社団法人　情報処理学会　研究報告
IPSJ SIG Technical Reports

2007－ＡＶＭ－57（３）
2007／7／19

Awareness 映像通信のための機能的階層符号化

岩橋 政宏† 高橋 知将† 鈴木 貴之†

†長岡技術科学大学工学部電気系 〒940-2188 新潟県長岡市上富岡町 1603-1
E-mail: †{iwahashi, ttakahashi}@nagaokaut.ac.jp

あらまし 本報告では，アウェアネス通信のために機能の点で階層化された符号化データを送受信する新しいタイプの映像通信システムを提案する．ここで、アウェアネスは縲細かつ半透明な人物領域を意味し，これを受信側で表示するのに必要最小限のデータを「基本階層」として伝送する．それ以外のデータは「拡張階層」として、これを追加伝送することで通常の映像通信が可能となる．基本階層には、(1)人物領域（前景）、かつ、(2)人物縲細化のための低周波数帯域、かつ、(3)人物半透明化のための高位（MSB 側の）ビットプレーンが含まれている．背景画像は予め伝送しておく．上記(2),(3)を JPEG2000 国際標準の帯域分割およびビットプレーン分解により実現する．
キーワード アウェアネス, JPEG2000, 映像, 通信, 階層符号化

Functionally Layered Video Coding for Awareness Video Communication

Masahiro IWAHASHI† Tomoyuki Takahashi† Takayuki Suzuki†

† Faculty of Engineering, Nagaoka University of Technology, 1603-1 Kamitomioka, Nagaoka, Niigata, 940-2188
E-mail: † {iwahashi, ttakahashi}@nagaokaut.ac.jp

Abstract This paper proposes a new video coding system which transmits functionally layered data for "awareness" communication. The system displays "awareness", which is defined as a "blurred" and "semi-transparent" person region by transmitting minimum data necessary for displaying the awareness. The minimum data contains (1) person region and (2) lower frequency band signals of the region for displaying blurred person and (3) higher bit-planes of the band signals for displaying blurred and semi-transparent person. Scenery without person is previously transmitted. It becomes possible to reduce redundancy in data transmission since (2) and (3) above are implemented by the "band decomposition" and the "bit-plane decomposition" of the JPEG2000 (JP2K) international standard respectively.
Keyword awareness, JPEG 2000, video, communication, layered coding

1. INTRODUCTION

Recently, various kinds of video surveillance systems have been developed to prevent crimes and disasters. However, privacy invasion problems have been also arising [1,2]. So far, to balance the "privacy" and "surveillance", large variety of video communication systems have been proposed [3-9]. Especially, since a new concept referred to "awareness" is introduced in [5], awareness communication systems which can display existence of a person have been proposed [6-9]. However, wide band width for communication (data amount to be transmitted) is required since these methods transmit video signal first and then transform it to the awareness. In addition, huge computational load to detect contour of a person or to recognize a person is required [9,10-12].

The international standard MPEG-4 for visual coding supports encoding each of objects in a video signal [13,14]. It enables to transmit person region only and transforming it into the awareness at a receiver. This approach can reduce data amount for transmission. However redundancy remains in transmitting information for displaying the awareness.

This paper proposes a new awareness communication system which can reduce communication bit rate by means of transmitting minimum data for displaying the awareness. So far, there are various types of awareness. The awareness in this paper means a "blurred" and "semi-transparent" person region. Utilizing the JPEG 2000 (JP2K) international standard [15,16], this paper proposes a new type of communication system which can display the awareness at minimum communication bit rate and display normal video by transmitting additional data.

In details, in awareness communication mode, the system transmits minimum data which contains (1) pixels in person region, (2) higher bit-planes for displaying semi-transparent person, and (3) a part of band signals (bands) for blurred person region signal. It becomes possible to reduce redundancy of data transmission since (2) and (3) above are implemented by the "bit-plane decomposition" and "wavelet transform" of the JP2K respectively.

JP2K encodes each of bands and each of bit-planes of an input video signal by the EBCOT composed of the bit modeling and the arithmetic coding [15,16]. Therefore, the proposed system in awareness mode transfers the bands and the bit-planes relevant to (2) and (3) above. On the other hand, the system in video communication mode transfers other remaining bands and bit-planes. As a result, it becomes possible to reduce data amount necessary to be transmitted for awareness and video communication.

2. PROPOSED COMMUNICATION SYSTEM

2.1. Transmitter and Receiver

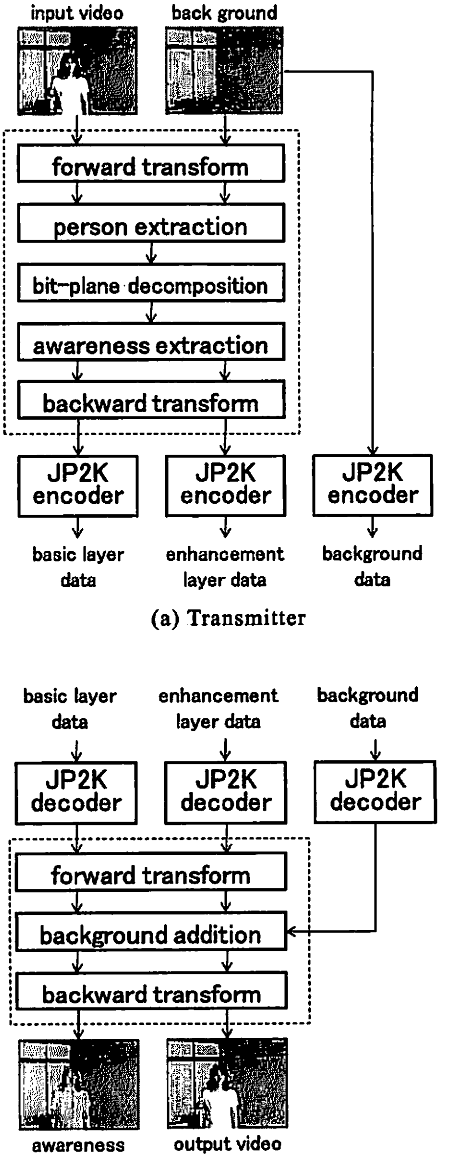
The transmitter illustrated in figure 1(a) encodes each of the "background" image signal and the "input video" signal by JP2K encoder after applying the "pre-processing" in the dotted line in the figure. The receiver in figure 1(b) synthesizes the "background" and the "video" after decoding each of them by JP2K decoder followed by the "post-processing" in the dotted line in the figure. It is our purpose to utilize the forward and backward wavelet transform, which are core technologies of the JP2K, at the pre- and post- processing so that the system can be implemented with relatively popular IP core and software sub-routines of the JP2K standard.

2.2. Forward / Backward Transform

The forward and backward wavelet transform at the pre- and post- processing are exactly same as the JP2K encoder and decoder. In case of using the 5/3 "lossless" filter, for example, a one *dimensional* input signal $X(n)$ is processed by

$$Y(2n+1) = X(2n+1) - \left\lfloor \frac{X(2n)+X(2n+2)}{2} \right\rfloor \quad (1)$$

$$Y(2n) = X(2n) + \left\lfloor \frac{Y(2n-1)+Y(2n+1)+2}{4} \right\rfloor \quad (2)$$



(a) Transmitter

(b) Receiver

Fig.1 Awareness video communication system.

and $Y(2n)$ as a low (L) frequency band and $Y(2n+1)$ as a high (H) frequency band are produced. Applying this processing vertically and horizontally to the two dimensional signal, four bands at the 1st stage 1LL, 1LH, 1HL and 1HH are produced. Furthermore, applying it to the 1LL recursively, four more bands 2LL, 2LH, 2HL and 2HH are generated. This procedure is repeated to S-th stage for $S>0$ [15,16].

In case of 9/7 "lossy" filter, similar bands are produced except the wavelet filter bank different from equations (1) and (2). The corresponding backward transform is defined as a inverse procedure of the forward transform [15,16]. Figure 2 indicates an example of the band decomposition. Bands of the background image are stored in a memory device as well as the n-th stage LL band for $s=1,2,3,\ldots,S$ for the person region extraction explained in figure 4.

2.3. Person Extraction

The "person extraction" in figure 1(a) subtracts "background" in figure 3(b) from "input video" in figure 3(a) to produce "subtracted" image in figure 3(c). The person region in figure 3(d), which is a binary image signal, is used to mask the bands in figure 3(e) to produce "masked" image in figure 3(f). As a result, encoded data amount of the non-person region in the bands is close to zero since JP2K outputs almost zero data for zero-padded region. It is possible to embed shape information into the tile header as overhead information [16]. The system in this paper does not require transmitting any overhead as shape information.

The person region with zero padded non-person region and the non-person region with zero padded person region are embedded into the basic layer and the enhancement layer in figure 1 respectively.

It is feasible to employ the contour extraction method such as SNAKES [10-12]. However, there is no influence on displaying awareness even though the contour is not exact as explained in later subsection *E*. The proposed system employs a simple method for extracting contour indicated in figure 4. It utilizes multi-resolution expression of an image signal by means of the JP2K wavelet transform.

2.4. Awareness Extraction

The bands are decomposed into bit-planes as same as the JP2K encoder. For the basic layer to display awareness, denoting "transparency" parameter $t=0,1,2,\ldots$, the pixel value is shifted to right by t [bit], namely multiplied by 2^{-t}. Parameter $t=0$ indicates normal video, $t=1$ is semi- (or half) transparent and $t=2$ is quarter transparent. As a result, the number of bit-planes to be encoded by the JP2K encoder is reduced by t and the data to be transmitted is also reduced.

For the basic layer, denoting "blurness" parameter $n=0,1,2,\ldots$, only LL band at n-th stage for $n>0$ and all the bands for $n=0$ is transmitted. Parameter $n=0$ indicates clear, $n=1$ is weakly blurred and $n=2$ is more blurred. The number of pixels to be encoded is all for $n=0$, 1/4 for $n=1$, 1/16 for $n=2$ and the data to be transmitted is also reduced.

As explained, the basic layer contains minimum data which are higher bit-planes and lower bands necessary for displaying the awareness. The enhancement layer contains the rest of the bit-planes and bands. This procedure enables to reduce bit rate for awareness and video communications.

2.5. Background Addition

The "background addition" in figure 1(b) in the receiver synthesizes the "input video" and the "background" after decoding data by the JP2K decoder. For displaying awareness at parameters (t, n), LL band at n-th stage is shifted to left by t [bit], namely multiplied by 2^{t}, and the "background" multiplied by $(1-2^{-t})$ is added. As a result, the output signal,

$$[output]=[person]\times 2^{-t}+[background]\times(1-2^{-t}) \quad (3)$$

is displayed at the receiver as the awareness. If $t=0$, [output] = [person] and if $t=1$, [output] = { [person]+ [background] }/2 = [semi-transparent]. If [person] = [background], always [output] = [background]. Therefore, there is no influence on displaying person as far as the extracted contour includes actual person region.

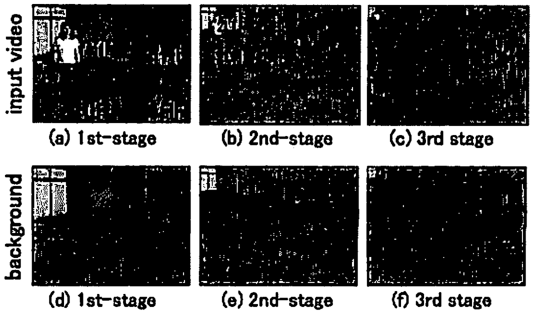


Fig.2 Band decomposition with the wavelet transform.

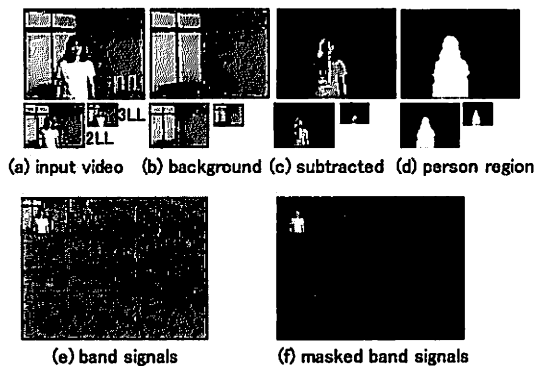


Fig.3 Zero padding on the noon-person region.

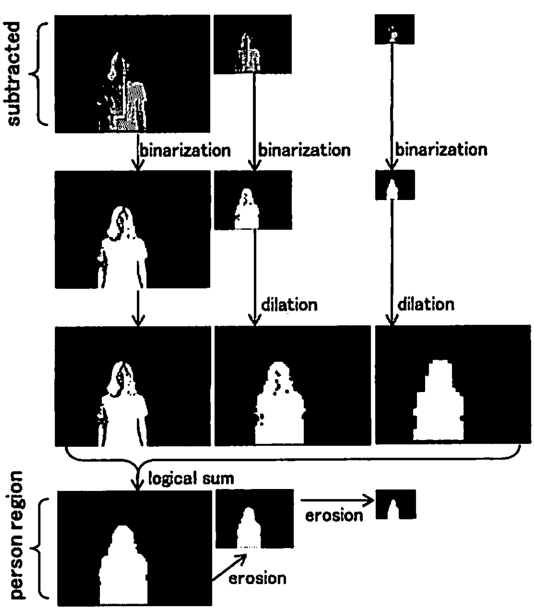


Fig.4 Person region determination with the wavelet transform.

3. EVALUATION OF THE SYSTEM

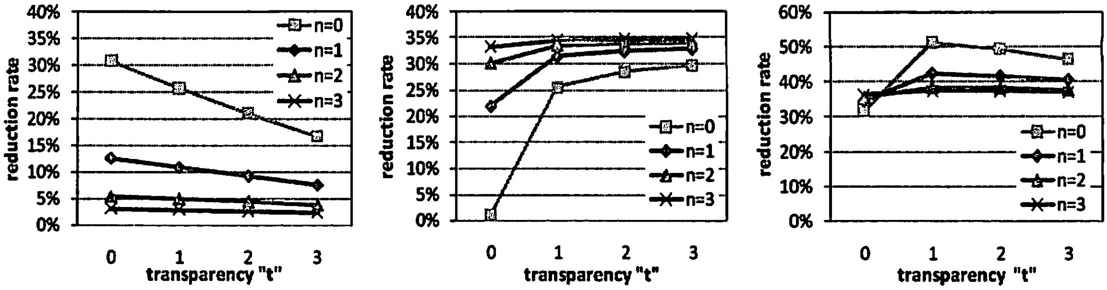
In this section, bit rate reduction ratio due to the procedure proposed in this paper for awareness communication is experimentally examined. Luminance of 600 [frames] of a video signal with 320 x 240 [pixel] and 15 [frame/sec] is used. "Lossless" 5/3 filter (not "lossy" 9/7 filter) of the JP2K is used so that the system is compared under no coding error caused by "lossy" coding.

Awareness at various parameter values (t, n) is indicated in figure 6. For this frame, encoded data of the person region at t=0 and n=0 was observed to be 22.8 [%] to the encoded data of whole region at t=0 and n=0. Figure 5 indicates data reduction rate which is defined as a ratio of data amount of the basic layer at various parameter values (t, n) to the encoded data amount of whole region at t=0 and n=0.

For this example, details of the person seems to be concealed for n>1. In this case, data reduction rate is less than 10 [%]. Optimum parameter value (t, n) is determined according to the user. However, this paper insists that the data to be transmitted can be reduced by more than 90 [%] for n>1 by transmitting minimum components to display awareness.

4. CONCLUSION

The awareness video communication system which can reduce data amount to be transmitted is proposed. The system transmits minimum components, higher bit-planes and lower bands generated by JP2K, to display awareness. It is confirmed that the system can reduce bit rate to 10 [%] at awareness communication mode. The system has a feature that it can be implemented by core technologies of the international standard JP2K.



(a) Basic layer (b) Enhancement layer (c) Basic layer + enhancement layer

Fig.5 Data reduction rate.

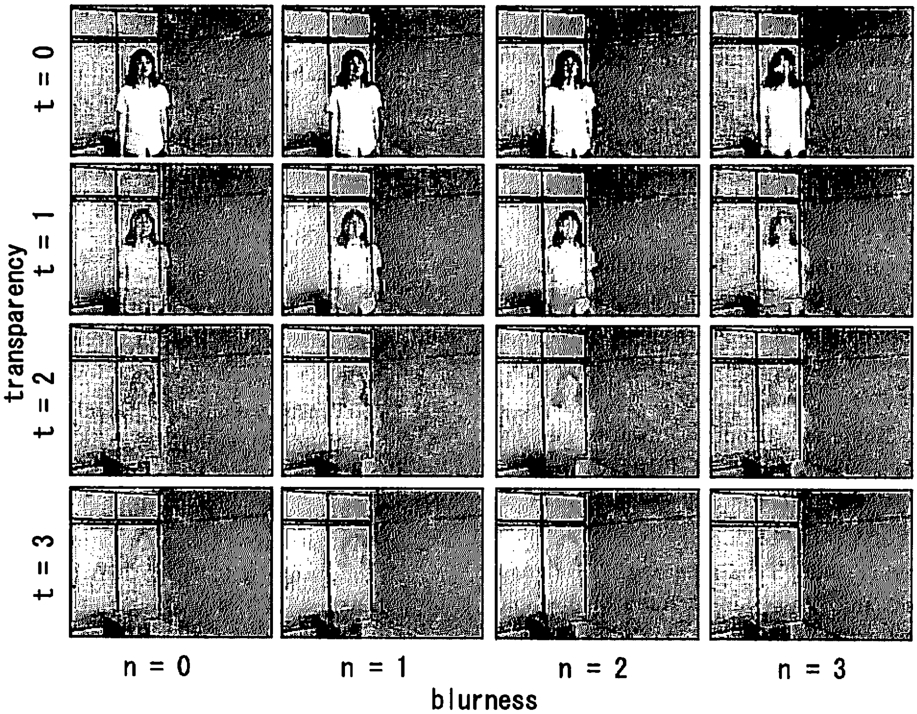


Fig.6 Displayed "awareness" at the receiver.

REFERENCES

[1] Gian Luca Foresti, Petri Mahonen, Carlo S. Regazzoni, "Multimedia video-based surveillance systems: requirements, issues, and solutions." Boston, Mass.: Kluwer Academic Publishers, 2000.

[2] J.Wang, W.Q. Yan, M.S. Kankanhalli, J. Ramesh and M.J.T. Reinders, "Adaptive monitoring for video surveillance", Proceedings of IEEE International Conferenceon Information, Communications and Signal Processing, Vol.2, pp.1139-1143, Dec. 2004.

[3] Stefan Junestrand, et al, "Private and Public Spaces - The Use of Video Mediated Communication in A Future Home Environment", Proc. CHI 2000, HAAG, ACM press, 2000.

[4] S. Hanaki, S. Tansuriyavong, S. Tsubaki, "Human Image Concealment Depending on Position for Video Awareness Systems", ECSCW03, pp.31-33, 2003.

[5] Uwe M. Borghoff, J. H. Schlichter, U. M. Borghoff, Johann H. Schlichter, Computer-Supported Cooperative Work: Introduction to Distributed Applications. Berlin; Tokyo: Springer, 2000.

[6] Michael Boyle, Christopher Edwards and Saul Greenberg: "The Effects of Filtered Video on Awareness and Privacy", Proceedings of the CSCW'00, ACM, 2000.

[7] M. Boyle and S. Greenberg, "Balancing Awareness and Privacy in a Video Media Space Using Distortion Filtration", in Proceedings of the Western Computer Graphics Symposium 2000, Panorama Mountain Village, BC, Canada, March26-29, 2000.

[8] Paul Dourish: "Extending Awareness Beyond Synchronous Collaboration", Workshop on Awareness in Collaborative Systems CHI'97, 1997.

[9] S. Hanaki, S. Tansuriyavong, S. Tsubaki, "Human Image Concealment Depending on Position for Video Awareness Systems", Supplement Proceedings of the Eighth European Conference on Computer Supported Cooperative Work, pp.31-33, Sep. 2003.

[10] M. Kass, A. Witkin and D. Terzopoulos , "Snakes: active contour models", Int. J. Comput. Vision 1 4, pp. 321-331, 1987.

[11] A. E. Hassanien, M. Nakajima, "Feature-Specification Algorithm based on Snake Model for Facial Image Morphing", IEICE Trans. Inf. & Syst., vol E82-D, no.2,pp.439-446, Feb. 1999.

[12] S. Araki, T. Matsuoka, N. Yokoya, H. Tanemura, "Real-Time Tracking of Multiple Moving Object Contours in a Moving Camera Image Sequence", IEICE Trans. Inf. & Systems,vol.E83-D, no.7, pp.1583-1591, July 2000.

[13] MPEG-4 video group, "Generic Coding of Audio-Visual Objects: part 2 – visual", ISO/IEC JTC1/ SC29/WG11, N1902, FDIS of ISO/IEC 14496-2, 1998

[14] T. Ebrahimi, C. Horne, "MPEG-4 Natural Video Coding – an Overview", signal processing: image communication, Elsevier, no.15, pp.365-385, 2000. ISO/IEC FCD15444-1, "JPEG2000 IMAGE CODING SYSTEM" March 2000.

[15] D. S. Taubman, M. W. Marcellin, "JPEG 2000 -Image Compression fundamentals, standards and practice", Kluwer Academic Publishers, 2002.

[16] S. Tansuriyavong, T. Suzuki, M. Takeda, M. Iwahashi, "Development of a Privacy Conscious Video Communication System Using JPEG2000 Technology", International Conference on Information, Communications and Signal Processing (ICICS), no.T2B.1, pp.696-699, Dec. 2005.