obtained. In fact, the curves show that the proposed method outperforms those in [6] and [7] by a large margin—behaviour observed also at lower rates.

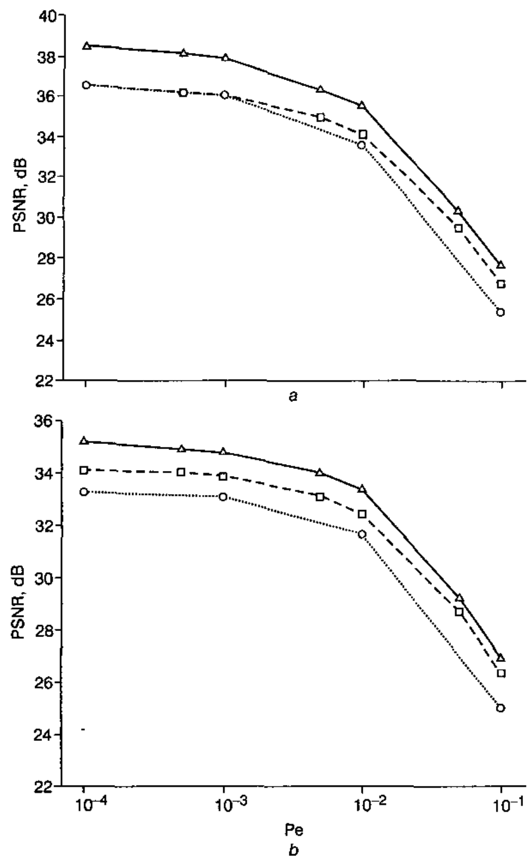

Fig. 2 *Proposed scheme performance for various BER at 1 bpp*
—△— proposed
---□-- proposed - no class.
···○··· A-RQ
a Lena
b Goldhill

Conclusion: We have proposed a novel scheme for joint source/channel coding of still images over BSCs. We have shown that with simple classification and bit allocation strategies, significant improvements can be achieved, thus narrowing the gap between tandem schemes such as [1] and purely joint source/channel coding approaches. In [4] a natural extension of [7] using a classification based on coefficient significance is presented. For moderate bit error probability their results are similar to the results in this Letter. However, the method in [4] still suffers from bad bit allocation when the channel is very noisy, thus producing worse results than our scheme (by around 1.6 dB for $Pe = 0.1$).

17 April 2002
Electronics Letters Online No: 20020759
DOI: 10.1049/el:20020759

A.L.A. da Cunha and W.A. Finamore (*Center for Telecommunication Studies, Pontifical Catholic University of Rio, Rua M. S. Vicente 255, Rio de Janeiro, Brazil*)

E-mail: weiler@cetuc.puc-rio.br

E.A.B. da Silva (*Universidade Federal do Rio de Janeiro, PEE/COPPE/DEL/EE, Rio de Janeiro, RJ - Brazil*)

References

1 SHERWOOD, P.G., and ZEGER, K.: 'Progressive image coding on noisy channels'. Proc. IEEE Data Compress. Conf., March 1997, pp. 72–81
2 FARVARDIN, N.: 'A study of vector quantization for noisy channels', *IEEE Tran Inf. Theory*, July 1990, **36**, (4), pp. 799–809
3 WANG, MIN, and FISCHER, T.R.: 'Trellis-coded quantization designed for noisy channels', *IEEE Trans. Inf. Theory*, November 1994, **40**, (5), pp. 1792–1802
4 ABOUSLEMAN, G.P.: 'Adaptive image coding with robust channel-optimized trellis-coded quantization'. Proceedings, 29th Applied Imagery Pattern Recognition Workshop, 2000, 1999, pp. 109–114
5 CAI, JIANFEI, and WEN CHEN, CHANG: 'Uniform threshold tcq with block classification for image transmission over noisy channel', *IEEE Trans. Circuits Syst. Video Technol.*, January 2001, **11**, (1), pp. 105–110
6 CHEN, QING, and FISCHER, T.R.: 'Image coding using robust quantization for noisy digital transmission', *IEEE Trans. Image Process.*, April 1998, **7**, (4), pp. 496–505
7 LAM, TUYET-TRANG, ABOUSLEMAN, G.P., and LINA, L.J.: 'Image coding with robust channel-optimized trellis-coded quantization', *IEEE JSAC*, 2000, **18**, (6), pp. 940–951
8 JOSHI, R., JAFARKHANI, H., KASNER, J., FISCHER, T., FARVARDIN, N., MARCELLIN, M., and BAMBERGER, R.: 'Comparison of different methods of classification in subband coding of images', *IEEE Trans. Image Process*, 1997, **6**, (11), pp. 1473–1486
9 VAISHAMPAYAN, V.A., and FARVARDIN, N.: 'Optimal block cosine transform image coding for noisy channels', *IEEE Trans. Commun.*, March 1990, **38**, (3), pp. 327–336
10 DE QUEIROZ, R.L., and TRAN, T.D.: 'Lapped transforms for image compression' *in* RAO, K.R., YIP, P. (Eds.): The Handbook on Transforms and Data Compression' (CRC Press, Oct. 2000), Chap. 5

MPEG-4 video object-based rate allocation with variable temporal rates

Jeong-Woo Lee, A. Vetro, Yao Wang and Yo-Sung Ho

In object-based coding, bit allocation is performed at the object level and temporal rates of different objects may vary. The proposed algorithm deals with these two issues when coding multiple video objects (MVOs). The proposed algorithm is able to successfully achieve the target bit rate, effectively code arbitrarily shaped MVOs with different temporal rates, and maintain a stable buffer level.

Introduction: The problem that we consider in this Letter is the object-based rate-distortion (R-D) encoding of multiple video objects (MVOs) for MPEG-4 video coding. In a recent paper by Vetro *et al.* [1], models that estimate the distortion for coded as well as non-coded frames have been proposed. The frame-based optimisation process determines if it is better to code more frames with lower quality or fewer frames with higher quality. However, this algorithm is not directly applicable to the coding of MVOs. In [2], an algorithm that considered the trade-off in spatial and temporal quality for coding MVOs was presented; however, it did not address the possibility of coding the objects with varying temporal rates, i.e. every object in the current time is either coded or all objects are skipped. In this Letter we propose a framework that supports object-based coding with different temporal rates, aiming to improve the coding efficiency of an object-based coder.

Object-based coding with variable frameskip: When each video object (VO) has a different frameskip, we define the constraints on the rate and buffer. Let M denote the set of VOs, and L denote the complete set of time indices. $M(l)$ denotes the set of coded VO at time index l ($t = t_l$). The constraint on the rate can then be written as

$$\sum_{l \in L} \sum_{j \in M(l)} R_j(t = t_l) \leq R_{budget} \quad (1)$$

where $R_j(t = t_l)$ is the rates used for jth VO at time index l. (1) essentially says that the total rates for the coded VOs at all time instants within the specified time interval must be less than the calculated bit rate budget over that time interval.

To ensure that buffer overflow and underflow are avoided at every coded time, we have a set of buffer constraints which are given by

$$B_{i+l_0} + \sum_{j\in M(l)} R_j(t = t_{l_0} : t_l) < B_{\max}; \quad \forall l \in L \tag{2}$$

$$B_{i+l_0} + \sum_{j\in M(l)} R_j(t = t_{l_0} : t_l) - (l - l_0)R_{drain} > 0; \quad \forall l \in L \tag{3}$$

where B_{i+l_0} is the current buffer level, $B_{\max}$ is the maximum buffer size, and R_{drain} is the rate at which the buffer drains per time instant. $R_j(t = t_{l_0}:t_l)$ is the number of bits used for jth VO at time from l_0 to l.

The proposed constraints are valid for arbitrary frameskip factors. With these constraints, it is possible to formulate a problem that aims to minimise the overall distortion. Such a problem could be solved by searching over all valid combinations of frameskip factors and quantisation parameters (QP) within a specified period of time. However, the complexity is very high. Not only are there many combinations of frameskip factors and quantisation for each VO, but we must also track the individual time instants that each object will be coded.

Rate control: We consider the rate control for a restricted object-based framework. We refer to this framework as being restricted since a decision on the frameskip factor of a particular VO is made locally, i.e. without considering the various combinations of frameskip factors. With this framework, the proposed algorithm first determines the set of VOs to be coded at the next coding time, and then considers the bit allocation for each VO to be coded.

In frame-based coding, if the target bits are less than the header bits which are used to code the motion, shape and header information, then the encoder is forced to skip all VOs. However, the proposed object-based algorithm allows the encoder to code a portion of VOs because the bit allocation is performed at the object level. To support uniform quality from frame to frame, we restrict the current QP to the previous QP of the same VO [2]. In addition, we set the initial target based on the current buffer level; therefore, the buffer check is necessary to prevent buffer overflow and underflow.

In addition to the bits used for texture and motion, an object-based coder must consider the significant amount of bits that are used to code the shape information. In [3], the significance of this problem due to the high percentage of shape bits has been shown for low bit-rate coding conditions. Therefore, it is necessary to consider a buffer control strategy that determines the number of frames to be skipped. The frameskip rate, f_s, is increased, as any VO cannot be coded at the current time index. It should be noted that the frameskip rate for the proposed algorithm is smaller than that for the frame-based algorithm. This is because the part of all VOs can be coded at the object-based bit allocation. Let M_L denote all the complete set of partially coded VO. If M_L is defined as an empty set, therefore, the frameskip factor is increased.

The set of VOs to be coded at a particular time instant is selected based on the total distortion associated with each frame, which includes both coded distortion due to the quantisation error and non-coded distortion due to the skipped VOs. The distortion models presented in [2] for frame-based coding are used in this work for object-based coding.

The coded distortion for the jth VO is modelled by

$$D_c(Q_{j,i}) = a \cdot 2^{-2R_j(t_i)} \cdot \sigma^2_{z_{j,i}} \tag{4}$$

where $\sigma^2_{z_{j,i}}$ is the variance of the jth coded VO at time index i, $R_j(t_i)$ is the average rate for texture data, $Q_{j,i}$ denotes the quantisation parameter.

The non-coded distortion at time index k for the jth VO that was last coded at time index i is modelled by

$$D_s(Q_{j,i}, k) = D_c(Q_{j,i}) + E_j\{\Delta^2 z_{i,k}\} \tag{5}$$

where $E_j\{\Delta^2 z_{i,k}\}$ denotes the expected interpolation error between time index i and k.

The optimal set of VOs to be coded, $M^*(i+f_s)$, are those that satisfy

$$d_{M^*(i+f_s)}(Q, f_s) = \min_{M(i+f_s)\subset M_L} |d_{M(i+f_s)}(Q, f_s)| \tag{6}$$

$$d_{M(i+f_s)}(Q, f_s) = \sum_{j\in M(i+f_s)} D_{j,c}(Q_{j,i+f_s}) + \sum_{j\notin M(i+f_s)} D_{j,s}(Q_{j,i}, f_s) \tag{7}$$

The initial target bits, T, for the current time index is allocated based on the initial assumption that every VO is coded at the current time index [1]. If the jth VO is not coded at the previous coded time index, $\tilde{T}_{p,j}$ is assigned the coded bits determined from the actually coded time index. The target bits are then scaled according to the current buffer occupancy. To improve the coding efficiency of each VO, we now consider distributing the available bits to the objects to be coded. Let T_{hdr} be the number of header bits of the previous VOs belonging to the subset $M(i+f_s)$. To guarantee that the target for the jth VO is always larger than $T_{j,hdr}$, we propose to use the following equation for distributing bits to each object:

$$T_j = (T - T_{hdr}) \cdot (w_m \mathrm{MOT}_j + w_v \mathrm{VAR}_j) + T_{j,hdr} \tag{8}$$

The QP for each object is then calculated based on a quadratic rate-quantiser model, as described in [3].

Results: To evaluate the performance, we compared the proposed algorithm with the standard MPEG-4 reference VM5 algorithm. Fig. 1*a* shows the R-D curves. It is clear from the plots that the proposed algorithm has a better performance than the VM5 algorithm. This is because the proposed algorithm skips VOs based on buffer constraints as well as the minimum distortion. Fig. 1*b* shows the ratio of the actual coded bits to the target bits. We can observe that the actual bits are well matched to the target bits in the proposed algorithm, while there are significant fluctuations in the VM5 algorithm.

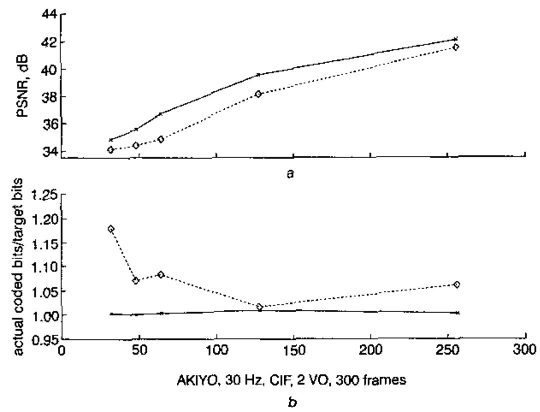

Fig. 1 *Performance comparisons*

a Comparison of R-D curves for AKIYO
b Ratio of actual coded bits to target bits
—— proposed algorithm
----- VM5 algorithm

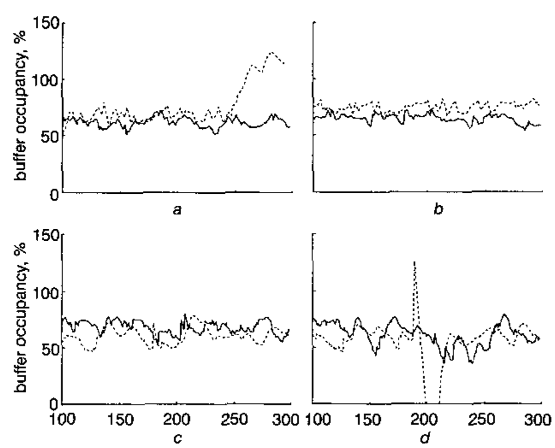

Fig. 2 *Comparisons of buffer occupancy*

a 32 kbit/s
b 64 kbit/s
c 128 kbit/s
d 256 kbit/s
—— proposed algorithm
----- VM5 algorithm

Fig. 2 shows the buffer occupancy from 32 to 256 kbit/s. As shown in the Figure, the buffer occupancy for the proposed algorithm is quite stable over the broad range of testing conditions. It was found from the result that the buffer has little chance of overflow or underflow. As shown in Figs. 2*a* and *d*, however, the buffer in the VM5 algorithm experiences at least one overflow.

Conclusions: We have proposed an MPEG-4 video object-based rate allocation method with variable temporal rates. The proposed algorithm enables the encoder to code arbitrarily shaped MVOs with different temporal rates where the actual coded bits are well matched to the target bits.

26 March 2002
Electronics Letters Online No: 20020752
DOI: 10.1049/el:20020752

Jeong-Woo Lee and Yo-Sung Ho (*K-JIST, 1 Oryong-dong, Puk-gu, Kwangju 500-712, Korea*)

E-mail: jeongwoo@kjist.ac.kr

Anthony Vetro (*MERL-Mitsubishi Electric Research Laboratories, Murray Hill, NJ, USA*)

Yao Wang (*Polytechnic University, Brooklyn, NY, USA*)

References

1 VETRO, A., WANG, Y., and SUN, H.: 'Rate-distortion optimized video coding considering frameskip'. Proc. ICIP, Thessaloniki, Greece, October 2001, pp. 534–537

2 LEE, J.W., VETRO, A., WANG, Y., and HO, Y.S.: 'Object-based rate allocation with spatio-temporal trade-offs'. Proc. VCIP'02, San Jose, CA, USA, January 2002

3 VETRO, A., SUN, H., and WANG, Y.: 'MPEG-4 rate control for multiple video objects', *IEEE Trans. Circuits Syst. Video Technol.*, 1999, **9**, pp. 186–199

Adaptive image enhancement for retinal blood vessel segmentation

Tusheng Lin and Yibin Zheng

Retinal blood vessel images are enhanced by removing the non-stationary background, which is adaptively estimated based on local neighbourhood information. The result is a much better segmentation of the blood vessels with a simple algorithm and without the need to obtain *a priori* illumination knowledge of the imaging system.

Introduction: Retinal blood vessel images can be used as a biological characteristic for personal identification in advanced security applications. It can also provide information for ocular fundus survey, diagnosis, and quantitative analysis. In all of the applications, threshold selection for segmentation is still a difficult problem [1] since the processing results vary greatly with images and no universal algorithm exists. Existing methods of segmentation, including Otsu's method, have advantages and limitations [2]. Consider a typical retinal blood vessel image shown in Fig. 1*a*, where the histogram is not bimodal. The segmentation result by Otsu's method is poor (Fig. 1*b*) with serious confusion between blood vessels and background. The method of two-dimensional matched filter [3] is able to achieve local enhancement but does not perform well in global preprocessing. It is also time consuming due to high complexity in matching templates. Digital subtraction angiography (DSA) [4] must take two exposures in order to subtract the mask image from the contrast image, and must perform image registration prior to subtraction.

We introduce a novel efficient and adaptive image enhancement method in which the background image is estimated and subtracted from the original image. The resulting image is free from background interference and suitable for subsequent segmentation.

Proposed method: The retinal blood vessel image $g(x, y)$ is modelled as having two components: the blood vessels $f_o(x, y)$ and the background $f_b(x, y)$:

$$g(x, y) = f_o(x, y) + f_b(x, y)$$

The blood vessel image $f_o(x, y)$ is obtained by subtracting the background from the original image, i.e.

$$f_o(x, y) = g(x, y) - f_b(x, y)$$

To estimate the background image $f_b(x, y)$, blood vessels in $g(x, y)$ are suppressed directly by neighbourhood intensity information as the following:

(i) Label a pixel as blood vessel pixel if its grey level is greater than or equal to the mean value of its eight neighbours. Change the value of the pixel to the minimum of neighbour; otherwise, keep its original value. The mathematical relationships for this step are given by

$$\bar{g}(x, y) = \frac{1}{(2r-1)(2c-1)} \sum_{m=y-r}^{y+r} \sum_{n=x-c}^{x+c} g(m, n)$$

$$g_1(x, y) \begin{cases} g(x, y) & \text{if } g(x, y) < \bar{g}(x, y) \\ g_{\min} & \text{if } g(x, y) \geq \bar{g}(x, y) \end{cases}$$

where $g(x, y)$ is the central (current) pixel, $\pm c$ and $\pm r$ are the width and height of the neighbourhood, $g_{\min}$ is the minimum grey level in the neighbourhood, $\bar{g}(x, y)$ is the mean grey level of the neighbourhood, $g_1(x, y)$ is the output image.

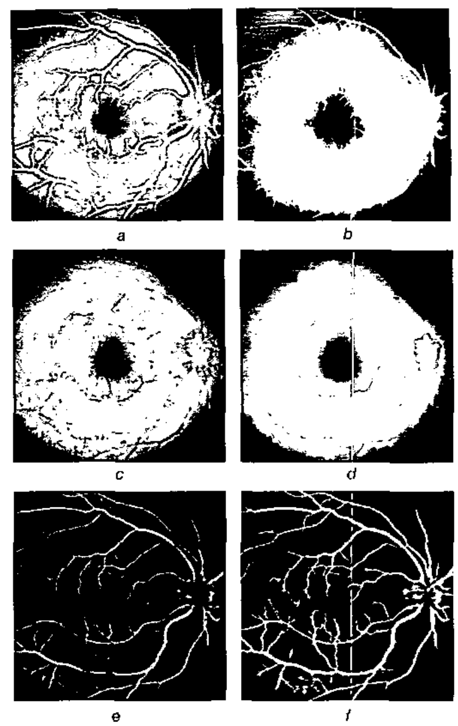

Fig. 1 *Example of retinal blood vessel image enhancement*
a Original image
b Segmentation result of applying Otsu's method to original image
c Background image after suppressing blood vessels
d Background image after smoothing blood vessel contours
e Retinal blood vessel image after background subtraction
f Segmentation result of applying Otsu's method to enhanced image

(ii) This processing tends to introduce negative bias to blood vessel pixels (Fig. 1*c*). Therefore we smooth the contour of blood vessels in the background estimate $g_1(x, y)$ by threshold averaging. Specifically, we calculate the sum of neighbourhood pixels with grey levels that are greater than a threshold $t_{\min}$, then average over all pixels in the neighbourhood. The threshold $t_{\min}$ is determined as the minimum of local average values from four stable corner neighbourhoods of the ocular fundus image. Let $g_2(x, y)$ be the output image. The mathematical expression for this algorithm is

$$g_2(x, y) = \begin{cases} t_{\min} & \text{if } \forall g_1(m, n) \leq t_{\min} \\ \frac{1}{(2r-1)(2c-1)} \sum_{m=y-r}^{y+r} \sum_{n=x-c}^{x+c} [g_1(m, n) \mid g_1(m, n) > t_{\min}] & \text{otherwise} \end{cases}$$

(iii) Subtract the estimated background image $g_2(x, y)$ from the original image $g(x, y)$ to obtain the enhanced blood vessel image $f_o(x, y)$.