Multi-Frame Motion-Compensated Prediction for Video Transmission

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

Doktor-Ingenieur

vorgelegt von

Thomas Wiegand

Erlangen, 2000

Als Dissertation genehmigt von
der Technischen Fakultät der
Universität Erlangen–Nürnberg

Tag der Einreichung:	13. Juni 2000
Tag der Promotion:	7. Juli 2000
Dekan:	Prof. Dr. Harald Meerkamm
Berichterstatter:	Prof. Dr.-Ing. Bernd Girod
	Prof. Dr.-Ing. Heinz Gerhäuser

Berichte aus der Kommunikations- und Informationstechnik

Band 18

Thomas Wiegand

Multi-Frame Motion-Compensated Prediction for Video Transmission

D 29 (Diss. Universität Erlangen-Nürnberg)

Shaker Verlag
Aachen 2000

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Wiegand, Thomas:
Multi-Frame Motion-Compensated Prediction for Video Transmission /
Thomas Wiegand. Aachen : Shaker, 2000
(Berichte aus der Kommunikations- und Informationstechnik ; Bd. 18)
Zugl.: Erlangen-Nürnberg, Univ., Diss., 2000
ISBN 3-8265-7764-7

Printed in Germany.

ISBN 3-8265-7764-7
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

To Ela.

Acknowledgments

I would like to thank Prof. Bernd Girod for giving me the opportunity to work in his group and for providing stimulating ideas during my doctoral studies. I would also like to thank Prof. Heinz Gerhäuser for his interest in my work and for reviewing my thesis. Furthermore, many thanks to Prof. Johannes Huber for his support.

During my doctoral studies, I had the pleasure to interact with a large number of excellent researchers and skilled engineers. I am very thankful to my colleagues and friends Peter Eisert, Joachim Eggers, Niko Färber, Markus Flierl, Uwe Horn, Eckehard Steinbach, Klaus Stuhlmüller, and Jonathan Su for the many stimulating discussions, joint work, and proof-reading. I would also like to thank Barry Andrews, Prof. Robert Gray, Paul Ning, and Gary Sullivan for their support.

I would like to thank the many others around the Telecommunications Laboratory in Erlangen who made my stay here so enjoyable. In addition I would like to thank Ursula Arnold for her invaluable support.

This thesis is dedicated to my wife, Gabriela, who gave me the strength and piece of mind to pursue this project.

Contents

Glossary

Notation

Operators

$\mathrm{d}\cdot/\mathrm{d}x$	derivation of function with respect to x
$E\{\cdot\}$	expectation

Basic Definitions

b	bit-stream
b'	received bit-stream
x, y	discrete pixel positions in the image
t	discrete temporal image index
$\boldsymbol{l}$	vector of the pixel position
$\boldsymbol{\mathcal{A}}_k$	set of pixel positions
$\boldsymbol{S}_k$	block of pixels
$s[\boldsymbol{l}]$	luminance component of the pixel
$s_{Cb}[\boldsymbol{l}], s_{Cr}[\boldsymbol{l}]$	chrominance component pixel
$\boldsymbol{s}[\boldsymbol{l}]$	vector valued color pixel
$s'[\boldsymbol{l}]$	decoded luminance pixel
$\hat{s}[\boldsymbol{l}]$	motion-compensated luminance pixel
$u[\boldsymbol{l}]$	motion-compensated prediction error luminance pixel
$u'[\boldsymbol{l}]$	decoded version of $u[\boldsymbol{l}]$
$v[\boldsymbol{l}]$	transmission error luminance pixel
m_x, m_y	spatial displacements in horizontal and vertical direction, respectively
m_t	picture reference parameter
$\boldsymbol{m}$	motion vector
J_{MODE}	Lagrangian cost function for macroblock mode decision
D_{REC}	distortion term for the macroblock after reconstruction
R_{REC}	bit-rate associated to the macroblock
λ_{MODE}	Lagrange parameter for macroblock mode decision
J_{MOTION}	Lagrangian cost function for motion estimation

D_{DFD}	distortion term corresponding to motion compensation error
R_{MOTION}	bit-rate associated to the motion vector
$\lambda_{\mathrm{MOTION}}$	Lagrange parameter for motion estimation

Optimization Using Lagrangian Techniques

$\boldsymbol{S}_k$	source sample
$\boldsymbol{\mathcal{S}}$	set of source samples
O_{kl}	coding option
$\boldsymbol{\mathcal{O}}_k$	set of coding options
I_k	chosen coding option
$\boldsymbol{\mathcal{I}}$	set of chosen coding options
$D(\boldsymbol{\mathcal{S}},\boldsymbol{\mathcal{I}})$	distortion when quantizing $\boldsymbol{\mathcal{S}}$ using $\boldsymbol{\mathcal{I}}$
$R(\boldsymbol{\mathcal{I}})$	bit-rate associated to the set of chosen coding option $\boldsymbol{\mathcal{I}}$
R_c	rate constraint
λ	Lagrange parameter
$J(\boldsymbol{\mathcal{S}},\boldsymbol{\mathcal{I}}\vert\lambda)$	Lagrangian cost function

Comparison to Other Encoding Strategies

C_{INTRA}	cost value for the INTRA macroblock mode
$\mu_{\boldsymbol{\mathcal{A}}_k}$	mean value of the pixels in the area
C_{INTER}	cost value for the INTER mode
$\boldsymbol{m}_k^F$	resulting motion vector from integer-pixel motion search
ξ	bias value to give a preference towards choosing the SKIP mode
$C_{\mathrm{INTER+4V},l}$	cost value for the INTER+4V mode after searching the set

Statistical Model for the Prediction Gain

m	reference frame index
D_m	minimal MSE value for a block when referencing frame m
L_m	logarithmic distortion corresponding to D_m
$\mathcal{X}$	random variable associated to distortion
$\mathcal{Y}$	random variable associated to distortion minimization
$\mathcal{X}^M$	M-tuple of random variables
$f_{\mathcal{X}^M}$	probability density function of $\mathcal{X}^M$
$\boldsymbol{C}$	covariance matrix
$c_{n,m}$	element of the covariance value

$\boldsymbol{\mu}$	vector of mean values
μ	mean value
σ^2	variance
ρ	correlation factor
Δ	difference between expected values
$\mathcal{N}$	jointly Gaussian probability density function
K	number of dyadic minimizations
α	factor

Affine Multi-Frame Motion-Compensated Prediction

M	overall number of reference pictures at the encoder
K	number of decoded reference pictures
N	number of warped pictures at the encoder
M^*	chosen number of reference pictures
$\boldsymbol{a}$	affine motion parameter vector
a_i	affine motion parameter
c_i	coefficients of the orthonormalized affine motion model
Δ	quantizer step size for the affine motion parameters
L	number of initial translational motion vectors for a cluster
g_x, g_y	image intensity gradients in x and y direction
x', y'	shifted discrete spatial positions in the image
$\boldsymbol{A}^z, \boldsymbol{B}^z$	matrices for gradient approximation

Fast Motion Estimation for Multi-Frame Prediction

D'_{DFD}	approximate distortion term
p	norm order
Θ	threshold for search space reduction
Ω	factor for early termination of the motion search
Ξ	threshold for sub-sampling of the block
Γ	ratio between two values of average computation time

Error Resilient Video Transmission

p	probability of packet loss
q	probability that packet is correctly received
s'_l	the lth decoded version of the transmitted video frame at the decoder

p_l	probability of the occurrence of s_l' at the decoder
D_{DFD}^l	distortion for the MCP error when referencing s_l'
$\mathcal{D}_{\mathrm{DFD}}$	random variable for distortion for the MCP error at the decoder
D_{REC}^l	distortion for the reconstruction error when referencing s_l'
$\mathcal{D}_{\mathrm{REC}}$	random variable for distortion for the reconstruction error at the decoder
D_{ERR}	error modeling term approximating the expected transmission error distortion
κ	weight of D_{ERR} when incorporated into Lagrangian coder control
Δt	time interval between the current and a past reference frame
f_D	Doppler frequency
f_c	carrier frequency
R_t	total transmission bit-rate
T_s	modulation symbol interval
m	number of bits per symbol
N	number of symbols per block
K	number of information symbols per block
$N_{\max}$	maximum number of symbols per block
r	code rate

Abbreviations and Acronyms

ACK	acknowledgment message
AWGN	additive white Gaussian noise
CIF	Common Intermediate Format
CR	conditional replenishment
DCT	discrete cosine transform
DECT	Digital Enhanced Cordless Telecommunications
DFD	displaced frame difference
ECVQ	entropy-constrained vector quantization
ETSI	European Telecommunications Standards Institute
FD	frame difference
FEC	forward error correction
FS	full search
GOB	group of blocks
GSM	Global System for Mobile Communications
HP-MC	half-pixel motion compensation
IP-MC	integer-pixel motion compensation
ITU-T	Intern. Telecommunications Union - Telecommunications Standardization Sector
MCP	motion-compensated prediction
MFP	multi-frame prediction
NACK	negative acknowledgment message
NOS	norm-ordered search
PDF	probability density function
POS	probability-ordered search
PSNR	peak signal-to-noise ratio
QCIF	Quarter Common Intermediate Format
RS	Reed-Solomon (code)
RWER	residual word error rate
MSE	mean squared error
SAD	sum of absolute differences
SSD	sum of squared differences
SSR	search space reduction
TMN-9	test model near-term, version 9
TMN-10	test model near-term, version 10