A Scalable Motion-Compensated Subband Image Coder

Kenji Tsunashima, Joseph B. Stampleman, and V. Michael Bove, Jr.

Abstract—**Treating the temporal direction of subband-coded images has in the past involved a choice between temporal subband decomposition, which allows scalable (partial bitstream) decoding but may reduce the coding efficiency, or motion-compensated coding, which can be more efficient but is not scalable. We propose a motion-compensated subband coder which preserves the scalability of the representation, and demonstrate its application in conjunction with a novel bit allocation process.**

Keywords—**Image compression, subband coding, motion compensation, digital video, bit allocation.**

I: Introduction: Scalability

It is increasingly becoming recognized that the importance of representing video digitally lies not strictly in compression or immunity to noise, but also (and perhaps more significantly) in the ability of a digital representation to decouple the form in which the information is produced and stored or transmitted from the manner in which it is ultimately used.

A critical manifestation of this property is in the form of scalability.[1] Concisely stated, scalability means that useful images result when only a portion of the bit stream is decoded; this situation might occur because of a desire to transmit a signal over channels of differing capacity without re-encoding, a need to support decoders of differing computational complexity (perhaps down to software-only decoding on a general-purpose microprocessor) and differing display resolutions, or an application in which processing resources may flexibly be reallocated (*e.g.* displaying four images reduced to quarter-frame size with the same hardware that would decode one at full-frame size).

Spatiotemporal subband coders are scalable through the simple means of discarding entire subbands; depending on the quantization scheme employed it may also be possible to transmit some components in a partial (more coarsely quantized) form, so as to permit finer control of bandwidth.[2] Several researchers have instead treated the temporal dimension through motion-compensated predictive coding.[3],[4] In this case it must be noted that scalable decoding would break the basic assumption of predictive coders that the encoding process can track the state of the decoder so as to generate a correct prediction error signal. It may be possible to set up a predictive coder in such a way that the error resulting from a decoder's disregarding some of the components is not visually objectionable, but this approach appears difficult to optimize.

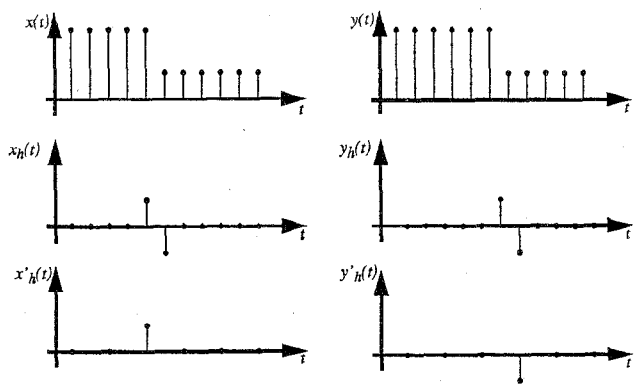

Fig. 1. Interaction between decimation and motion: $y(t)$ is $x(t)$ delayed by one sample, however, the decimated highpass subbands $y'_h(t)$ and $x'_h(t)$ are very different.

II: Theory

Ideally, we might wish to put each spatial subband into a separate motion compensation loop. Each subband of a predicted frame would be predicted from only the corresponding subband in the previous frame. Unfortunately for this simple approach, subband decomposition is a space-variant process, and the effect of decimation is that translational motion in an image does not result in simple translations in the resulting subbands.

A simplified example may help illustrate this phenomenon (Fig. 1). Let $x(t)$ be a piecewise-constant one-dimensional function that changes value at $(t = 0)$, and let the filter bank (for clarity) be the Haar transform. The lowpass subband will reflect the pairwise average values of the signal, whereas the highpass band $x_h(t)$ will be all zeros except for a doublet at the point in at which the value of the input signal changes. The latter is then subsampled to form $x'_h(t)$, which will be zero everywhere except for one value that will be either positive or negative, depending upon sampling phase. For this example, let us assume that the value will be positive. Next, consider what happens for a shifted version of the original signal: $y(t) = x(t-1)$. The

Paper approved by Barry G. Haskell, the Editor for Image Communications Systems of the IEEE Communications Society. Manuscript received March 13, 1992; revised October 30, 1992.

This research was performed at the Media Laboratory of the Massachusetts Institute of Technology, Cambridge, Massachusetts, USA, and was supported by the Television of Tomorrow and Movies of the Future research consortia.

K. Tsunashima is now with Mitsubishi Consumer Electronics America Inc., Somerset, New Jersey, USA.

J. Stampleman is now with Apple Computer Inc., Cupertino, California, USA.

IEEE Log Number 9401591.

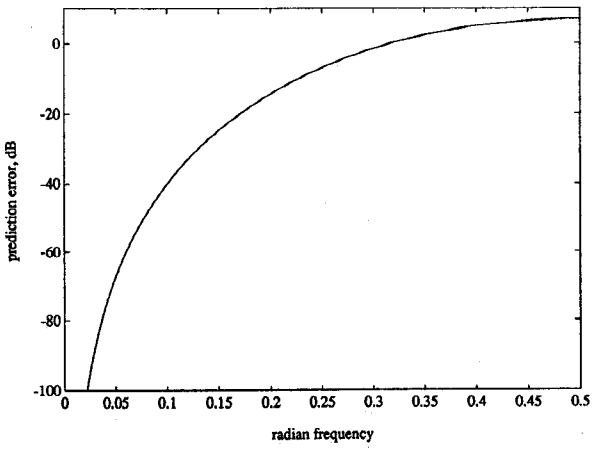

Fig. 2. Error of using half-band subband to predict the next frame of a cosine moving at a rate of one sample per frame.

lowpass subband will look virtually the same as it did for $x(t)$, but the one nonzero value in the decimated highpass band, $y'_h(t)$, will change sign. A similar analysis holds for two-dimensional signals.

For a more rigorous analysis, consider the effect of a shifting sinusoidal input on a two-band subband decomposition with an ideal lowpass/highpass filter pair. Since the filters are ideal, they will not alter the input in any way if it is within the bandpass region of the filter. Let the input signal be

$$x(t) = \cos(2\pi ft). \tag{1}$$

Assuming f is within the passband of a filter, the decimated output of that filter is

$$x'_h(t) = \cos(4\pi ft). \tag{2}$$

We will again use $y(t)$ to denote a one-sample-shifted version of $x(t)$, so its decimated version is

$$y'_h(t) = \cos(2\pi f(2t-1)). \tag{3}$$

Since a one-sample shift corresponds to a half-sample shift for the same signal at half resolution, $\tilde{y}'_h(t)$, a prediction for $y'_h(t)$ is formed by using bilinear interpolation to shift $x'_h(t)$ by half a sample. This operation corresponds to a straightforward use of motion compensation on the subbands, and we can now write the error of the prediction as $e(t)$:

$$\tilde{y}'_h(t) = \frac{x'_h(t) + x'_h(t-1)}{2} = \frac{\cos(4\pi ft) + \cos(4\pi f(t-1))}{2}, \tag{4}$$

$$e(t) = y'_h(t) - \tilde{y}'_h(t) = \cos(2\pi f(2t-1)) - \frac{\cos(4\pi ft) + \cos(4\pi f(t-1))}{2}. \tag{5}$$

Through trigonometry, $e(t)$ can be simplified, and then $e^2(t)$ can be integrated in order to obtain the mean-squared error of the prediction:

$$e(t) = 2\sin^2(\pi f)\cos(2\pi f(2t-1)), \tag{6}$$

$$\bar{e}^2(t) = 2f\int_0^{1/2f} e^2(t)dt \tag{7}$$

$$\bar{e}^2(t) = 4\sin^4(\pi f)\left(\frac{1}{2} + 2f\left(\frac{\sin(4\pi(1-f)) + \sin(4\pi f)}{8\pi}\right)\right) \tag{8}$$

$$\bar{e}^2(t) = 2sin^4(\pi f). \tag{9}$$

It is notable that this MSE expression is a function of frequency. See Fig. 2.

III: Scalable Motion Compensation

Noting the problems inherent in straightforward application of motion compensation to the subband structure, we suggest instead upinterpolating sampled subbands (to remove the effects of the decimation), applying the motion compensation, and then resampling. This idea forms the basis for our scalable motion-compensated subband coder.

The subband structure we use consists of seven spatial frequency bands as shown in Fig. 3. This recursive structure is formed by applying a quadrature mirror filter (QMF) separably in the horizontal and vertical directions.[5] The structure of the coder (Fig. 4) aims to preserve scalability by motion compensating the images at full size, half size and quarter size. The lowest subband (subband 0) is applied to an independent motion compensation DPCM loop. The reconstructed lowest subband signal is used together with reconstructed subbands 1-3 to rebuild a mid-resolution picture. Motion compensation is also applied to the reconstructed mid-resolution picture. The motion-compensated mid-resolution picture is decomposed into subbands to obtain prediction signals for subbands 1-3. The same motion compensation procedure is applied to encode subbands 4-6. To reconstruct the low resolution picture on the receiving end, only the encoded data corresponding to subband 0 and motion vectors are needed. In addition to subband 0, the encoded data corresponding

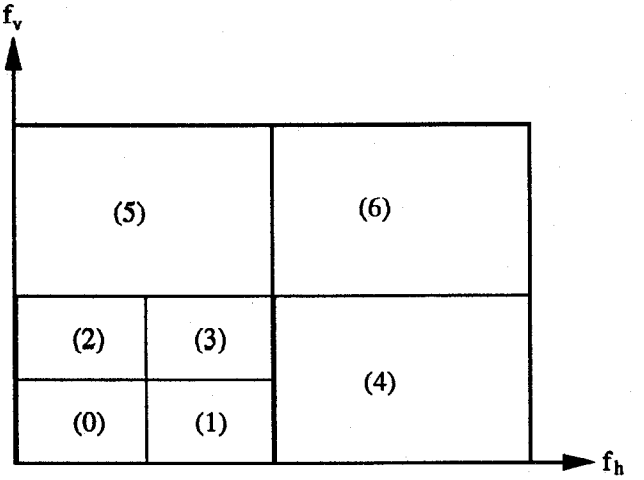

Fig. 3. Separable subband decomposition used in this paper.

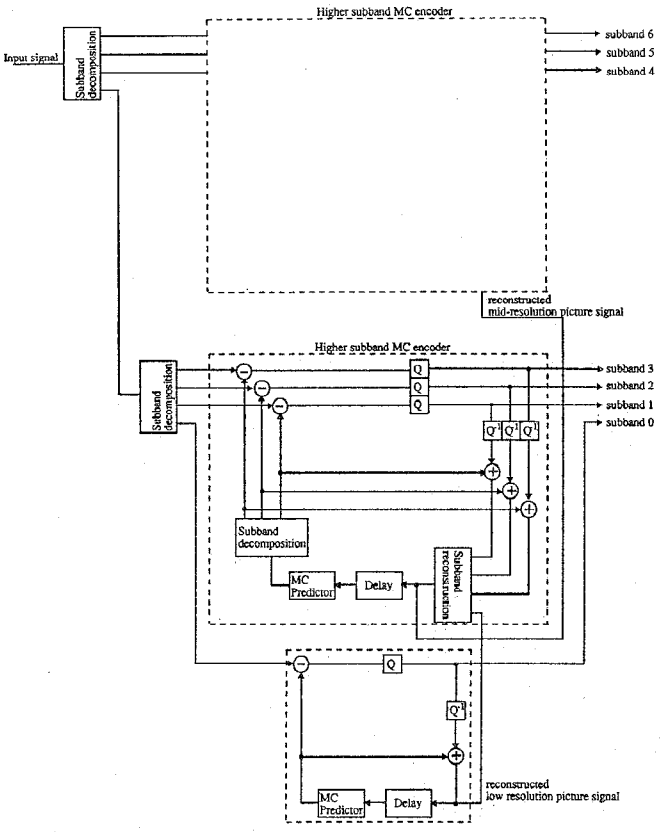

Fig. 4. Block diagram of the scalable motion-compensated subband coder used in this paper.

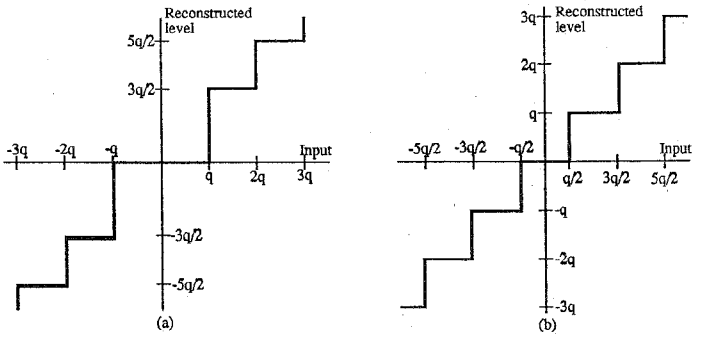

Fig. 5. Quantizer characteristics. (a) Quantizer with deadzone used for all subbands except baseband. (b) Quantizer used for baseband prediction error.

to subbands 1-3 are used to reconstruct a mid-resolution picture. Further, subbands 4-6 are added to obtain a full-resolution picture.

The form of this algorithm is related to a hierarchical pel-recursive structure proposed by Woods and Naveen,[6] though scalability was not their stated application, and they did not interpolate the images to increase accuracy as will be shown in a later section. Gharavi [3] for scalability reasons suggests predictively coding individual subbands without interpolation. We also report a variant of our scalable algorithm, which works with a spatiotemporal subband structure, in [7].

IV: Bit Allocation

In this section, we describe the bit allocation method for our scalable motion-compensated subband coder. Reference [8] explains the use of Lagrange multipliers to maximize the subband coding gain within a constrained number of bits. Using this method, we calculate the optimum bit allocation for each subband based on its variance. Here, we introduce a novel bit allocation scheme that directly determines the appropriate quantizer step size from the variance and the distortion parameter which is derived by using a method to be described in the next section. Given the distortion parameter, the target distortion for each subband is set by dividing the distortion parameter by the factor indicated in Table 1.

The coefficients shown in Table 1 are determined based on the following reasoning: for a bit allocation obtained through the Lagrange multipliers method, the calculated distortion is proportional to the size of the subband, *e.g.*, the distortion of subband 6 is four times as large as that of subband 3. Since the baseband (subband 0) is considered to be visually important, we put an appropriately large weighting on the baseband component.

All subbands except the lowest frequency subband (baseband) in both intra-frame-coded frames and prediction frames are quantized by a scalar quantizer with a center deadzone. The characteristics of this quantizer are shown in Fig. 5(a). For the lowest frequency subband in an intra-frame, the 2D-DPCM method is employed, with pixel configuration as shown in Fig. 6. The prediction value x is computed by

$$x = \frac{1}{2}A + \frac{1}{4}(C + D), \tag{10}$$

and its error is quantized by a uniform scalar quantizer as shown in Fig. 5(b). The same quantizer is also used to quantize the motion-compensated prediction error of the baseband signal.

Given a certain distribution of subband data, we can obtain the quantizer step size q for which the quantization noise is equal to the target distortion level. We assume the probability density function for high frequency sub-

Table 1. Weighting factors for the subbands. Note that the chrominance subbands at a given level are half the size (in each dimension) of the corresponding luminance subbands.

subband	*weighting*	
	luminance	chrominance
(00)	4.0	4.0
(01)	1.0	1.0
(02)	1.0	1.0
(03)	1.0	1.0
(04)	0.25	0.25
(05)	0.25	0.25
(06)	0.25	0.25

bands and the DPCM-encoded baseband can be well approximated by the generalized Gaussian distribution given below [9] :

$$p(x) = \left[\frac{\alpha\eta(\alpha,\beta)}{2\Gamma(1/\alpha)}\right] exp\{-[\eta(\alpha,\beta)|\,x\,|]^{\alpha}\} \qquad (11)$$

where

$$\eta(\alpha,\beta) = \beta^{-1}\left[\frac{\Gamma(3/\alpha)}{\Gamma(1/\alpha)}\right]^{\frac{1}{2}}, \qquad (12)$$

and α – which is always positive – is the shape parameter, β is the standard deviation, and $\Gamma(\cdot)$ is the gamma function. Sometimes – especially for low bit rates – the variance of a high frequency subband falls below the assigned distortion. In this case, the above method cannot give the step size, and we adopt the following approximation. Outside the center dead zone, we regard the distribution as slowly varying and approximate the average quantization noise by

$$d = \frac{q^2}{12}. \qquad (13)$$

We then compute the step size from the target distortion d using the following equation:

$$q = \sqrt{12d}. \qquad (14)$$

V: Rate Control

The distortion parameter used in the previous section is determined both from the relationship between the bit rate and the distortion (to be given below) and from the buffer fullness which is evaluated in the manner described in the MPEG Video Simulation Model.[10] We estimate the distortion for subband k, d_k, under the bit rate R_k by the following equation:

$$d_k = \varepsilon \cdot 2^{-2R_k} \cdot \sigma_k^2 \qquad (15)$$

where ε is constant and σ_k^2 is the variance of subband k.

Suppose that the bit rate for subband k in frame n is R_{kn} when the distortion parameter d_{0n} is used. The bit rate R_n for the entire image n is given by

$$R_n = \frac{1}{N}\sum_k n_k R_{kn}, \qquad (16)$$

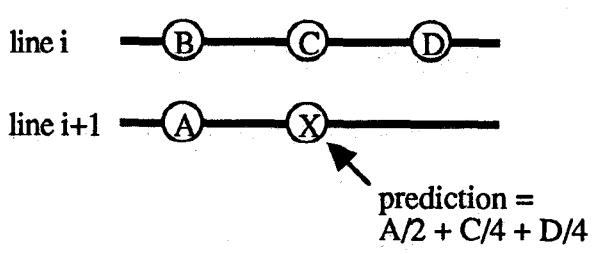

Fig. 6. Pixel configuration used for DPCM coder.

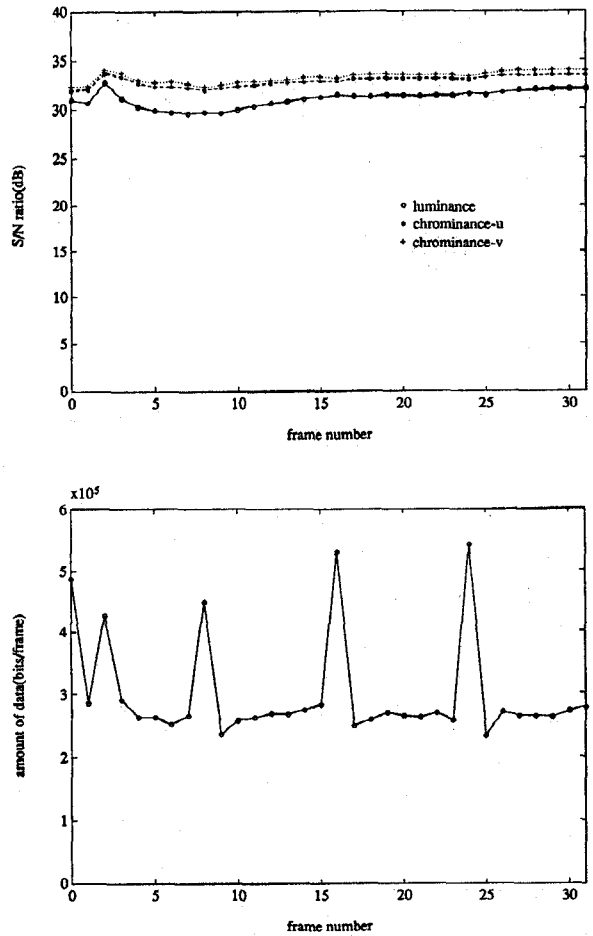

Fig. 7. Results for "Mobile & Calendar" sequence with target bit rate of 9Mbps. (a) Variation in signal-to-noise ratio. (b) Amount of coded data for each frame.

where N is the total number of pixels in a frame and n_k is the number of values in the subband k. To achieve the target bit rate R_t, we add the constant value ΔR to each subband:

$$R_t = \frac{1}{N}\sum_k n_k\,(R_{kn} + \Delta R)\,. \qquad (17)$$

From (16) and (17), ΔR can be expressed by

$$\Delta R = R_t - R_n. \qquad (18)$$

Therefore, by using the distortion parameter d_{1n} given by

$$d_{1n} = d_{0n} \cdot 2^{-2\Delta R} \qquad (19)$$

$$= d_{0n} \cdot 2^{-2(R_t - R_n)}, \qquad (20)$$

the frame n can be encoded at the target bit rate R_t.

Assuming that the characteristics of a picture do not change very much between two neighboring frames, we apply the distortion parameter d_{1n} together with variances for the frame n to calculate quantizer step sizes which are used to encode the frame $n+1$, except just after an intra-coded frame (admittedly the statistics of the subband-coded error signal can change more rapidly than those of the original

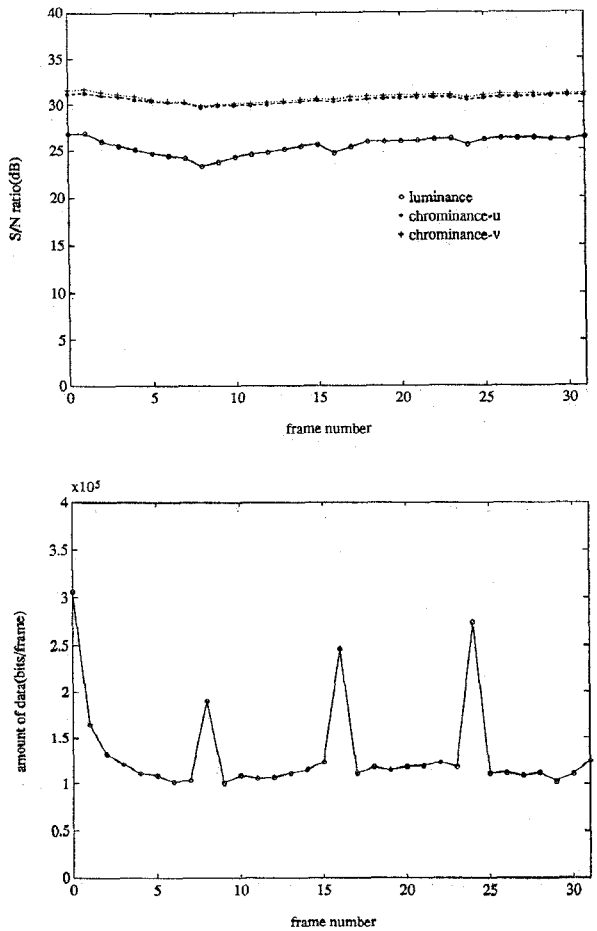

Fig. 8. Results for "Mobile & Calendar" sequence with target bit rate of 4Mbps. (a) Variation in signal-to-noise ratio. (b) Amount of coded data for each frame.

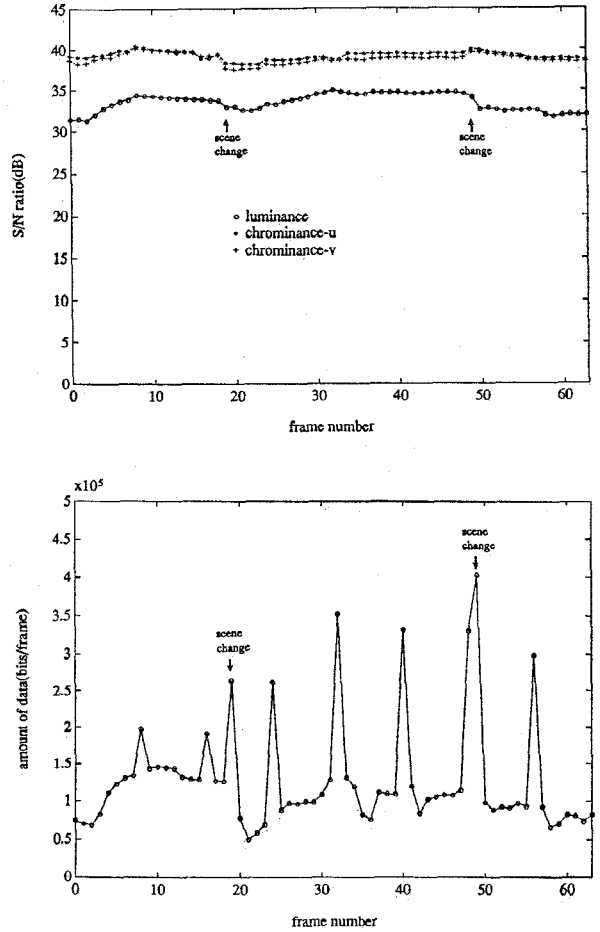

Fig. 9. Results for "Table Tennis" sequence with target bit rate of 9Mbps. Two scene changes are included in this sequence. (a) Variation in signal-to-noise ratio. (b) Amount of coded data for each frame.

images; we have also tried using the variances from the current frame, though this introduces added coding delay and in our experiments has produced minimal change in image quality).

Because the above equations do not necessarily give an accurate estimate for a bit rate especially for large ΔR, we restrict the change of the distortion parameter calculated from (20) to the range given by

$$\frac{2}{3}d_{0n} < d_{1n} < \frac{3}{2}d_{0n}. \tag{21}$$

Considering applications which require random access, forward/reverse searches, and reverse playback, we would like to insert an intra-frame periodically. The target bit rates for an intra-frame R_i and for a prediction frame R_p should be determined to equate their distortions. From (15), this condition can be expressed by the following equation:

$$2^{-2R_{ki}} \cdot \sigma_{ki}^2 = 2^{-2R_{kp}} \cdot \sigma_{kp}^2 \tag{22}$$

where the subscript i stands for an intra-frame and p for a predicted frame. Therefore, the following value is constant regardless of bit rates:

$$R_{ki} - R_{kp} = log_2 \frac{\sigma_{ki}}{\sigma_{kp}}. \tag{23}$$

Under the assumption that the statistical characteristics do not change drastically within a group of pictures that consists of N_i intra-frames and N_p predicted frames, we have

$$R_i - R_p = constant \tag{24}$$

and

$$N_i R_i + N_p R_p = N_t R \tag{25}$$

where R is the target average bit rate, and N_t is the total number of frames in a group of pictures, *i.e.*,

$$N_t = N_i + N_p. \tag{26}$$

To obtain the estimate for the constant in (24), we use the bit rate of the encoded data for the intra-frame R_{ie} and the first predicted frame R_{pe} in a group of pictures. We encode these frames with the same quantizing stepsize so that the distortion is almost the same for both the intra-frame and

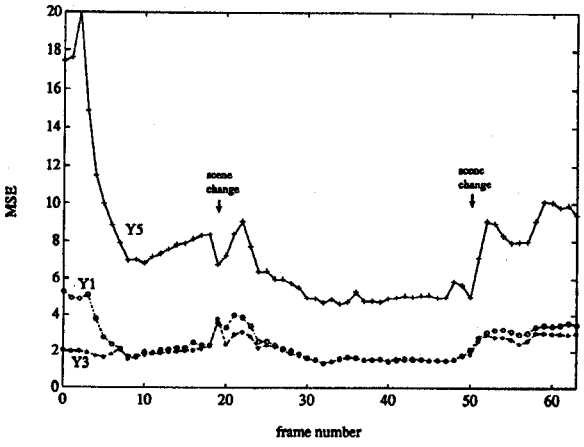

Fig. 10. Mean-squared errors for subbands 1, 3, and 5 (see Fig. 3) of the luminance signal for the experiment in the preceding figure.

the predicted frame. Using R_{ie} and R_{pe}, we set

$$R_i - R_p = R_{ie} - R_{pe}. \tag{27}$$

From (25) and (27), the target bit rate for the intra-frame R_i and the predicted frame R_p can be computed by

$$R_i = R + \frac{N_p}{N_t}(R_{ie} - R_{pe}) \tag{28}$$

$$R_p = R + \frac{N_i}{N_t}(R_{ie} - R_{pe}). \tag{29}$$

The target bit rates R_i and R_p are updated at the end of the first predicted frame after an intra-frame. The target bit rate which is used to compute the distortion parameter is modified to control the average bit rate according to the buffer fullness which (as noted above) is evaluated in the same manner as in the MPEG Video Simulation Model. The buffer fullness measure B_r is evaluated at the end of every frame by adding the number of bits generated for the frame and subtracting the target number of bits which is the product of the number of pixels in a frame and the target bit rate. Using the buffer fullness B_r, the target bit rate which is used for (20) is produced by

$$intra\ frame : R_t = R_i\left(C_1 - C_2\frac{B_r}{C_b}\right), \tag{30}$$

$$predicted\ frame : R_t = R_p\left(C_1 - C_2\frac{B_r}{C_b}\right), \tag{31}$$

where C_b is the output buffer capacity, and C_1 and C_2 are constants. We choose C_1 and C_2 so that $R_t = R_i = R_p$ for $B_r/C_b = 1/4$ and $R_t = 1.3R_i = 1.3R_p$ for $B_r = 0$. That is, we set $C_1 = 1.3$ and $C_2 = 1.2$.

VI: Experimental Results

Software simulation of the above algorithm has been carried out on several standard test sequences at CCIR-601

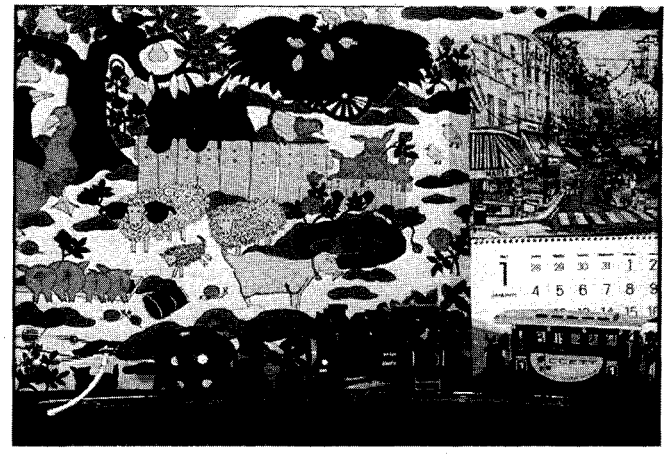

Fig. 11. Original frame from "Mobile & Calendar" sequence used in experiments.

Fig. 12. Frame from "Mobile & Calendar" coded at target bit rate of 9Mbps.

Fig. 13. Frame from "Mobile & Calendar" coded at target bit rate of 4Mbps.

TABLE 2. AVERAGE SIGNAL-TO-NOISE RATIOS FOR THE SCALABLE MOTION COMPENSATED CODER OPERATING ON TWO DIFFERENT SEQUENCES AT TWO DIFFERENT BIT RATES.

target bit rate(Mbps)			9.0	4.0
mobile & calendar	S/N (dB)	Y	31.1	25.5
		U	32.9	30.5
		V	33.3	30.8
	average bit rate(Mbps)		8.99	3.98
flower garden	S/N (dB)	Y	34.4	29.0
		U	35.9	32.2
		V	35.9	33.5
	average bit rate(Mbps)		8.91	4.00

TABLE 3. AVERAGE SIGNAL-TO-NOISE RATIOS WHEN THE ALGORITHM IS MODIFIED TO REDUCE THE COMPUTATIONAL COMPLEXITY (SEE TEXT).

target bit rate(Mbps)			9.0	4.0
mobile & calendar	S/N (dB)	Y	30.4	24.8
		U	32.6	30.0
		V	32.9	30.2
	average bit rate(Mbps)		9.05	4.03
flower garden	S/N (dB)	Y	34.0	28.5
		U	35.6	31.9
		V	35.6	33.2
	average bit rate(Mbps)		8.95	4.01

resolution. The subband structure shown in Fig. 3 is produced with the 9-tap QMF pair from [11]. To make the horizontal and vertical resolution of the chrominance components approximately equal, they are vertically filtered and decimated by a factor of two using the decimation filter specified in the MPEG draft.[10] Both fields of an interlaced frame are coded together. An intra-frame is inserted every 8 frames, *i.e.*, $N_i = 1$, $N_p = 7$. Motion estimation is performed using the hierarchical block matching method as in [12]: after one level of subband decomposition, the baseband signal is used with a block size of 8×8 in order to give an initial estimate which is then refined by searching the original image with a block size of 16×16. The resulting set of vectors (with half-pixel accuracy) is used at all three resolution levels of the motion compensated coder. As shown in Fig. 4, the lowest level image (subband 0) is interpolated back up to the full original size, warped using these vectors, and then reduced back by a factor of four in each dimension to use as the lowest level prediction image. This prediction plus subbands 1, 2, and 3 is then interpolated to full size, motion compensated, and then decomposed again to provide the predictions for subbands 1, 2 and 3.

TABLE 4. AVERAGE SIGNAL-TO-NOISE RATIOS FOR THE NONSCALABLE MOTION COMPENSATED CODER (SUBBAND CODING THE FULL-FRAME ERROR SIGNAL).

target bit rate(Mbps)			9.0	4.0
mobile & calendar	S/N (dB)	Y	31.8	26.2
		U	33.2	30.8
		V	33.6	31.0
	average bit rate(Mbps)		8.94	3.94
flower garden	S/N (dB)	Y	34.7	29.2
		U	36.2	32.3
		V	36.1	33.4
	average bit rate(Mbps)		8.88	3.98

The variance of the motion estimation error signal for a high frequency subband is compared with that of the original subband signal. For the baseband signal, this comparison is made between the motion estimation error signal and the output signal of the 2D-DPCM. The signal with the smaller variance is selected and coded. For all subbands, the output signal of the quantizer is coded with an arithmetic coder which is based on the JPEG Q-coder.[13].

Fig. 7 shows the variation of the signal-to-noise ratio (a) and the amount of encoded data for each frame (b) for the first thirty-two frames of the "Mobile & Calendar" test sequence. The SNR is evaluated from mean square error (MSE) by

$$SNR = 20 \log_{10} \frac{255}{\sqrt{MSE}}. \tag{32}$$

The target bit rate of this simulation is 9Mbps. Peaks which appear periodically in Fig. 7(b) correspond to the intra-frames. Results at a rate of 4Mbps are shown in Fig. 8. The average SNR and the actual average bit rate for these simulations are shown in Table 2 together with the results obtained for the test sequence "Flower Garden". To demonstrate the stability of the bit rate control, we encoded a portion of the test sequence "Table Tennis" which includes two scene changes (Fig. 9). Fig. 10 shows the MSE for subbands 1, 3, and 5 of the luminance signal. Subbands 1 and 3 have a similar MSE as expected. Images from the "Mobile & Calendar" sequence are illustrated in Figs. 11, 12, and 13.

It may seem that the above application of motion compensation is unnecessarily demanding of computation and memory, and a simpler if less accurate alternative exists in applying the motion compensation to the lower-resolution images rather than interpolating them to full size. In our experiments we also compared a variant of the algorithm in which the lowest-resolution image is interpolated one level and motion compensated (giving 1/4-pixel accuracy), and the mid- and full-resolution images are motion compensated at 1/2-pixel accuracy without further interpolation (the corresponding accuracies for the original algorithm are 1/2, 1/4, 1/8, from high to low resolution respectively). Bit allocation and rate control were the same as above. The interpolation of the smallest image was found necessary to give reasonable coding gains from the motion estimation. Results for this method are given in Table 3. Also, for comparison, we applied the motion compensation nonscalably: the full-resolution image was motion compensated and the errors were subband decomposed and coded. These results are shown in Table 4.

VII: Conclusions

We have described and analyzed a method for using motion compensation in conjunction with scalable subband coding. In our experiments, the SNR cost of this scalability was quite small, though it is apparent that more investigation is needed for a full exploration of the quality/complexity/scalability tradeoffs involved. We also proposed a bit rate control method that regulates the bit rate through the target distortion for each subband. The results show that this method has the ability to achieve both good SNR and stable bit rate.

Acknowledgment

The authors wish to thank our many co-workers who provided useful comments on this research, and developed the tools to make it possible.

References

[1] V. M. Bove, Jr. and A. B. Lippman, "Scalable Open-Architecture Television," *SMPTE Journal*, vol. 101, pp. 2-5, Jan. 1992.

[2] A. Singh and V. M. Bove, Jr., "Multidimensional Quantizers for Scalable Video Compression," *IEEE Journal on Selected Areas in Communications*, vol. 11, no. 1, pp. 36-45, Jan. 1993.

[3] H. Gharavi, "Subband Coding Algorithms for Video Applications: Videophone to HDTV-Conferencing," *IEEE Trans. Circuits Syst. for Video Technology*, vol. 1, no. 2, pp. 174-183, June 1991.

[4] F. Lallauret and D. Barba, "Motion compensation by block matching and vector post-processing in sub-band coding of TV signals at 15Mbit/s," in *Proc. SPIE*, vol. 1605, pp. 26-36, Nov. 1991.

[5] M. Vetterli, "Multi-Dimensional Sub-Band Coding: Some Theory and Applications," *Signal Processing*, vol. 6, no. 2, pp. 97-112, Feb. 1984.

[6] J. W. Woods and T. Naveen, "Subband Encoding of Video Sequences," in *Proc. SPIE*, vol. 1199, pp. 724-732, Nov. 1989.

[7] J. B. Stampleman, "Scalable Video Compression," S.M. thesis, Massachusetts Institute of Technology, 1992.

[8] N. S. Jayant and P. Noll, *Digital Coding of Waveforms*, Englewood Cliffs, NJ: Prentice-Hall, pp. 490-494, 1984.

[9] N. Tanabe and N. Farvardin, "Subband Image Coding Using Entropy-Coded Quantization over Noisy Channels," in *Proc. ICASSP'90*, Apr. 1990.

[10] ISO-IEC JTC1/SC2/WG8, *MPEG Video Simulation Model Three (SM3)*, 1990.

[11] E. H. Adelson, E. Simoncelli, R. Hingorani, "Orthogonal Pyramid Transforms for Image Coding," in *Proc. SPIE*, vol. 845, pp. 50-58, Oct. 1987.

[12] H. Hölzlwimmer, A. V. Brandt, W. Tengler, "A 64Kbits Motion Compensated Transform Coder Using Vector Quantization with Scene Adaptive Codebook," in *Proc. ICC'87*, June 1987.

[13] ISO/JPEG, *JPEG Technical Specification, Revision 8*, JPEG-8-R8, Aug. 1990.

Kenji Tsunashima was born in Osaka, Japan in 1960. He received his B.S. and M.S. degrees in Electronic Engineering from Kansai University, Osaka, Japan in 1983 and 1985 respectively. He joined Mitsubishi Electric Corporation in 1985, and since then, he has been engaged in working on VCR development at the corporation's Consumer Electronics Laboratory. From 1990 to 1992 he was a research affiliate at the Media Laboratory at the Massachusetts Institute of Technology, Cambridge MA.

Joseph B. Stampleman was born in Montreal, Canada in 1967. He recieved B.S. degrees in Applied Mathematics and EECS from Massachusetts Institute of Technology in 1989, and a Master's degree in Media Arts and Sciences from MIT in 1992. He began working on image compression at the MIT Media Laboratory as an undergraduate in 1986. In 1992 he completed his Master's degree work at the Media Laboratory and is now working in the Video Architecture Group at Apple Computer Inc.

V. Michael Bove, Jr. holds a B.S.E.E., an S.M. in Visual Studies, and a Ph.D. in Media Technology, all from the Massachusetts Institute of Technology. In 1989 he was appointed to the faculty of MIT, where he is currently associate professor of media technology, working in the MIT Media Laboratory. He is the current holder of the Sony Corporation Career Development Professorship. His research includes advanced digital television systems, design of video processing hardware and software, and scene modeling. He holds patents on inventions relating to video recording and hardcopy, and has served on several professional and government committees.