<table>
<thead>
<tr>
<th>Title</th>
<th>Awareness and Privacy Balanced System Using Functional Layered Video Coding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Tansuriyavong, Suriyon; Suzuki, Takayuki; Chokchaitam, Somchart; Iwahashi, Masahiro</td>
</tr>
<tr>
<td>Citation</td>
<td>冲縄工業高等専門学校紀要 = Bulletin of Okinawa National College of Technology(4): 23-28</td>
</tr>
<tr>
<td>Issue Date</td>
<td>2010-03</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/20.500.12001/18672">http://hdl.handle.net/20.500.12001/18672</a></td>
</tr>
<tr>
<td>Rights</td>
<td>冲縄工業高等専門学校</td>
</tr>
</tbody>
</table>
Awareness and Privacy Balanced System Using Functional Layered Video Coding

Suriyon TANSURIYAVONG\textsuperscript{1}, Takayuki SUZUKI\textsuperscript{2}, Somchart CHOKCHAITAM\textsuperscript{3} and Masahiro IWASHI\textsuperscript{4}

\textsuperscript{1}Okinawa National College of Technology, Nago, Okinawa, 905-2192 Japan.
\textsuperscript{2}THine Electronics,Inc., Chiyoda, Tokyo, 100-0005 Japan
\textsuperscript{3}Thammasat University, Rangsit, Pathum-Thani 12121, Thailand
\textsuperscript{4}Nagaoka University of Technology, Nagaoka, Niigata, 940-2188 Japan

ABSTRACT

This paper proposes a new video communication system which transmits “awareness” at very low bit rate in usual, and displays a details “video” by sending additional information on demand. The “awareness” means information without “privacy related details and it expresses “existence” or “motion” of a person, and so forth. The proposed system decomposes a person’s region into (a) basic layer which contains awareness and (b) enhancing layer which contains other details information. This processing is implemented by basic constituent technology of the JPEG2000 such as wavelet transform and bit plane decomposition. The basic layer contains (1) person’s region, (2) a part of bit plane, and (3) a part of bandwidth image. So that, the awareness information is transmitted at very low bit rate. It is confirmed by quantitative measurement that bit rate can be reduced to 1/10 compared with exiting method. The system can be applied to use in the filed of welfare, for example, monitoring and nursing-care for an elderly or a patient from distance.

Keywords: awareness, layered, video coding, JPEG2000, privacy

1. INTRODUCTION

Recently, there are an always-connected video awareness which have been proposed to support communication between remote site\cite{1,2}.

However, an always-connected video brings a privacy issues. According to above proposals, several video communication system have been also proposed in the term of protecting a "privacy" of the person in the scene while providing his/her "awareness" information\cite{3-7}. However, those systems are implemented a video "recognition" such as moving object detection, object's contour extraction, and video "compression" such as the MPEG or H.261,H.264 separately, which made a system become large and complicated. In the way of focusing on providing "awareness", the data transfers bit rates of those systems are very high.

This paper proposes a “privacy balanced” video communication system balancing awareness and privacy base on technology of JPEG2000. Person region is extracted into (a) basic layer which contains awareness and (b) enhancing layer which contains other details information by using the “multi-resolution expression” of Wavelet transform, and “bit-plane decomposition” technique, which these techniques are a constituent technology of JPEG2000\cite{8-10}. The basic layer is focused to convey “awareness” information, it contains (1) person’s region, (2) a part of bit plane, and (3) a part of bandwidth image. Therefore, the awareness information is transmitted at very low bit rate.

In the proposed method, a part of the compression is utilized for “recognition” of an object contributing for reducing system complexity and shortening development term of the system.

2. AWARENESS AND PRIVACY BALANCED VIDEO COMMUNICATION SYSTEM

\begin{figure}
\centering
\includegraphics[width=\textwidth]{fig1.png}
\caption{The configuration of awareness and privacy balanced video communication system.}
\end{figure}
Awareness and privacy balanced video communication system is a system that provides awareness information, while protects a privacy of a person in the scene. The system works not only as a conventional video communication tool but also serves as an always-connected video awareness system to support communication of person between remote sites. In this system, the person who stays far from camera at the sending site is automatically displayed as an awareness (such as semi-transparent, blurred, or contoured) at the receiving site, in order to protect his/her privacy.

This system has functions as shown in Fig.1. The useful cases of using this system are described as below.

1. In the case when the person who wants to communicate to each other, he/she come close to the camera. In this case, at the receiving site, a details “video” image is displayed and person can communicate to each other like as a conventional video communication system.

2. In the case when the person who has no intention to talk right that time, but he/she is ready to response if someone calls. In this case he/she stays far from camera, then only “awareness” information (“existence” or “motion” e.g.) is sent to another site without “privacy” related details, in very low bit rate.

The displaying of awareness is the so-called “an awareness communication”. It is a new type of communication means [11]. In the case that person who has intention not to be talked, or not to be observed, Displaying of awareness has a merit to protect a privacy of that person.

At the sending site, we use the fixed camera. Therefore, the background image is almost fixed, so, it is not necessary to send background image in real-time. Firstly, we send the background image and store it in the memory at the receiving site.

We create a bit stream of data that corresponding to person region, and embed the awareness information into tile’s header (tile part header). We encode and send only the data that corresponding to person region in order to reduce transfer bit rate.

At the receiving site, the tile part header of bit stream from the sending site is analyzed, and then the tile information and type of awareness information are extracted. By combining this information with the background image, which once has been stored in the memory, the reconstruction of awareness with background image is carried out.

3. SIGNAL PROCESSING BY THE PROPOSED METHOD

3.1 Person Region Extraction Using “Multi-Resolution Expression” in JPEG2000

Generally, using the background image subtraction to extract the foreground image usually raises a hole in the region, in particular, when the background image is similar to the foreground image. The problem discussed here is how to extract the foreground image with the minimum implementation cost. In order to cope with this issue, the proposed method employed the multi-resolution expression by wavelet transform, which already included in technology of JPEG2000.

Figure 2 shows the using multi-resolution expression to extract moving object. We execute wavelet transform for the result image that obtained from the background subtraction. Once the wavelet transforms has been executed, the vertical and horizontal direction of the low-pass and high-pass, total 4 images are created. For one stage of wavelet transform, the resolution of each image is scaled to a half size. We pick up the low-pass image, and repeat the processing stage of wavelet transforms to the low-pass image. Therefore, according to the number of processing stage:N, the low-pass images in difference size are created. Finally, restore each of them to the original size, and then reconstruct them into one image as an original by the logical OR operation.

![Fig.2: Moving person extraction by using multi-resolution expression (the proposed method).](image)

As shown in Fig.3, the result of moving person extraction using multi-resolution expression is able to reduce an undesired hole in the obtained region; the effectiveness of this method is confirmed. After that, the labelling process is performed, and the amount pixel in the object region is calculated, as the “total pixel in the person region”. We use this value to set or do not set the transparency rate(T), as shown in Table1.

There are several ideas in order to control the transparency rate. For instance, using multi camera as a stereo vision, to calculate the position of the person, and then control the transparency rate, and so forth. However, in this paper we use single camera for computationally feasible and relatively inexpensive.

![Fig.3: Moving Object extraction results. Undesired holes in the object region are reduced by the proposed method.](image)
Table 1: Transparency of the person is depending on “total pixel in the person region”.

<table>
<thead>
<tr>
<th>Distance (non-transparent)</th>
<th>Total pixel in the person region: A[pixel]</th>
<th>Transparency rate: ( T )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Near</td>
<td>( A &gt; 15000 )</td>
<td>( T = 2^n )</td>
</tr>
<tr>
<td>Far (transparent)</td>
<td>( A = 15000 )</td>
<td>( T = 2^n ) Where, ( n = 1, 2, \ldots, 8 ) (( n ) is set by user)</td>
</tr>
</tbody>
</table>

3.2 Transparency Rate Control using “Bit-plane decomposition” in JPEG2000

In the JPEG2000 compression algorithm, each frame is decomposed into adjacent small rectangular called "tile". The proposed system classifies each tile into "object" or "background". In this paper, all the tile which contains a part of (or all) the person region are categorized into "object". Identifier of "object or not" is embedded into CME (comment and extension) header at tile-part header in the JPEG2000 bit-stream composition as shown in Fig.4. This system encodes and sends only the tiles that correspond to object region, in order to reduce transfer bit rate of video data.

This procedure brings an advantage that the system does not directly extract contour of an object. A viewer can see a transparent object with its contour. Namely, the system does not require any complicated contour extraction procedure such as Ref.[12-14].

4. EXPERIMENTAL RESULTS

Experimental results in this paper were given under the environment listed in Table2.

<table>
<thead>
<tr>
<th>Table2: System Environment</th>
</tr>
</thead>
<tbody>
<tr>
<td>OS</td>
</tr>
<tr>
<td>CPU</td>
</tr>
<tr>
<td>RAM</td>
</tr>
<tr>
<td>Software</td>
</tr>
<tr>
<td>Movie</td>
</tr>
<tr>
<td>DWT</td>
</tr>
</tbody>
</table>

4.1 Person region extraction

First, we evaluate the person region extraction process. Fig.5(a) shows image frame number 100,320,400 and 490 as a representative examples of input image. The person regions are successfully extracted as shown in Fig.5 (b). There are no undesired holes in the region. It is confirmed that the proposed method is effective to reduce a hole in the extracted object.

![Input video image](image)

![Person-region extracted by the proposed method.](image)

Fig.5: Moving Person extraction results. Undesired holes in the object region are reduced by the proposed method.

4.2 Bit rate of sending only person region

Next, we evaluated total bit rate to be transferred to a receiver by sending only person region.

Figure6 shows transfer bit rate of the existing and the proposed method. The existing method send all region of image, so, the transfer bit rate is always about 7-8[Mbps]. In the contrast, the proposed method send only the region of interest; the bit rate is about 3.6[Mbps] as a maximum.

\[ \text{Transparent output signal} = \text{Object signal} \times T + \text{Background signal} \times (1 - T) \]
Figure 7 shows compression ratio of sending only person region and sending all region. It shows that bit rate is reduced in proportion to the size of person region. It is confirmed that the system is able to reduce bit rate less than a half of existing method.

\[ \text{Compression ratio} = \left( \frac{\text{Bit rate of sending only person region}}{\text{Bit rate of sending all region}} \right) \times 100\% \]

Figure 8 shows effectiveness of sending only apart of bit plane (5 bit plane).

4.3 Bit rate of sending only a part of bit plane

Total bit rate of sending only a part of bit plane is evaluated. In this case, we set both the proposed method and the existing method to send all bandwidth and all background images in the same condition.

As show in Fig. 8, the proposed method (sending only a part of bit plane) can reduce bit rate to 1/4 compare to the existing method (sending all bit plane).

Figure 9: Compression ratio of sending a part of bit plane and sending all bit planes.

Figure 9 shows compression ratio of sending only a part of bit plane and sending all bit plane. It shows that bit rate is reduced in proportion to the number of bit plane. It is confirmed that reducing bit plane is effective to reduce bit rate.

4.4 Bit rate of sending only a part of bandwidth

Total bit rate of sending only a part of bandwidth is evaluated. In this case, all bit plane and whole background image is set as the same condition for both proposed and existing method.

Figure 10 shows bit rate of sending only a part of bandwidth. It is confirmed that sending only a part of bandwidth is also able to reduce bit rate about 1/4 of the existing method.

Figure 8: Effectiveness of sending only apart of bit plane(5 bit plane)
Figure 11: Compression ratio of sending a part of bandwidth and sending all bandwidth.

Figure 11 shows that bit rate is reduced in proportion to the stage of wavelet transform. It is confirmed that the system is able to reduce bit rate less than 60% of the existing method.

4.5 Combination of compression ratio

Figure 12 shows a combination of compression ratio for sending awareness and normal video. It is confirmed that bit rate for sending awareness can be reduced to 1/10 compared with sending normal video.

Fig.12: Combination of compression ratio for sending awareness and normal video.

5. CONCLUSIONS

This paper proposes a new video communication system which transmits “awareness” at very low bit rate in usual. The proposed system decomposes a person’s region into (a) basic layer which contains awareness and (b) enhancing layer which contains other details information. This processing is implemented by basic constituent technology of the JPEG2000 such as wavelet transform and bit plane decomposition. The basic layer contains (1) person’s region, (2) a part of bit plane, and (3) a part of bandwidth image. It is confirmed by quantitative measurement that bit rate can be reduced to 1/10 compared with exiting method.

It is our future work to adapt quantization and bit truncation, namely lossy compression, to the system for more bit rate reduction.

In addition, the variations of displaying awareness image are shown in Fig.13-14.

6. REFERENCES


