10-25-2004

Joint source-channel decoding for MPEG-2 video transmission

Liuguo Yin

Weigang Chen

Jianhua Lu

Chang Wen Chen

Joint source-channel decoding for MPEG-2 video transmission

Liuguo Yin
Weigang Chen
Jianhua Lu
Chang Wen Chen

Joint Source-Channel Decoding for MPEG-2 Video Transmission

Liuguo Yin[*a], Weigang Chen[a], Jianhua Lu[a], Chang Wen Chen[b]
[a]Department of Electronic Engineering, Tsinghua University, Beijing, China, 100084
[b]Depart of Electrical and Computer Engineering, Florida Institute of Technology, USA

ABSTRACT

Joint source-channel coding schemes have been proven to be effective ways for reliable multimedia communications. In this paper, a joint source-channel decoding (JSCD) scheme that combines the hidden Markov source (HMS) estimation and low-density parity-check (LDPC) coding is proposed for the standard MPEG-2 video transmission. The LDPC code of the proposed scheme has a near-Shannon-limit error-correcting capability, while the HMS estimator may accurately extract the residual redundancy within the MPEG-2 video stream without any prior information. Furthermore, with a joint iterative decoding algorithm, the estimated source redundancy may be well exploited by the LDPC decoder, and the channel decoding feedback may refine the subsequence HMS estimation, thereby effectively improving the system performance. On the other hand, we also show that the proposed JSCD scheme has approximately the same computation complexity as that of the standard decoding scheme. Moreover, it is worth noting that the proposed scheme is based on separation encoding schemes, which is very convenient to be applied to existing multimedia transmission systems.

Keywords: Joint source-channel coding (JSCC), MPEG-2, residual redundancy, hidden Markov source (HMS) estimation, low-density parity-check (LDPC) code

1. INTRODUCTION

Future wireless networks are expected to support high-quality broadband multimedia services, which may mix data, text, speech, audio, image and video. Due to the limited bandwidth of the wireless channels, the multimedia signals have to be compressed by source coding scheme, such as MPEG-1, MPEG-2, MPEG-4, and H.26X etc. However, the compressed data is very sensitive to channel errors, and the wireless channels are usually severely impaired due to multi-path fading, shadowing, inter-symbol interference, and noise disturbances. Therefore, great challenge arises when multimedia signals are transmitted over wireless channels.

Current multimedia communication systems designed according to Shannon's separation theorem are now being reconsidered, since in practical situations the redundancy of a dynamic multimedia source is not always completely removed by source coding, and the wireless channel is always a correlated channel, with unanticipated fadings. Joint source-channel coding (JSCC) schemes that systematically consider imperfect source compression and channels with

---

[*] ylg00@mails.tsinghua.edu.cn; phone (86)10-6279-4447; fax 8610-6277-0528.

Multimedia Systems and Applications VII, edited by Chang Wen Chen, C.-C. Jay Kuo,
Anthony Vetro, Proceedings of SPIE Vol. 5600 (SPIE, Bellingham, WA, 2004)
0277-786X/04/$15 · doi: 10.1117/12.580529

burst errors are thereby proposed to provide more effective ways for wireless multimedia transmission [1]. Currently there are mainly three types of JSCC scheme. The first one is the joint source-channel bit allocation scheme, which dynamically allocates source bit rate according to channel conditions [2][3]. The second one is based on unequal error protection codes, in which different channel codes are applied to different bits according to their respective importance on the reconstructed video frame [4][5]. The third one is the joint source channel-decoding (JSCD) scheme, which exploits a hidden Markov source estimation to extract the residual redundancy within the multimedia stream to improve the channel decoding [6][7].

In this paper, a joint source-channel decoding (JSCD) scheme that combines the hidden Markov source (HMS) estimation and low-density parity-check (LDPC) coding is proposed for the standard MPEG-2 video transmission. The LDPC code of the proposed scheme has a near-Shannon-limit error-correcting capability, while the HMS estimator may accurately extract the residual redundancy within the MPEG-2 video stream without any prior information. Furthermore, with a joint iterative decoding algorithm, the estimated source redundancy may be well exploited by the LDPC decoder, and the channel decoding feedback may refine the subsequence HMS estimation, thereby effectively improving the system performance. On the other hand, we also show that the proposed JSCD scheme has approximately the same computation complexity as that of the standard decoding scheme. Moreover, it is worth noting that the proposed scheme is based on existing separation encoding and the MPEG standards, which is very convenient to be utilized in the existing multimedia communication systems.

The remainder of this paper is organized as follows. Section 2 describes the proposed joint source-channel decoding scheme for MPEG-2 transmission. Then, simulations and the results are presented in Section 3. Section 4 concludes this paper with summary and some discussion.

2. THE PROPOSED JOINT SOURCE-CHANNEL DECODING SCHEME

The multimedia transmission system model with the proposed joint source-channel decoding scheme is shown in Figure 1. In this system model, the modulation module adopts the binary phase shift keying (BPSK) modulation and the channel module is assumed to be Additive White Gaussian Noise (AWGN) channel, since burst errors of the wireless channel may be transformed to be random errors with an appropriate interleaving scheme. The key components are the JSCD modules, each of which will be discussed in detail below.

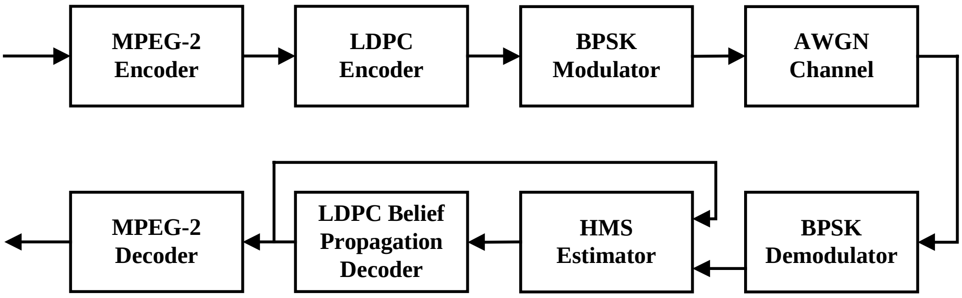

Fig. 1. Multimedia transmission system model with the proposed joint source-channel decoding scheme.

2.1 MPEG-2 standard and residual redundancy

In this system, we assume the multimedia source is first encoded by the standard MPEG-2 codec. The Moving Pictures Experts Group (MPEG) is part of the International Standards Organisation (ISO), and defines standards for digital video and digital audio, i.e. MPEG-1, MPEG-2, MPEG-4 and MPEG-7, each of which provides different levels and profiles to support special applications in an optimal way [9][10]. MPEG-2 video is an ISO/IEC standard that specifies the syntax and semantics of an enclosed video bit stream. In this paper, we limit our discussions to this part.

In MPEG-2 video standard, encoding of multimedia information, such as video is achieved by using two main techniques: spatial and temporal compression, applied to remove the spatial and temporal redundancies in the source. Spatial compression involves the analysis of a picture to determine the redundant information in that picture using a DCT, which converts the information in the picture to be encoded in the frequency domain, such as by discarding frequencies that are not visible to the human eye. Furthermore, more compression can also be achieved by using another technique called run length coding. Temporal compression is achieved by using motion compensated prediction to exploit redundant temporal information that has no changed from picture to picture. The key idea of motion compensated prediction is to encode a video frame based on other video frames temporally close to it. Therefore, MPEG-2 defines three picture types: I (Intra frame) pictures, P (Predictive) pictures, and B (Bi-directional) pictures, which are formed into a group of pictures (GOP). That is, temporal compression is based on GOP by motion compensated prediction technique.

MPEG-2 video coding scheme can achieve significant compression for video transmission over channels with limited bandwidth. However, redundant information in multimedia sources cannot be completely removed by the MPEG-2 video coding scheme. Therefore, some residual redundancy remains in the standard MPEG-2 coded bit stream. If these redundancies are fully exploited by the channel decoding, better error-correcting performance may be obtained.

In this paper, the residual redundancy is estimated with a HMS estimation algorithm, which does not need any prior information about the source. Furthermore, the extracted residual redundancy is feed forward to the LDPC decoder with an extrinsic information form, which may be fully exploited and makes the iterative decoding of LDPC codes more effective.

2.2 The HMS estimation algorithm

HMS estimation algorithms are practical solutions to learn about source through the source output. For many years, they have been successfully used in prediction, recognition, and identification systems [8]. An HMS estimation used to exploit the source redundancy was firstly proposed by Garcia-Frias [6]. In the paper [7], a modified HMS algorithm is applied to a two-state Markov source with unknown transition parameters. By simulation, it has been proved that the modified HMS estimation algorithm can accurately extract the source redundancy even over noise interference. In this paper, the hidden Markov source estimation algorithm is further utilized to exploit the residual redundancy within the dynamic MPEG-2 stream.

We can define s_k to be the *k-th* state of the hidden Markov source, $e_k = \{s_{k-1} = i; \quad u_k; \quad s_k = j\}$ be the *k-th* transition branch of the source model, and $p\{u, s_j | s_i\}$ be the transition probability, where $u \in \{0, 1\}$. Moreover, assume that the initial parameters of the HMS are $\lambda = \{p\{u, s_j | s_i\}, \pi, \quad u \in \{0,1\}\}$, where π is the initial state of the HMS model. Then, the forward-backward equations for a modified Baum-Welch algorithm can be expressed as follows:

$$
\begin{aligned}
\alpha_k(i) &= P\{\, s_k = i \,|\, R_1^k, \lambda \,\} \\
&= \frac{\sum_{s_{k-1}}\sum_{u_k} P\{s_{k-1} = j, u_k, s_k = i, R_1^{k-1}, R_k \,|\, \lambda\}}{P\{R_1^{k-1} \,|\, \lambda\} \cdot P\{R_k \,|\, R_1^{k-1}, \lambda\}} \\
&= \frac{\sum_{s_{k-1}}\sum_{u_k} \alpha_{k-1}(j) \cdot P\{u_k, s_k = i \,|\, s_{k-1} = j, \lambda\} \cdot P\{R_k \,|\, u_k\}}{P\{R_k \,|\, R_1^{k-1}, \lambda\}}
\end{aligned}
\tag{1}
$$

Where $\alpha_k(i)$ is the forward equation, and

$$
\begin{aligned}
P\{R_k \,|\, R_1^{k-1}, \lambda\} &= \sum_{s_{k-1}}\sum_{u_k}\sum_{s_k} P\{s_{k-1} = j, u_k, s_k = i, R_k \,|\, R_1^{k-1}, \lambda\} \\
&= \sum_{s_{k-1}}\sum_{u_k}\sum_{s_k} \alpha_{k-1}(j) \cdot P\{u_k, s_k = i \,|\, s_{k-1} = j, \lambda\} \cdot P\{R_k \,|\, u_k\}
\end{aligned}
\tag{2}
$$

Likewise, $\beta_k(i)$, which is the backward equation, may be given by

$$
\begin{aligned}
\beta_k(i) &= \frac{P\{R_{k+1}^K \,|\, s_k = i, \lambda\}}{P\{R_{k+1}^K \,|\, R_1^k, \lambda\}} \\
&= \frac{\sum_{u_{k+1}}\sum_{s_{k+1}} P\{u_{k+1}, s_{k+1} = j, R_{k+1}, R_{k+2}^K \,|\, s_k = i, \lambda\}}{P\{R_{k+1} \,|\, R_1^k, \lambda\} \cdot P\{R_{k+2}^K \,|\, R_1^{k+1}, \lambda\}} \\
&= \frac{\sum_{u_{k+1}}\sum_{s_{k+1}} P\{u_{k+1}, s_{k+1} = j \,|\, s_k = i, \lambda\} \cdot P\{R_{k+1} \,|\, u_{k+1}\} \cdot \beta_{k+1}(j)}{P\{R_{k+1} \,|\, R_1^k, \lambda\}}
\end{aligned}
\tag{3}
$$

As a result, the probability for the *k-th* transition branch may be computed as,

$$
\begin{aligned}
P(e_k) &= P\{s_{k-1} = i, u_k, s_k = j \,|\, R_1^K, \lambda\} \\
&= \frac{P\{s_{k-1} = i, u_k, s_k = j, R_1^K \,|\, \lambda\}}{P\{R_1^K \,|\, \lambda\}} \\
&= \frac{P\{s_{k-1} = i, R_1^{k-1} \,|\, \lambda\}}{P\{R_1^{k-1} \,|\, \lambda\}} \cdot \frac{P\{u_k, s_k = j \,|\, s_{k-1} = i, \lambda\}}{P\{R_k \,|\, R_1^{k-1}, \lambda\}} \\
&\quad \cdot \frac{P\{R_k \,|\, u_k\} \cdot P\{R_{k+1}^K \,|\, s_k = j, \lambda\}}{P\{R_{k+1}^K \,|\, R_1^k, \lambda\}} \\
&= \frac{\alpha_{k-1}(i) \cdot P\{u_k, s_k = j \,|\, s_{k-1} = i, \lambda\} \cdot P\{R_k \,|\, u_k\} \cdot \beta_k(j)}{P\{R_k \,|\, R_1^{k-1}, \lambda\}}
\end{aligned}
\tag{4}
$$

Note that, in the branch of the HMS model, the *k-th* branch, e_k, is determined by the initial state, S_i, the final state, S_j, and the associated output, u. Thereby, the branch probability $P\{s_i, u, s_j | \lambda\}$ for the hidden Markov source may be expressed as

$$P\{s_i, u, s_j | \lambda\} = \frac{1}{K} \sum_{k=1}^{K} P(e_k | \lambda), \quad s_i, s_j \in \{S_0, S_1\}, u \in \{0, 1\} \tag{5}$$

Accordingly, the transition probability is re-computed as

$$P^{'}\{u, s_j | s_i, \lambda\} = \frac{P\{s_i, u, s_j | \lambda\}}{\sum_u \sum_{s_j} P\{s_i, u, s_j | \lambda\}} \tag{6}$$

Thereby, the HMS model parameters are re-estimated as $\lambda^{'} = \left\{p^{'}\{u, s_j | s_i, \lambda\}, \pi, \quad u \in \{0,1\}\right\}$. It is worth noting that, when the block size is large enough, degradation due to improper estimation of the initial state is negligible. Therefore, re-estimation of the initial state is omitted in this study.

By recursive procedures mentioned above, accurate estimation of transition probabilities can be achieved. Then, with the obtained model parameters, source redundancy for each information bit can be safely extracted as extrinsic information, which may be expressed with the log-likelihood ratio form:

$$LLR_e(u_k) = \ln \frac{\sum_{s_{k-1}} \sum_{s_k} \alpha_{k-1}(i) \cdot P\{u_k = 1, s_k = j | s_{k-1} = i, \lambda^{'}\} \cdot \beta_k(j)}{\sum_{s_{k-1}} \sum_{s_k} \alpha_{k-1}(i) \cdot P\{u_k = 0, s_k = j | s_{k-1} = i, \lambda^{'}\} \cdot \beta_k(j)} \tag{7}$$

2.3 The joint source-channel decoding algorithm

The JSCD scheme proposed in this paper is based on the turbo iterative decoding principle, where the residual redundancy estimator and the belief-propagation decoder are taken as two uncorrelated subdecoders in the JSCD scheme. Specifically, the proposed JSCD algorithm can be expressed in the form given below:

1) Perform the first HMS estimation with the received information sequence R_1^K as described in Section 2.2. The extrinsic information for every information bits is transferred to the belief-propagation decoder;
2) Apply the belief-propagation algorithm for some iterations, which is described in [15], and then compute the extrinsic information output of the belief-propagation decoding;
3) Compute the hard decision output of the LDPC decoding. If the hard decision output is not an LDPC code word and the iteration times is less than the maximum number of iterations allowed, re-estimation is initiated with the extrinsic information of the belief-propagation decoding for better HMS model parameters and more accurate extrinsic information, then go to step 2); else, exit from the joint decoding procedure and output the hard decision result.

It is noted that we combine one HMS estimation with a suitable number of belief-propagation iterations in order to improve the effectiveness, which also reduces the computation complexity. In the following section, some computer experimental simulations are performed to verify the performance of the proposed scheme.

3. SIMULATION RESULTS

In this section, we evaluate the performance of the standard-based JSCD scheme by computer simulation. The simulation system architecture has been shown in Figure 1.

The multimedia source is first source encoded into bit stream by the standard MPEG-2 codec [16]. Some parameters of the MPEG-2 codec are summarized in Table 1.

Table 1. MPEG-2 codec Setting parameters.

Parameter	**Setting**
Profile	**main**
Level	**low**
Frame Rate	**30 f/s**
Frame Size	**352×288**
Color Subsampling	**4:2:0**
Bit Rate	**4 Mbit/s**

The encoded bit stream from MPEG-2 codec is channel encoded by a rate 1/2, $(8192, \mu, \nu)$ irregular LDPC code, with the variable node degree distribution $\mu(x) = 0.27684x + 0.28342x^2 + 0.43974x^8$ and the check node degree distribution $\upsilon(x) = 0.01568x^5 + 0.85244x^6 + 0.13188x^7$, respectively [15].

After BPSK modulation, AWGN channel transmission, and soft demodulation, the received sequences are decoded with the proposed JSCD scheme. Moreover, the standard decoding (SD) scheme with the belief-propagation algorithm is also investigated to give a performance comparison. Furthermore, in the proposed JSCD scheme, there is no need to perform just one HMS estimation in every belief-propagation iteration. In our simulation, one HMS estimation is combined with 10 belief-propagation iterations.

In the simulation, a sequence of 21 successive frames of the Foreman test sequence is chosen as the multimedia source, which are encoded with the standard MPEG-2 codec. The output MPEG-2 bit stream is then encoded into 673 irregular LDPC codewords with a length of 8192 bits. For each individual channel condition specified by E_b/N_0, simulation is performed 80 to 100 times and the average performance of the test sequence over the channel is obtained.

The peak signal-to-noise ratio (PSNR) vs. E_b/N_0 curves for the two decoding schemes are plotted in Figure 2. It is shown that in each individual channel condition, the PSNR value of the Foreman test sequence with proposed joint source-channel decoding scheme is higher than that with the standard decoding scheme. The improvement is obvious when the channel E_b/N_0 is low. This is because the proposed scheme more effectively extracts the residual redundancy of the MPEG-2 bit stream.

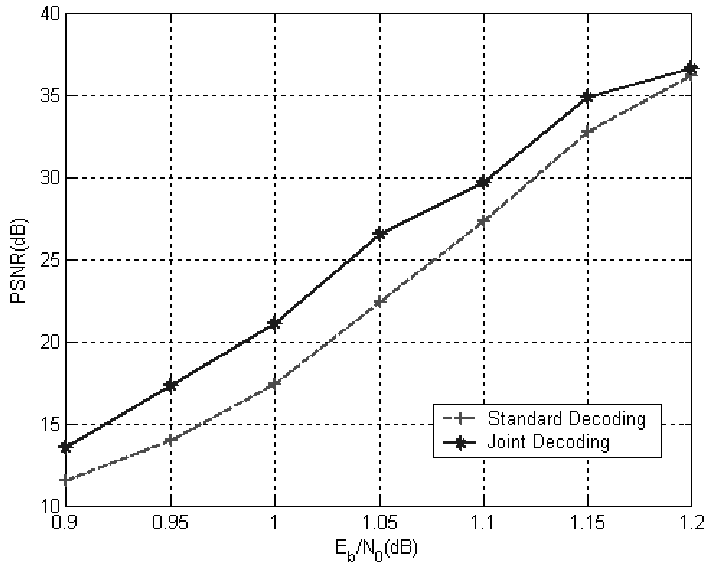

Fig. 2. PSNR Performance with the proposed joint source-channel decoding scheme and the standard decoding scheme (BIAWGN channel).

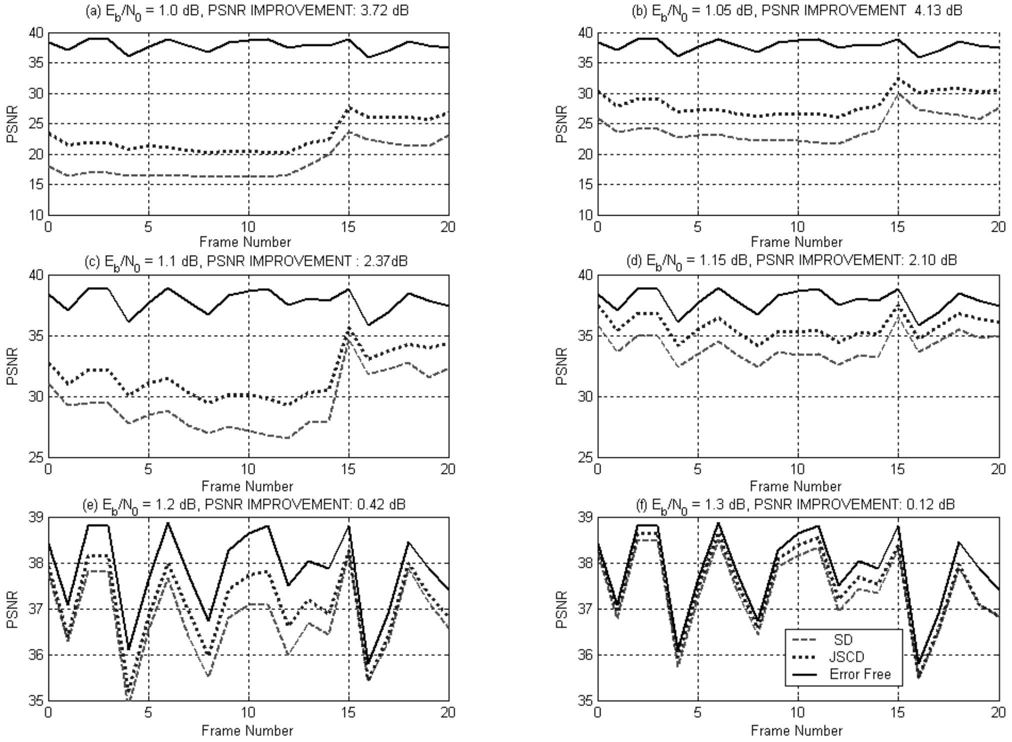

Fig. 3. PSNR Performance of the JSCD scheme and the standard iterative decoding scheme over dynamic MPEG-2 source and AWGN channels.

Furthermore, the PSNR performances of the two decoding schemes over dynamic MPEG-2 frames and different channel conditions are presented in Figure 3. In the simulation, the MPEG-2 codec adapts fixed bit rate encoding, thus the quality of every frame is not constant but varying just as the error-free curve shows in subfigure (a) to (f). The length of GOP is 15, and the PSNR values of frame 0 and frame 15 are higher than the others in the most cases. From subfigure (a) to (f), the PSNR is increasing with the improvement of channel condition, which is specified by E_b/N_0. When the channel is very bad, neither of the two schemes can decode correctly in this situation. There is significant quality loss in these cases. However, the proposed JSCD scheme still outperforms the standard decoding by several dB in PSNR. When E_b/N_0 increases, the JSCD scheme consistently achieves better performance than the standard decoding scheme. For example, it is showed in subfigure (b) that the PSNR difference is about 4.13 dB, when $E_b/N_0=1.05dB$. When the channel condition is further improved, both of the two schemes can decode correctly. In this situation, the improvement is just a little since the PSNRs of both decoding scheme are very close to performance limit. As showed in subfigure (f), when $E_b/N_0=1.3dB$, the PSNR difference is only 0.12 dB, while the PSNR of the proposed JSCD scheme is within 0.1 dB away from that of error-free transmission. Therefore, it is concluded that the proposed scheme is more robust than the standard decoding scheme, since it can effectively improve the PSNR over dynamic media source and different channel conditions, which is key to video coding and transmission for wireless applications.

On the other hand, simulation results also show that, with the aid of estimated residual redundancy, the average number of belief-propagation iterations of the proposed JSCD scheme is reduced compared to that of the standard decoding scheme, as listed in table 2. When the channel condition is bad, such as $E_b/N_0=0.9dB$ and $E_b/N_0=1.0dB$, the reduced number of belief-propagation iteration is big, say, to be 25 and 9, respectively. These reduced iterations may well compensate for the computations consumed by the HMS estimations in the proposed scheme, so the proposed scheme has approximately the same computation complexity as the standard decoding scheme.

Therefore, the proposed JSCD scheme achieves better performance in PSNR and robustness than the standard decoding scheme with approximately the same computation complexity, thus providing a more effective way for wireless multimedia transmissions.

Table 2. Average number of iteration for the standard decoding scheme and the proposed JSCD scheme with 10 iteration /HMS.

E_b/N_0 scheme	0.9 dB	1.0 dB	1.05 dB	1.1 dB	1.15 dB	1.2 db	1.3 dB
JSCD Scheme	69.80	35.94	29.10	25.12	22.50	20.97	18.68
SD Scheme	94.68	45.39	33.92	27.71	24.09	21.91	19.31

4. CONCLUSIONS

In this paper, a joint source-channel decoding scheme is proposed for MPEG-2 transmission. This scheme combines a hidden Markov source estimation algorithm and a low-density parity-check coding with an iterative estimation/decoding scheme. We have shown that the proposed scheme achieves much better PSNR performance than the separation decoding with approximately the same computational complexity. Since the encoding process is unchanged, the proposed JSCD scheme can be conveniently adopted for existing multimedia communication systems.

ACKNOWLEDGEMENT

This work was supported in part by the National Science Foundation (60328103) and the Tsinghua Ph.D. Candidate Research Innovation Foundation (092406618).

REFERENCES

1. S. Vembu, S. Verdu, and Y. Steinberg, "The source channel theorem revisited," *IEEE Trans. on Inform. Theory*, vol. 41, no.1, pp. 44–54, Jan. 1995.
2. J. Ribas-Corbera and S. Lei, "Rate control in DCT video coding for low-delay communications," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, pp. 172-185, Feb. 1999.
3. S. Aramvith, I-M. Pao, and M. T. Sun, "A rate control scheme for video transport over wireless channels," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, pp. 569-580, May 2001.
4. M. J. Ruf and J. W. Modestino, " Operational rate-distortion performance for joint source and channel coding of image," *IEEE Trans. on Image Processing*, vol. 8, pp. 305-320, March 1999.
5. J. Cai and C. W. Chen, "Robust joing source-channel coding for image transmission over wireless channels," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, pp. 962-966, Sep. 2000.
6. J. Garcia-Frias and J. D.Villasenor, "Joint turbo decoding and estimation of hidden Markov sources," *IEEE J. Select. Areas Commun.*, vol. 19, no. 9 pp. 1671-679, Sept. 2001.
7. L. Yin, J. Lu, and Y. Wu, "Combined Hidden Markov Source Estimation and Low-Density Parity-Check Coding: A Novel Joint Source-Channel Coding Scheme for Multimedia Communications," *Wireless Communications and Mobile Computing*, (*Special Issue: Multimedia Over Mobile IP*.), vol. 2, no. 6, pp. 643-650, Sep. 2002.
8. L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition" *Proceedings of the IEEE.* , vol. 77 , no. 2 , pp. 257 – 286, Feb. 1989.
9. ISO/IEC 13818-1:2000 *Information technology -- Generic coding of moving pictures and associated audio information: Systems*
10. ISO/IEC 13818-2:2000 *Information technology -- Generic coding of moving pictures and associated audio information: Video*
11. R. G. Gallager, *Low-density parity-check codes*, M.I.T. press, 1963.
12. D. J. MacKay and R. M. Neal, "Near Shannon limit performance of low density parity check codes," *Electronics Letters*, vol. 32, no. 18, pp. 1645-1646, Aug. 1996.
13. D. J. C. Mackay, "Good Error-Correcting Codes Based on Very Sparse Matrices", *IEEE trans. on information theory*, vol. 45, no.2, pp.399-431, Mar. 1999.
14. T. Richardson, M. Shokrollahi, R. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," *IEEE trans. on information theory*, vol. 47, no. 2, pp. 619-637, Feb. 2001.
15. J. L. Fan, *Constrained Coding and Soft Iterative Decoding,* Kluwer Academic Publishers New York, 2001.
16. ISO/IEC TR 13818-5:1997 Information technology -- Generic coding of moving pictures and associated audio information -- Part 5: Software simulation.