

**Geometry-Adaptive
Multi-View Coding Techniques
for Image-based Rendering**

Der Technischen Fakultät der
Friedrich-Alexander-Universität Erlangen-Nürnberg
zur Erlangung des Grades

Doktor-Ingenieur

vorgelegt von

Marcus Andreas Magnor

Erlangen, 2000

Als Dissertation genehmigt von
der Technischen Fakultät der
Friedrich-Alexander-Universität
Erlangen-Nürnberg

Tag der Einreichung: 13. Oktober 2000

Tag der Promotion: 21. November 2000

Dekan: Prof. Dr.-Ing. Harald Meerkamm

Berichterstatter: Prof. Dr.-Ing. Bernd Girod

Prof. Dr. Günther Greiner

Berichte aus der Kommunikations- und Informationstechnik

Band 21

Marcus Magnor

**Geometry-Adaptive
Multi-View Coding Techniques
for Image-based Rendering**

D29 (Diss. Universität Erlangen-Nürnberg)

Shaker Verlag
Aachen 2001

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Magnor, Marcus:

Geometry-Adaptive Multi-View Coding Techniques for
Image-based Rendering/Marcus Magnor.

Aachen : Shaker, 2001

(Berichte aus der Kommunikations- und Informationstechnik; Bd. 21)

Zugl.: Erlangen-Nürnberg, Univ., Diss., 2000

ISBN 3-8265-8315-9

Copyright Shaker Verlag 2001

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the publishers.

Printed in Germany.

ISBN 3-8265-8315-9

ISSN 1432-489X

Shaker Verlag GmbH • P.O. BOX 1290 • D-52013 Aachen

Phone: 0049/2407/9596-0 • Telefax: 0049/2407/9596-9

Internet: www.shaker.de • eMail: info@shaker.de

Acknowledgments

The work at hand is the product of much more than just one person's labor. First of all, I wish to thank my supervisor Bernd Girod for pointing me towards a fascinating topic and for giving me the academic liberty to develop and pursue my own ideas in due course. Sincere thanks go to my office mate Matthias Teschner who was always there to help, in professional as well as personal matters. I am indebted to my colleague Peter Eisert whose scientific expertise and computer skills made possible much of my own work in the first place. Ursula Arnold will always be remembered for her cheerful personality and for taking the burden of bureaucracy off the academic mind. All colleagues at the Telecommunications Institute in Erlangen may be thanked for the congenial and companionable atmosphere I had the pleasure to enjoy for three years. Above all, however, I wish to express my deepest-felt gratitude for my parents whose loving support constitutes the *conditio sine qua non* of my thesis.

Contents

Notation	xi
Abbreviations and Acronyms	xii
1 Introduction	1
2 Background and Related Work	5
2.1 The Plenoptic Function	6
2.1.1 Parameterization	6
2.1.2 Light Field Rendering	8
2.1.3 Sampling Criterion	10
2.1.4 Disparity Compensation	14
2.2 Image-Based Rendering	16
2.2.1 Viewpoint-Centered IBR Methods	16
2.2.2 Object-Centered IBR Methods	17
2.2.3 Coding Techniques in IBR	19
2.3 Compression Concepts	20
2.3.1 Intra-Image Redundancy Reduction	21
2.3.2 Inter-Image Redundancy Reduction	23
2.3.3 Stereo and Multi-View Coding	24
2.3.4 Geometry Compression	25
3 Block-Adaptive Coder	29
3.1 Video Coding Concepts	29
3.1.1 Block-based Coding	30
3.1.2 Block Coding Mode Selection	31
3.1.3 Residual Error Coding	31
3.2 Adaptation to Light-Field Compression	32
3.2.1 Disparity-Compensated Prediction	32
3.2.2 Multi-Reference Prediction	32
3.2.3 Disparity Compensation Accuracy	33
3.3 Block-Adaptive Coder Design	35

3.4	Performance Evaluation	37
3.4.1	Rate-Distortion Performance	37
3.4.2	Reconstruction Quality	40
3.4.3	Decoding and Rendering	41
4	Hierarchical Disparity-Compensated Coder	43
4.1	Disparity Estimation and Compensation	43
4.1.1	Disparity Map Generation	44
4.1.2	Distant Disparity Compensation	45
4.2	Multi-Directional Disparity Compensation	45
4.2.1	BACKWARD compensation	47
4.2.2	FORWARD compensation	47
4.2.3	FORWARD-BACKWARD compensation	48
4.3	Prediction Performance	49
4.3.1	Image Prediction Modes	49
4.3.2	Disparity-Map Block Size	50
4.3.3	Disparity Accuracy	51
4.3.4	Overlapped Block Disparity Compensation	52
4.3.5	Summary	53
4.4	Hierarchical Image Coding Order	53
4.4.1	Coding Order for Two-Plane Parameterized Light Fields	54
4.4.2	Coding Order for Arbitrarily Arranged Light-Field Images	55
4.5	Performance Evaluation	56
4.5.1	Rate-Distortion Performance	57
4.5.2	Reconstruction Quality	59
4.5.3	Decoding and Rendering	60
5	Model-Aided Coder	63
5.1	Scene Geometry Reconstruction	63
5.2	Geometry Compression	65
5.2.1	Embedded Mesh Coding	66
5.2.2	EMC Coding Performance	68
5.3	Model-Aided Prediction	71
5.4	Performance Evaluation	72
5.4.1	Rate-Distortion Performance	73
5.4.2	Reconstruction Quality	74
5.4.3	Decoding and Rendering	74
6	Progressive Texture Coder	77
6.1	Texture Map Generation	77
6.1.1	Model Approximation	78

6.1.2	Texture Map Optimization	79
6.1.3	Image Conversion	81
6.2	Texture Coding	82
6.2.1	Sparse Texture-Map Interpolation	82
6.2.2	4-D Subband Decomposition	85
6.3	Progressive Texture Coding	86
6.4	Performance Evaluation	86
6.4.1	Rate-Distortion Performance	87
6.4.2	Reconstruction Quality	89
6.4.3	Decoding and Rendering	89
7	Coder Comparison	91
7.1	Overall Coding Efficiency	91
7.2	Inter-Image Similarity	94
7.3	Model-Aided and Progressive-Texture Coding	95
7.3.1	Geometry Accuracy	99
7.3.2	Calibration Accuracy	100
7.3.3	Variable Texture	101
7.4	Summary	102
8	Conclusions	103
A	Appendix	105
A.1	Calibrated Image Acquisition	105
A.2	Multi-View Image Sets	106
B	Titel, Inhaltsverzeichnis, Einleitung, Zusammenfassung	111
B.1	Titel	111
B.2	Inhaltsverzeichnis	111
B.3	Einleitung	114
B.4	Zusammenfassung	117
	References	119

Notation

α	angular distance
$\Delta s, \Delta t$	pixel size in horizontal and vertical direction
Δz	difference in scene-depth
λ	wavelength, Lagrangian multiplier
Φ	plenoptic function
$\hat{\Phi}$	discretized plenoptic function
θ, ϕ	polar and azimuthal angles
$A^{\text{geo,txt}}$	geometry, texture-map triangle area
a	camera lens aperture
b	imaging distance
D	image distortion
d	disparity
d_S	disparity along S direction
d_T	disparity along T direction
f	focal length
g	distance camera to object center
$H_{0,1}^{(0,1)}$	1-D Haar wavelet coefficient
$I_i^{(j)}$	4-D Haar wavelet coefficient
I	discrete image
\hat{I}	predicted image
$I(x, y)$	intensity value of image I at position (x, y)
N_S, N_T	number of pixels in S and T direction
N_U, N_V	number of images in U and V direction
N_Z	number of scene-depth levels
Q	quantizer step size
R	coding bit-rate
$R^{\text{geo,txt}}$	relative geometry, texture-map triangle area
r	radius
S, T	horizontal and vertical image-plane directions
s, t	horizontal and vertical image-plane coordinates
U, V	horizontal and vertical recording-plane directions
u, v	horizontal and vertical recording-plane coordinates
x, y, z	Cartesian spatial coordinates

Abbreviations and Acronyms

2-D	Two-Dimensional
3-D	Three-Dimensional
4-D	Four-Dimensional
BAC	Block-Adaptive Coder
bpp	bits per pixels
dB	decibel
DCT	Discrete Cosine Transform
EMC	Embedded Mesh Coding
GBR	Geometry-Based Rendering
H.263	Video coding standard of the ITU for low bit-rates
HDC	Hierarchical Disparity-Compensated Coder
HVS	Human Visual System
IBR	Image-Based Rendering
ITU	International Telecommunications Union
JPEG	Joint Picture Experts Group
LFR	Light Field Rendering
MAC	Model-Aided Coder
MCP	Motion-Compensated Prediction
MPEG-1,2,4	Video coding standards of the Motion Picture Experts Group
MSE	Mean Squared Error
PSNR	Peak-Signal-to-Noise Ratio
PTC	Progressive Texture Coder
RD	Rate-Distortion
RGB	Red-Green-Blue color space
SLF	Surface Light Field
SPIHT	Set-Partitioning-In-Hierarchical-Trees still-image coder
SSE	Summed Squared Error
VDTM	View-Dependent Texture Mapping
YCbCr	Luminance-Chrominance color space