Modeling and Optimization of Video Transmission Systems

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

Doktor-Ingenieur

vorgelegt von

Klaus Stuhlmüller

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.-Ing. Harald Meerkamm
Berichterstatter:	Prof. Dr.-Ing. Bernd Girod
	Prof. Dr.-Ing. Johannes Huber

Berichte aus der Kommunikations- und Informationstechnik

Band 17

Klaus Stuhlmüller

Modeling and Optimization of Video Transmission Systems

D 29 (Diss. Universität Erlangen-Nürnberg)

Shaker Verlag
Aachen 2000

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Stuhlmüller, Klaus:
Modeling and Optimization of Video Transmission Systems /
Klaus Stuhlmüller. Aachen : Shaker, 2000
(Berichte aus der Kommunikations- und Informationstechnik ; Bd. 17)
Zugl.: Erlangen-Nürnberg, Univ., Diss., 2000
ISBN 3-8265-7763-9

Printed in Germany.

ISBN 3-8265-7763-9
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

I would like to thank Prof. Dr.-Ing. Bernd Girod for giving me the opportunity to work in his group in a stimulative and supportive atmosphere. I would also like to thank Prof. Dr.-Ing. Johannes Huber for his support and for spending his leisure time in reading my thesis and making it readable for channel coding experts.

I am very grateful to my colleague Niko Färber. His comments were very precious and large parts of this thesis are the result of joint work with him.

Many thanks to Alexander Stenger for reviewing this thesis, Thomas Wiegand for his never-failing flood of ideas and helpful comments, Michael Link for contributing the *lower transmission layers*, Uwe Horn for contributing the scalable video codec and to all colleagues at the Telecommunications Laboratory of the University Erlangen-Nuremberg for the cooperation and the excellent atmosphere that made work both interesting and fun.

Special thanks to Alexandra for her comprehension and for giving me the *licence to work*.

Contents

Glossary

Acronyms

ADSL	asymmetric digital subscriber line
AMR	adaptive multi rate
ARQ	automatic repeat request
BCH	Bose-Chaudhuri-Hocquenghem
BL	base layer
block	8×8 (luminance) pixel
BOP	block of packets
CIF	common intermediate format, 352×288 pixel
dB	decibel
DCT	discrete cosine transform
DDF	distortion-distortion function
DFT	discrete Fourier transform
DMC	discrete memoryless channel
DPCM	differential pulse code modulation
DR	distortion-rate
EDGE	enhanced data rates for GSM evolution
EL	enhancement layer
EREC	error resilient entropy coding
ET	error tracking
FEC	forward error correction
fps	frames per second
GByte	giga byte, 1024 MByte
GOB	group of blocks
GOP	group of pictures
GSM	Global System for Mobile communications

IETF	Internet Engineering Task Force
Inter	encoding with reference to previous frame
Intra	encoding without reference to previous frame
ISDN	Integrated Services Digital Network
ISO	International Standardization Organization
ITU-T	Int. Telecom. Union - Telecom. Standardization Sector
i.i.d.	independent identically distributed
kbps	1000 bit/second
kByte	kilo byte, 1024 byte
macroblock	16×16 (luminance) pixel
Mbps	1000 kbps
MByte	mega byte, 1024 kByte
MC	motion compensation (motion compensated prediction)
MSE	mean squared error
OBMC	overlapped block motion compensation
OSI	open-system-interconnection
PACET	pixel-accurate error tracking
PSD	power spectral density
PSNR	peak-signal-to-noise ratio
Q	quantization
QCIF	quarter common intermediate format, 176×144 pixel
RD	rate-distortion
RS	Reed-Solomon
RTP	real time protocol
SCUEP	scalable video transmission using unequal error protection
SNR	signal-to-noise ratio
SPIHT	set partitioning in hierarchical trees
SSE	summed squared error (for a macroblock)
UEP	unequal error protection
UMTS	universal mobile-telecommunication system
VLC	variable length code
YCrCb	color-space for digital images

Sets

$\mathcal{A}$	pixels in a certain macroblock, frame, etc.

Operators

$\lvert\cdot\rvert$	magnitude of scalar
$d\cdot/dx$	derivation of function with respect to x
$E\{\cdot\}$	expectation
ld	*logarithmus dualis*, logarithm to the base 2
$\Pr(\cdot)$	probability

Variables

Video Encoder

β	Intra rate
β^*	optimal Intra rate
β_{eff}	effective (measured) Intra rate
D_e	distortion at output of video encoder
$D_{\mathrm{e},measure}$	distortion D_e achieved by measurement
$D_{\mathrm{e},model}$	distortion D_e computed by the model
$D_{\mathrm{e}P}$	D_e for Inter coding at a fixed rate
ΔD_{eIP}	difference between D_e for Intra and Inter coding
D_{Gauss}	distortion of a Gaussian source
D_0	distortion offset of DR function
D_{0P}	D_0 for Inter coding
ΔD_{0IP}	difference between D_0 for Intra and Inter coding
J	Lagrangian cost function for mode decision
λ	Lagrangian multiplier for mode decision
$PSNR_\mathrm{e}$	PSNR at output of video encoder
QP	quantizer parameter
QP_0	quantizer parameter of present frame
QP_{ref}	quantizer parameter of reference frame
R	data rate
R_e	data rate at output of video encoder
R_0	rate offset of DR function
R_{0P}	R_0 for Inter coding
ΔR_{0IP}	difference between R_0 for Intra and Inter coding
θ	multiplier of DR function
θ_P	θ for Inter coding
$\Delta\theta_{IP}$	difference between θ for Intra and Inter coding

Global Video Decoder Model

$\alpha[t]$	power transfer after t time steps
D_{d}	distortion at output of video decoder
D_{v}	distortion at video decoder caused by transmission errors
$d[x, y; t]$	decoded video sequence
$e[x, y; t]$	encoded video sequence
$F(\omega_x, \omega_y)$	frequency response, Fourier transform of $f[x, y]$
$F_1(\omega)$	frequency response, Fourier transform of $f_1[k]$
$f[x, y]$	impulse response of loop filter
$f_1[k]$	impulse response of 1-D loop filter
γ	leakage
$H_t(\omega_x, \omega_y)$	Fourier transform of $h_t[x, y]$
$h_t[x, y]$	impulse response of video decoder after t time steps
	(relates introduced error $u[x, y]$ at $t = 0$ to propagated error $v[x, y; t]$)
$o[x, y; t]$	input (original) video sequence
ω	spatial frequencies
ω_x, ω_y	spatial frequencies
$PSNR_{\mathrm{d}}$	PSNR at output of video decoder
$PSNR^*_{\mathrm{d}}$	optimal PSNR at output of video decoder
$\Phi_{uu}(\omega_x, \omega_y)$	PSD of $u[x, y]$
σ_f^2	filter strength
σ_g^2	spectral shape of introduced error
σ_u^2	energy of introduced error
$\sigma_{u_0}^2$	error sensitivity
$\sigma_v^2[t]$	energy of propagated error after t time steps
T_β	Intra update interval
t	temporal index, frame number
U	process introducing errors in decoder
$u[x, y]$	error introduced in decoder after error concealment
$u[x, y; t]$	error introduced in decoder after error concealment at time t
$v[x, y; t]$	difference between encoded and decoded sequence
x	horizontal pixel position
y	vertical pixel position

Pixel Accurate Error Tracking

$d_{c,\max}$	maximal correlation distance
h	frame height (number of pixels per column)
μ	multiplier for transmission errors in mode decision
ρ	correlation coefficient between encoding and transmission error
$\sigma_d^2[x,y;t]$	distortion at decoder at pixel $(x,y;t)$
$\sigma_e[x,y;t]$	encoding error at pixel $(x,y;t)$
$\sigma_e^2[x,y;t]$	distortion at encoder at pixel $(x,y;t)$
w	frame width (number of pixels per line)

Error Control Channel

B	state 'bad' of Markov model
G	state 'good' of Markov model
$G(\nu)$	gap distribution function
$g(\nu)$	gap density function
i	FEC block number
k	number of information symbols per block
κ	number of symbol errors per block
L_{B}	average burst length
$L_{\mathrm{B},DMC}$	average burst length of discrete memoryless channel
$L_{\mathrm{B},i}$	average burst length after interleaving
$\tilde{L}_{\mathrm{B}}$	normalized average burst length ($L_{\mathrm{B}}/L_{\mathrm{B},DMC}$)
m	bits per symbol
n	number of symbols per block (after shortening)
n_{max}	maximum number of symbols per block
ν	gap length
P_{B}	symbol error rate
$P_{\mathrm{B},i}$	symbol error rate after interleaving
P_{D}	block error density function
$P_{\mathrm{D,DMC}}$	P_{D} of a discrete memoryless channel
P_{L}	loss (erasure) rate
P_{R}	residual word error rate after channel decoding
p_{BG}	transition probability B to G in Markov model

p_{GB}	transition probability G to B in Markov model
$\mathbf{\Psi}$	one step transition matrix
$\mathbf{\Psi}_i$	i step transition matrix
R_{c}	raw channel data rate, transmitted data bits per time
$R(n,\kappa)$	probability of $\kappa-1$ symbol errors following an error
r	relative code rate
r^*	optimal relative code rate
t_c	number of symbol errors that can be corrected
$t_{c,erasure}$	number of symbol errors that can be interpolated

Scalable transmission

AMF	auto-MSE-function
β_{BL}	base layer Intra rate
β_{EL}	enhancement layer Intra rate
C	capacity of erasure channel
D_{BL}	base layer distortion
D_C	least distortion of Gaussian source achievable over erasure channel
$D_{loss}(n,\kappa)$	distortion for κ out of n packets lost
D_S	distortion of Gaussian source achievable with R_S
D_{SCUEP}	overall distortion at decoder for SCUEP
D_{single}	overall distortion at decoder for single layer transmission
$D_{singlew/oFEC}$	overall distortion at decoder for single layer transmission without using FEC
$D_{\mathrm{v},GOP}$	average D_{v} within one group of pictures
$D_{\mathrm{v},I}$	average D_{v} for lost Intra frames
Δ_k	determinant of quadratic equation
ϵ_w	width of supported ranges of the bi-modal weighting function
k_I	number of info. symbols per block for *Info* field
k_{BL}	number of information symbols per block for base layer
k_{EL}	number of information symbols per block for enhancement layer
k_ξ	number of information symbols per block for layer ξ
L_{AMF}	stored length of *AMF*
l	packet length

l_{BL}	number of bytes per packet occupied by the base layer
l_{EL}	number of bytes per packet occupied by the enhancement layer
l_{ξ}	number of bytes per packet occupied by layer ξ
MSE_{BL}	base layer MSE
MSE_{EL}	enhancement layer MSE
N_{BOP}	number of symbols per block of packets (BOP)
P_{c}	critical loss rate
$P_{\mathrm{c,BL}}$	critical loss rate for base layer
$P_{\mathrm{c,EL}}$	critical loss rate for enhancement layer
$P_{\mathrm{R}\,P_{\mathrm{c}}}$	residual loss rate at channel loss rate P_{c}
$PSNR_{BL}$	base layer PSNR
$PSNR_{EL}$	enhancement layer PSNR
$PSNR_{CIF}$	PSNR measured in CIF format
$PSNR_{QCIF}$	PSNR measured in QCIF format
$PSNR_{int}$	PSNR integrated over a range of loss rates
$PSNR_{sjt}$	PSNR evaluated by informal subjective tests
R_{BL}	base layer data rate
R_{BL0}	R_{BL} to which enhancement layer model is normalized
R_{EL}	enhancement layer data rate
$R_{header\mathrm{EL}}$	header data rate of enhancement layer
R_S	source data rate of Gaussian source
$R_{0P\,R_{\mathrm{BL\,0}}}$	R_{0P} for R_{BL0}
$\Delta R_{0P\,R_{\mathrm{BL}}}$	difference in R_{0P}
$\Delta R_{0IP\,R_{\mathrm{BL\,0}}}$	ΔR_{0IP} for R_{BL0}
$\Delta\Delta R_{0IP\,R_{\mathrm{BL}}}$	difference in ΔR_{0IP}
s	scaling factor for lattice quantization
σ^2	variance, energy of Gaussian source
τ	temporal distance
ϑ	percentage of bits for base layer
$w()$	weighting function
w_0	value of weighting function at $P_{\mathrm{L}} = 0$
w_c	value of weighting function at $P_{\mathrm{L}} = P_{\mathrm{c}}$
ξ	layer number
ζ	distribution of FEC between layers