Advanced Wireline/Wireless Visual Communications

Xinhua Zhuang

Multimedia Communications and Visualization Laboratory
Dept. of Computer Engineering and Computer Science
University of Missouri-Columbia
Columbia, MO 65211, USA
http://meru.cecs.missouri.edu
Presentation Overview

- Multimedia Communications and Visualization Laboratory (MCVL)
- Videoconferencing over Ethernet
- 56K Modem Accessible Distance Learning
- SLCCA Family of Wavelet Image/Video Codecs
- Joint Source-Channel Coding
- Wireless Video Communication System
- Conclusions
Facilities

- State-of-the-art $1.6 million facility funded from external sources:
  - Hardware $700K NASA
  - Hardware $400K SGI
  - Hardware $200K NSF
  - Software $300K MU-SGI Varsity Program
- Established Sept. 1997
- Personnel - 6 faculty, 30 graduate students
Xinhua Zhuang, Director, Professor, Peking University
Kannappan Palaniappan, Associate Professor, Ph.D., University of Illinois at Urbana-Champaign
Yunxin Zhao, Associate Professor, Ph.D., University of Washington at Seattle
C. Shyu, Assistant professor, Ph.D., Purdue University
M. Jurczyk, Assistant professor, Ph.D., Stuttgart University, Germany
D. Ho, Assistant professor, Ph.D., Chinese University of Hong Kong
Alumni

Students hired by:
- Cisco
- David Sarnoff
- Microsoft
- IBM
- Intel
- Lucent
- Texas Instruments

In 1999, Dr. Yan Huang was promoted as Manager of Bell Lab in China
Hardware Equipment

- Video and image server
  - Onyx2, 6 x R10000 CPUs
  - Mass storage 400 GB 6-way striped RAID array

- Multiprocessor systems
  - Four Octanes (SSI and MXI), 2 x R10000 CPUs
  - Real-time Compression and Digital Video boards

- Visualization cluster
  - Eight O2 workstations (WebForce and Video)

- High-end multimedia PCs

- High speed networking
  - 100 BaseT, ATM,
  - Gigabit research network
Current Research Topics

- Video over IP (scalable, error resilient IP video streaming)
- ATM wireless video (scalable, error resilient low complexity ATM video streaming)
- Software-only multicast-capable interactive videoconferencing over Ethernet
- 56k modem accessible Web broadcast for distance learning
- Web conferencing
- Scalable, error resilient SLCCA+ image codec beyond JPEG2000
- Automatic spatio-temporal video object segmentation
- Prototype client-server system for Web-based biomedical text and image retrieval over Internet
- Intelligent satellite hyper-spectral transmission over wireless ATM
Current Research Topics

- Enhanced multimedia telehealth for hearing disabilities
- Intelligent virtual endoscopy simulator for performing remote robotic-assisted minimally invasive medical procedures
- Gesture-controlled panoramic map browser
- Secured Web-based negotiation support system for e-commerce
- Content-based lung image database management and search using statistical similarity and visualization tools
- Adaptive multi-user CDMA receivers
- Single-user channel estimation and equalization
- Network traffic control and performance evaluation
- Spoken language processing, noise cancellation and multiuser separation and detection
- Adaptive co-channel speech separation for design of listening assistive devices
Current Research Topics

- Remote sensing land cover and land classification
- Distributed large data set visualization using Interactive Image SpreadSheet (IISS) tool
- Data mining using robust clustering and 3D visualization
- Terrain and cloud surface visualization for visual simulation (FlyBy)
- Fast geometric modeling, rendering, and scene-based visualization of video data
Videoconferencing over Ethernet

- The Local Area Network is fast becoming a viable network choice for delivering organizational videoconferencing system.
- The future of networked conferencing lies with the Internet.
- Software-only implementation of videoconferencing systems on personal computers is now feasible with increasingly available computing powers and innovative high-efficiency, low-complexity video codec.

  - Advantages:
    - Cost-effective
    - Easily deployed
    - Easily upgradeable
Videoconferencing over Ethernet

Main features of MCVL videoconferencing tool
- Support standard videoconferencing frame sizes (QCIF, CIF) and more (e.g., VGA)
- Support real-time interactive video at 15 fps
- Support synchronized audio/video communication
- Coding delay is well controlled under 150 ms
- High picture quality to enable lip-reading
- Support unicast and multicast
- Efficient bandwidth usage
  - For QCIF format, the total bandwidth usage is controlled under 350 kbps
  - For CIF format, the total bandwidth usage is controlled under 600 kbps
Distance learning (DL) is the acquisition of skills and knowledge through electronic communications, which allow the student and instructor to be separate in either time or space, or both.

The market capital for DL is projected to reach more than 100 billion US dollars by 2005.

Two categories of DL
- Asynchronous (e.g., Web-based course server)
- Synchronous (e.g., live broadcasting, real-time interactive communication)

Being able to access course materials through 56 kbps connections provides high flexibility in deploying a DL system.
MCVL provides Web-based video broadcasting technology for high quality, low bit rate video transmission which outperforms ITU-T standard H.263+.

- Software only implementation
- Low-medium motion video sequences (typical in DL):
  - 5 fps at 24 kbps, or 10 fps at 48 kbps
  - PSNR = 28-32 dB for the luminance component, and 36-40 dB for the chrominance components.
- Decoding time:
  - Averaging 35 ms per frame on 195 MHz R10000 CPU (comparable to 350 MHz Pentium II).

MCVL interactive video conferencing tool is being adapted to provide the technological backbone for real-time interactive DL systems.
SLCCA Image Coding

The tremendous success of advanced wavelet image coding:
- EZW: Embedded Zerotree Wavelet (Shapiro, 1993)
- MRWD: Morphological Representation of Wavelet Data (Servetto, Ramchandran, and Orchard, 1995)
- SPIHT: Set Partitioning in Hierarchical Trees (Said and Pearlman, 1996)
- SLCCA: Significance-Linked Connected Component Analysis (Chai, Vass, and Zhuang, 1997)

Attributed to innovative data organization and representation strategies
Main features of SLCCA include:

- Biorthogonal wavelet decomposition forming a pyramid of subbands
- Exploiting within subband clustering property of wavelet transform by *connected components* of significant fields
- Exploiting cross-scale dependency by *significance-linkage* among cross-scale clusters
- Bit-plane encoding of significance map and significant magnitudes with scale/space-variant higher order Markov source modeling
SLCCA Image Coding

- The significant coefficients (coefficient with a large magnitude) within a subband are more geometrically clustered than random patterns.

- Organizing and representing significant coefficients in connected components or clusters avoids the encoding of a large number of insignificant coefficients.
Encoded significance map. Black and gray pixels are encoded significant and insignificant coefficients in clusters, respectively.
If a child coefficient is significant, it is likely that its parent is also significant.

By marking a parent as having a *significance-link*, the cost of encoding positional information of its child cluster is reduced.
SLCCA Image Coding

- Quantization of magnitudes of significant coefficients with a uniform scalar quantizer
- Organizing and representing magnitudes in bit-planes leads to high probability of 0's in more significant bit-planes producing low entropy
- Adaptive arithmetic coding further takes advantage of local statistical distribution leading to even higher compression
- The context used to define conditional probability models at each significant coefficient is related to the significance status of its eight neighbors and its parent yielding a total of 18 possible models
Performance comparison (PSNR, [dB]) of advanced wavelet image codecs for the 512 x 512 standard test image “Lena”

<table>
<thead>
<tr>
<th>Compression Ratio</th>
<th>64:1</th>
<th>32:1</th>
<th>16:1</th>
<th>8:1</th>
</tr>
</thead>
<tbody>
<tr>
<td>EZW</td>
<td>30.23</td>
<td>33.17</td>
<td>36.28</td>
<td>39.55</td>
</tr>
<tr>
<td>MRWD</td>
<td>31.09</td>
<td>34.12</td>
<td>37.17</td>
<td>40.33</td>
</tr>
<tr>
<td>SPIHT</td>
<td>31.09</td>
<td>34.11</td>
<td>37.21</td>
<td>40.41</td>
</tr>
<tr>
<td>SLCCA</td>
<td><strong>31.25</strong></td>
<td><strong>34.28</strong></td>
<td><strong>37.35</strong></td>
<td><strong>40.47</strong></td>
</tr>
</tbody>
</table>
Objective performance improvement (PSNR, [dB]) of SLCCA over set partitioning in hierarchical trees (SPIHT) for eight 512 x 512 standard natural test images “Lena,” “Barbara,” “Baboon,” “Couple,” “Man,” “Boat,” “Tank,” and “Goldhill”

<table>
<thead>
<tr>
<th>Compression Ratio</th>
<th>32:1</th>
<th>16:1</th>
<th>8:1</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLCCA Improvement</td>
<td>0.17</td>
<td>0.19</td>
<td>0.15</td>
</tr>
</tbody>
</table>
Objective performance comparison (PSNR, [dB]) of SLCCA versus JPEG for 512 x 512 standard test image “Lena”

<table>
<thead>
<tr>
<th>Compression Ratio</th>
<th>64:1</th>
<th>32:1</th>
<th>16:1</th>
<th>8:1</th>
</tr>
</thead>
<tbody>
<tr>
<td>JPEG</td>
<td>27.79</td>
<td>31.42</td>
<td>34.84</td>
<td>37.94</td>
</tr>
<tr>
<td>SLCCA</td>
<td>31.25</td>
<td>34.28</td>
<td>37.35</td>
<td>40.47</td>
</tr>
</tbody>
</table>
SLCCA Image Coding

SLCCA CR=130:1

JPEG CR=130:1

SLCCA CR=80:1

JPEG CR=80:1
VSLCCA Video Coding

- Video Significance-Linked Connected Component Analysis (VSLCCA): High performance low bit rate wavelet video codec

- Wavelet transform with advanced data organization and representation strategies can be successfully applied in standard block-based motion-compensated video coding when motion-compensated error frames are made coherent

- A few successful attempts in wavelet-based low bit rate video coding: Sarnoff Corporation's zerotree entropy (ZTE) coder providing slightly inferior performance than H.263 standard
Building blocks of VSLCCA

- Fine-tuned motion estimation
- Exhaustive overlapped block motion compensation
- Wavelet transform of motion-compensated error frames
- SLCCA data organization and representation
- Adaptive arithmetic coding with scale/space-variant high order Markov source modeling
Motion estimation is similar to that of H.263 Recommendation:

- Full search block matching algorithm
- 16 x 16 non-overlapping macroblocks
- 15 pixels search range
- Half pixel refinement by using bilinear interpolation
- Each macroblock is split into four 8 x 8 blocks, and one motion vector per block is determined
- Based on two thresholds zero, one, or four motion vectors per macroblock are used
- Motion vectors are encoded by adaptive arithmetic coder
Exhaustive overlapped block motion compensation (OBMC):

- Block size: 8 x 8 pixels
- One motion vector per block is allowed
- Each predicted block is composed as the weighted sum of as many as nine possibly overlapped blocks translated from the previous reconstructed frame by using the motion vectors associated with the current block and its eight neighboring blocks
- Raised cosine window with 4 pixels overlap is used as weighting function
VSLCCA Video Coding

- First frame of each of the eight standard MPEG-4 test sequences
Performance comparison of H.263, ZTE, and VSLCCA (PSNR, [dB]) on MPEG-4 test sequences

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Frame Rate (fps)</th>
<th>Bit Rate (kbps)</th>
<th>H.263</th>
<th>ZTE</th>
<th>VSLCCA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Akiyo</td>
<td>5.0</td>
<td>8.87</td>
<td>34.61</td>
<td>34.61</td>
<td>35.55</td>
</tr>
<tr>
<td>Container Ship</td>
<td>5.0</td>
<td>10.04</td>
<td>30.79</td>
<td></td>
<td>31.13</td>
</tr>
<tr>
<td>Hall Monitor</td>
<td>5.0</td>
<td>8.18</td>
<td>30.36</td>
<td>30.25</td>
<td>31.51</td>
</tr>
<tr>
<td>Mother &amp; Daughter</td>
<td>5.0</td>
<td>9.54</td>
<td>33.50</td>
<td></td>
<td>33.87</td>
</tr>
<tr>
<td>Coast Guard</td>
<td>7.5</td>
<td>46.03</td>
<td>29.74</td>
<td>29.20</td>
<td>29.82</td>
</tr>
<tr>
<td>Foreman</td>
<td>7.5</td>
<td>46.77</td>
<td>31.91</td>
<td></td>
<td>32.26</td>
</tr>
<tr>
<td>News</td>
<td>7.5</td>
<td>49.25</td>
<td>35.10</td>
<td>35.17</td>
<td>35.41</td>
</tr>
<tr>
<td>Silent Voice</td>
<td>7.5</td>
<td>49.94</td>
<td>35.94</td>
<td></td>
<td>36.10</td>
</tr>
</tbody>
</table>
Three-Dimensional Significance-Linked Connected Component Analysis (3D-SLCCA): High performance high-to-low/error-resilient wavelet codecs

- Computationally expensive motion estimation is replaced by temporal filtering: Wavelet decomposition is extended to include the time domain
- SLCCA algorithm is extended to 3-D data structure
3D-SLCCA Video Coding
Non-Coding Delay Critical Applications

- 3-D wavelet transform extending 2-D wavelet transform to include the time domain
- Instead of block-based motion estimation/compensation, Daubechies 9/7 biorthogonal wavelet is used for temporal filtering
- Daubechies 9/7 biorthogonal wavelet is also used for spatial decomposition
3D-SLCCA Video Coding
Non-Coding Delay Critical Applications

- 3-D connected component analysis with volume thresholding
- Adaptive arithmetic coding with scale/space/time-variant high order Markov source modeling for subband-wise bit-plane encoding
3D-SLCCA Video Coding
Non-Coding Delay Critical Applications

- Applications: Non-interactive video such as Internet video distribution, video broadcast, video-on-demand, etc.
  - Low computational complexity
  - High performance
  - Highly scalable
  - Moderate robustness against channel error propagation
  - Large coding delay
  - Large memory requirements
3D-SLCCA Video Coding
Coding Delay Critical Applications

H-SLCCA: 3D-SLCCA encoder using Haar transform

- Temporal lowpass and highpass subbands are obtained as the sum and difference of two consecutive input frames (Haar wavelet)
- Dyadic spatial wavelet decomposition is separately applied for each temporal subband
- SLCCA technique is separately applied for each temporal subband
Applications: Interactive video such as teleconferencing and videophony, etc.
- Very low computational complexity
- Moderate to high performance
- Highly scalable
- High robustness against channel error propagation
- Small coding delay
- Small memory requirements
3D-SLCCA Video Coding

- Objective performance comparison with 3D-SPIHT for the “Football” sequence (9/7 Daubechies wavelet for temporal filtering in both 3D-SLCCA and 3D-SPIHT)

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Bit Rate [bpp]</th>
<th>0.3</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D-SPIHT</td>
<td>27.90</td>
<td>34.20</td>
<td></td>
</tr>
<tr>
<td>3D-SLCCA</td>
<td>28.14</td>
<td>34.51</td>
<td></td>
</tr>
</tbody>
</table>

- Average PSBR increase: 0.28 dB
3D-SLCCA vs. MPEG-2

Objective performance comparison with MPEG-2 at 30 frames-per-second coded at 15 Mbps (9/7 Daubechies wavelet for temporal filtering in 3D-SLCCA)

<table>
<thead>
<tr>
<th>Algorithm Sequence</th>
<th>MPEG-2</th>
<th>3D-SLCCA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flower Garden</td>
<td>36.97</td>
<td>37.31</td>
</tr>
<tr>
<td>Mobile Calendar</td>
<td>34.31</td>
<td>33.68</td>
</tr>
<tr>
<td>Football</td>
<td>36.99</td>
<td>37.70</td>
</tr>
<tr>
<td>Table Tennis</td>
<td>36.79</td>
<td>37.65</td>
</tr>
</tbody>
</table>

One 195 MHz R10000 CPU of SGI Octane Workstation

<table>
<thead>
<tr>
<th>Function</th>
<th>Encoder Time</th>
<th>Decoder Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelet Transform</td>
<td>62.5</td>
<td>68.0</td>
</tr>
<tr>
<td>Quantization and Postprocessing</td>
<td>56.5</td>
<td>-</td>
</tr>
<tr>
<td>SLCCA Coding</td>
<td>60.2</td>
<td>72.1</td>
</tr>
<tr>
<td>Total</td>
<td>179.3</td>
<td>140.1</td>
</tr>
</tbody>
</table>

MPEG-2 Encoder: 1209.7 seconds
MPEG-2 Decoder: 31.26 seconds
3D-SLCCA vs. MPEG-2

Original 77th frame

MPEG-2

3D-SLCCA
Joint Source-Channel Coding

- Reasons for error sensitivity of advanced wavelet image codecs
  - Significance map: Location of significance coefficients is recursively specified
  - Variable length coding: Synchronization between the encoder and decoder may get lost even in case of a single bit error
Joint Source-Channel Coding

- Main components of the proposed joint source-channel codec
  - Error-resilient packetization
  - Bit-plane-wise unequal error protection (Reed-Solomon codes)
  - Hierarchical synchronization
  - Significance-link reduction
Reed-Solomon Codes

- Over Galois field $GF(2^m)$, we use $m=8$
- Non-binary code: 8 bits constitute a symbol
- Block code $(k,n)$
- Block size $n$ ($n < 2^m$), $k$ data symbols, $n-k$ parity symbols
- Error correction performance
  - Unknown error location: $(n-k)/2$ symbols per block
  - Known error location (erasure error): $(n-k)$ symbols per block
- Low computational complexity
Transmission Error Protection

- Assume the average bit error rate of the channel is known
- Number of erroneous symbols per block is calculated
- RS code is used to correct all erroneous symbols per block

Number of parity symbols per block = 2*(Number of erroneous symbols per block)
Joint Source-Channel Coding
Error Resilient Packetization

- RS codes over GF(256) limit the packet size to 255 symbols
- The first packet contains vital information needed for decoding
  - Image size
  - Number of wavelet scales
  - Quantizer step size
  - Size of the source file for each bit-plane
  - Explicit seed information
- The first packet also contains the first few most significant bit-planes
- For the rest of the bit-planes, each packet contains source stream only belonging to a particular bit-plane (but one bit-plane may be placed into several packets)
Joint Source-Channel Coding
Bit-Plane-Wise Unequal Error Protection

- Contribution of each bit-plane to the root mean-squared error reduction for the “Lena” image coded at 0.5 bpp

- Despite smaller size, more significant bit-planes contain most of the information

<table>
<thead>
<tr>
<th>Bit-plane</th>
<th>Size of Source File</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>8</td>
<td>27</td>
</tr>
<tr>
<td>7</td>
<td>50</td>
</tr>
<tr>
<td>6</td>
<td>400</td>
</tr>
<tr>
<td>5</td>
<td>1189</td>
</tr>
<tr>
<td>4</td>
<td>4117</td>
</tr>
<tr>
<td>3</td>
<td>1694</td>
</tr>
<tr>
<td>2</td>
<td>3339</td>
</tr>
<tr>
<td>1</td>
<td>3138</td>
</tr>
<tr>
<td>0 (LSB)</td>
<td>2420</td>
</tr>
</tbody>
</table>
Joint Source-Channel Coding
Bit-Plane-Wise Unequal Error Protection

- More significant bit-planes are allocated more parity bits
- Reed-Solomon codes are used

\[ C_j = \left\lfloor \frac{C}{1 + \alpha \exp\left(\frac{j-n}{2}\right)} \right\rfloor \]

for \( j=0, \ldots, n-1 \).

- \( C \): Number of parity symbols to protect the vital packet
- \( C_j \): Number of parity symbols for the \( j \)-th bit-plane
- \( n \): Number of bit-planes
- \( \alpha \): Code parameter
Joint Source-Channel Coding
Bit-Plane-Wise Unequal Error Protection

Three different codecs designed for BER = 0.005:

<table>
<thead>
<tr>
<th>Codec</th>
<th>Parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Codec 0</td>
<td>Equal error protection, $\alpha=0$</td>
</tr>
<tr>
<td>Codec I</td>
<td>$\alpha=0.5$</td>
</tr>
<tr>
<td>Codec II</td>
<td>$\alpha=1$</td>
</tr>
<tr>
<td>Codec III</td>
<td>$\alpha=2$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bit-Plane</th>
<th>Codec 0</th>
<th>Codec I</th>
<th>Codec II</th>
<th>Codec III</th>
</tr>
</thead>
<tbody>
<tr>
<td>First Packet</td>
<td>(207,255)</td>
<td>(188,255)</td>
<td>(171,255)</td>
<td>(157,255)</td>
</tr>
<tr>
<td>6</td>
<td>(207,255)</td>
<td>(197,255)</td>
<td>(188,255)</td>
<td>(180,255)</td>
</tr>
<tr>
<td>5</td>
<td>(207,255)</td>
<td>(202,255)</td>
<td>(198,255)</td>
<td>(194,255)</td>
</tr>
<tr>
<td>4</td>
<td>(207,255)</td>
<td>(205,255)</td>
<td>(204,255)</td>
<td>(203,255)</td>
</tr>
<tr>
<td>3</td>
<td>(207,255)</td>
<td>(207,255)</td>
<td>(207,255)</td>
<td>(208,255)</td>
</tr>
<tr>
<td>2</td>
<td>(207,255)</td>
<td>(208,255)</td>
<td>(210,255)</td>
<td>(211,255)</td>
</tr>
<tr>
<td>1</td>
<td>(207,255)</td>
<td>(209,255)</td>
<td>(211,255)</td>
<td>(213,255)</td>
</tr>
<tr>
<td>0</td>
<td>(207,255)</td>
<td>(210,255)</td>
<td>(212,255)</td>
<td>(214,255)</td>
</tr>
</tbody>
</table>
Joint Source-Channel Coding
Hierarchical Synchronization

- Bit-plane level
  - Restart the arithmetic coder after each bit-plane
  - Transmit the size of the source file of each bit-plane
  - Insert synchronization characters into the bitstream between bit-planes

- Scale/subband level
  - Restart the arithmetic coder after each scale/subband
  - Insert synchronization characters into the bitstream between scales/subbands

- If an uncorrectable error occurs, then decoding can continue at the next synchronization point (bit-plane, scale, or subband)
Transmission errors may propagate from coarse to fine scales through significance-linkage.

Reduce the number of significance-links and increase the explicit seed positions.

<table>
<thead>
<tr>
<th>Coding Pattern</th>
<th>Number of Links</th>
<th>Number of Seeds</th>
<th>PSNR [dB]</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Links</td>
<td>463</td>
<td>10</td>
<td>37.35</td>
</tr>
<tr>
<td>50% Links</td>
<td>234</td>
<td>233</td>
<td>37.26</td>
</tr>
<tr>
<td>0% Links</td>
<td>0</td>
<td>473</td>
<td>37.18</td>
</tr>
</tbody>
</table>

Noiseless performance for the “Lena” image coded at 0.5 bpp.
Joint Source-Channel Coding
Link Reduction

All Links

50% Links

0% Links
Joint Source-Channel Coding

- Standard 512 x 512 test image “Lena”
- Total bit rate (source rate and channel rate) at 0.25 bpp, 0.5 bpp, and 1.0 bpp
- Binary symmetric channel
- PSNR are the average value of 100 simulations
Effectiveness of synchronization

Average PSNR improvement is 1.01 dB (as much as 2 dB)
Joint Source-Channel Coding

- Binary symmetric channel
- The channel error rate is used for the codec design
- Noiseless SLCCA: PSNR=34.28 dB, PSNR=37.35 dB, and PSNR=40.47 dB at 0.25 bpp, 0.5 bpp, and 1.0 bpp, respectively

<table>
<thead>
<tr>
<th>Algorithm/Rate [bpp]</th>
<th>0.25</th>
<th>0.5</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed Scheme</td>
<td>33.83</td>
<td>36.94</td>
<td>39.93</td>
</tr>
<tr>
<td>Sherwood &amp; Zeger</td>
<td>33.16</td>
<td>36.35</td>
<td>39.34</td>
</tr>
<tr>
<td>Man, Kossentini, and Smith</td>
<td>32.59</td>
<td>35.77</td>
<td>-</td>
</tr>
<tr>
<td>Tanabe &amp; Farvardin</td>
<td>30.94</td>
<td>33.90</td>
<td>36.71</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Algorithm/Rate [bpp]</th>
<th>0.25</th>
<th>0.5</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed Scheme</td>
<td>32.59</td>
<td>35.60</td>
<td>38.70</td>
</tr>
<tr>
<td>Sherwood &amp; Zeger</td>
<td>31.91</td>
<td>34.95</td>
<td>38.03</td>
</tr>
<tr>
<td>Man, Kossentini, and Smith</td>
<td>31.52</td>
<td>34.14</td>
<td>-</td>
</tr>
<tr>
<td>Tanabe &amp; Farvardin</td>
<td>29.96</td>
<td>32.38</td>
<td>35.44</td>
</tr>
</tbody>
</table>

BER=0.001
0.66 dB average PSNR increase over Sherwood and Zeger’s RCPC/CRC technique
Equal error protection only gives good performance when the BER used to design the codec is close to the actual BER.

Codec I, II, and III result in 2.16 dB, 3.26 dB, and 4.27 dB average PSNR improvement over Codec 0, which uses equal error protection.
Joint Source-Channel Coding

<table>
<thead>
<tr>
<th>Function</th>
<th>Encoder Time</th>
<th>Decoder Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelet Transform</td>
<td>0.46</td>
<td>0.55</td>
</tr>
<tr>
<td>Quantization</td>
<td>0.07</td>
<td>0.11</td>
</tr>
<tr>
<td>Conditioned Dilation</td>
<td>0.14</td>
<td>-</td>
</tr>
<tr>
<td>Postprocessing</td>
<td>0.28</td>
<td>-</td>
</tr>
<tr>
<td>Arithmetic Coding</td>
<td>0.55</td>
<td>0.74</td>
</tr>
<tr>
<td>RS Coding</td>
<td>0.09</td>
<td>0.14</td>
</tr>
<tr>
<td>Total</td>
<td>1.59</td>
<td>1.54</td>
</tr>
</tbody>
</table>

- RCPC/CRC decoding: 6.02 seconds
- Low computational complexity is especially important in power-constrained mobile scenario
Wireless Visual Communication System

Mobile Networking Model

- Fixed host (FH): Video source, resource rich
- Mobile host (MH): Video receiver, resource limited
- Mobile support station (MSS): Designated fixed host, gateway between wireless and wireline networks
Limitations of Mobile Wireless Scenario

Characteristics of wireless connections
- Much lower bandwidth than wireline networks
- Much higher transmission error rate than wireline networks
- Large variations in network connections (bandwidth, error rate, delay, etc.) due to roaming
- Frequent disconnections

Mobile host
- Limited battery power
- Very heterogeneous hardware capabilities (Personal Digital Assistant (PDA), Laptop)
- Very heterogeneous networking capabilities
System Design

- Due to the low bandwidth error-prone wireless connections, **application-driven** and **integrated system design** is highly desirable for efficient multimedia communications
  - The application shall play central part in the communications. Several network functionalities are implemented as part of the application
  - Strong interaction between different network layers
We propose a three-layer system

- **Application layer**
  - Source codec
  - Channel codec
  - Packetizer

- **Data link layer**
  - H.223 Recommendation to multiplex data, audio, and video streams

- **Physical layer**
Mobile Proxy Server

- Located at MSS
- Provide efficient and effective video communications: Adapting the video to both the network capabilities and hardware resources of mobile hosts

- Functionalities:
  - Protocol translation
  - Video transcoding
  - Video streaming
  - Error control coding
  - Caching and retransmission
  - Maintaining (or registering) the capabilities of the mobile hosts
Between the video server and MSS, either TCP/IP or UDP/IP may be used (depending on the video server).

Between MSS and MH, H.223 Recommendation is used: Multiplexing protocol for low bit rate multimedia communications:

- Support quality-of-service parameters (e.g., bandwidth): No contention among data, audio, and video streams.
- Short packet size (120 bytes) => lower delay.
- Small protocol overhead (<5 bytes) => less bandwidth.
- Support for IP.
Mobile Proxy Server - Video Transcoding

Transcoding is required in order to shape the video traffic to the hardware resources of mobile clients.

Transcoding steps:

- Download MPEG encoded video from the server
- Decode the video file
- Reformat according to the preferences of the mobile hosts
- Reencode the video by using 3D-SLCCA
- Apply error control
Mobile Proxy Server - Video Streaming

- When video server such as web server does not support streaming, the entire MPEG movie must be first downloaded to the MPS.
- Then the MPS performs video streaming needed for transmission of video to the resource poor MH.
- Video streaming enables downloading, decoding, and displaying being carried out simultaneously.
- Video streaming requires real-time handling of transmission errors.
Most of the packet losses appear over the wireless link between the MSS and MH.

Video is cached at the MPS.

In the case of retransmission, packets are only transmitted from the cache located at the MPS and not from the video server.

Since the MPS is located close to the mobile client, time delay due to retransmission is significantly reduced.
Mobile Proxy Server - Maintaining Capabilities of MH

- The MPS maintains the capabilities of each mobile host
  - Resolution
  - Color depth
  - Maximal frame rate
  - Available bandwidth

- These properties are used during transcoding to construct the video best suitable for the mobile host
Source Coding

- Three-dimensional significance-linked connected component analysis
  - Temporal wavelet decomposition is applied on two consecutive input frames to obtain temporal lowpass and temporal highpass subbands (Haar wavelet)
  - SLCCA data organization and representation strategy is applied for each temporal subband separately
  - Adaptive arithmetic coding with space/scale-variant higher order Markov source modeling is used for encoding each temporal subband

- Desirable properties for wireless video
  - Low computational complexity
  - High coding efficiency
  - Scalable video representation in terms of resolution, frame rate, and quality
  - Prevention of transmission errors from propagation (spatial and temporal)
Transmission Error Protection

- Adding parity bits to the source bitstream
- Tradeoff between bandwidth increase and error resilience
- Three-layer error protection mechanism
  - Bit-plane-wise synchronization between encoding and decoding
  - Unequal forward error correction
  - Retransmission
Transmission Error Protection

- **Synchronization**
  - Bit-plane-wise transmission in 3D-SLCCA
  - Insert synchronization characters into the bitstream between bit-planes
  - Optional synchronization characters may be inserted into the bitstream between subbands as well

- **Unequal forward error correction**
  - Reed-Solomon codes over GF(256)
  - Protect each bit-plane according to its importance

- **Retransmission**
  - Apply only for the most important parts of the bitstream
Transmission Error Protection

Tested codecs:
- Codec 0: No forward error correction
- Codec 1: Average BER = 0.001 (designed for H.223 Level 1)
- Codec 2: Average BER = 0.01 (designed for H.223 Level 2)

<table>
<thead>
<tr>
<th>Bit-Plane</th>
<th>Codec 0</th>
<th>Codec 1</th>
<th>Codec 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 (MSB)</td>
<td>(120,120)</td>
<td>(104,120)</td>
<td>(80,120)</td>
</tr>
<tr>
<td>5</td>
<td>(120,120)</td>
<td>(104,120)</td>
<td>(80,120)</td>
</tr>
<tr>
<td>4</td>
<td>(120,120)</td>
<td>(110,120)</td>
<td>(84,120)</td>
</tr>
<tr>
<td>3</td>
<td>(120,120)</td>
<td>(110,120)</td>
<td>(84,120)</td>
</tr>
<tr>
<td>2</td>
<td>(120,120)</td>
<td>(110,120)</td>
<td>(84,120)</td>
</tr>
<tr>
<td>1</td>
<td>(120,120)</td>
<td>(114,120)</td>
<td>(88,120)</td>
</tr>
<tr>
<td>0 (LSB)</td>
<td>(120,120)</td>
<td>(114,120)</td>
<td>(88,120)</td>
</tr>
</tbody>
</table>

Applied RS codes
Transmission Error Protection

- Resulting bit-rates for the “Foreman” sequence

<table>
<thead>
<tr>
<th>Codec</th>
<th>Rate [kbps]</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Codec 0</td>
<td>48.05</td>
<td>48.05</td>
</tr>
<tr>
<td>Codec 1</td>
<td>41.89</td>
<td>48.10</td>
</tr>
<tr>
<td>Codec 2</td>
<td>33.13</td>
<td>48.15</td>
</tr>
</tbody>
</table>
Five different channel conditions are tested

- Correlated Rayleigh fading model
- Two mobile speeds
- Two signal-power-to-noise-power (E/N) conditions

<table>
<thead>
<tr>
<th>Channel Symbol</th>
<th>Channel Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>C0</td>
<td>Noiseless Channel</td>
</tr>
<tr>
<td>C1</td>
<td>Rayleigh, 1.4 km/h, E/N = 24 dB, Average BER = 0.00089</td>
</tr>
<tr>
<td>C2</td>
<td>Rayleigh, 1.4 km/h, E/N = 14 dB, Average BER = 0.0093</td>
</tr>
<tr>
<td>C3</td>
<td>Rayleigh, 14 km/h, E/N = 24 dB, Average BER = 0.0011</td>
</tr>
<tr>
<td>C4</td>
<td>Rayleigh, 14 km/h, E/N = 14 dB, Average BER = 0.0098</td>
</tr>
</tbody>
</table>
Both low motion and high motion sequences are tested

Results for the “Foreman” sequence sampled at 7.5 frames-per-second coded at 48k bits-per-second

<table>
<thead>
<tr>
<th>Channel</th>
<th>Codec 0</th>
<th>Codec 1</th>
<th>Codec 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>C0</td>
<td>28.77</td>
<td>28.06</td>
<td>27.40</td>
</tr>
<tr>
<td>C1</td>
<td>27.32</td>
<td>27.43</td>
<td>26.85</td>
</tr>
<tr>
<td>C2</td>
<td>22.78</td>
<td>25.24</td>
<td>24.94</td>
</tr>
<tr>
<td>C3</td>
<td>26.15</td>
<td>26.84</td>
<td>26.75</td>
</tr>
<tr>
<td>C4</td>
<td>18.33</td>
<td>22.15</td>
<td>23.18</td>
</tr>
</tbody>
</table>

Objective performance (PSNR [dB]) of the three codecs among five different channel conditions
Performance Evaluation

Timing results for SGI Octane workstation for accessing the 10 second-long “Foreman” sequence

<table>
<thead>
<tr>
<th>Function</th>
<th>Time [sec]</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPEG Decoding</td>
<td>3.45</td>
</tr>
<tr>
<td>Reformatting</td>
<td>0.37</td>
</tr>
<tr>
<td>3D-SLCCA Encoding</td>
<td>3.98</td>
</tr>
<tr>
<td>Transmission Error Protection</td>
<td>0.25</td>
</tr>
<tr>
<td>Total</td>
<td>8.05</td>
</tr>
</tbody>
</table>

Mobile Proxy Server

<table>
<thead>
<tr>
<th>Function</th>
<th>Time [sec]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transmission Error Correction</td>
<td>0.35</td>
</tr>
<tr>
<td>3D-SLCCA Decoding</td>
<td>2.72</td>
</tr>
<tr>
<td>Total</td>
<td>3.07</td>
</tr>
</tbody>
</table>

Mobile Host
Conclusions

- High performance image codec: SLCCA
- High performance and versatile video codec: 3D-SLCCA
  - 3D-SLCCA outperforms MPEG-2 and is seven times faster than MPEG-2 for encoding
  - For error resilient coding, 3D-SLCCA outperforms H.263+
- Both SLCCA and 3D-SLCCA provide desirable features:
  - Scalability
  - Low computational complexity
  - Symmetric computational complexity
  - Robustness against transmission error propagation
Conclusions

- High performance error resilient image codec
  - Hierarchical resynchronization, unequal error protection, packetization, and significance-link reduction
  - Provides graceful performance degradation for mismatched error rate

- Wireless Video Communications: Application-driven integrated system design
  - Application layer research (joint source coding and channel coding)
  - Video is adapted to both the networking capabilities and hardware resources of individual receivers

- More information: http://meru.cecs.missouri.edu
System Design