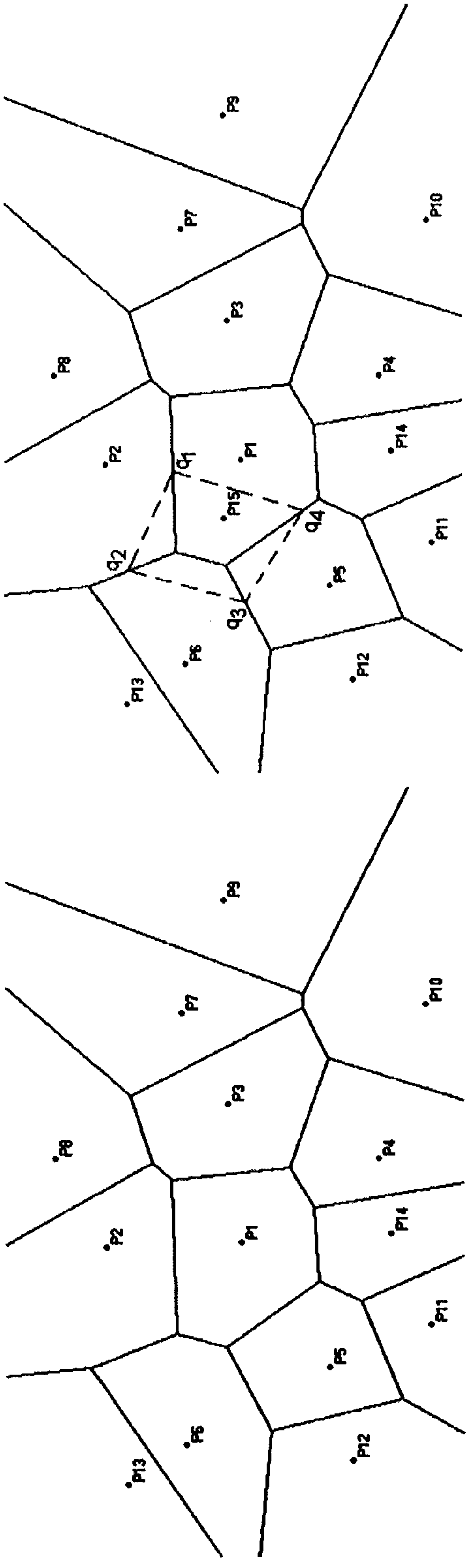


Fig. 3. OVD for 14 point generators (left) and the new Voronoi region defined when the point P15 is inserted.

It is worthwhile mentioning here that the generation of a GVD based on the digital image of obstacles is, in fact, an AGVD due to the discrete nature of the pixels.

2.2.3. The robot configuration. The robot initial configuration in the workspace is identified and provided by the GVM. The visiting station configurations, which are target configurations to be reached by the robot, are specified by the user with the mouse clicking at the desired position on the free configuration space and indicating the required orientation in a dialog box. The TPM transforms the data structure of the roadmap into a graph and computes the maximal clearance shortest feasible path (MCSFP) along the sequence of all visiting stations. The MCSFP is the shortest path along the Voronoi edges between two robot configurations. The MCSFP applies the Dijkstra's algorithm on the Voronoi graph. Nevertheless, the MCSFP is not the actual shortest path in the collision-free configuration space $\mathcal{C}_f$, although it can be interpreted as the safest. The MCSFP is not a complete smooth path as it contains several turning points, which may cause navigation errors along the path. Therefore, an improvement is further achieved by computing the *visibility pathway* (VP), which is the shortest path with the minimum number of turning points that can be computed based on the visibility principle applied to the set of vertices defining the MCSFP.

Although the VP is already a much better path compared to the MCSFP, it is not yet the shortest path homotopically speaking to the set of possible paths between two positions in the free configuration space $\mathcal{C}_f$. Figure 6 shows the AGVD over the image representation of the obstacles, the MCSFP and VP for two visiting stations.

2.3. *Navigation control module*

The NCM is responsible for controlling all the robot functionalities as reading the sensorial data provided by the robot sensors, control the radio, arm/gripper, IO turrets, motors, etc. The data structure of the visibility pathway for a sequence of visiting stations are read by the NCM which makes the robot execute the navigation and configurations roles.

To properly control the robot navigation, the NCM need to know the relationship between the actual scale of the workspace arena and the workspace window displayed by the GVM. In other words, the length/pixel scale has to be set by the user for the x and y coordinates. This is done through the configuration menu of the RoTA system.

The NCM is currently adapted to the Khepera[24] Lab robot, which is a disc-like robot with 5 cm of diameter and containing 8 infrared proximity sensors distributed around its perimeter. The communication between the host computer and the robot is wireless using a radio base and a radio turret under the SERCOM protocol. The robot energy can be supplied by either an umbilical wire connection or by its own battery. The umbilical connection causes a lot of problems while the battery supply is far better when using the wireless communication between the robot and the host computer.

2.3.1. Recharging stations. Khepera's battery recharging procedure requires human intervention. However, for the purpose of a further application, we have assumed that a set of recharging stations (RS) can be specified by the user in

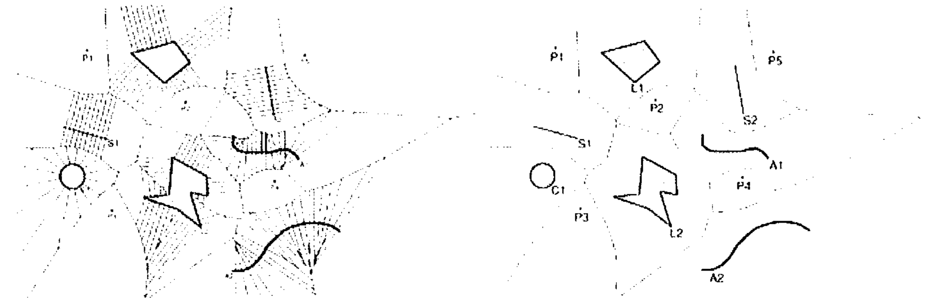


Fig. 4. GVD for a set of obstacles including the invalid Voronoi edges (left) and omitting the invalid ones (right).

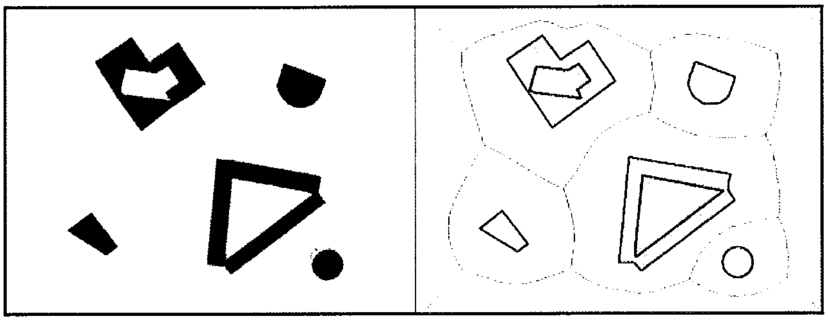


Fig. 5. Image of obstacles including two with holes (left) and their GVD (right).

the workspace. Essentially, a RS is a special configuration on the workspace where the robot can recharge autonomously its batteries. A RS is not considered as an obstacle during the roadmap generation; in other words, the robot or an obstacle can occlude a recharging station. When this occurs, the RS is not considered as a valid one at that moment, as it is not available.

When the robot battery drops to a minimum charge value the NCM halts all robot activities, if any, and requests the TPM to verify which RS is the closest to the robot, and plans its navigation according to the visibility pathway length from the current robot configuration to the nearest RS. Once this is done, the robot navigates to this recharging station. If the robot was executing a task at that time, after the robot's battery has been recharged, the system replans the trajectory from the current RS to continue executing the previous requested task.

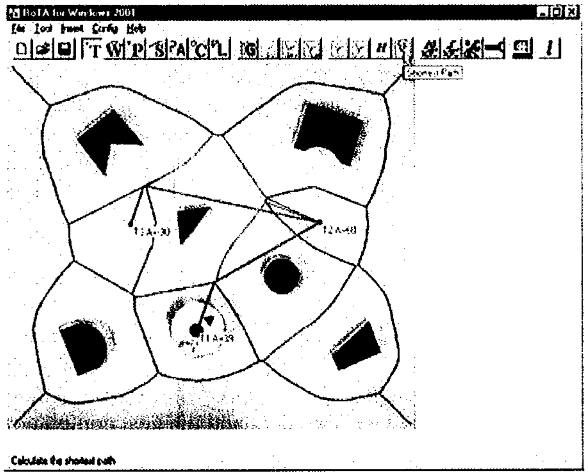


Fig. 6. Visibility pathway for two visiting stations (T2 and T3). The robot orientation to be reached at the stations are specified by the angle A.

2.4. The RoTA system

The integration of the three modules gives origin to a path planning system called RoTA. Figure 7 shows a screen dump of the RoTA system main window. The workspace has a set of obstacles with different geometries and the robot is identified by the circle-triangle coupled landmarks. A set of two visiting stations are seen as T2 and T3, with T1 being the initial

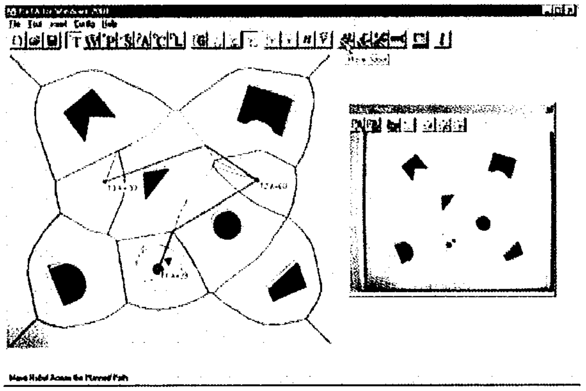


Fig. 7. Screen dump of RoTA system. The obstacles, the robot and the visiting stations are seen on the system window and the real-time image of the workspace is displayed on the image viewer window.

configuration of the robot. Each station is characterized by the position and orientation that the robot should match. The MCSFP and the visibility pathway connecting the visiting stations are shown. On the right-hand side the real-time GVM visualization window can be seen. Through this window the user can visualize the actual robot navigation in the workspace through the screen.

Once the set of visiting stations has been specified the system performs all planning and starts the robot navigation.

3. EXPERIMENTS WITH KHEPERA

To validate the RoTA system, a set of three navigation experiments has been done using the Khepera[24] robot. The first set of experiments consisted in measuring the deviation occuring during 20 navigation paths between the configuration chosen by the user as a goal to be reached by the robot and the actual configuration reached by the robot after its navigation. The robot configuration is always measured by the GVM. The length/pixel scale rate was set to 1.87, which provides an approximate arena of 84 cm x 93 cm according to the workspace image that was cropped to 450 x 500.

3.1. Multiple visiting stations

The first experiment concerns robot navigation to accomplish the task of visiting two stations that are chosen randomly by the user on the workspace. The configuration reached at each visiting station is measured and the result compared to the goal configuration values that were specified by the user. No intruder obstacle or recharging request were considered during this experiment. The results are shown in the Table I. In Fig. 8 the deviations between the goal and reached configuration coordinates are shown. Figure 9 illustrates the

Table I. Deviation measurements for the configuration coordinates.

Configuration variables	x	y	θ
Mean (95% confidence)	0.0657	0.1453	0.1462
Lower Bound	0.0318	0.0630	0.0232
Upper Bound	0.0996	0.2276	0.2690
Std. deviation	0.0724	0.1758	0.2628

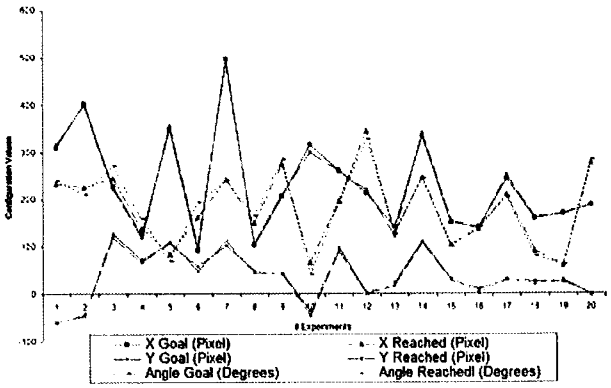


Fig. 8. Deviation measured between the goal and reached configurations for two stations.

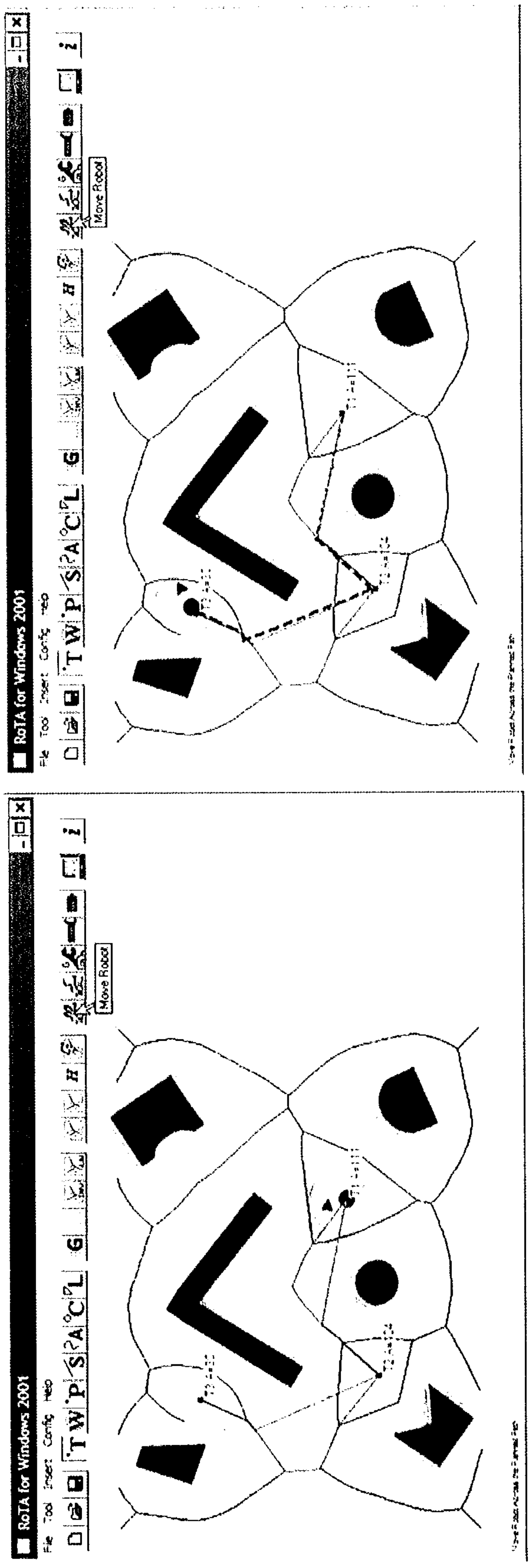


Fig. 9. Visibility pathway (full line on the left) and the actual robot navigation path (dashed line on the right).

visibility pathway and the actual navigation path done by the robot for a single experiment.

3.2. Recharging stations

The second set of experiments consisted of measuring the configuration reached by the robot, when while navigating it was requested to visit a recharging station, versus the actual configuration of the recharging station that was previously set by the user. For the purpose of this experiment, only two recharging stations were considered, one at the upper left corner (position at (50, 50) and orientation of 0° degree) and the other at the bottom right corner (position at (450, 350) and orientation of 180° degree) of the workspace.

As currently the robot does not provide a sensor for battery charge reading, the recharging command was simulated by the user. Once the user request for recharging, the NCM halts the robot navigation, the TPM computes the visibility pathway towards the closest recharging station and the NCM controls the robot navigation to it. After the robot reaches the RS, it recomputes the path to accomplish the previous task. Table II shows the results, Fig. 10 shows the deviations observed when the robot had to be recharged and Fig. 11 illustrates the visibility pathway and the actual robot navigation to the recharging station for a single experiment.

3.3. Replanning due to an intruder obstacle

The third set of experiments was done considering the same deviation approach as in 3.1, but now an intruder obstacle was introduced during the robot navigation. Actually, the intruder obstacle was posed in such way that the robot proximity sensors detect it along the forward navigation direction. Once an intruder is detected, the NCM halts the robot navigation, the GVM computes the new workspace scenario and the

Table II. Deviation measurements due to recharging.

Configuration variables	x	y	θ
Mean (95% confidence)	0.1029	0.2085	0.1157
Lower Bound	0.0415	0.0182	0.0032
Upper Bound	0.1642	0.3988	0.2335
Std. deviation	0.1309	0.4065	0.2518

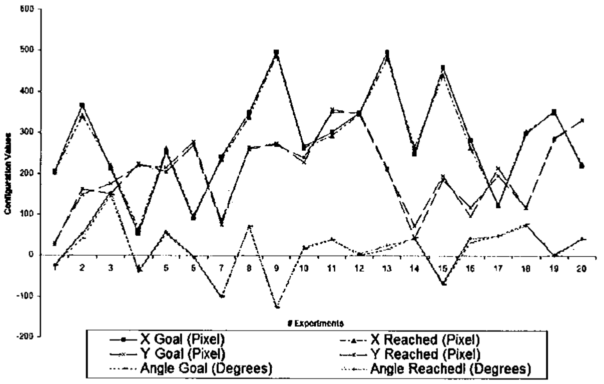


Fig. 10. Deviation measured for the configuration coordinates when recharging took place.

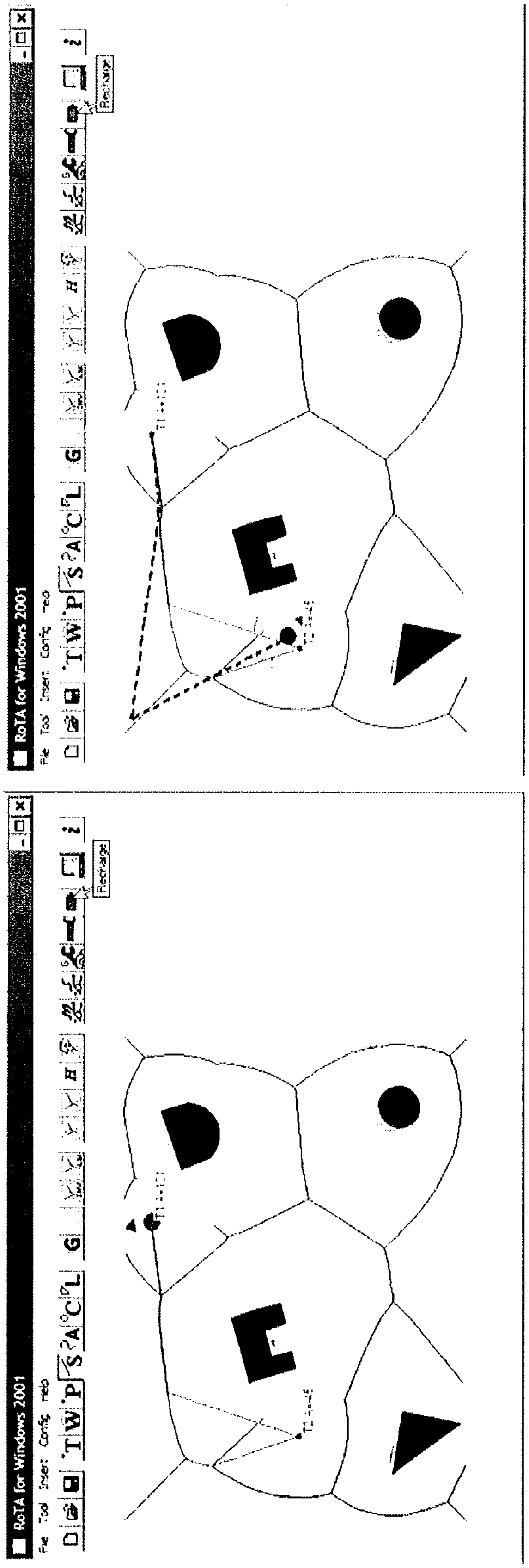


Fig. 11. Visibility pathway (full line on the left) and the actual robot navigation path (dashed line on the rigth) to the recharging station and then to goal target.

Table III. *Deviation measurements when replanning took place.*

Configuration variables	x	y	θ
Mean (95% confidence)	0.0363	0.5778	0.0029
Lower Bound	0.0096	−0.2563	0.0160
Upper Bound	0.0631	1.4115	0.0421
Std. deviation	0.0572	1.7818	0.0278

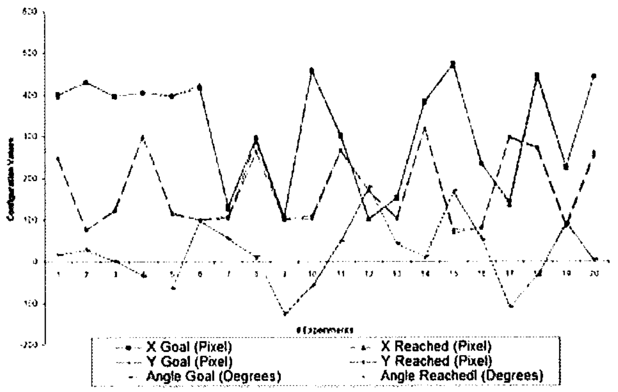


Fig. 12. Deviation measured for the configuration coordinates due to an intruder obstacle.

TPM replans the trajectory to continue the robot navigation to fulfill the previous task.

The computation time for the trajectory replanning is currently the same as for the planning in the presence of a new obstacle in the workspace. Although the modification of the Voronoi diagram due to the intruder seems to be only local, the computation of the new trajectory has to take into account the goal configuration to be reached by the robot and this requires a non-local computation. We are working to reduce the computation time for trajectory replanning by defining a semi-global computation of the roadmap.

The set of experiments realized provide a measure for when a trajectory replanning had to be done. Table III presents the results, Fig. 12 shows the deviations between the goal and reached configuration coordinates due to the replanning to avoid the intruder obstacle and Fig. 13 illustrates the visibility pathway and the actual replanned path when an intruder obstacle was detected, as the result of a single experiment.

4. COMMENTS AND CONCLUSIONS

In this paper we have presented the current state of development of a path planning system called RoTA. This system is composed of three main modules, namely: i) the global vision module (GVM), ii) the trajectory planning module (TPM), and iii) the navigation control module (NCM). In summary, the GVM is responsible for capturing the workspace image, process it to identify the obstacles and the robot configuration (position and orientation) and to provide the minimum geometrical structure of the obstacles and robot to the TPM. Here we have assumed that the obstacles can be generalized polygons and circular shapes with possible holes. Provided with the geometrical data of

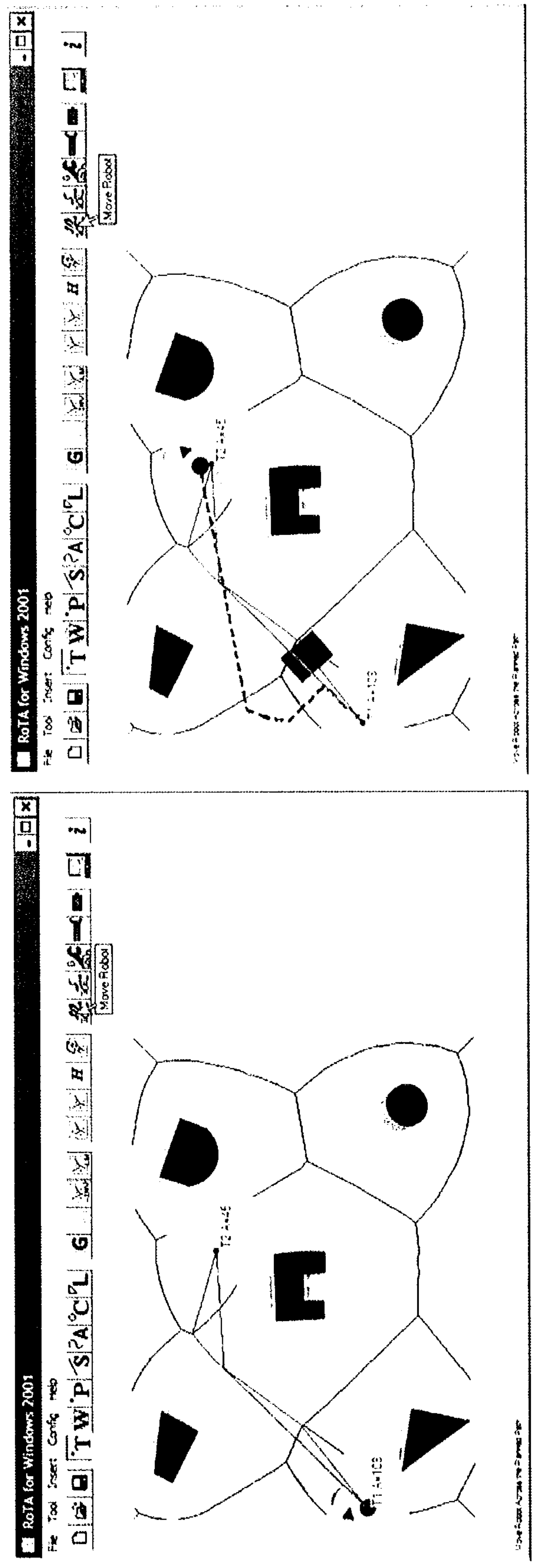


Fig. 13. Visibility pathway (full line on the left hand side) and the actual replanned robot navigation path (dashed line on the right hand side) due to intruder detection (right).

the obstacles, the TPM generates the roadmap based on the approximate generalized Voronoi diagram technique.

The user can indicate on the RoTA workspace screen the configuration of a set of visiting stations where the robot shall navigate to. The TPM computes the maximal clearance shortest feasible path between the initial and goal configurations, taking account of the sequence of all visiting stations. The TPM also computes the visibility pathway, which is a shorter path with minimal number of links connecting the visiting stations. The NCM controls the navigation along the visibility pathway and all sensorial data provided by the robot. Currently, the mini-robot Khepera (disc-like robot) has been used as our robotic platform for validation. When the proximity sensors detect an unexpected intruder obstacle, the NCM halts the navigation and then requests the GVM to read the new workspace scenario and to provide the data to the TPM, that replans the trajectory restarting from the current robot configuration, taking care of the sequence of all previous unvisited stations.

The NCM also takes care of the robot's battery charge. For the experiments in section 3 two battery recharge stations were located at two corners (upper-left, bottom-right) of the workspace arena. Therefore, when the robot's battery supply drops to a low limit level (simulated by the user), the NCM halts any robot activity and immediately replans its trajectory, moving the robot to the nearest recharging station. After it has been recharged, the system replans the trajectory to continue the execution of an incomplete task or just departs for a new one. As currently only one robot is at work, when it has to be recharged, no task can be accomplished during recharging. Therefore, we are working on the introduction of a second robot to assume the previous robot duties. It will complete the previous robot task and becomes a new active robot until it has to be recharged, changing the robots again.

The three set of experiments have shown that the path planning system is quite robust. In other words, if we take into account the length/pixel scale the deviations observed in the Tables I, II and III they are of the order of few millimeters, which compared to the robot size means that the goal configurations were essentially reached in the workspace. Inasmuch, the deviations observed between the goal and reached configurations in these experiments may have partially arisen also from exogenous causes due to some mechanical parts of the robot, to small errors in the length/pixel scale and even to the luminosity of the workspace. The Figs. 8, 10 and 12 reflect the accuracy of the RoTA system and the Figs. 9, 11 and 13 illustrates the actual navigation path done by the robot (dashed line) and the theoretical one (visibility pathway). One approach to minimize the deviation between the actual configuration reached by the robot and the specified goal configuration is to check if the actual configuration agrees with a threshold value for the position and orientation. It not, the GVM acts and the system computes a new trajectory and makes the robot navigate to the goal target. Although the results obtained so far consider a Lab robot, the technique developed may be similarly applied to a larger scale robot and workspace.

There is quite a lot of further work to be done, for instance: introduce more than one camera to cover a much wider workspace arena, include an additional robot in the workspace, consider a more realistic workspace with a dynamic scenario, include a video camera turret and other sensors on the robot to support the navigation, intruder identification and other task accomplishments. Although the visibility pathway is already an improved path with respect to the MCSFP, the visibility pathway is not yet the shortest path between two configurations and it is not a smooth curve. We are current working on smoothing out the visibility pathway and on computing the homotopic shortest path[25] between two robot configurations in the free configuration space.

5. Acknowledgements

W. L. Roque thanks the CNPq/CTPetro for the financial support (grant 522635/95-4). D. Doering thanks FAPERGS for the scholarship. Both authors thank Bruno Alves for his help with the experiments and Manoel M. Silveira for his help with the statistics.

References

1. United Nations, *World Robotics* (The International Federation of Robotics, United Nations Report, Geneva, Switzerland, 2001).
2. J.-P., Latombe, *Robot Motion Planning*. (Kluwer Academic Publishers, New York, 1991).
3. H. Choset and J. Burdick, "Sensor-Based Planning: The Generalized Voronoi Graph" (Part I). "Sensor-Based Planning: Incremental Construction of the Generalized Voronoi Graph" (Part II), *IEEE Proceedings of the International Congress on Robotics and Automation* (1995) **Vol. 2**, pp. 1643–1648.
4. D. Blanco, B. L. Boada, L. Moreno and M. A. Salichs, "Local Mapping from On-Line Laser Voronoi Extraction", *IEEE International Conference on Intelligent Robotics and Systems* (2000) **Vol. 1**, pp. 103–108.
5. R. Mahkovic and T. Slivnik, "Constructing the Generalized Local Voronoi Diagram from Laser Range Scanner Data", *IEEE Transaction on Systems, Man, and Cybernetics* **Part A, 30**, 710–719 (2002).
6. C. Arcelli and G. Sanniti di Baja, "Computing Voronoi Diagrams in Digital Pictures", *Pattern Recognition Letters* **4**, 383–389 (1986).
7. S. K. Parui, S. Sarkar and B. B. Chaudhuri, "Computing the Shape of a Point Set in Digital Images", *Pattern Recognition Letters* **14**, 89–94 (1993).
8. M. Melkemi and D. Vandorpe, "Fast Algorithm for Computing the Shape of a Set of Digital Points", *IEEE International Conference in Image Processing* (1994) **Vol. 1**, pp. 705–709.
9. P. Preparata and M. I. Shamos, *Computational Geometry: An Introduction* (Springer-Verlag, New York, 1985).
10. L. F. Costa and R. M. Cesar Jr., *Shape Analysis and Classification: Theory and Practice* (CRC Press, Boca Raton, USA, 2001).
11. F. Mokhtarian and R. Suomela, "Robust Image Corner Detection Through Curvature Scale Space", *IEEE Trans. on Pattern Analysis and Machine Intelligence* **20**, 1376–1381 (1998).
12. M. Trajkovic and M. Hedley, "Fast Corner Detection", *Image and Vision Computing* **16**, 75–87 (1998).
13. F. Deschênes and D. Ziou, "Detection of Line Junctions and Line Terminations Using Curvilinear Features", *Pattern Recognition Letters* **21**, 637–649 (2000).
14. L. Wen-Yu, L. Hua and Z. Guang-Xi, "A Fast Algorithm for Corner Detection Using the Morphologic Skeleton", *Pattern Recognition Letters* **22**, 891–900 (2000).
15. D.-M. Tsai, H.-T. Hou and H.-J. Su, "Boundary-Based Corner Detection Using Eigenvalues of Covariance Matrices", *Pattern Recognition Letters* **20**, 31–40 (1999).

16. P. Sauer, "On the Recognition of Digital Circles in Linear Time", *Computational Geometry: Theory and Applications* **2**, 287–302 (1993).
17. P. Damaschke, "The Linear Time Recongnition of Digital Arcs", *Pattern Recognition Letters* **16**, 543–548 (1995).
18. N. S. V. Rao, "On Fast Planning of Suboptimal Paths Amidst Polygonal Obstacles in Plane", *Theoretical Computer Science* **140**, 256–289 (1995).
19. O. Takahashi and R. Schilling, "Motion Planning in a Plane using Generalized Voronoi Diagrams", *IEEE Trans. Robotics and Automatation* **5**, 143–150 (1989).
20. A. Okabe, B. Boots and K. Sugihara, *Spatial Tesselation Concepts and Applications of Voronoi Diagrams* (John Wiley & Sons, New York, 1992).
21. F. Aurenhammer, "Voronoi Diagrams: A Survey of a Fundamental Geometric Data Structure", *ACM Computing Survey* **23**, 345–405 (1991).
22. K. Sugihara and M. Iri, "Construction of the Voronoi Diagram for One Million Generators in Single-Precision Arithmetic", *Proceedings of the IEEE* **80**, 1471–1484 (1992).
23. W. L. Roque and D. Doering, "Constructing Approximate Generalized Voronoi Diagrams from Digital Images of Generalized Polygons and Circular Objects", *Proceedings of the 11th International Conference in Computer Graphics, Visualization and Computer Vision* (Ed. V. Skala) (2003) pp. 119–125.
24. F. Mondada, E. Franzi and A. Guignard, "The Development of Khepera", *Proceedings of the First International Khepera Workshop* (1999) pp. 7–13.
25. E. Efrat, S. G. Kobourov and A. Lubiw, "Computing Homotopic Shortest Paths Efficiently", *LNCS 2461* (Eds. R. H. Möhring and R. Raman) (2002) pp. 411–423.

EFFICIENT GAMMA-RAY SIGNAL DECOMPOSITION ANALYSIS BASED ON ORTHONORMAL TRANSFORMATION AND FIXED POLES

Sergio Zimmermann, Dionisio Doering

Lawrence Berkeley National Laboratory*
Berkeley, CA 94720

ABSTRACT

Gamma-ray energy tracking is a new technique for the detection of gamma radiation. In such scheme, the individual interactions of the gamma rays with the germanium detectors are described by the energy, position and interaction time. Signal decomposition is the name of the procedure used to estimate the three-dimensional positions of the interactions based on pulse-shape analysis of the signals on the two-dimensional segments deposited on the faces of the detector. The present signal decomposition algorithm is computational intensive. For GRETINA, a detector being built based on this concept and that covers just a quarter of a sphere, 140 quad-processors are required to decompose 20,000 gamma interactions per second. In order to reduce the computational cost, we have conceptualized that the segments waveforms are generated by amplitude modulated discrete-time unit impulse. Projection of these waveforms into a more suitable basis reduces the computation costs while optimizing the same cost function. In this article we describe the framework for such projection, and we provide an example. For the present example, the computational cost was reduced by a factor of 5 times.

Index Terms— Signal analysis, transfer functions, transformation, pole assignment, reduced order systems

1. INTRODUCTION

Gamma-ray energy tracking is a new technique for the detection of gamma radiation. In such detector, the individual interactions of the gamma-rays with the germanium (Ge) detectors are identified by their energies, positions and interaction times [1]. GRETINA, a detector based in this concept, is under construction in the US [2], and AGATA, a similar detector, is under construction in Europe. GRETINA consists of 28 close-packed Ge crystals. Each crystal is a cylinder with flat tapered shape (see Fig. 1), and the segments are deposited on the surfaces forming a matrix of 6 × 6 longitudinal and transverse segments.

Fig. 2 shows examples of nine typical waveforms on the segments of the detector. The gamma-rays are interacting with the detector at a specific location below the B4 segment. The segments shown are the one that collects the charge (B4) and its eight nearest neighbors. The neighbors detect induced charge. A total of 16 measured shapes are plotted in gray and the calculated waveforms for this given position are plotted in black. Observe that the measured signals include noise.

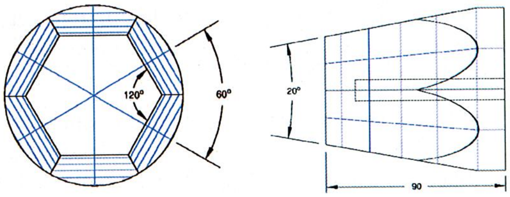


Fig. 1. Sketch of the segmentation of the Ge detector

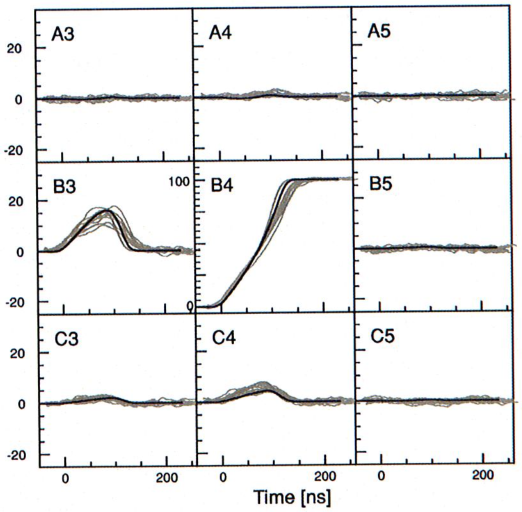


Fig. 2. Waveforms for interactions below segment B4

Signal decomposition is the name given to the process of finding the position of the interaction based on the signals collected on the segments. The fact that the waveforms can be the result of multiple interactions at various locations, each with different amplitudes, increases the complexity of the problem. The algorithm must separate (or "decompose") a pulse shape resulting from one or more interactions with the Ge crystal. [3]. The present decomposition algorithm is computational intensive and, for GRETINA, it requires 140

---

* This work has been supported by the Director, Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

quad-processors to decompose approximately 20,000 gamma-rays per second.

In order to reduce the computational cost one can project the information into a more suitable basis which requires the calculation of a smaller number of parameters. To find such basis, we have conceptualized that the sampled pulse-shapes on the segments are generated by $a\delta[n]$, an amplitude modulated discrete-time unit impulse, that is acting over a transfer function $H(z)$. So, what we are sampling are the parameters b_i of an FIR filter multiplied by a,

$$H(z) = a\left(b_0 + b_{-1}z^{-1} + \ldots + b_{-p+1}z^{-p+1}\right). \tag{1}$$

Actually, from this standpoint, the original data is projected into a basis composed of z^{-i}. We can now research for compact ways to represent these FIR filters. We will demonstrate that we can find a new description that supports the optimization of the same cost function used originally with comparable results.

The procedure to run the signal decomposition on the new basis is the following. First, a new basis is selected and the grid functions are projected on the new basis. When a new gamma-ray interaction is detected, the segment waveforms are also projected into the new basis. Then the procedure for signal decomposing is applied.

2. SIGNAL DECOMPOSITION ALGORITHM

A version of the algorithm works in the following way. Using simulations we calculate the waveforms of the interaction of the gamma rays in a 2 × 2 × 2 mm grid inside the volume of the detector (for an example, refer to the black waveforms in Fig. 2), resulting in approximately 3000 points. Let's define the vector D_l, where l is one of the 3000 positions

$$D_l = [d_{lA0}\ d_{lA1}\ \ldots\ d_{lA5}\ d_{lB0}\ \ldots\ d_{lF5}]^T \tag{2}$$

composed by the p samples of the calculate waveforms

$$d_{lk} = [d_{lk,0}\ d_{lk,-1}\ \ldots\ d_{lk,-p+1}],$$

(normally $p = 50$) for the 36 segments $k = $ A0,A1,..., B0,...,F5. We will name the set of all D_l as grid functions. Define $\hat{Y}_m$ as the linear combination of the grid functions

$$\hat{Y}_m = \sum_{i=1}^{N} c_{m,i} D_l\ ,$$

where N is the number of gamma-ray interactions that contributed to the signals and l can be any point in the grid. The approach used to decompose the measured waveform Y is the minimization of the least square error

$$e_{LS,m} = \min_{\hat{Y}_m}\left\|Y - \hat{Y}_m\right\|_2 = \min_{\hat{Y}_m}\left([Y - \hat{Y}_m]^T[Y - \hat{Y}_m]\right) \tag{3}$$

by the selection of $\hat{Y}_m$ to minimize $e_{LS,m}$. Y is assembled as a vector in the same way as D_l in (1).

For illustration purpose, we will now apply this algorithm for the case where Y is the result of two gamma-ray interactions. Then

$$\hat{Y}_{ij} = c_{ij,1}D_i + c_{ij,2}D_j\ . \tag{4}$$

where $i, j = 1, \ldots, 3000$ any combination of two grid functions in any i and j positions in the grid [4]. Substituting (4) in (3), and setting

$$\frac{\partial e_{LS,ij}}{\partial c_{ij,1}} = \frac{\partial e_{LS,ij}}{\partial c_{ij,2}} = 0$$

we arrive to two equations with two unknowns $c_{ij,1}$ and $c_{ij,2}$. The 2nd derivative confirms that this is a minimum. Then

$$C_{ij} = [\mathbf{D}_{ij}^T\mathbf{D}_{ij}]^{-1}\mathbf{D}^T Y$$

where $\mathbf{D}_{ij} = [D_i\ , D_j]$ is a matrix composed of two vectors of the grid functions, and we select $C_{ij} = [c_{ij,1}\ , c_{ij,2}]^T$ as the one that minimizes the error (3) within all possible combinations of two grid functions. The grid functions $\mathbf{D}_{ij}$ and gains C_{ij} that minimize the collection of errors $e_{LS,ij}$ select the position of the gamma-ray interactions and distribution of relative amplitudes between them. This procedure can be easily adapted to handle other numbers of gamma-ray interactions with the crystal.

For these two interactions in one segment, we have $\sim L^2/2$ or ~5 million points for grid search. Therefore, it is very important to reduce the computational cost of the process of finding the position of the interaction of the gamma-ray with the crystal.

3. REPARAMETRIZATION

Let $\mathbf{U}$ be a $p \times q$ orthonormal basis, where p is the number of samples of the waveforms and q is the order of the basis. We will assume that $q < p$ to reduce the computational complexity. The grid functions D_l can be projected into this new basis to minimize the least square error

$$e_{LS} = \min_{\tilde{Y}_m}\left([D_i - \tilde{D}_i]^T[D_i - \tilde{D}_i]\right)$$

such that $\tilde{D}_i = \mathbf{U}\tilde{\Delta}_i$ and $\tilde{\Delta}_l = [\mathbf{U}^T\mathbf{U}]^{-1}\mathbf{U}^T D_l$. For the demonstration of a similar problem refer to [4]. We will be using tilde (~) to denote vectors projected into the new basis. Observe that in general $D_l - \tilde{D}_l = E_l$ is non-zero. However, by selecting a new suitable basis E_l can be made small. The same reasoning can be employed to project the sampled waveforms into this new basis,

$$\tilde{Y} = \mathbf{U}\tilde{\Psi}\ , \tag{5}$$

$\tilde{\Psi} = [\mathbf{U}^T\mathbf{U}]^{-1}\mathbf{U}^T Y$, and $Y - \tilde{Y} = E$ is non-zero.

The next step is to demonstrate that the optimization of the cost function (2) is approximately the same in the new basis $\mathbf{U}$, depending on how well the sampled waveforms and grid functions are represented in the new basis. We will demonstrate this for the same case as (4). All other cases (from 1 to several interactions can be demonstrated in the same way).

Theorem: The optimization of the cost function (1) in the original basis is the same as the optimization of the cost function in the new basis, as long as the projection of the new signal in the new basis reproduces the original signal.

Proof: Let
$$\hat{\tilde{Y}}_{ij} = \mathbf{U}[\tilde{c}_{ij1}\tilde{\Delta}_i + \tilde{c}_{ij2}\tilde{\Delta}_j], \tag{6}$$
$\hat{\tilde{Y}}_{ij}$ a linear combination of $\tilde{\Delta}_i$ and $\tilde{\Delta}_j$. Then, similar to (3),

$$\tilde{e}_{LS,ij} = \min_{\hat{\tilde{Y}}_{ij}}\left(\left[\tilde{Y} - \hat{\tilde{Y}}_{ij}\right]^T\left[\tilde{Y} - \hat{\tilde{Y}}_{ij}\right]\right) \tag{7}$$

and we need to minimize this error function. We replace $\tilde{Y}$ and $\hat{\tilde{Y}}_{ij}$ by (5) and (6) respectively, and by definition of orthonormal basis $\mathbf{U}^T\mathbf{U} = \mathbf{I}$. Now setting

$$\frac{\partial \tilde{e}_{LS,ij}}{\partial c_{ij,1}} = \frac{\partial \tilde{e}_{LS,ij}}{\partial c_{ij,2}} = 0$$

this results in two equations with two unknowns. The solution of these equations generate

$$\tilde{C}_{ij} = \left[\tilde{\mathbf{\Delta}}_{ij}^T\tilde{\mathbf{\Delta}}_{ij}\right]^{-1}\tilde{\mathbf{\Delta}}_{ij}^T\tilde{Y}$$

where $\tilde{\mathbf{\Delta}}_{ij} = \left\lfloor\tilde{\Delta}_i, \tilde{\Delta}_j\right\rfloor$ is a matrix composed of two vectors of the grid functions projected in the new basis, and we select $\tilde{C}_{ij} = \left[\tilde{c}_{ij,1}, \tilde{c}_{ij,2}\right]^T$ as the one that minimizes the error (7). But since we assumed that the sampled waveforms and the grid functions projected on the new basis reproduces the original signal, then $D_l - \tilde{D}_l = 0$ and $Y - \tilde{Y} = 0$, it follows that $Y = \tilde{Y}$, $\tilde{e}_{LS,ij} = e_{LS,ij}$ and $\tilde{C}_{ij} = C_{ij}$. *ced*

One can demonstrate that if the new basis represents well the original signals (but not exactly), i.e., $\tilde{Y} \to Y$ and $\tilde{D}_i \to D_i$, then $\tilde{e}_{LS,ij} \to e_{LS,ij}$ and $\tilde{C}_{ij} \to C_{ij}$.

4. NEW BASIS: FIXED POLES

We will now describe a set of candidate basis to re-parameterize the sampled waveforms and grid functions. As stated before, one can understand the original information (sampled waveforms and grid functions) as a set of impulse responses that described the instantaneous charge deposited into the crystal. This concept renders a description of the original information as impulse responses with FIR filters descriptions (see Eq. (1)). Of course, this does not apply to the segment that integrates the charge (channel B4 in Fig. 2). It has a transition and then stays at a constant value different than zero and equal to the gamma-ray energy deposit in the crystal at that point. So, we need to pass it through some filter that brings the signal back to the baseline. We can do this by using the following:

$$y(n) = E(x) - x(n), \tag{8}$$

where $E(x)$ is calculated automatically by the front-end electronics and is the estimate of the energy deposited in the crystal (for an example of such shape, refer to Fig. 6). Now, all sample data and grid functions can be represented as FIR filters. These functions are slow varying and then slowly decay toward zero (of course, if we disregarding the noise). This type of FIR filter can be represented in a more compact form by transfer functions of the type

$$H(z) = \frac{b_0 + b_{-1}z^{-1} + \ldots + b_{-q+1}z^{-q+1}}{(1 - a_{-1}z^{-1})\ldots(1 - a_{-q+1}z^{-q+1})}.$$

The poles of the filter are preselected based on *a priori* information of the grid functions and then kept fixed in these locations. By only adjusting the zeros, this projection has a linear-in-the-parameter characteristic, which allows finding the sampled data as the best fit of the linear combination of the grid functions in the new basis. This compact representation has been previously study using Laguerre [5], Lagrange (both with just one fixed pole), Kautz (with one set of two complex conjugated poles), and fixed poles adaptive filters [6] (where all poles can be located in their optimum position) and forming an orthonormal basis. For references about all these methods, see [6].

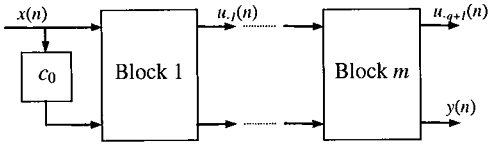


Fig 3. Block structure for transfer function description

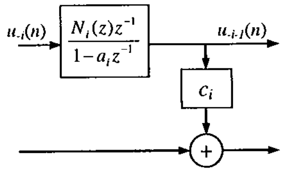


Fig 4. Block with real-value pole

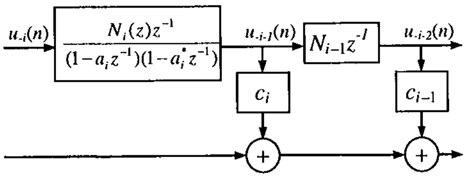


Fig 5. Block with complex-conjugated pole pair

Fig. 3-5 show a block diagram of the fixed pole and the parameters that are adjusted. Reference [5] describes a few procedures to obtain these optimum pole locations based on *a priori* information. The new basis $\mathbf{U} = [U_1, U_2, \ldots U_q]$, where the original data is projected is formed by the vectors

$$U_i = [u_i(n), u_i(n-1), \ldots, u_i(n-p)]^{\mathrm{T}}$$

where $u_i(n)$ are the samples of the impulse response of one of the taps of the structure described in Fig. 3-5, and $N_i(z)$ are selected to generate an orthonormal basis.

5. PERFORMACE COMPARISON

We will now compare the performance of the fitting algorithm for the original algorithm (Section 2) and new basis (Section 3). For the new basis we have selected a Laguerre polynomial (Fig. 4) [5] with $i = 10$ parameters and the real pole at $a = 0.5$.

Fig. 6 shows waveforms for the contact that collects the charge (similar to B4 in Fig. 2) after applying (8). The signal was generated as a linear combination of two elements of the grid functions ($Y_{12} = 3D_1 + 2D_2$, shown as a solid line). The best fit for the original and Laguerre basis are shown by circles and crosses respectively. Observe that the best fit for Y_{12} for two random grid functions using the original and the Laguerre basis are very close, but the Laguerre basis required the estimation of just 10 parameters (instead of 50).

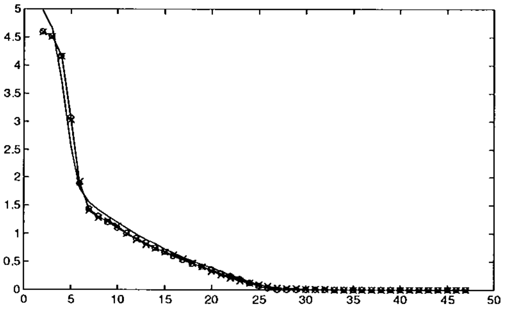


Fig. 6. Measured signal (solid line), and best fit for original (circle) and Laguerre (cross) basis

The second experiment was conducted with 10 calculated waveform signals $Y_{ij} = c_iD_i + c_jD_j$ where the grid functions D_l (from a total of 100 grid function elements) and the gains c_l were selected randomly and the gains normalized to $c_i + c_j = 1$. We added noise to the signals equivalent to 50 keV$_{\text{rms}}$ (the equivalent of 5 bits rms in 13 bits). We run the algorithm using the Laguerre basis to confirm that it can identify the D_l and c_l used to generate the 10 signals Y_{ij}. In all 10 cases, the algorithm was able to recognize the proper D_l. Also, the mean square error between the set gains c_l and the estimated gains with the algorithm was 0.1%. Tab. 1 shows an example of a selected and estimated gains using Laguerre for one of the Y_{ij}.

Tab. 1. Gain comparison selected and estimated

	Selected	Laguerre
c_1	0.7	0.7034
c_2	0.3	0.2955

The last experiment compares the least square errors for the original and Laguerre basis, with the signal $Y = cD$. We added 50 keV of noise to the signal. We apply both algorithms and we record the square errors for 12 of the grid searches. We then sort one algorithm for the square errors and sort the other for the same sequence. Fig. 8 shows the square error of these 12 points. Observe that they are similar, but the square error for the original algorithm (cross) is bigger than the new (circles). The Laguerre basis, with less parameters to adjust, reduces the noise on the data.

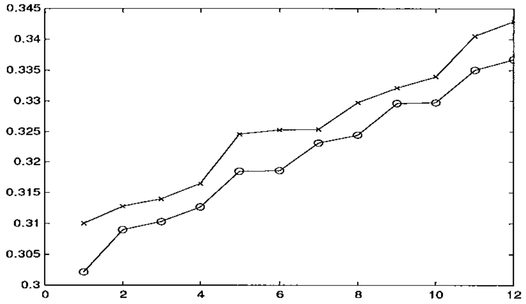


Fig. 7. Square errors between original (cross) and new algorithm (circle)

6. CONCLUSION

We have described an algorithm to reduce the computational cost of the signal decomposition for a Ge detector. This cost reduction is achieved by projecting the data on a more suitable basis allowing a more compact description of the information. If the new basis is orthonormal and offers a good description of the original signals, then the square error and relative gains are approximately the same as in the original basis. Because of the characteristic of the waveforms, pre-filters using poles are good candidates for the new basis. The simulations illustrate these claims.

7. REFERENCES

[1] Lee, I.Y., "Gamma-ray tracking detectors," *Nucl. Inst. Method A,* Elsevier, Maryland Heights, MO, vol. 422, pp. 195-200, Feb. 1999.

[2] Lee., I.Y, *et.al.*, "GRETINA: A gamma ray energy tracking array," *Nuclear Physics A*, Elsevier, Maryland Heights, MO, Vol. 746, pp. 255-259, Dec. 2004.

[3] Radford, D., "*Signal Decomposition*," GRETINA Software Workshop, Lawrence Berkeley Nat. Lab., Berkeley, June 2004.

[4] Johnson, Jr., C. R., *Lectures on Adaptive Parameter Estimation*, Englewood Cliffs, Prentice-Hall, NJ, 1988.

[5] Wahlberg, B., "System identification using Laguerre models," *IEEE Trans. Automat. Contr.*, Piscataway, NJ, Vol. 36, pp. 551-562, May 1991.

[6] Williamson, G., Zimmermann, S., "Globally Convergent Adaptive IIR filters Based on Fixed Pole Locations," *IEEE Trans. Signal Processing*, Piscataway, NJ, Vol. 44, June 1996.

Data Acquisition and Trigger System of the Gamma Ray Energy Tracking In-Beam Nuclear Array (GRETINA)

John T. Anderson, Renato Brito, Dionisio Doering, Todd Hayden, Bryan Holmes,
John Joseph, Harold Yaver, Sergio Zimmermann, *Senior Member, IEEE*

Abstract– The Gamma Ray Energy Tracking In-Beam Nuclear Array (GRETINA), capable of determining the energy and position (within 2mm) of each gamma-ray interaction point and tracking multiple gamma-ray interactions, has been designed. GRETINA will be composed of seven detector modules, each with four highly pure germanium crystals. Each crystal has 36 segments and one central contact instrumented by charge sensitive amplifiers. Two custom designed modules, the Digitizer/Digital Signal Processing (DSP) and the Trigger Timing & Control (TTC), compose the electronics of this system. The Digitizer/DSP converts the analog information with 14-bit analog to digital converters (operating at 100 Msamples/sec, and digitally processes the data to determine the energy and timing information of the gamma interactions with the crystal. Each Digitizer/DSP is controlled by and sends trigger information to the Trigger Timing & Control system through a bidirectional Gbit link. Presently four different trigger algorithms are planned for the trigger system and can be selected for trigger decision. In this paper the details of the electronics and algorithms of the GRETINA data acquisition and trigger system are presented and the performance is reviewed.

I. Introduction

We have designed the Gamma Ray Energy Tracking In-Beam Nuclear Array (GRETINA). GRETINA is based on germanium detectors and it will be capable of determining the energy and position (within 2mm) of the points of interaction of the gamma-rays with the germanium crystal and of tracking multiple gamma-ray interactions [1], [2]. GRETINA is composed of seven detector modules, each with four high purity germanium crystals (see Fig. 1), comprising a quarter or 1-π of a sphere. The detector module components include charge sensitive amplifiers [3] assembled inside the detector enclosure to instrument each of the 36 segments and the central contact. The gamma ray interaction with the germanium crystal induces charge on the segments and central contact. The amplifiers integrate this charge and drive an analog voltage to the GRETINA front-end electronics. There are plans to construct the full 4-π detector module array, GRETA [4], and the electronics system foresees this possible scenario.

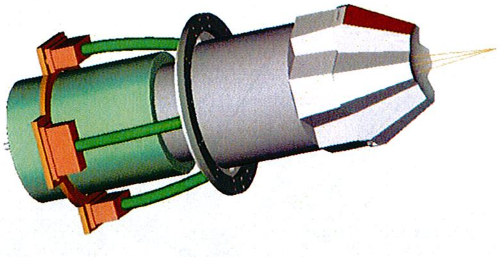


Fig. 1. GRETINA Detector Module

The detector modules are supported by a mechanical structure composed of two quasi-hemispheres shells that surround the target chamber. This shell structure is capable of supporting 21 detector modules. The structure allows rotation for detector mounting (through a gear box in the end of the axels) and translation to access the target chamber (through railroad cars). Hexapods connect the structure to the railroad cars. Fig. 2 shows a sketch of the support structure, with all possible positions instrumented with a detector module.

In this paper the details of the electronics and algorithms of the GRETINA data acquisition and trigger system will be presented and the performance will be reviewed. In addition, grounding and filtering techniques used to achieve the 14-bit analog to digital conversion (ADC) performance will be discussed, as well as transmission line techniques for the very low bit error rate of the gigabit links.

II. System Architecture

Fig. 3 shows a block diagram of the GRETINA Electronics and Computing Systems. The oblong shape on the left represents the detector modules and its crystals. Charge sensitive amplifiers instrument the segments and central contact. Fifteen meters of shielded twisted pair cable connect the pre-amplifier outputs to the digitizer modules. Two custom designed modules, the Digitizer/Digital Signal Processing (Digitizer/DSP) and the Trigger Timing & Control (TTC), compose the electronics of this system. Four Digitizer/DSP modules instrument one crystal: each master Digitizer/DSP interfaces with the TTC system and controls three slave

---

Manuscript received November 15, 2007. This work was supported by the Director, Office of Science, Office of Nuclear Physics, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Renato Brito, was with Lawrence Berkeley National Laboratory. He is now with the Department of Electrical Engineering, Federal University of Rio Grande do Sul, Porto Alegre, RS 90035 Brazil.

D. Doering, B. Holmes, J. Joseph, H. Yaver, S. Zimmermann are with Lawrence Berkeley National Laboratory, Berkeley, CA 94720.

John T. Anderson, Todd Hayden are with Argonne National Laboratory, Argonne, IL 60439.

Corresponding author: Sergio Zimmermann, phone 510-495-2964, e-mail: szimmermann@lbl.gov.

Digitizer/DSP modules. A digital bus in the front panel allows the digitizers serving a crystal to synchronize among themselves for clock and trigger information. A very simple communication protocol based on a single master controlling the operations is used.

The master digitizer monitors the crystal central contact. If a gamma ray deposits charge in the crystal above the programmable threshold of the leading edge discriminator (LED), the master Digitizer/DSP recognizes the event and reads the segment hit pattern using the front panel bus. In parallel the Master digitizer estimates the energy of the central contact signal using a fast algorithm. It then assembles the trigger information (time stamp (TS) of the LED, central contact energy and segment hit pattern) and sends it to the TTC over one pair of a bidirectional 1Gb/sec serial link. A buffer records the timestamps of all recent discriminator decisions.

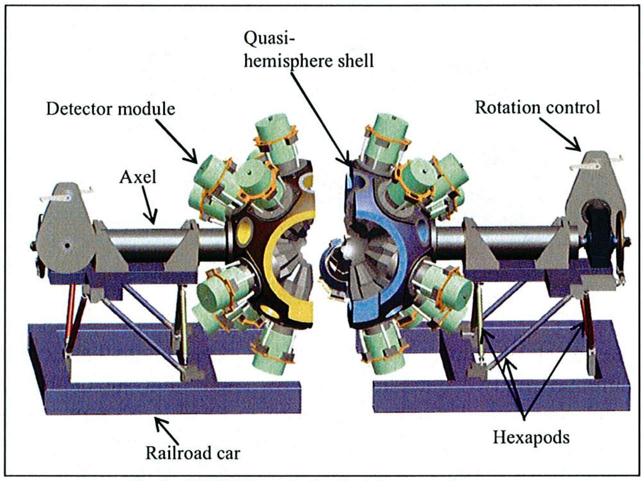


Fig. 2. GRETINA Mechanical Support Structure

The serial link connecting the TTC and Digitizer/DSP modules is implemented using the National Semiconductor DS92LV18 Serializer/Deserializer (SerDes). It transmits 20 bits per word at a 50 MHz rate, where 18 bits are available to transmit trigger and control information. The link has stable (within the requirements) and predictable latency. The predictable latency allows the same twisted pair to both transmit control information and provide the 50MHz master clock to each Digitizer/DSP. The trigger information exchange between the master trigger and Digitizer/DSP employs a synchronous protocol and the master trigger module determines the synchronism. The protocol consists of an endlessly repeating series of 20 command frames transmitted every 2 μs that allows the TTC system to regularly synchronize and control activities at the Digitizer/DSPs. Conversely, all trigger information from the Digitizer/DSP is transmitted to the TTC within this 2 μs. A synchronous implementation like this is easier to implement and maintain.

The whole TTC system is assembled around the same hardware module. The firmware is configured for two different functions: the Router module and the Master TTC. The Router routes all information between the Master TTC and the Digitizer/DSP modules and assists in fast multiplicity trigger decisions. The Master TTC collects all messages from the Routers plus additional information from auxiliary trigger modules and uses it to make a global trigger decision (refer to Fig. 4). Once a global trigger decision is made, the TTC system sends a trigger decision command to the Routers for distribution to all master Digitizer/DSP modules and auxiliary detectors. Each master Digitizer/DSP identifies a match between the timestamp embedded within the trigger decision message and the saved LED timestamps, and requests all slave Digitizer/DSP to transfer the data from the circular buffer into its own readout FIFO. Later, the VME readout CPUs in the digitizer crates read the event data from each Digitizer/DSP FIFO, assemble the crystal event, and send the data to the network switch. The switch routes the events to the farm where they are processed. The processing uses the segment information to estimate the position (r, θ, z) and energy of the interaction points. Additional processing establishes the tracks by connecting the individual interaction points.

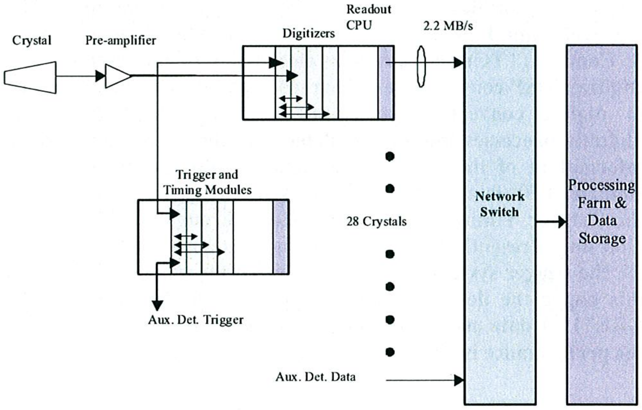


Fig. 3. GRETINA Electronics and Computing Systems

We have extensively tested the performance of the DS92LV18 SerDes for bit error rate (BER). We have measured BER better than 10^{-16}, which corresponds to less than one error per day for GRETINA (four per day for GRETA). This component is a very simple SerDes that adds minimum protocol (an important characteristic for the constant latency). We observed that adding external FPGA logic to execute the DC balancing of the communication and replacing the cable from unshielded CAT5 to a good quality shielded CAT5 significantly decreased the BER from one error every few hours to the level reported above. LVDS drivers with programmable pre-emphasis are used in all modules to compensate for cable losses.

Each Digitizer/DSP and TTC module is implemented using two field programmable gate arrays (FPGA): one is a smaller FPGA, which controls the VME interface and has a steady configuration, and the other is a larger FGPA which executes the module specific algorithms and is easily re-configured through VME.

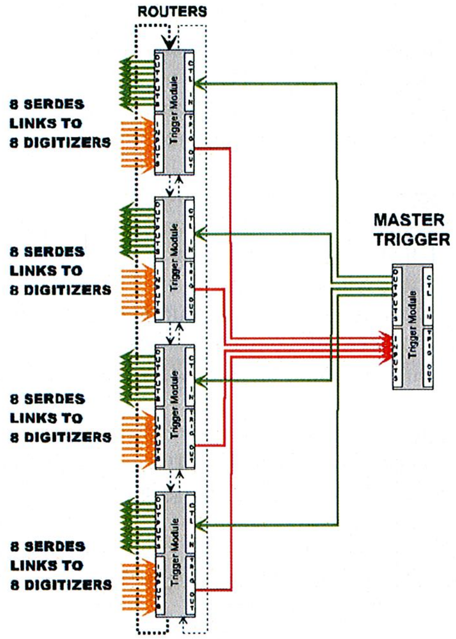


Fig. 4. TTC Router and Master Connection

III. DIGITIZER/DSP

The Digitizer/DSP module samples the crystal segment and central contact analog information using 14-bit ADC operating at 100 Msamples/s. The ADC used is the AD6645 from Analog Devices. Fig. 5 shows a picture of the module and Fig. 6 shows its block diagram. The Digitizer/DSP has a total of 10 analog inputs. The ADCs are connected to an FGPA, which digitally processes the conversion and executes the following operations: leading and constant fraction discrimination, trapezoidal filtering, energy determination, and pole/zero cancellation.

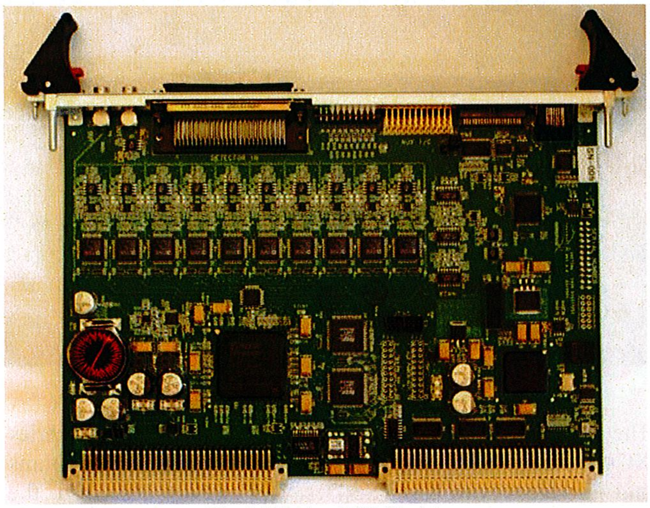


Fig. 5. Digitizer/DSP Module

Table I describes the digital processing algorithms and Fig. 7 shows a block diagram of these algorithms and how they interconnect. All processing occurs at a clock rate of 100 MHz, which is generated inside the FPGA by multiplying the 50 MHz master clock distributed from the TTC. This results in approximately 20 giga-operations/s. While the processing occurs, the raw data is stored in 40 µsec circular buffers (designed around the FPGA block RAMs).

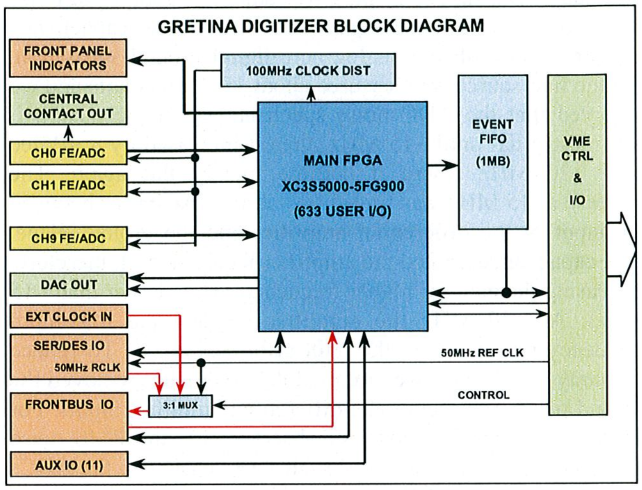


Fig. 6. Digitizer/DSP Block Diagram

TABLE I
DIGITAL SIGNAL PROCESSING

Leading Edge Discrimination

$y(n)=x(n)-x(n\text{-}k)$ (differentiation)

$y(n)=(x(n)+x(n\text{-}2))+x(n\text{-}1)<<1$ (×4, Gaussian filtering)

Threshold comparison → LED time

Constant Fraction Discrimination

$y(n)=x(n)-x(n\text{-}k)$ (differentiation)

$y(n)=(x(n)+x(n\text{-}2))+x(n\text{-}1)<<1$ (×2, Gaussian filtering)

$y(n)=x(n\text{-}k)<<ab-x(n)$ (constant fraction)

Zero crossing comparison → CFD time

Trapezoidal filter and energy determination [5]

$y(n)=y(n\text{-}1)+((x(n)+x(n\text{-}2m\text{-}k))-(x(n\text{-}m)+x(n\text{-}m\text{-}k))$

Maximum tracking → energy

Pole-Zero cancellation

$I(n)=I(n\text{-}1)+x(n)$

$y(n)=x(n)+I(n)/t$ (where t is the pre-amp time constant)

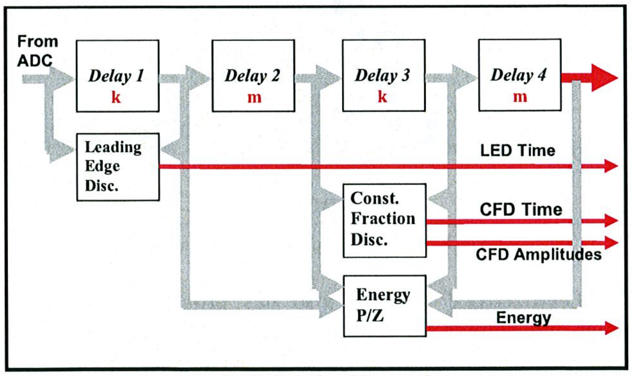


Fig. 7. Digital Signal Processing

The present noise performance of the ADC inputs at full digital processing rate is approximately 1.5 RMS counts for all inputs (refer to Fig. 8). Achieving this performance in a board with digital power supplies and heavy digital activity required special attention. We have filtered the +5V with a π-filter (observe the inductor on the left portion of Fig. 5). Also, the analog inputs are all implemented using balanced differential techniques. During prototyping we found two sources of noise due to board layout: one associated with a stray capacitance and the other with the analog and digital ground separation. For the first source we measured about 15 RMS counts and we observed that the ADC noise spectrum was not constant: it was mostly flat until ~15 MHz; then it increased by 8 dB/dec until ~30 MHz where it became mostly flat again. The increase at 15 MHz was traced to a stray capacitance between the input of the differential amplifier and the ground plane. This capacitance limited the amplifier feedback and, therefore, increasing the gain for higher frequencies (i.e., larger than ~15 MHz). At ~30 MHz the amplifier reached the maximum frequency response and the gain stabilized again. To reduce this stray capacitance we modified the layout and removed the ground plane right below the differential amplifier inputs. For the second source of noise we observed that the RMS noise performance was not similar for all ADC channels: the ADCs on the left portion of the board (i.e., close to theπ-filter, refer to Fig. 5) had approximately 1.5 RMS counts while the ones on the right had approximately 3 RMS counts. We then changed the layout and cut the ground plane around the analog section to create an analog ground mostly free of digital switching. We connected the analog ground together with the digital ground under the ADCs (as suggested on the datasheet) and also close to the π-filter. With these two layout changes we now obtain approximately the same noise performance to all channels.

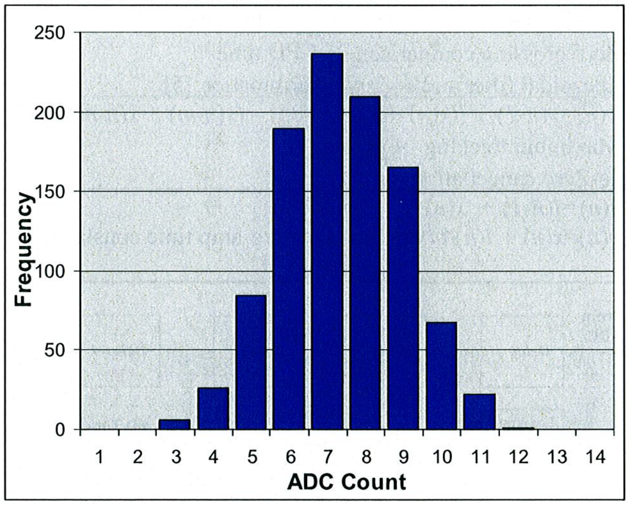


Fig. 8. Noise Distribution: σ= 1.5 counts

The Digitizer/DSP module also has programmable digital inputs and outputs for interfacing with other units and front panel LEDs for status. All the processing described above is done in a XC3S5000 FPGA from Xilinx. Presently about 50% of the FPGA is used, allowing for further increase in algorithm complexity.

IV. TRIGGER, TIMING AND CONTROL

Fig. 9 and 10 show the picture and block diagram of the TTC module, respectively. The TTC is presently configured for multiplicity algorithm, and we intend to add three more in the next few months. When completed, the trigger algorithms will be:

a) Multiplicity: Uses the LED detection of the crystals central contacts. The TTC generates a trigger when it detected that the sum of LEDs crosses some threshold within a time window.
b) Energy: uses the energy of a central contact or the sum of energies of central contacts. When this energy falls within some programmable window the TCC generates a trigger.
c) Pattern distribution: It also uses central contact energy as described above. The TTC triggers when it detects coincidence of gamma-rays energy above threshold in any two pre-programmed crystals.
d) Auxiliary detector trigger: The TTC receives an external trigger within the 40 μsec pipeline depth of the Digitizer/DSP.

We call these algorithms prompt triggers, because a trigger is generated as soon as the proper condition is detected. A delayed coincidence trigger can further refine the meaning of these prompt triggers. When this option is enabled, the TTC generates a trigger command when it detects two prompt trigger conditions (e.g., two multiplicity triggers) in two time windows within the allowed overall trigger window. Finally, observe that all the trigger parameters are configurable through VME.

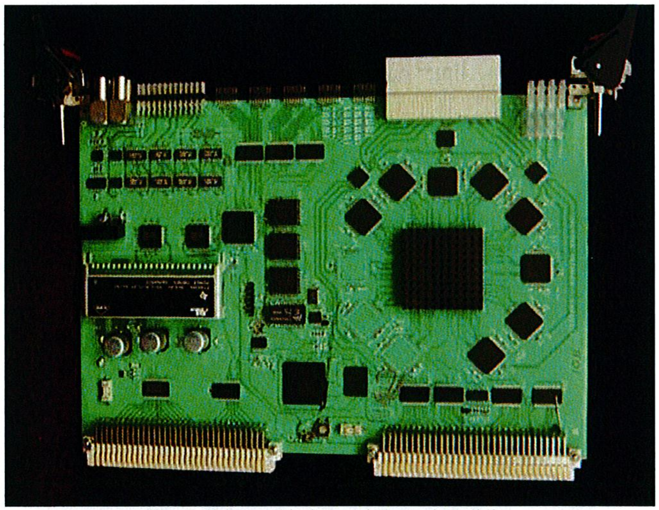


Fig. 9. Trigger, Timing and Control Module

The main FPGA in the TTC module is the Virtex 4 series XC4VLX80 FPGA. This part executes the trigger algorithms and interfaces with eight SerDes. As already described, the SerDes components transfer information between the Digitizers/DSP. The base multiplicity algorithm presently uses only 15% of the resources inside this FPGA, which allows for further development of the trigger algorithms.

The high-speed switching of the SerDes has edge transition times in the order of 100 psec range. These high-speed signals require special attention during layout, and are more critical on the TTC (when compared with the Digitizer/DSP module), since it has eight SerDes and not all can be mounted close to the connector. Carefully matched, impedance controlled differential traces are used. Additionally, the number of vias was minimized: the component side transmits trigger information, and it does not have any vias, while the solder side is used to receive trigger information, and it has just one via per trace of the differential signal. To further improve signal fidelity, a mixed stackup is used where the dielectric under the component and solder layers of the board is Rogers 4350, which has lower losses than FR4.

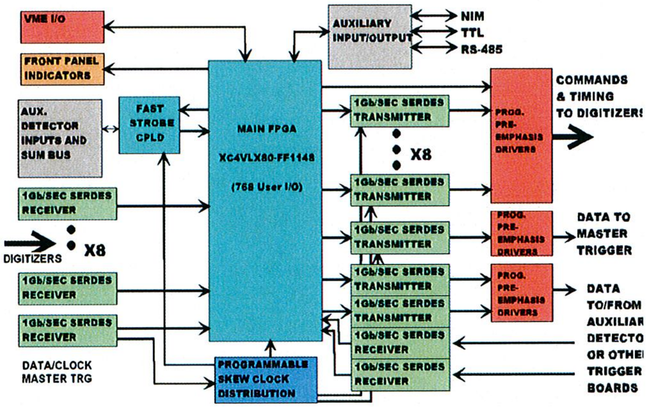


Fig. 10. Trigger, Timing and Control Block Diagram

The electronics requirements state that the Digitizers/DSPs shall sample all crystals within a 2 ns clock phase window. To synchronize all these modules, all CAT5 cables between the TTC and the Digitizer/DSP are approximately of the same length. However, small cable and component delay differences may generate delays that exceed this limit. To compensate for this, the TTC uses programmable clock phase parts (Cypress Semiconductor CY7B992 "RoboClock") to adjust the phase of the transmit clock to each SerDes. This, in turn, adjusts the phase of the sampling clock of the Digitizer/DSP at the other end of the cable.

GRETINA may provide trigger to auxiliary detectors. Some of these detectors use old technology and they do not have digital pipelines; the trigger decision has to be done when the signals are actually traveling on cables, and the sampling has to start when they arrived at the auxiliary electronics. Only the multiplicity algorithm will be used when interfacing with this type of auxiliary detectors. We have implemented a parallel path that by-passes the SerDes to meet the more stringent timing requirements. The SerDes add a delay of approximately 70 ns to serialize and de-serialize the data. This new path uses a dedicated available pair of the CAT5 cable to transmit from the Digitizer/DSP to the TTC system the status of the central contact LED detection. To determine multiplicity, the TTC makes the sum of the LEDs using front panel connectors and CPLDs. The partial sums travel from the Routers to the Master Trigger module, which executes the final sum and determines the multiplicity. We estimate that this processing will take less than 250 ns. Also, it is interesting to highlight that this requirement also imposed one of the criteria to select the ADC: the latency of the AD6645 is just three clock cycles, which allows for faster LED determination.

The TTC module also has programmable digital inputs and outputs for interfacing with other units (e.g., auxiliary detectors & NIM electronics) plus front panel LEDs for status.

V. CABLE AND CONNECTOR SELECTION

We have dedicated substantial effort in selecting the proper cabling for this electronics. One example already discussed is the CAT5 cable for the SerDes. We will now describe the selection of another very important cable, the one that connects the detector pre-amplifiers to the Digitizers/DSPs, which carries the crystal interaction information. Reducing crosstalk between the signals in this interconnection is very important for estimating the position (r, θ, z) of the interactions points: excessive crosstalk decreases the position resolution of the gamma ray interaction. The requirement of total crosstalk in the detector module is less than 0.2% and, therefore, the rest of the electronics crosstalk has to be substantially less than this (required of $< 0.04\%$). To avoid ground loops we decided that the input of the Digitizers/DSPs would be differential and that we would employ twisted pairs. We tested several cables. We obtained the best performance with twisted individually shielded pairs. The crosstalk between adjacent pairs was $< 0.04\%$ for a 15 m cable span. For comparison, a similar cable, but non-twisted (used for LVDS transmission), had a crosstalk of ~2.5%. This crosstalk is explained by the magnetic field created in one pair passing through the shielding and magnetically coupling to an adjacent pair. The shielding between pairs is a thin aluminum layer, and it does not attenuate the magnetic field significantly for these speeds. These tests considered a constant rise time (10-90%) of 30 nsec, well within the expected rise times in GRETINA.

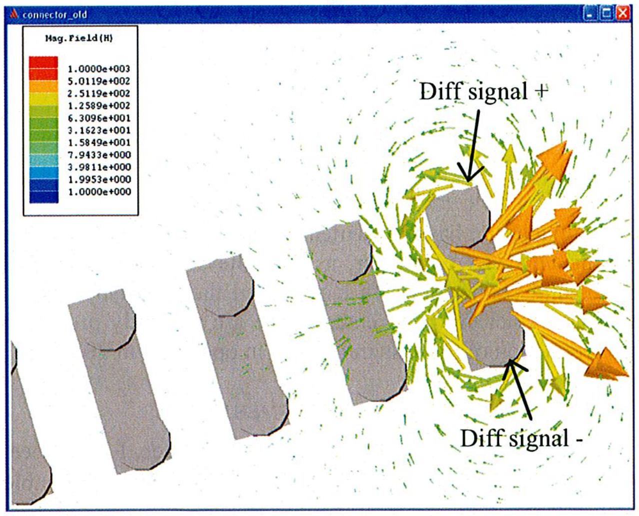


Fig. 11. Magnetic Coupling in 0.1" Spacing Connectors

For the connector between the cable and the digitizer we have also tested several connector types, and we also observed crosstalk between signals. We traced the crosstalk, again, to the magnetic field of one signal coupling into the next signal. This can be visualized in Fig. 11. This figure shows the results of Maxwell simulations [from Ansoft] where the arrows represent the direction and the magnitude of the magnetic field. These simulations were done considering that the positive input of the differential signals are connected on the top row and the negative on the bottom row, forming these differential connections. The offending signal uses the differential connection on the right. One can observe that the magnetic field propagates well into the adjacent differential connections, causing crosstalk. The connector simulated has 0.1" pin spacing. Ways to mitigate this effect are threefold: (a) increasing the distance between differential connections, (b) selecting the direction of the pairs (ideally in an angle of 90°), and (c) short circuiting unused connector pairs (to create an opposing magnetic field that decreasing the offending field). For example, the simulations and subsequent tests have demonstrated that skipping two differential connections (on the right of the offending differential connection) and using the third for the next signal decreased the crosstalk by ~10×, and that short circuiting these two unused set of pins reduce the crosstalk on the third pair by a further ~2.5×. These tests were run with a 100 MHz sinusoidal signal. To meet the crosstalk mitigation strategies outlined above we used a 100 pin Double Density, Subminiature D type connector manufactured by ITT-Cannon. Fig. 12 shows a sketch of how we connected the differential pairs (represented by back circles) and its individual shielding (represented by gray circles). The unused pins were short circuited in pairs. With this arrangement the measured crosstalk was below what we can measure using 14-bits ADC of the Digitizer/DSP (i.e., crosstalk < 0.025%). The signal used had a constant rise time (10-90%) of 30 nsec.

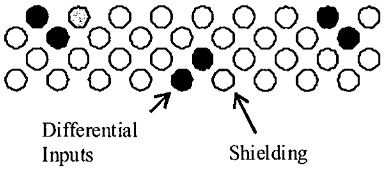


Fig. 12. IIT-Cannon Connector Pin Assignment

VI. SUMMARY AND CONCLUSIONS

In this paper we have presented the DAQ and trigger system for the GRETINA detector, as well the selection of critical cables and connectors. We have tested the prototype system and it has met the requirements. We are presently preparing the production of the system to enter in operation in 2011.

ACKNOWLEDGMENT

We thank the GRETINA Advisory Committee, I-Yang Lee, Augusto O. Macchiavelli and David Radford for the valuable suggestions.

REFERENCES

[1] I. Y. Lee, "Gamma-ray tracking detectors," *Nucl. Inst. Methods A*, vol. 422, pp. 195-200, Feb. 1999.
[2] M. A. Deleplanque, *et.al.*, "GRETA: utilizing new concepts in γ-ray detection," *Nucl. Inst. Methods A*, vol. 430, pp. 292-310, July 1999.
[3] F. S. Goulding and D. A. Landis, "Semiconductor Detector Spectrometer Electronics," in *Nuclear Spectroscopy and Reactions Part A*, J. Cerny, Ed. New York: Academic Press, 1974, pp. 413-481.
[4] "The Future of Gamma-Ray Spectroscopy: GRETA, the Gamma-Ray Energy Tracking Array," fsunuc.physics.fsu.edu/~gretina/GRETA_WP_Jan07_4.pdf, Dec. 2006.
[5] V. T. Jordanov and G. F. Knoll, "Digital synthesis of pulse shapes in real time for high resolution radiation spectroscopy," *Nucl. Inst. Methods A*, vol. 345, pp. 337-345, June 1994.

CONTAGEM DE CÉLULAS SANGUÍNEAS USANDO O ESPAÇO DE ESCALAS.

D. Doering*, A. Schuck Jr.*, L. V. Guimarães*

* UFRGS/Engenharia Elétrica, Porto Alegre, Brasil

e-mail: ddoering@eletro.ufrgs.br

Abstract: In this paper it is proposed a novel method of blood cell counting based on the scale-space blob detector. The method is applied to gray scale images. The blob detector is calculated for a set of defined scales, generating a set of binary images where white pixels are defined as blobs. The set of scales is the only inferred parameter. The lifetime of the blobs are calculated using a "and" logic. After that, the noisy blobs are removed and the cells are detected. The results show that the proposed method is effective on automatic detecting cells.

***Palavras Chaves*: células sanguíneas, contagem, espaço de escala.**

Introdução

As técnicas reconhecidas e utilizadas para a contagem de células são: contagem manual e contagem automática por citômetro de fluxo [1, 2]. Apesar da utilização destes equipamentos sofisticados a contagem manual ainda é muito requisitada. Como alternativa para a contagem manual tem-se o uso de técnicas de processamento e reconhecimento de imagem digital. A automação da contagem de células do sangue através de técnicas de processamento de imagens apresenta as seguintes vantagens: a) aumenta o número de amostras examinadas por dia, quando comparada ao exame manual; b) armazenamento das amostras em forma de imagens (mídia digital); c) um maior número de células é contado e d) o exame através de processamento de imagens tem custo mais baixo que o exame por citômetro de fluxo.

Características das imagens de células do sangue como baixo contraste, distribuição não uniforme da luminosidade e o formato toroidal da hemácia tornam o processo de binarização, ou separação fundo-célula, uma tarefa difícil. Da mesma forma, a aglomeração das células exige um processo de decomposição destas partículas. Outras características como tipo de doença e idade do paciente, por exemplo, fazem da inferência de um parâmetro no processo um fator muito dependente destas características. Portanto, várias técnicas como correlação e componentes conectados [3], morfologia matemática [4, 5], aproximação poligonal [6] e modelo geométrico [7-9], tem sido aplicadas para realizar as tarefas de binarização e decomposição de forma automática e robusta.

Neste trabalho propõe-se uma nova técnica de segmentação e contagem de células baseada no operador detector de bolhas definido em espaço de escalas [10, 11]. Com esta técnica o único parâmetro de inferência é o conjunto de escalas usado para se calcular o detector de bolhas.

Fundamentos Teóricos

O espaço de escalas [10, 11] de uma imagem pode ser visto como um conjunto de imagens onde sucessivamente se retiram os detalhes da imagem original de acordo com o aumento da escala. A representação de uma imagem em uma dada escala, $L(x, y, t) \in \Re$, é definida por:

$$L(x,y,t) = f(x,y) * g(x,y,t) \tag{1}$$

onde $f(x, y) \in \Re$ é a imagem original, * significa o operador convolução, t é a escala e $g(x, y, t) \in \Re$ é o núcleo gaussiano dado por:

$$g(x, y; t) = \frac{1}{(2\pi t)} e^{-\left(x^2+y^2\right)/2t} \tag{2}$$

sendo a escala (t) igual à variância (σ^2) da gaussiana bidimensional.

A vantagem do uso do espaço de escalas é que existem operadores para a detecção de bordas (*edges*), cristas (*ridges*) e bolhas (*blobs*). Estes operadores baseiam-se nas derivadas das imagens no espaço de escalas. O detector de *blobs* [12] é definido através do laplaciano da imagem:

$$\nabla^2 L = L_{xx} + L_{yy} \tag{3}$$

onde $\nabla^2 L$ é o laplaciano da imagem, L_{xx} é a derivada segunda de L em relação a x e L_{yy} é a derivada segunda de L em relação a y. Para imagens com regiões escuras em um fundo claro, por exemplo lâminas de células sangüíneas (Figura-1), o operador detector de *blobs* segue a seguinte regra para classificar tais regiões como sendo *blobs*:

$$\begin{aligned} I(x,y) &= 0, & \nabla^2 L &< 0 \\ I(x,y) &= 255, & \nabla^2 L &\geq 0 \end{aligned} \tag{4}$$

onde I = 0 representa o fundo da imagem e I = 255 representa regiões classificas como *blobs*.

Em [13] Lindeberg e Eklundh mostram, através do esboço primário em espaço de escala, diferentes formas de como relacionar as estruturas (bordas, cristas ou bolhas) em diferentes escalas e como avaliar a sua importância. Um dos métodos apresentados para classificar a importância de uma estrutura é o tempo de vida útil. De acordo com este conceito quanto maior for o tempo de vida útil de um elemento maior será a sua importância. Neste trabalho aplicou-se este conceito *pixel* a *pixel* nas imagens dos *blobs,* onde o uso da lógica *and* entre imagens obtidas em duas escalas sucessivas mostra a persistência destes *pixels*.

Materiais e Métodos

As contagens das células foram realizadas de acordo com a seguinte seqüência de passos:

1. Leitura da imagem original e conversão para tons de cinza.
2. Cálculo do laplaciano da imagem em diferentes escalas, detecções dos *blobs* e a escolha das escalas mais adequadas para o item 3.
3. Cálculo dos *pixels* com tempo de vida útil igual a cem por cento.
4. Eliminação de *blobs* inválidos.
5. Contagem e representação das células sangüíneas.

Os experimentos foram realizados com dez imagens retiradas do banco de dados do Laboratório de Processamento de Sinais (Lapsi) da Universidade Federal do Rio grando do Sul (UFRGS). Elas estão no formato *Joint Photographic Experts Group* (JPEG) com tamanho de 618 x 646 *pixeis* e 24 bits de cores.

Inicia-se o algoritmo com a leitura da imagem e sua conversão para tons de cinza. Esta conversão é realizada através da média aritmética dos valores das componentes de cor vermelha (R), verde (G) e azul (B), dado por:

$$f(x,y)=\frac{R(x,y)+G(x,y)+B(x,y)}{3} \qquad (5)$$

onde f é a imagem representada em tons de cinza (Figura 1). O uso da média aritmética para esta conversão é justificado por minimizar a variação de contraste nas diferentes imagens.

A partir da imagem representada em tons da cinza, aplicou-se o operador detector de *blobs* (4) para as escalas 4, 6, 8, 12, 16 e 256 (Figura 2 (a) – (f) respectivamente). Esta seqüência de escalas foi escolhida após serem realizados testes com as diversas combinações de diferentes escalas. As escolhas das escalas apropriadas levam em consideração o resultado obtido também com as etapas posteriores. No gráfico da Figura 3 são apresentados os erros percentuais e o percentual de falsos positivos obtidos para as seguintes combinações:

I. t = 4, 6, 8, 12, 16 e 256;
II. t = 4, 8, 16 e 256;
III. t = 4, 8, 16, 32 e 256;
IV. t = 4, 8, 16, 64 e 256;
V. t = 4, 8, 16, 128 e 256;
VI. t = 2, 4, 8, 16 e 256,

Foram considerados, para o cálculo do erro percentual, como valores corretos de contagem os obtidos através de contagem manual realizada por especialistas (Tabela 1).

Após a detecção dos *blobs* nas diferentes imagens realizou-se a lógica *and* entre as diferentes imagens dos *blobs*. O resultado desta operação é visto na Figura 4-a.

A partir do resultado da operação *and*, obteve-se a imagem dos *blobs* das células adicionado de alguns pontos de ruído. Foi então realizada uma etapa de eliminação de objetos que não correspondem a *blobs* de células. Este processo foi realizado em duas partes: primeiramente, eliminou-se os *blobs* que tocavam a borda da imagem, eliminando assim efeitos de ruídos de borda. Em seguida, foi feita à eliminação de objetos com tamanho inferior a 250 *pixels*. Este parâmetro foi fixado após verificar-se que para uma das imagens o menor tamanho de *blob* detectado foi de 263 *pixels*. Este resultado é visto na Figura 4-b.

Finalizando o algoritmo foi então realizada a contagem do número de *blobs* restantes, usando para isso a função *bwlabel* do Matlab 5.3 para vizinhos conectados de 4. O número de *blobs* é considerado igual ao número de células detectadas. Para cada *blob* foi desenhado um círculo em seu centróide para facilitar a visualização do resultado do algoritmo (Figura 5). Os resultados das contagens de células, para as imagens analisadas com este algoritmo, com a contagem manual e com os resultados obtidos pelo algoritmo proposto em [14] são vistos também na Tabela 1. Na Figura 6 são vistos os resultados deste algoritmo para outras 3 imagens.

Resultados

A seguir são vistos os resultados obtidos com o método proposto no item anterior.

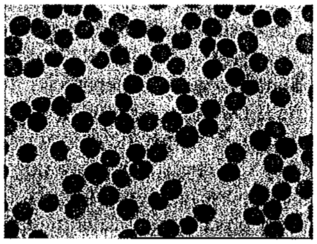


Figura 1- Imagem das células em tons de cinza (imagem #9 da Tabela 1).

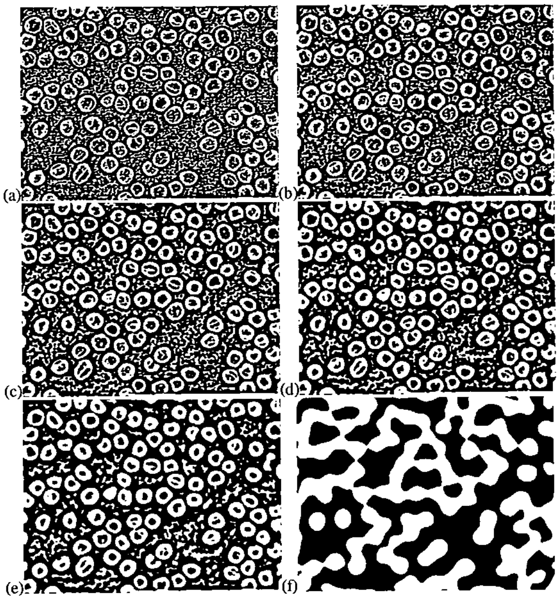


Figura 2- Imagem resultante da aplicação do operador detector de *blobs* nas escalas 4 (a), 6 (b), 8 (c), 12 (d), 16 (e) e 256 (f).

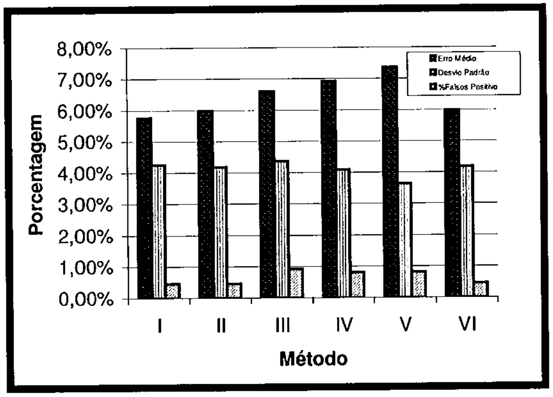


Figura 3- Erro médio, desvio padrão e percentual de falsos positivos obtidos com as diferentes combinações de operadores detectores de *blobs*.

Tabela 1: Número das imagens (Img. (#)), contagens obtidas com o método proposto (Cont. 1), contagem manual (Cont. 2) e contagem com o método descrito em [14] (Cont. 3).

Img. (#)	Cont. 1	Cont. 2	Cont. 3
1	89	91	90
2	80	90	84
3	83	95	88
4	87	86	91
5	98	100	103
6	89	90	97
7	93	101	102
8	83	91	90
9	86	92	89
10	91	95	98

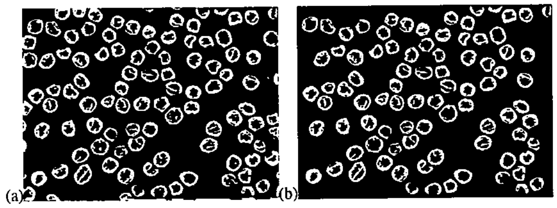


Figura 4- Em (a), imagem obtida após a execução da etapa 3 (Cálculo dos *pixels* com tempo de vida útil igual a cem por cento) para a imagem da Figura 1. Em (b), imagem dos *blobs* restantes após a execução da etapa 4 (eliminação dos blobs inválidos) para a imagem da Figura 1.

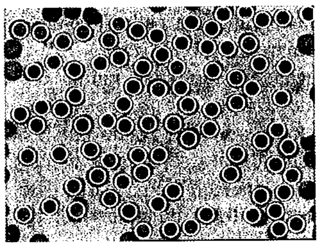


Figura 5- Imagem obtida depois de realizado todo o processamento do método proposto para a imagem da Figura 1.

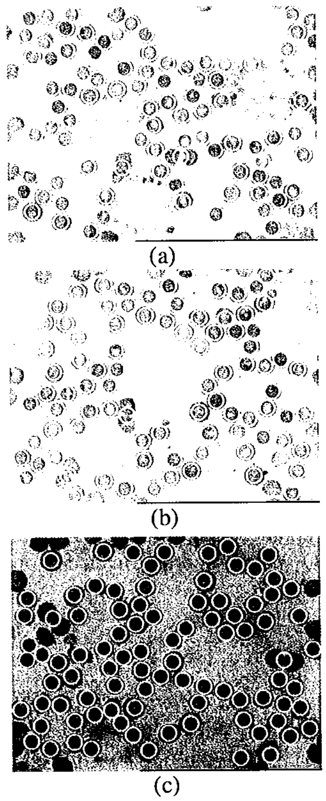


Figura 6- Resultados obtidos para as imagens #1(a), #4(b) e #10(c) da Tabela 1.

Discussões e Conclusões

Neste trabalho é apresentado um novo método de segmentação e contagem de células sanguíneas. Este método baseia-se nos operadores do espaço de escalas, especificamente no detector de bolhas.

No método apresentado define-se apenas um parâmetro de inferência, que é o conjunto de escalas que é usado para detectar as bolhas. Este parâmetro é único para um determinado equipamento, uma vez que o método é robusto em relação ao contraste, como pode ser visto na Figura 6. Por este parâmetro variar apenas com a mudança de equipamento, torna-o mais rápido do que o método apresentado em [4], onde é necessário o cálculo do histograma de cada imagem para então se definir um limiar para a binarização.

Nas imagens apresentadas, muitas das células estão agrupadas, de forma que a binarização por limiar identifica todo o conjunto, sendo então necessário um processo adicional para separar as células. Em [4] foi usado métodos morfológicos para esse fim. Neste trabalho, não se utilizou qualquer método para separar as células, uma vez que um pequeno número delas se manteve agrupado após o processamento da imagem. Isto mostra mais um ponto positivo do método apresentado, uma vez que não é necessário um processo adicional para separar as células.

O erro apresentado pela contagem do algoritmo, em relação a contagem manual, é de 5,77% e desvio padrão de 4,25% (Figura 3). Isto se deve ao fato de que as bolhas que tocam a borda são desconsideradas, como descrito anteriormente, evitando assim erros de borda. A solução para este problema é a mudança no método de aquisição. Neste trabalho foram fotografadas partes de uma lâmina, por limitação do equipamento. Mas usando outro equipamento que fotografe a lâmina inteira, não haverá células posicionadas junto à borda. Com isso elas não serão desconsideradas e farão com que o erro absoluto do método diminua.

Após a apresentação do método e das discussões pode-se afirmar que o método apresentado neste trabalho realiza a contagem de células sanguíneas de forma automática e robusta.

Agradecimentos

Nossos agradecimentos são direcionados à CAPES e ao CNpQ pelo suporte financeiro parcial deste trabalho.

Referências

[1] BROWN B. A., *Hematology, Principles and Procedures*, Lea & Febiger, 1988.

[2] National Committee for Clinical Laboratory Standards, *Reticulocyte Counting by Flow Cytometry; Proposed Guideline*, NCCLS document H44-P, NCCLS, 1993.

[3] FIGUEIRO, T. SCHUCH, N., GUIMARÃES, L., SUSIN, A., A Method for Automatic Detection of Blood Cells on Images using Image Correlation and Connected Components, *WICCGPI SIBGRAPI* 2003.

[4] KOYA, T. K.; GUIMARÃES, L. V. ET AL., "Automatic reticulocyte counting indigitized images - step one", *International Congress of the World Apheresis Association, Boletim da Sociedade Brasileira de Hematologia e Hemoterapia*, 1998.

[5] Y. D. MA, R. L. DAI, L. LIAN, Z. F. ZHANG, "An Counting and Segmentation method of Blood Cell Image with Logical and morphological Feature of Cell", *ICONIP-2001, 8th International Conference on Neural Information Process,* 2001.

[6] F. KUBO AND K. SAKAUE, "A Decomposition Algorithm for Particle Image using Polygonal Approximation", *Trans. IEICE*, J71(D): 1852-1854, 1988. (in Japanese).

[7] L. V. GUIMARÃES, A. A. SUZIM AND J. MAEDA, "A Circle Similarity Algorithm for an Automatic Circular Decomposition of Blood Cell Images", *Optical Review*, 8(6):436-443, 2001.

[8] L. V. GUIMARÃES, A. A. SUZIM AND J. MAEDA, "A New Automatic Circular Decomposition Algorithm Applied to Blood Cell Images", *Proceedings. of IEEE Int. Conf. on Bio-Informatics & Biomedical Eng.,* pp.277-280, 2000.

[9] N. MALPICA, C. O. DE SOLÓRZANO, J. J. VAQUERO, A. SANTOS, I. VALLCORBA, J.M.GARCÍA-SAGREDO AND F. DEL POZO, "Applying Watershed Algorithms to the Segmentation of Clustered Nuclei", *Cytometry*, 28:289-297, 1997.

[10] WITKIN, A. P., "Scale-space Filtering". *In Proc. 8th Int. Joint. Conf. Art. Intel.* 1019-1022, 1983.

[11] LINDEBERG, T., "Scale-Space: A Framework for handling image structures at multiple scales", *Proc. CERN School of Computingm Netherlands,* 1996.

[12] LINDEBERG, T., "Detecting salient blob-like image structures and their scales with a scale-space primal sketch: A method for focus-of-attention", *International Journal of Computer Vision*, 11:283-318, 1993.

[13] LINDEBERG, T., EKLUNDH, J. "Scale-Space primal sketch: construction and experiments", *Image and Vision Computing*, 10(1), 1992.

[14] KOYA, T. K., *Sistema automático de Contagem de Células Sangüíneas*, Dissertação de Mestrado, 2001.

SEGMENTAÇÃO DE CAVIDADES EM IMAGENS DE ECO-CARDIOGRAFIA EM ESPAÇO DE ESCALAS

D. Doering*, A. Schuck Jr.*

* UFRGS/Engenharia Elétrica, Porto Alegre, Brasil

e-mail: ddoering@eletro.ufrgs.br

Abstract: It is proposed in this paper a novel method for segmentation of the cardiac chambers based on scale-space and binarization by threshold. The threshold is calculated using the interior chamber image histogram, selected by the user. Four equations are evaluated for both the threshold calculus and the performance of the whole method. After then, the scale space is applied to the images, each scale-space image is binarized and an "and" logic is performed between the binarized scale-space images. At last, the chamber is enhanced with a Canny edge detector. The quality of the results obtained shows the adequancy of the method to echo-cardiographic images.

***Palavras Chaves*: Eco-cardiografia, segmentação, espaço de escala.**

Introdução

A eco-cardiografia é uma técnica amplamente utilizada em diagnósticos de doenças cardíacas [1]. Com o intuito de aumentar a qualidade do diagnóstico e diminuir sua dependência da experiência dos técnicos, é desejável a extração de parâmetros a respeito do coração (deformação do ventrículo esquerdo [2], fração de fluxo ejetado e débito cardíaco [3]) de forma automática ou semi-automática. Para isso é necessária a segmentação automática das cavidades cardíacas.

As imagens de ultra-som vêm acompanhadas de diversos tipos de degradações, incluindo, ruído impulsivo, distorções e não homogeneidade das intensidades [4], o que não permite o uso de técnicas convencionais para a realização da segmentação dessas imagens. Para a solução desta tarefa são usadas técnicas como *wavelets* [5]; correção das inhomogeneidades das intensidades [4]; redução de ruído impulsivo [6]; o uso de espaço de escalas [3] e combinações entre múltiplas escalas [2].

Neste trabalho propõe-se a segmentação de cavidades em imagens de eco-cardiografia baseada na representação das imagens em espaço de escalas, binarização por limiar das imagens nas diferentes escalas, e a combinação linear das imagens binarizadas. O valor do limiar é calculado com base no histograma do interior da cavidade que se deseja segmentar.

Fundamentos Teóricos

A representação de uma imagem em espaço de escala [7, 8] é formada por um conjunto de imagens que representam a imagem original em diferentes escalas. Matematicamente, dada uma função bi-dimensional $f(x, y) \in \Re$, a sua representação em espaço de escala é $L(x, y, t) \in \Re$, definida por:

$$L(x,y,t) = f(x,y) * g(x,y,t) \qquad (1)$$

onde * é o operador convolução, t é a escala e $g(x, y, t) \in \Re$ é o núcleo (*kernel*) gaussiano dado por:

$$g(x, y; t) = \frac{1}{(2\pi t)} e^{-\left(x^2+y^2\right)/2t} \qquad (2)$$

sendo a escala (t) igual à variância (σ^2).

A eq. (1) produz um conjunto de imagens suavizadas por uma gaussiana bidimensional de largura variável. Pode-se re-definir esta expressão no domínio da freqüência por:

$$L(x, y; t) = F^{-1}\{F[g(x, y; t)]F[f]\} \qquad (3)$$

onde F é o operador DFT (*Discrete Fourier Transform*) bidimensional da imagem e F^{-1} é o operador DFT inverso.

Lindeberg e Eklundh [9] argumentam que o espaço de escalas é uma ferramenta bem definida para lidar com as estruturas da imagem, que naturalmente ocorrem em diferentes escalas. Para extrair informações das imagens, são definidos operadores em espaço de escalas, tais como detectores de bordas (*edges*), cristas (*ridges*) e bolhas (*blobs*) [8]. Em [9] são mostradas diferentes formas sobre como relacionar as estruturas em diferentes escalas e como avaliar a sua importância. Um dos métodos apresentados para classificar a importância de uma estrutura é o tempo de vida útil. Segundo este conceito quanto maior for o tempo de vida útil de um elemento (a presença destes elementos nas diferentes imagens do espaço de escalas) maior será a sua importância. Neste trabalho aplicou-se este conceito *pixel* a *pixel* nas imagens do espaço de escala.

Gonzales e Woods [10] apresentam diversas técnicas de segmentação de imagens usando um limiar extraído de um histograma. Neste trabalho usa-se a limiarização global simples para a segmentação das cavidades, com o histograma retirado apenas do interior da cavidade. Este procedimento é definido por:

$$s(x,y)=\begin{cases}1 & f(x,y)>th\\ 0 & f(x,y)\le th\end{cases} \qquad (6)$$

onde th é o valor de limiar escolhido para a binarização e s é o valor do *pixel* da imagem segmentada, sendo que 1 representa a parede da cavidade e zero o interior da mesma. O procedimento para a escolha do valor de th é descrito na seção seguinte.

Materiais e Métodos

A segmentação das cavidades e a extração de suas bordas foram realizadas de acordo com o seguinte algoritmo:

1. Seleção de uma região no interior da cavidade de interesse pelo usuário.
2. Cálculo do histograma da região selecionada no item 1.
3. Determinação do limiar para a binarização da imagem.
4. Representação das imagens em espaço de escala.
5. Binarização das imagens nas diferentes escalas.
6. Cálculo dos *pixels* que possuem tempo de vida útil igual a 100%.
7. Segmentação da cavidade.
8. Cálculo dos *pixels* de borda da cavidade.

A validação do algoritmo proposto foi realizada com o uso de oito imagens na vista denominada 4 câmaras, obtidas do banco de dados do Dr. Iran Castro do Instituto de Cardiologia do Rio Grande do Sul. Tais imagens foram obtidas com o uso do equipamento HP Sonos 2500. Seu tamanho é de 640 x 480 *pixels* com 256 cores em níveis de cinza, sendo que as imagens foram recortadas para ficarem quadradas com tamanho final igual a 480 x 480 *pixels*. Nestas imagens podem ser encontradas cavidades com a borda aberta, ou seja, situações onde não há fronteira entre duas cavidades ou cavidades com a borda fechada (conforme a Tabela 1). Os programas foram desenvolvidos usando o Matlab versão 5.3.

A determinação do limiar para ser utilizado na binarização das imagens (item 3 do algoritmo) foi baseada no histograma (item 2) da região selecionada (item 1) pelo usuário. O valor do limiar utilizado para segmentar a imagem foi obtido após a avaliação de quatro diferentes fórmulas, sendo elas descritas pelas seguintes equações:

$$th=\frac{th_1+th_2}{2} \qquad (5)$$

$$th=th_1 \qquad (6)$$

$$th=th_2 \qquad (7)$$

$$th=th_{-3dB} \qquad (8)$$

onde th é o limiar escolhido; th_l é o valor onde o histograma atinge o máximo valor; th_2 é o valor da intensidade do *pixel* mais claro da região e th_{-3dB} é o nível de cinza correspondente a –3dB do valor máximo do histograma (th_l). Isto é indicado na Figura 1. O algoritmo completo de segmentação foi implementado para estas quatro binarizações, e os melhores resultados foram os obtidos com as equações (5) e (8), optando-se por usar (5) neste trabalho. Os limiares utilizados para cada uma das imagens são vistos também na Tabela 1.

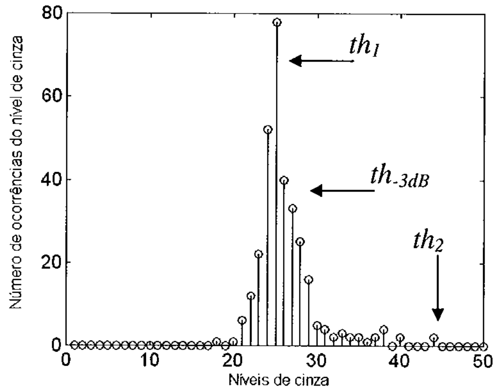


Figura 1: Exemplo dos pontos escolhidos para a determinação do limiar para a binarização da imagem nº9 da Tabela 1.

Para a representação das imagens em espaço de escala (item 4), foram usadas as escalas 2^j para j = 1,2,...8. O limite de escala 2^8 foi estabelecido a partir da observação de que valores de escalas maiores não influenciam o resultado do item 6.

O próximo passo foi a binarização de todas as imagens do espaço de escala usando o limiar escolhido. Este procedimento é exemplificado na Figura 2 para quatro escalas diferentes da mesma imagem.

A seguir determinaram-se os *pixels* da imagem com tempo de vida útil igual a 100%, ou seja, que são brancos em todas as imagens do espaço de escala. Este cálculo é obtido fazendo-se uma lógica E (*And*) com todas as imagens. Isto é mostrado nas Figuras 3-(a) e 3-(b).

A segmentação da cavidade (item 7) foi feita em duas partes. Primeiro fez-se a rotulação de todas as cavidades por vizinhança (pontos pretos que se encontram ligados pertencem à mesma cavidade, obtidos no item 6) e a seguir, selecionou-se aquela região que contivesse a região selecionada para o histograma. Isto é mostrado na Figura 4.

A seguir, na imagem da cavidade segmentada, aplicou-se o algoritmo de Canny para a determinação das bordas desta cavidade. Nas Figuras 5, 6 e 7 são mostrados exemplos dos resultados obtidos com o procedimento apresentado.

Resultados

A seguir são vistos os resultados obtidos com o método proposto no item anterior.

Tabela 1: Imagens utilizadas, o tipo de cavidade encontrada (Aberta ou Fechada) e o limiar calculado para a sua binarização.

Imagem (#)	Tipo	Limiar
1	Fechada	37
2	Fechada	36
3	Fechada	27
4	Fechada	29
5	Aberta	34
6	Fechada	34
7	Aberta	34
8	Aberta	35
9	Aberta	34

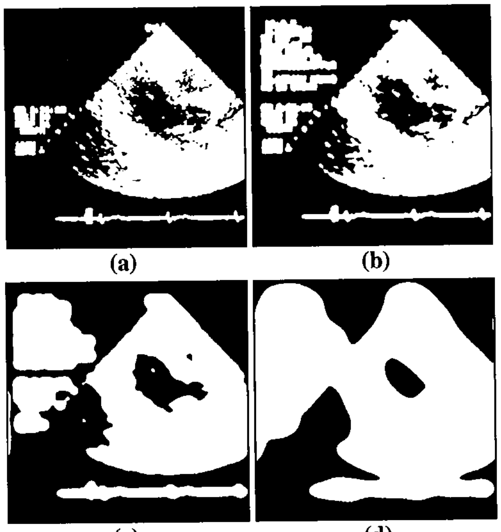
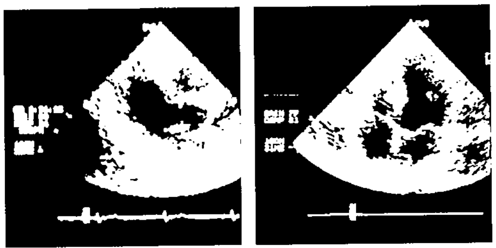


Figura 2: Representação da imagem #6 em espaço de escalas, binarizadas pela equação (5), para as escalas 1 (a), 4 (b), 32 (c) e 256 (d).



(a) (b)

Figura 3: As regiões em branco correspondem a *pixels* da imagem que possuem tempo de vida útil igual a cem por cento: (a) Imagem #6 da Tabela 1. (b) Imagem #3 da Tabela 1

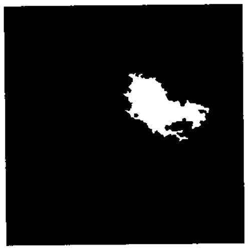
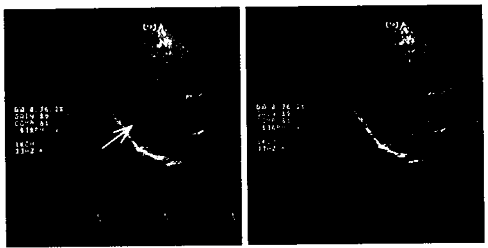
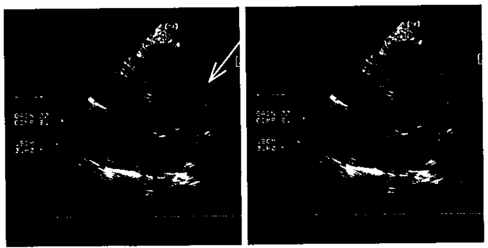


Figura 4: Cavidade segmentada com o algoritmo proposto para a imagem #6 da Tabela 1.



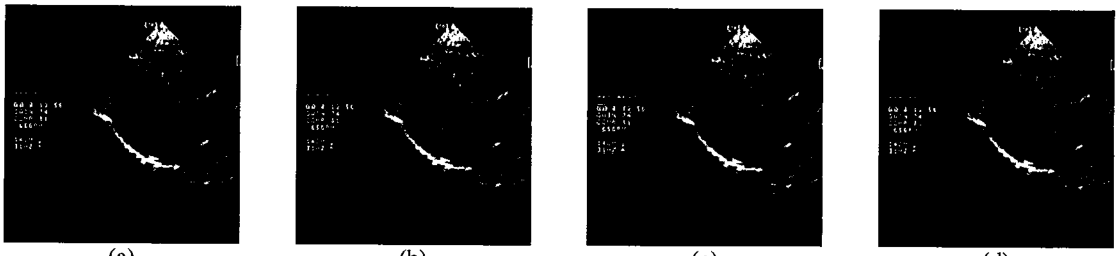
(a) (b)

Figura 5: Exemplo de resultado obtido com cavidade fechada (imagem #6 da Tabela 1): (a) Imagem original com a indicação da cavidade selecionada. (b) Imagem original com as bordas da cavidade de interesse destacadas.



(a) (b)

Figura 6: Exemplo de resultado obtido com cavidade fechada (imagem #3 da Tabela 1): (a) Imagem original com a indicação da cavidade selecionada. (b) Imagem original com as bordas da cavidade de interesse destacadas.



(a) (b) (c) (d)

Figura 7: Exemplo de resultado obtido com os diferentes métodos de cálculo de limiar (imagem #1 da Tabela 1), (a) equação (5), (b) equação (6), (c) equação (7), (d) equação (8).

Discussões e Conclusões

Para a discussão dos resultados obtidos com o método proposto nesse trabalho, existem alguns pontos que merecem ser destacados, tais como:

a) Os resultados para o cálculo do limiar usando (5) para as diferentes imagens foram mostrados na Tabela 1. O limiar para esse conjunto de cavidades varia desde 27 até 37. Como pode ser observado, a binarização de uma imagem com diferentes valores de *threshold* varia significativamente (Figura 7). Esta característica se deve à alta sensibilidade da segmentação em função do limiar aplicado. Com base neste fato pode-se afirmar que não é possível determinar um valor ótimo para este parâmetro, sendo necessária a indicação de uma região no interior da cavidade para então poder iniciar o algoritmo proposto. Esse procedimento exige a intervenção do usuário, mas ela é mínima e os resultados obtidos a justificam. Para a segmentação automática das cavidades, é necessário um procedimento adicional como o descrito em [3] que indique uma região no interior da cavidade de interesse.

b) Conforme relatado em [2], o uso de método de segmentação de imagens de ultra-som baseadas em limiar não apresenta bons resultados. Porém, a segmentação destas imagens usando a técnica apresentada mostra que a união do espaço de escalas com a segmentação por limiar é capaz de gerar resultados satisfatórios, como visto nos resultados apresentados.

Neste trabalho foi apresentado um novo método de segmentação de cavidades de imagens de eco-cardiografia. Este método é baseado na representação das imagens em espaço de escalas e sua segmentação é baseada em binarização por limiar. O limiar é calculado com base no histograma do interior da cavidade que se deseja segmentar. Foram apresentados resultados obtidos com o método aplicado em imagens reais utilizadas em diagnósticos cotidianos. Com base nos resultados e nas discussões apresentadas pode-se afirmar que este novo método de segmentação de imagens de eco-cardiografia detecta as cavidades e suas bordas de forma correta.

Agradecimentos

Nossos agradecimentos são direcionados à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo suporte financeiro parcial do projeto.

Referências

[1] CASTRO, I., Cardiologia: Princípios e Prática, Artmed, 1999.

[2] LIN, N., YU, W., DUNCAN, J. S., "Combinative multi-scale level set framework for echocardiographic image segmentation", *Medical Image Analisis* 7:529-537, 2003.

[3] ZUCCOLOTTO, M., OLABARRIAGA, S., SCHUCK JR., A., 'Segmentação de cavidades em imagens cardiográficas utilizando espaço de escalas', IICLAEB, 2002.

[4] XIAO, G., BRADY, M., NOBLE, J., ZHANG, Y., "Segmentation of ultrasound B-mode images with intensity inhomogeneity correction", *IEEE Transactions on Medical Imaging* 21(1):48-57, 2002.

[5] KASPERSEN, J. H., LANGO, T., LINDESETH, F., "Wavelet-based edge detection in ultrasound images", *Ultrasound in Med & Biol.*, 27(1):89-99, 2001.

[6] MONTAGNAT, J., SERMESANT, M., DELINGETTE, H., MALANDAIN, G., AYACHE, N., "Anisotropic filtering for model-based segmentation of 4d cylindrical echocardiographic images". *Pattern Recognition Letters*, 24:815-828, 2003.

[7] WITKIN, A. P., 'Scale-space Filtering'. *In Proc. 8th Int. Joint. Conf. Art. Intel.* 1019-1022, 1983.

[8] LINDEBERG, T., '*Scale-Space: A Framework for handling image structures at multiple scales*', Proc. CERN School of Computing Netherlands, September, 1996.

[9] LINDEBERG, T., EKLUNDH, J. "*Scale-Space primal sketch: construction and experiments*", Image and Vision Computing, 10(1), 1992.

[10] GONZALES, R.C., WOODS, R.E., "Limiarização", *Processamento de Imagens Digitais*, São Paulo: Edgard Blücher, pp. 316-326, 1992.

Constructing Approximate Voronoi Diagrams from Digital Images of Generalized Polygons and Circular Objects

Waldir L. Roque and Dionísio Doering
Instituto de Matemática
Universidade Federal do Rio Grande do Sul
91509-900, Porto Alegre, Brazil
[roque,ddoering]@mat.ufrgs.br

ABSTRACT

In this paper we present the geometrical construction of an approximate generalized Voronoi diagram for generalized polygons and circular objects based on their minimum geometrical structure that are extracted from the object's digital image. The construction is done in $O(n)$ time complexity, where n is the number of single points defining the set of objects. An application of this technique has been done for mobile robot path planning.

Keywords

Voronoi diagrams, Generalized polygons, Digital images.

1. INTRODUCTION

The geometric construction of Voronoi diagrams has an extensive literature (see [Aur91,Oka92] and references therein). The image-based construction of Voronoi diagrams for a set of digital points has been treated in [Par93,Bor86,Mel92] and for extended digital shapes in [Arc86,Mel94,Sud99]. In these papers the main approach to compute the Voronoi diagram is based on labeling the connected components of the objects and in the application of the morphological operation of shape dilation. By this means, the Voronoi edges are found between two adjacent objects when two different labels are met. The computational cost is of the order of $O(n^2)$, where n^2 is the image size.

The construction of the Voronoi diagram for a set of digital shapes is actually an approximate diagram due to the fact that the objects are constituted of pixels, which have a discrete structure. In this paper, we discuss an alternative approach to construct the Voronoi Diagram for generalized polygons and circular objects based on their minimum geometrical structures, which are extracted from their digital images. The principle is quite simple. A simple polygon is fully characterized by its ordered sequence of vertices and an arc segment can be approximated by the set of vertices that form its polygonal line approximation. Therefore, a generalized polygon, whose edges are straight lines and arc segments, can ultimately be characterized by a set of vertices, too.

WSCG SHORT PAPERS Proceedings
WSCG'2003, February 3-7, 2003, Plzen, Czech Republic.

In the section 2 a brief introduction to the geometric construction of planar Voronoi diagrams is given, in the section 3 the image processing required to identify the minimum geometrical structure of generalized polygonal objects and circles is presented; section 4 briefly report the application to robot path planning and section 5 is devoted to conclusions.

2. VORONOI DIAGRAM

In what follows, we provide a brief introduction to the geometrical construction of planar Voronoi diagrams.

2.1 Ordinary Voronoi diagrams

The planar ordinary Voronoi diagram (OVD) [Oka92] is defined as a partition of the plane into regions according to the principle of the *nearest neighbor*. More precisely, let consider

$P = \{p_1, p_2, \ldots, p_n\}$ a set of non-collinear points in the plane and let consider $d(p, p_i)$ the Euclidean distance from a point $p \neq p_i$ to a point p_i. The Voronoi region $R(p_i) \equiv R_i$ generated by the point p_i is defined as

$$R_i = \{p; d(p, p_i) \leq d(p, p_j), \forall p_i \neq p_j\} \tag{1}$$

.

The Voronoi diagram $V(P)$ for a set of points $P = \{p_1, p_2, \ldots, p_n\}$, is defined as the union of all Voronoi regions $V(P) = \bigcup_{i=1}^{n} R(p_i)$. The points p_i are called Voronoi generators, the edge common to two Voronoi regions is called a Voronoi edge and the vertices where three or more Voronoi edges meet are called Voronoi vertices. We say that a Voronoi generator p_i is adjacent to p_j when their Voronoi regions share a common edge. According to this definition, the Voronoi diagram is such that any point on the edge of two neighboring regions is equidistant from the corresponding Voronoi generators.

2.2 Generalized Voronoi diagram

The planar OVD can be extended to objects like straight lines and arcs segments or polygons and circles. The diagram for these objects is called the generalized Voronoi diagram (GVD).

2.2.1 GVD for lines and arcs

Let $L = \{L_1, L_2, L_3, \ldots, L_n\} \subseteq R^2$ be a set, where L_i can be a straight line or arc segment, such that $L_i \cap L_j \neq 0$, for $i \neq j$. Let us define the distance from a point p to L_i by the shortest distance between p and a point p_i on L_i:

$$d_s(p, L_i) = \min_{x_i} \| \mathrm{x} - \mathrm{x}_i \|; \mathrm{x}_i \in L_i \tag{2}$$

where x and x_i are the position vectors of p and p_i, respectively. The Voronoi region $R(L_i)$ is given by

$$R(L_i) = \{p; d_s(p, L_i) \leq d_s(p, L_j), j \neq i, j \in I_n\} \tag{3}$$

The union of all Voronoi regions $V(L) = \bigcup_{i=1}^{n} R(L_i)$ generates the line Voronoi diagram for the set L.

2.2.2 GVD for polygons and circles

An extension of the line Voronoi diagram for the case of polygonal (simple or generalized[1]) objects can be done considering the generation of the Voronoi diagram for areas. Let $A = \{A_1, A_2, \ldots A_n\}$ in R^2 be a set of areas. Assuming that the areas are connected closed sets with unity Euler number and that they do not intersect each other, we define the distance from a point p to A_i as the shortest distance from p to p_i on A_i as follows:

$$d_s(p, A_i) = \min_{\mathrm{x}_i} \|\mathrm{x} - \mathrm{x}_i\|; \mathrm{x}_i \in A_i, \tag{4}$$

where x and x_i are the position vectors of p and p_i, respectively. According to this distance, we may define the Voronoi regions $R(A_i)$ associated to each area as

$$R(A_i) = \{p; d_s(p, A_i) \leq d_s(p, A_j), j \neq i, j \in I_n\}. \tag{5}$$

The area Voronoi diagram is the set $V(A) = \bigcup_{i=1}^{n} R(A_i)$. The area Voronoi diagram can be seen as the diagram for generalized polygons, where the generalized polygons can be represented by their area contours. Note that the area Voronoi diagram subsumes the line and the ordinary Voronoi diagrams.

2.3 Computational generation of the GVD

Several algorithms have been proposed to generate the planar Voronoi diagram for a set of objects [Aur91]. An interesting algorithm was proposed by Sugihara and Iri [Sug92], based on an incremental construction, to generate the OVD with average running time complexity of $O(n)$, where n is the number of Voronoi generators, that is also stable to numerical errors.

The construction of the OVD based in [Sug92] starts with a trivial diagram for three generators and adds up a new generator one by one at a time. To generate the new Voronoi region $R(p_l)$ for p_l one need first to identify the generator p_i, in whose region $R(p_i)$ the new generator p_l is contained in. Then one draw the perpendicular bisector between p_l and p_i until it intersects the edges of $R(p_i)$. The bisector intersects the edges of $R(p_i)$ in two points. Let call q be one of them. Now take the perpendicular bisector of p_l and p_j, starting from q till it intersects another edge of $R(p_j)$. This procedure should be followed for all Voronoi generators adjacent to p_l until we get back to the region $R(p_i)$. Now, removing the edges enclosed by the closed sequence of bisectors the Voronoi region for the new generator p_l is found. Figure 1 shows the OVD for a set of 14 Voronoi generators. Figure 2 illustrates the procedure to

---

[1] Generalized polygons are polygons whose edges are straight lines or arc segments.

generate the new Voronoi region $R(p_{15})$, when the generator P15 is inserted.

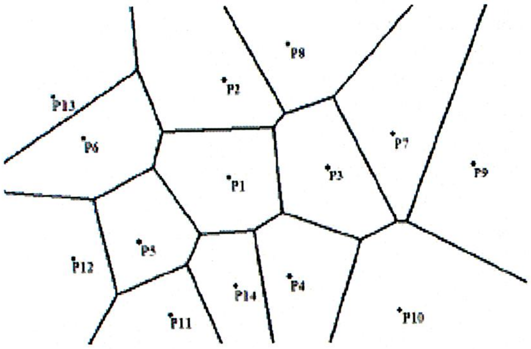


Figure 1: OVD for 14 Voronoi generators.

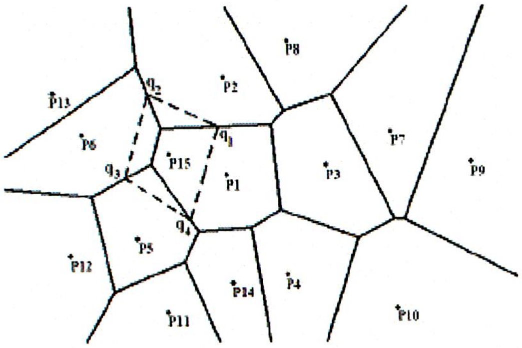


Figure 2: Construction of Voronoi region $R(p_{15})$ for the new generator P15.

2.3.1 *Constructing approximate GVD*

An extension of the incremental type algorithm described above can be done to construct an approximate generalized Voronoi diagram (AGVD) [Sug93] for straight lines, arc segments, generalized polygons and circles. For that, let us consider a line or arc segment approximated by a sequence of n points with a constant small displacement δ between them. In other words, we are doing a polygonal approximation of the segments with a variable number of vertices. The OVD algorithm can be applied to generate the AGVD for these objects if the Voronoi edges that cross the line or arc edges be invalidated and omitted, remaining only the Voronoi edges of adjacent objects. The same approach can be applied to generalized polygons and circles.

Figure 3 shows the AGVD for a set of generalized polygonal objects in a bounded square region. In this picture all Voronoi edges can be seen and the resolution displacement parameter δ was fixed, but it can be arbitrarily adjusted at the expense of computational cost. Figure 4 shows the corresponding AGVD omitting the invalid Voronoi edges.

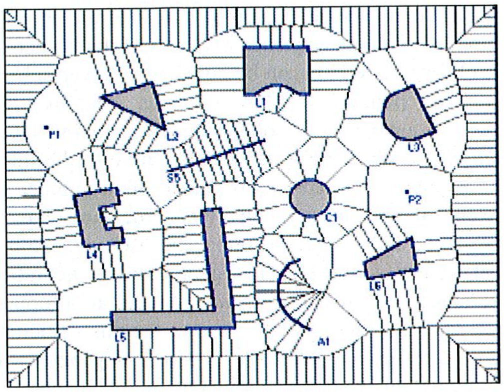


Figure 3: AGVD for a set of objects, including the invalid Voronoi edges in a bounded region.

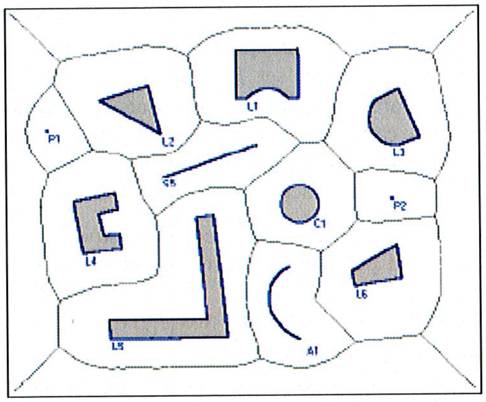


Figure 4: AGVD for objects omitting the invalid Voronoi edges.

3. GEOMETRICAL STRUCTURE FROM DIGITAL IMAGES

The construction of the GVD for a set of objects in a digital picture is actually an approximate diagram due to the fact that the objects are constituted of pixels, which correspond to a discrete structure. Remind that, in a 8-neighborhood scheme, the distance from the central pixel to its neighbors is either 1 or $\sqrt{2}$.

Several algorithms have been proposed to compute the Voronoi diagram from a set of digital points [Par93, Mel92] or for extended digital shapes [Arc86, Bor86, Sud99]. The main approach used by them to compute the Voronoi diagram was based on the principle of labeling the objects and applying the morphological operation of shape dilation to grow the objects. The Voronoi edge is formed between two adjacent objects when two different labels are met.

According to the AGVD algorithm construction discussed in section 2.3, an alternative approach to construct the diagram for a set of digital objects is to find out the pixels that form the border of the objects. Provided with the sequence of pixels of the border of an object the AGVD can be computed. Nevertheless, as was pointed in the previous section, the computational cost becomes high due to the large number of single points. One alternative to diminish the computational cost is to get rid of a number of pixels of the border, but this procedure has a drawback and requires caution because while doing that we may loose the actual geometrical structure of the object. For instance, if some vertices are dropped out. Therefore, in order to preserve the this structure it is necessary to identify the minimal geometrical structure that characterize each one of the objects.

3.1 Vertex Detection

Let Q be a simple polygon and $V=\{v_0,v_1\ldots,v_{n-1};\mathrm{mod}(n)\}$ be its ordered sequence of vertices. A simple polygon can be minimally characterized by the set V. The edges of a simple polygon are straight line segments. Generalized polygons can also be characterized by its ordered sequence of vertices, however now the arc edges connecting two vertices must be specified.

To identify the minimum geometric structure of a generalized polygon P, from its digital image, one need to detect the pixels that form its ordered sequence of vertices. Let $S=\{I_1,I_2,\ldots,I_n\}$ be a digital image containing a set of generalized polygons. To get the pixels that form the border of a generalized polygon from its digital image I_k, first the digital image is segmented by threshold. In our case, we binarize the image as we are assuming that the objects are black on a white background. Once binarized, we apply the edge detection algorithm based on the standard second-order Laplacian operator, with mask $M=\begin{bmatrix}0 & -1 & 0\\ -1 & 4 & -1\\ 0 & -1 & 0\end{bmatrix}$, to find the list L containing the border pixels of all objects $I_k, k=1,\ldots,n$.

To obtain the ordered sequence of border pixels for each individual object a modified version of the contour following algorithm given in [Cos01], based on the 8-connected neighbor approach, is applied on the list L. The modified algorithm takes care of the chain-code direction in a clockwise manner to properly find the next contour pixel along the edge. After completion of this algorithm, each object has been labeled and their ordered sequence of pixels have been obtained.

To be able to get the ordered sequence of vertex pixels of an object, a corner detection algorithm shall be applied on the set of pixels that form each object. There are several algorithms in the literature for corner detection (see [Cos01] and references therein). Recently, Tsai et al. [Tsa99] proposed an algorithm, with an average running time complexity of $O(n)$, for corner detection that is simple to implement, robust to noise and sensible to identify both convex and nonconvex vertices. The algorithm relies in the analysis of the covariance matrix eigenvalues of a digital curve segment. It takes the sequence of border pixels of an object $P=\{(x_i,y_i), i=1,2,\ldots,n\}$, where the pixel p_{i+1} is the neighbor of $p_i, \mathrm{mod}(n)$ and (x_i,y_i) is the Cartesian coordinate of the pixel. A region of support around the pixel p_i is defined as $S_k(p_i)=\{p_i; j=i-k, i-k+1,\ldots,i+k-1\}$, where k is an integer number that defines the length of the support region. The covariance matrix is given by

$$C=\begin{pmatrix}c_{11} & c_{12}\\ c_{21} & c_{22}\end{pmatrix}, \tag{6}$$

where

$$c_{11}=\left(\frac{1}{2k+1}\sum_{j=i-k}^{i+k}x_j^2\right)-c_x^2, \tag{7}$$

$$c_{22}=\left(\frac{1}{2k+1}\sum_{j=i-k}^{i+k}y_j^2\right)-c_y^2, \tag{8}$$

$$c_{12}=c_{21}=\left(\frac{1}{2k+1}\sum_{j=i-k}^{i+k}x_j y_j\right)-c_x c_y, \tag{9}$$

$$c_x=\left(\frac{1}{2k+1}\sum_{j=i-k}^{i+k}x_j\right), \tag{10}$$

$$c_y=\left(\frac{1}{2k+1}\sum_{j=i-k}^{i+k}y_j\right). \tag{11}$$

The covariance matrix is Hermitian with real eigenvalues, which are given by:

$$\lambda_L=\frac{1}{2}\left(c_{11}+c_{22}+\sqrt{(c_{11}-c_{22})^2+4c_{12}{}^2}\right), \tag{12}$$

$$\lambda_S=\frac{1}{2}\left(c_{11}+c_{22}-\sqrt{(c_{11}-c_{22})^2+4c_{12}{}^2}\right). \tag{13}$$

The analysis of the eigenvalue λ_S shows that a corner is detected when its value is greater than a predefined threshold value. Each corner is separated by at least k pixels. It has been experimentally observed that pixels on a straight line have their λ_S value very close to zero and λ_S much greater correspond to a corner. After these observations we have chosen the region of support parameter with $k = 7$ and the threshold value for detecting a corner (vertex) as 10^{-1}. According to Tsai's paper, circular shapes would have $\frac{\lambda_S}{\lambda_L} \approx 1$. However, unfortunately, our experiments have shown that the algorithm is not reliable to identify circles as was claimed.

To characterize digital arc edges in generalized polygons, the eigenvalue λ_S is used to identify a minimum set of vertices in order to approximate the arc segment by a polygonal curve. Let $A = \{e_1 = v_1, e_2, \ldots, e_m = v_m\}$ be the set of vertices that characterize an arc edge of a generalized polygon P, then its geometric structure will be given by the ordered sequence of vertices $\{v_0, \ldots, v_k = e_k, e_{k+1}, \ldots, e_{k+m} = v_{k+m}, v_{k+m+1}, \ldots, v_{n-1}\}$.

3.2 Identifying Digital Circles

The geometric structure of a circle can be minimally characterized by its center $C = (x_0, y_0)$ and radius r. To recognize digital circles an algorithm was proposed by Sauer [Sau93] with linear time complexity $O(n)$, where n is the number of pixels of the digital function. In short, the algorithm is as follows: Let $C = (r, x_m, y_m)$ be a digital closed curve, where (x_m, y_m) is the centroid coordinate and r its radius mean value, then to the digital curve be a digital circle the following relation has to be fulfilled:

$$p_y - 0.5 - y_m \le \pm\sqrt{r^2 - (p_x - x_m)^2} \le p_y + 0.5 - y_m \tag{14}$$

where (p_x, p_y) are the pixel coordinates. Essentially, the identification of a circle translates to the computation of the centroid and the radius under a threshold value. The implementation of this algorithm was done to decide if the sequence of pixels of the border of an object form a digital circle.

The identification of the ordered sequence of vertices V of a generalized polygon, and of the center and radius of a circle, provide the minimal characterization of their geometrical structure and from that it allows its full reconstruction.

3.3 The Minimum Geometric Data Structure

Based on the geometric data structure extracted from the digital image of the objects, the AGVD can be constructed according to the algorithm provided in section 2.3. The geometrical data structure of the objects are given in the following format:

P (x,y)

S (x0,y0) (x1,y1)

C (xc,yc) r

A N (x1,y1)...(xn,yn)

L N (x1,y1)...(xn,yn),

where P specifies the Cartesian coordinate (x, y) of a single pixel (point); S corresponds to a straight line segment from pixel $p_0 = (x_0, y_0)$ to pixel $p_1 = (x_1, y_1)$; C gives the pixel coordinates $C = (x_c, y_c)$ of the center of a circle of radius r; A gives the sequence of n vertices from $p_1 = (x_1, y_1)$ to $p_n = (x_n, y_n)$ of the polygonal line that forms an arc segment; finally, L gives the sequence of n vertices that forms a generalized polygon.

Figure 5 exhibits, on the left side, the actual digital image of a set of generalized polygonal objects, captured by a CCD camera disposed on the top center of an arena, and on the right side the corresponding AGVD. There are 4 simple polygons, 2 generalized polygons and 1 circular object in the picture. The total number of pixels that form the border of all objects is $n = 1661$. The application of the algorithm to identify the minimal geometrical structure reduces this number of pixels (or points) to $m = 60$. The number of Voronoi generators to construct the AGVD, according to the current value of the resolution displacement parameter δ, is given by $p = 141$. We can see that by this approach only 8.4% of the total number of pixels were sufficient to construct the AGVD.

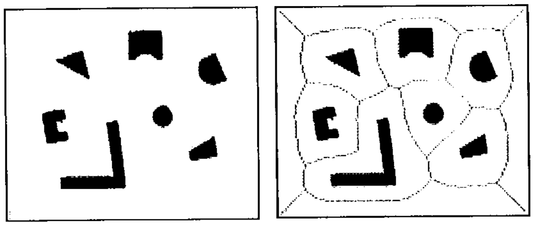


Figure 5: AGVD for a digital image with generalized polygons and a circular object.

Considering only the application of the algorithm for the AGVD construction base in the number of pixels, we can see that there would be a gain of 91.6% in computational time. Nevertheless, the overall computation cost to construct the AGVD from the digital image changes only after the computation of the total number of border pixels of the objects. In the full pixel approach the AGVD could be generated straightway from the set of all pixels that form the borders. On the other hand, in the reduced pixel approach, the AGVD construction is based on the identification of the minimal geometrical structure, which needs the application of the vertex detection algorithm of section 3.1, to reduce the number of pixels.

Full Pixel		Reduced Pixel	
# Pixels	Time(ms)	# Pixels	Time(ms)
260	2690	24	110
393	5600	33	270
559	10660	47	380
691	15600	55	490
763	18790	66	660

Table 1: The table illustrates the gain obtained by the reduced pixel approach over the full pixel approach.

Therefore, despite of both algorithms have a linear average running time complexity, the overall computational cost depends on the hidden constant factors, which for the AGVD using the minimal geometrical structure still provides a smaller average running time than just computing it considering the full pixel approach. This fact has been observed comparing the CPU time for the AGVD construction based on both approaches. Table 1 gives the time versus pixel number relation for both approaches. The linear fit of each approach shows that the AGVD construction based on the reduced pixel proposed here is better than the full pixel approach by an order of 38%. Figure 6 shows the graph of the linear fit curves taking into account only the slopes.

4. ROBOT APPLICATION

The generalized Voronoi diagram technique has been applied in many different areas. One of these applications can be seen in the field of mobile robotics, where the problem of collision-free path planning plays a central role for the robot safe navigation.

If we consider a global vision system composed of a single CCD camera posed on the top center of an arena, where the arena is seen as the workspace and the objects correspond to the obstacles, the AGVD provides a roadmap with maximal clearance from the obstacles. In addition, in a closed 2D workspace the AGVD is fully connected and any configuration (position and orientation) in the free configuration space C_{free} can be accessed by a robot navigating on the roadmap and then departing to reach the specified configuration. The Voronoi roadmap can be seen as a graph, therefore the shortest path between two configurations in C_{free} can be computed.

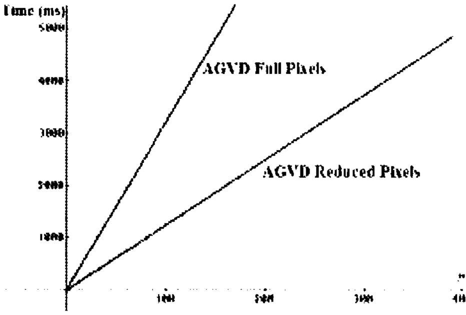


Figure 6: Linear fit curves for the AGVD construction based on the full pixels and reduced pixels approaches.

Considering a disk-like robot with radius r and taking into account that the Voronoi edges are the maximal clearance paths for the robot collision-free navigation, this will provide a natural threshold for the resolution displacement parameter as at most $\delta < 2r$. The Voronoi edges that are invalidated and omitted in the AGVD would, otherwise, lead to the robot collision with the obstacles.

Based on the approach described in the previous sections, a Global Vision module has been developed to provide a roadmap. In addition to this module, a Trajectory Planning module and a Navigation Control module have been developed and integrated in a mobile robot path planning system (see [Roq02] for further details).

5. CONCLUSION

In this paper we have shown that based on the geometrical structure that are extracted from the digital images of generalized polygons and circular objects, the construction of the approximate generalized Voronoi Diagram can be done with a running time complexity of $O(n)$, where n is the number of pixels defining the objects. The approach proposed here reduces the number of pixels improving the gain in computational time of the order of 38%, as pointed out in the previous section. In addition, this algorithm is also robust to numerical errors [Sug92]. The reduction in the number of pixels depends on the number of vertices representing the objects and on the resolution displacement parameter δ , whose value can be arbitrarily adjusted at the expense additional of computational cost.

An application of the reduced pixel approach has been done for robot path planning, where a global vision module was developed to capture the robot workspace image, identify the geometrical structure of the obstacles and the robot configuration, and finally generate the AGVD roadmap [Roq02].

6. ACKNOWLEDGMENTS

The authors thank Prof. Rudnei Dias da Costa for some comments on the paper, Bruno Alves for his help and the Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul – FAPERGS, for partial financial support.

7. REFERENCES

[Arc86] Arcelli, C. and Sanniti di Baja, G., Computing Voronoi Diagrams in Digital Pictures, Pattern Recognition Letters, 4, pp. 383-389, 1986.

[Aur91] Aurenhammer, F., Voronoi Diagrams: A Survey of a Fundamental Geometric Data Structure. ACM Computing Survey, 23, pp.345-405, 1991.

[Bor86] Borgefors, G., Distance Transformation in Digital Images, Computer Vision Graphics and Image Processing, 34, pp. 344-371, 1986.

[Cos01] Costa, L. F. and Cesar Jr., R. M., Shape Analysis and Classification: Theory and Practice, CRC Press, Boca Raton, USA, 2001.

[Mel92] Melkemi, M. and Chassery, J. M., Edge-Region Segmentation Process based on Generalized Voronoi Diagram Representation. IEEE International Conference on Pattern Recognition, pp. 323-326, 1992.

[Mel94] Melkemi, M. and Vandorpe, D., Fast Algorithm for Computing the Shape of a Set of Digital Points, IEEE International Conference in Image Processing, 1, pp. 705-709, 1994.

[Oka92] Okabe, A., Boots, B. and Sugihara, K., Spatial Tesselation Concepts and Applications of Voronoi Diagrams. John Wiley & Sons, New York, 1992.

[Par93] Parui, S. K., Sarkar, S. and Chaudhuri, B. B., Computing the Shape of a Point Set in Digital Images, Pattern Recognition Letters, 14, pp. 89-94, 1993.

[Roq02] Roque, W. L. and Doering, D., Multiple Visiting Stations Trajectory (Re)Planning for Mobile Lab Robots Based on Global Vision and Voronoi Roadmaps, Proceedings of the 33rd International Symposium on Robotics, Stockholm, Sweden, 2002.

[Sau93] Sauer, P., On the Recognition of Digital Circles in Linear Time. Computational Geometry: Theory and Applications, 2, pp. 287-302, 1993.

[Sud99] Sudha, N., Nandi, S. and Sridharan, K., A Parallel Algorithm to Construct Voronoi Diagram and Its VLSI Architecture, Proceedings of IEEE International Conference on Robotics and Automation, pp. 1683-1688, Detroit, US, 1999.

[Sug92] Sugihara, K. and Iri, M., Construction of the Voronoi Diagram for One Million Generators in Single-Precision Arithmetic. Proceedings of the IEEE, 80, pp. 1471-1484, 1992.

[Sug93] Sugihara, K., Approximation of Generalized Voronoi Diagrams by Ordinary Voronoi Diagrams. CVGIP: Graphical Models and Image Processing, 55, 522-531, 1993.

[Tsa99] Tsai, D.-M., Hou, H.-T. and Su, H.-J., Boundary-Based Corner Detection Using Eigenvalues of Covariance Matrices. Pattern Recognition Letters, 20, pp. 31-40, 1999.