Generation of Panoramic Image from Aerial Video utilizing JP2K Wavelet for River Surveillance

Masahiro IWAHASHI [1], Sakol UDOMSIRI [1], Jyun WATANABE [1] and Sinji FUKUMA [2]

1 Nagaoka University of Technology
Nagaoka City, Japan,
{iwahashi, sakol_ud}@nagaokaut.ac.jp

2 University of Fukui
Fukui City, Japan

***Abstract*—This report discusses how to generate a panoramic image from video data for environmental surveillance of a river. Video signal is assumed to be taken in a sunny day from an airplane. It includes relatively small obstacles such as a cloud or reflection of the sun. It is our purpose to generate a panoramic image avoiding these obstacles. The best part of the frames for surveillance is selected utilizing JP2K wavelet transform so that total amount of processing can be reduced.**

I. INTRODUCTION

It has been required to develop a video watching system of a river to reduce damages caused by floods. Image signals observed by satellites have been widely used to check a river's state at a glance. These plural images along a river are connected each other to produce a panoramic wide view still image with various techniques [1-4]. This report discusses how to generate a panoramic image from aerial video data compressed by the motion JPEG 2000 (JP2K) [5]. Computational load can be reduced by sharing the wavelet transform between "recognition" and "compression".

Video data are assumed to be taken in a sunny day. They usually include some obstacles such as small clouds or reflection of the sun called "halation". Figure 1 and figure 2 (a) illustrate an obstacle remained in the panoramic image signal. These are removed by utilizing redundancy of video data. Since video data include 30 frames per second, a region on the ground usually corresponds to plural frames. The proposed method selects the best part of the frames for water region surveillance and connects them as illustrated in figure 3. As a result, the small obstacles are removed from panoramic image as indicated in figure 2 (b).

The best part of the frames is chosen according to occupation ratio of water region. It is recognized with the maximum likelihood (ML) estimation based on feature vector calculated by "Gabor" filter [6-7]. This report also replaces the "Gabor" by "wavelet" in JP2K so that a part of "compression" is shared with "recognition" as illustrated in figure 4. It is expected to contribute to reduction of system's redundancy and total hardware complexity.

II. BASIC IDEA OF THE PROPOSAL

A. OBSTACLE IN PANORAMIC IMAGE

A simple conventional method connects each of the frames in a video sequence to produce a panoramic image as illustrated in figure 1. The current frame covers the previous frames. If small obstacles in water area such as clouds or reflections of the sun are included, they remain in the resulting panoramic image. These obstacles should be removed so that state of the river such as color change due to muddy stream, aquatic plant and so on can be observed.

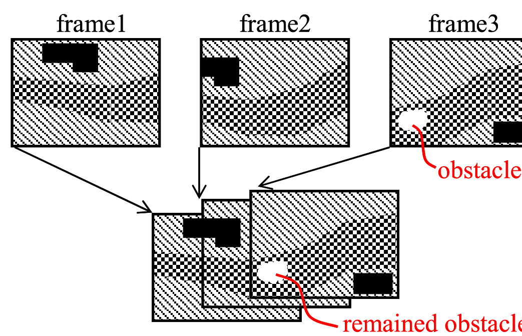

Figure 1 Small obstacles in the water region should be removed so that river's state can be clearly observed.

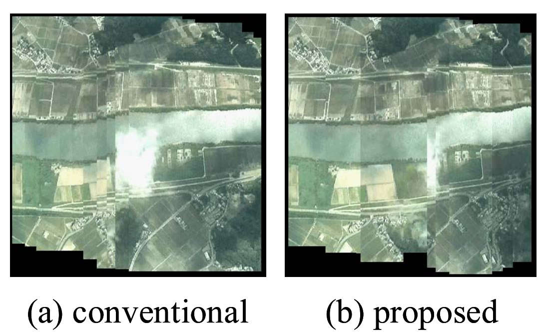

Figure 2 Clouds and halation are effectively removed from the panoramic image by the proposed method.

1-4244-0387-1/06/$20.00 ©2006 IEEE

B. UTILIZATION OF REDUNDANCY IN VIDEO

The proposed method selects the best of plural candidates in different frames for a designated grand region as illustrated in figure 3. Note that it is assumed that a region on the ground corresponds to plural frames in the video data. The selection is performed in respect of occupation ratio of water region as indicated in figure 7. The edge matching algorithm [4] is used for connection. As a result, small obstacles are removed from panoramic image.

C. SHARING WAVELET TRANSFORM BETWEEN RECOGNITION & COMPRESSION

Each pixel in the candidates is classified into “water” or not. This recognition is performed by the ML estimation and “Gabor” filter. This report replaces it by “wavelet” in JP2K so that it is shared by “compression” and “recognition” as illustrated in figure 4. It contributes to reduction of system’s total hardware amount.

III. DETAILS OF THE PROPOSAL

A. FEATURE VECTOR BY “GABOR”

For various classifications, the Gabor filter bank has been widely used to calculate feature vectors [6-8]. Its filter coefficients are defined as complex value by

$$h(x,y) = g(x,y)\exp\left\{ j\frac{\pi}{\sigma}\left(x\cos\theta + y\cos\theta\right)\right\} \tag{1}$$

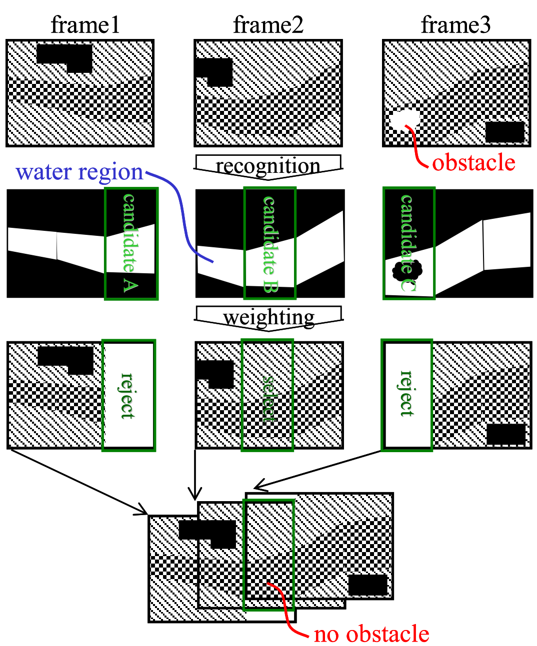

Figure 3 The proposed method selects the best of candidates to produce a panoramic image with reduced obstacles.

$$g(x,y) = \frac{1}{2\pi\sigma^2}\exp\left\{-\frac{\left(x^2+y^2\right)}{2\sigma^2}\right\}$$

$$\begin{aligned}\mathbf{G}(x,y) &= h(x,y) * g(x,y) \\ &= \sum_{\tau_y=-R}^{R}\sum_{\tau_x=-R}^{R} h\left(\tau_x,\tau_y\right)\cdot g\left(x-\tau_x, y-\tau_y\right)\cdot\end{aligned} \tag{2}$$

The feature vector **G**(*x,y*) for each pixel *p(x,y)* is calculated by a convolution of *p(x,y)* and *h(x,y)*. Components of the vector are 1) real part and imaginary part, or 2) amplitude and phase. Only the amplitude is used in some cases [8].

B. MAXIMUM LIKELIHOOD (ML) ESTIMATION

The ML estimation determines a class ω_i , *i=1,2,…, K*, for each pixel *p(x,y)* minimizing the conditional probability defined by

$$-2\ln P\left(\mathbf{G}(x,y)\,|\,\omega_i\right) = d_i^2(x,y) + \ln\left|C_i\right| \tag{4}$$

$$d_i^2(x,y) = \left(\mathbf{G}(x,y) - \mu_i\right)^t C_i^{-1}\left(\mathbf{G}(x,y) - \mu_i\right).$$

Notations **G**(*x,y*), *Ci* and μ_i denote feature vector, covariance matrix and average of the class ω_i of the teacher pixel area respectively [8]. One of the classes ω_i is set to be “river” area. The others can be “wood”, “rice field”, “house”, “road”, etc. In this application, it is important to determine “river” area or not. This binary decision makes the problem easier than automatic map generation.

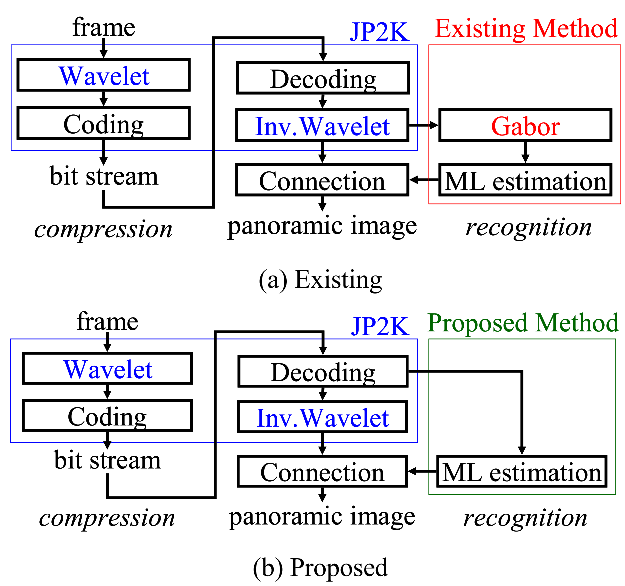

Figure 4 The proposed method utilizes “wavelet” of JP2K instead of “Gabor” for “recognition”.

C. FEATURE VECTOR BY “WAVELET”

The wavelet transform in JP2K referred to CDF 9/7 in [5] is examined. Experimental results for a model signal in figure 5 are summarized in table 1. The best case in each of table (a) - (d) are picked up and compared to the “Gabor” in figure 6. “Gabor 1” uses amplitude component only. Its dimension equals to $180/\theta$. “Gabor 2” uses real part and imaginary part. Its dimension is $360/\theta$. R=5 and σ =4 for both of them. The figure indicates that “wavelet” with three bands is equivalent to “Gabor” with dimension 4. Since “Gabor” has more freedom of parameters, it can be better than “wavelet” in general. However, there is no significant difference on the final panoramic image as illustrated in figure 1 and 8. It is because what is required to the filter is not high recognition rate but steady priority of the candidates as illustrated in figure 7.

Table 1 Rate [%] of correct recognition by the JP2K wavelet with one band or plural band signals.

(a) one band

band	rate [%]
LL	66.02
LH	**72.26**
HL	69.07
HH	70.18

(b) two bands

band	rate [%]
LL LH	58.48
LL HL	58.59
LL HH	65.92
LH HL	67.38
LH HH	**75.09**
HL HH	70.24

(c) three bands

band	rate [%]
LH HL HH	**75.83**
LL HL LH	46.07
LL HL HH	46.52
LL LH HH	46.40

(d) four bands

band	rate [%]
LL HL LH HH	**55.70**

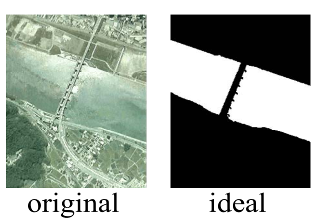

Figure 5 Image data for experiment of recognition.

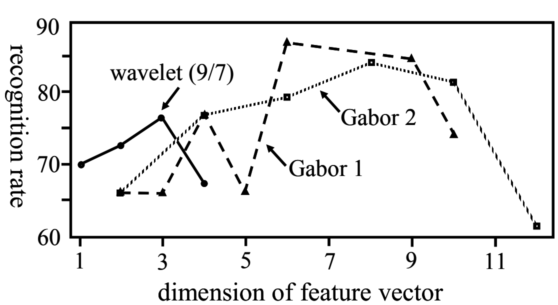

Figure 6 Comparison of recognition performance between “Gabor” and “wavelet” as feature vector.

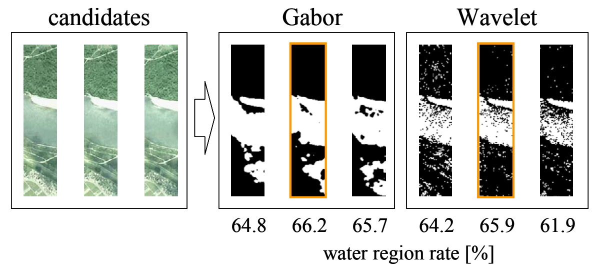

Figure 7 Occupation ratio of water region of candidates.

IV. EXPERIMENTAL RESULTS

Figure 2 and 8 summarize experimental results of producing panoramic image signal. Specifications of the experiment are indicated below.

[Video]; 240*320 pixels, 8 bit R,G,B, 30 frames per sec. AVI format. [Location]; "Shinano River” from Urasa to Niigata in Niigata prefecture. Video was taken from an airplane at 4000 feet height. Weather condition was fine. date: 15th June 2004. time: 13:00. [Processing]; Matlab 6.5.0, Pentium 2.4GHz, 512 MB RAM.

Effectiveness of the proposed method can be subjectively confirmed. In figure 2, small particles of the cloud are reduced by the proposed method. Figure 8 indicates that reflections of the sun are diminished.

The proposed method bases on the assumption that a region on the ground corresponds to plural frames in the video data and at least one of them have no effect by the obstacle. If it is not satisfied as indicated in figure 9 (b), there were too much clouds in this case, experiment results in failure. Figure 9 (a) illustrates another failed case when the airplane did not fly straightly.

As described in section II, the best one is considered to contain water region most. If there are some equivalent candidates, next priority is that image signals trough the center of the lens are preferable since the image through the edge is normally distorted as confirmed by figure 10.

V. CONCLUSIONS

This report proposed a generation method of a panoramic still image from aerial video data focusing on “river surveillance” application. It was confirmed that the proposed method can exclude relatively small obstacles such as clouds or reflection of the sun from the panoramic image of a river. A weighting procedure based on filter bank and ML estimation was introduced before connecting frame parts. It was also confirmed that a part of the JP2K international standard can be utilized for the system instead of conventional Gabor filter.

Whole view of the finally produced panoramic image can be utilized for timely observation of a river’s situation in emergency such as after heavy rain, flood, mud slide, and so on. It will help us quick evacuation and proper measures.

(a) Conventional.

(b) Proposed.

FIGURE 8 REFLECTIONS OF THE SUN WERE EFFECTIVELY REDUCED BY THE PROPOSED METHOD.

ACKNOWLEDGMENT

This research was supported by Ecology Science Co. Ltd. Nagaoka, Niigata, Japan.

References

[1] Zhingang Zhu, Allen R.Hanson, "Mosaic-Based 3D Scene Representation and Rendering", IEEE. Intern'l Conf. on Image Processing, vol.I, pp.633-636, Sept.2005.

[2] B.E.Pires, P.M.Q.Aguiar, "Featureless Global Alignment of Multiple Images", IEEE International Conference on Image Processing, vol.I, pp.57-60, Sept.2005.

[3] Richard Szeliski and Heung-Yeung Shum, "Creating Full View Panoramic Image Mosaics and Environment Maps," SIGGRAPH'97,pp.251-258,1997

[4] G. Borgefors, "Hierarchical Chamfer Matching: A Parametric Edge Matching Algorithm", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 10, No. 6, pp. 849-865, 1988.

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

[6] D.Dunn, W.E.Higgins, "Optimal Gabor Filter for Texture Segmentation", IEEE Trans. Image Processing, vol.4, no.7, pp.947-964, 1995.

[7] S.Mitani, H.Wu, T.Shioyama, "Car Detection with Gabor Filters," IEICE DII, vol.J83-DII, no.12, pp.2641-2651, Dec.2000.

[8] H.Tanaka, Y.Yoshida, K.Fukami, H.Nakano, "Texture Segmentation using Amplitude and Phase Information of Gabor Filters", IEICE Trans. D-II, vol.J84-DII, no.12, pp.2565-2576, Dec.2001.

(a) Rotation of the airplane.

(b) Too much clouds.

Figure 9 Failed cases .

(a) Center region of the frames were connected.

(b) Edge region of the frames were connected.

Figure 10 Images trough the center of the lens are preferable for avoiding distortion.