SICE-ICASE International Joint Conference 2006
Oct. 18-21, 2006 in Bexco, Busan, Korea

Panoramic Image Generation based on FFT Technique and its Hardware Realization

Ryo YONEMOTO[1], Toru UESUGI[1], Takao KAWAMURA[2] and Kazunori SUGAHARA[2]

[1]Graduate School of Engineering, Tottori University, Tottori, Japan
(E-mail: {yonemoto,uesugi}@ike.tottori-u.ac.jp)
[2]Department of Engineering, Tottori University, Tottori, Japan
(E-mail: {kawamura,sugahara}@ike.tottori-u.ac.jp)

Abstract: In most of the current teleconferencing systems, a full view image of a conference room and voice of speakers are bidirectionally exchanged between distant places. However, in teleconferencing situations, face images of speakers are also requested to perceive facial expression. To take both full view images of a conference room and face images of speakers, it is required to develop a multi camera system with human operations. In this paper, we develop a new video camera for teleconferencing without human operations in above mentioned situations. The proposed video camera outputs the images which composed with panoramic images generated with images from three NTSC video cameras and the facial images of speakers. In this paper, hardware realization of the Fast Robust Correlation technique proposed by Fitch et al. is described.

Keywords: Panoramic Image Generation, Fast Fourier Transform.

1. INTRODUCTION

The advancement of communication speed in digital networks makes it possible to exchange a large amount of information between distant places. Teleconferencing is one typical example of the practical applications in such fields.In most of the current teleconferencing systems, a full view image of a conference room and voice of speakers are bidirectionally exchanged between distant places. However, in teleconferencing situations, face images of speakers are also requested to perceive facial expression. To take both full view images of a conference room and face images of speakers, it is required to develop a multi camera system with human operations. Therefore we develop a video camera for teleconferencing to output the images which composed with panoramic images generated with images from three NTSC video cameras and the facial images of speakers. In this paper, "Fast Robust Correlation technique" proposed by Fitch et al. is adopted to generate panoramic images and its hardware realization in a FPGA chip is described [1].

2. HARDWARE CONFIGURATION OF THE PROPOSED SYSTEM

The proposed system is constructed with three NTSC video cameras and FPGA board constructed with NTSC video encoder/decoder LSI and memories as shown in Fig.1.

The proposed panoramic image generation function is realized on a FPGA chip as hardware circuits for easy constructions of embedded real-time systems. FPGA is LSI that has reconfigurable inner circuits. The operation clock applied to the FPGA is 48MHz. The Logic Element number is 20,060. These specifications are summarised in Table 1.

In this system, the memories are constructed as three banks and these banks can be read/written by FPGA simultaneously. By using three bank memories, inputting, processing and outputting data are accomplished at the same time as shown in Fig.2. Each frame of images from the NTSC video camera are decoded by the decoder LSI and the obtained digital data are stored in one bank of memory on the FPGA board. The proposed panoramic image generation function is applied to the frame stored in a bank of memory.

Table 1 FPGA specifications

Manufactured	ALTERA Co.Ltd.
Model number	EP1C20F400C7
Number of logic elements	20,060
Max. I/O port number	301
Package	400-Pin FineLine BGA
Package size	21×21 [mm]
Frequency	48[MHz]

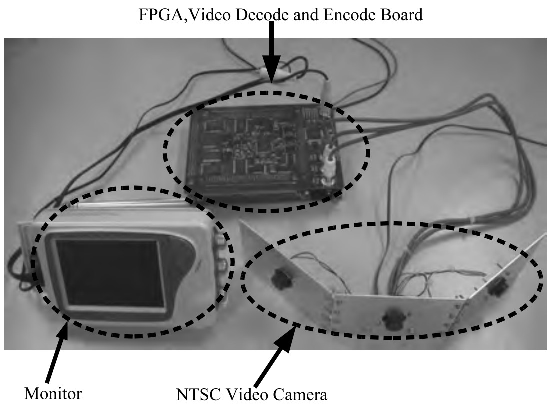


Fig. 1 System configuration

3. GENERATION OF PANORAMIC IMAGES

The input images from the three video cameras are compliant with the NTSC specifications, that is, they are

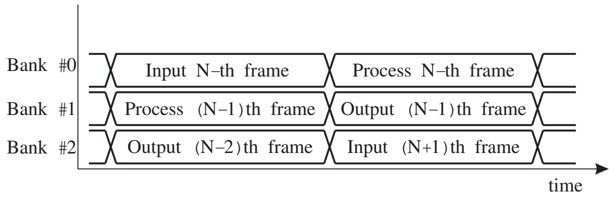


Fig. 2 Three bank memory

640 × 480 [pixels] size and 30 frames per second. The output image also should be compliant with the specifications, i.e. it should have the same size and the same frame rate. Considering these points, the output image is designed as shown in Fig.3.

As shown in the figure, the lower half area of the output image is arranged for the panoramic image and the extracted speaker face image is enlarged and is represented on the upper half of it. The panoramic image part in the output image is 640 × 180 [pixels] size and is generated with $3/8$ reduced 3 input images with two overlapped area in 40 pixel width as shown in Fig.4.

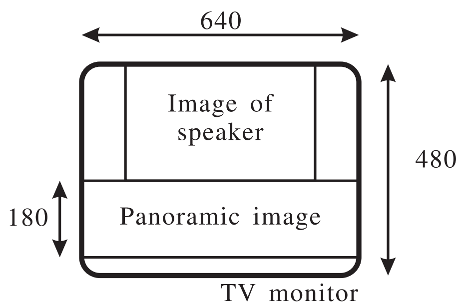


Fig. 3 Panoramic image generation

4. AUTOMATIC CORRECTION OF PANORAMIC IMAGE GENERATION

4.1 Automatic correction processing

The automatic correction is a process to detect the best position for smooth image combining. The process is used for the generation of panoramic images.

4.2 Fast Robust Correlation

Suppose the situation that two images f and g are combining and the corner point of the overlapped area corresponds to coordinate (m, n) of image f and to the original point of image g as shown in Fig.5. At this time, x and m are defined as $x = (x, y)$, $m = (m, n)$. The total difference between two images f, g is given as follows.

$$R(m) = \sum_x h\left(f(x) - g(x-m)\right) \cdot \alpha_f(x) \cdot \alpha_g(x-m) \quad (1)$$

Here, the functions $\alpha_f(x)$ and $\alpha_g(x)$ are the alpha mask functions of f and g correspondingly, and the function $h(r)$ should be an even function about r and not depend on magnitude relations of $f(x)$ and $g(x-m)$. One

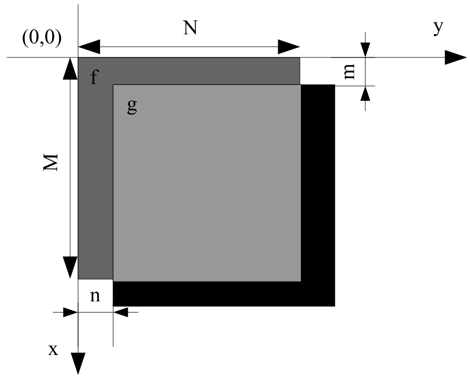


Fig. 5 Image comparison

of the most popular even function is $h(r) = r^2$, however,

$$h(r) \approx \sum_{p=1}^{P} b_p \left(1 - \cos\left(a_p \pi r\right)\right) \quad (2)$$

is adopted in this method. The expression of eq.(2) is well suited for calculation by using Fast Fourier Transform (FFT) technique.

Substitution of eq.(2) into eq.(1) gives following equation which expresses image differences $R(m)$ between f and g. In this equation, "$*$" represents convolution operation and $R(m)$ can be obtained by using Discrete Cosine Transform technique (DCT).

$$\begin{aligned} R(m) &= \sum_x \Bigg[\left(\alpha_f(x) \star \alpha_g(x)\right) \sum_{p=1}^{P} b_p \\ &- \sum_{p=1}^{P} b_p \Big(\left(\alpha_f(x) e^{j a_p \pi f(x)}\right) \\ &\star \left(\alpha_g(x) e^{j a_p \pi g(x)}\right) \Big) \Bigg] \end{aligned} \quad (3)$$

4.3 Image comparison in frequency domain

In this paper, development of the system with calculation of eq.(3) in the frequency domain is considered because the direct calculation cost of it in space domain becomes high. The block diagram to obtain $R(m)$ by using FFT technique is shown in Fig.4.

m which gives the minimum value among the obtained $R(m)$ is objective value. By using the obtained m, the automatic correction processing is accomplished to obtain a panoramic image by shifting g by m and by combining it with image f.

4.4 Hardware realization of FFT

For realization of the FFT as a hardware circuit, the butterfly operation plays an important role. Effective description of the butterfly operation gives small size circuits of FFT. The butterfly operation is shown in Fig. 7.

The flow of the butterfly operation is as follows.

1. The values of complex number $W_N = e^{-j2\pi/N}$ are calculated and stored in registers W_{Nr} and W_{Ni} be-

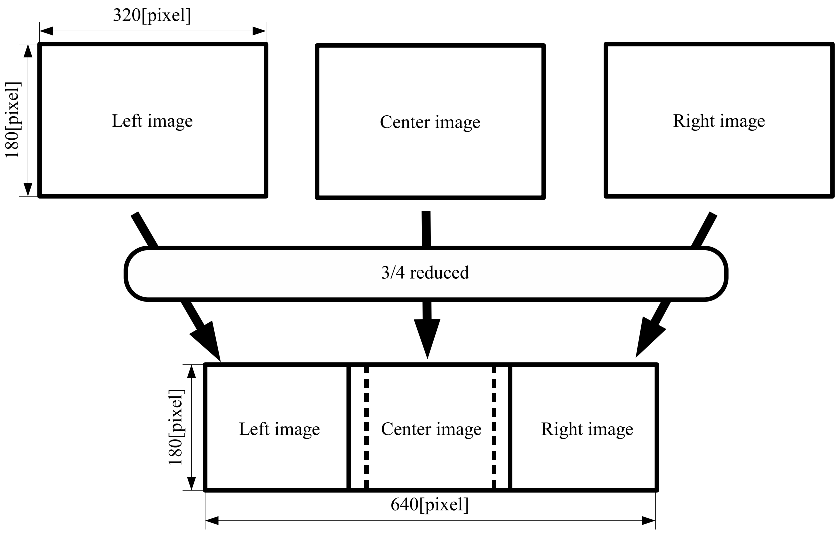


Fig. 4 Panoramic image generation

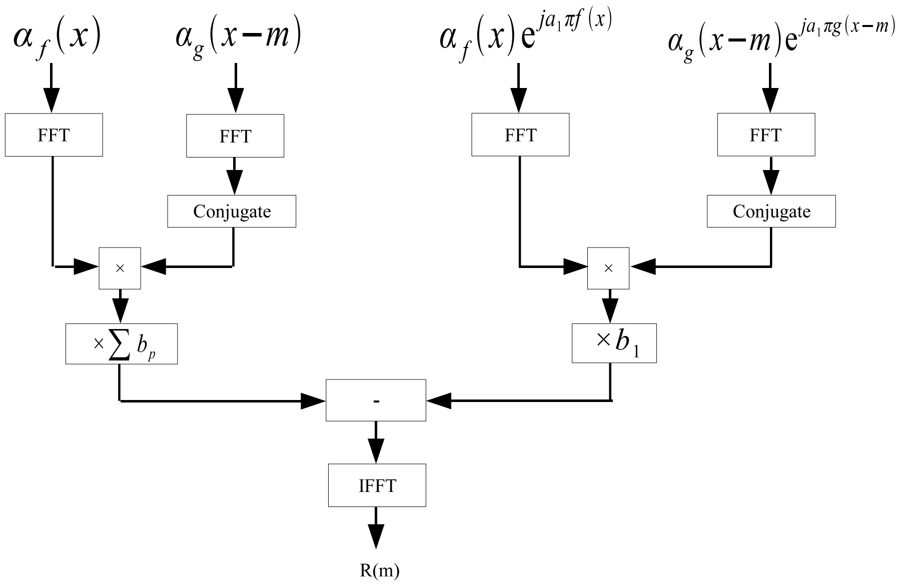


Fig. 6 Flow chart of image comparison

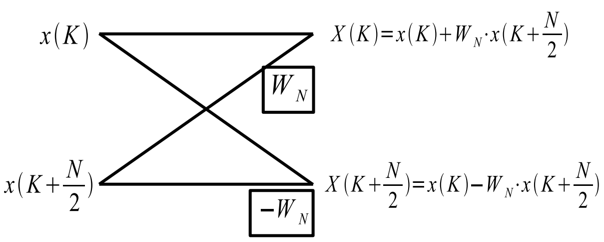


Fig. 7 Butterfly operation

fore starting the butterfly operations. Here, $W_N = W_{Nr} + jW_{Ni}$.

2. Real and imaginary parts of data are read from memories and stored into four registers such as x_0, x_1, x_2, x_3 in FPGA.
3. $(x_2 + jx_3)(W_{Nr} + jW_{Ni})$ is calculated from these data, and the result is preserved in the register t_0, t_1 temporarily.
4. $(x_0 + jx_1) + (t_0 + jt_1)$ is calculated, and the result is preserved in the register X_0, X_1 .
5. The values in the registers X_0, X_1 are stored in the outer memories.
6. $(x_0 + jx_1) - (t_0 + jt_1)$ is calculated, and the result is preserved in the register X_0, X_1 .
7. The values in the registers X_0, X_1 are stored in the outer memories.

5. EXPERIMENTAL RESULTS

To confirm the behaviours of the proposed panoramic image generation function, experiments to generate panoramic image are performed. Figure 8 shows the exaples of three input images obtained by NTSC video camera. The resultant panoramic image generated by the proposed system is shown in Fig. 9. In Fig. 9, we can confirm that the panoramic image is correctly generated by the proposed system.

Fig. 9 Generated panoramic image

6. CONCLUSION

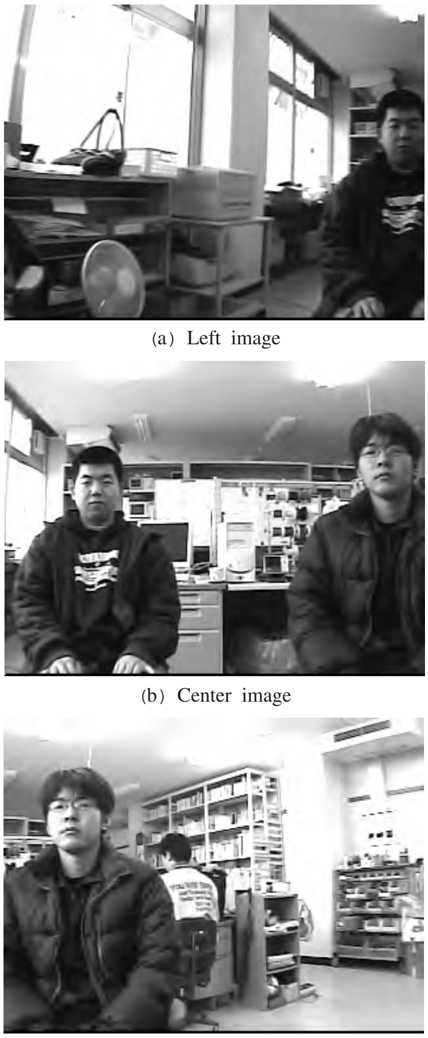
In this paper, realization of panoramic image generation function for teleconferencing video camera is described. The Fast Robust Correlation technique proposed by Fitch et al. is adopted for automatic correction processing technique and is realized as hardware circuits in FPGA.

REFERENCES

[1] Alistair J. Fitch, Alexander Kadyrov, William J. Christmas and Josef Kittler, "Fast Robust Correlation," IEEE (2005).

[2] Jayaram Bhasker, "A VHDL Primer," Prentice Hall, Inc. (1992).

[3] Altera, "Quartus II version 6.0 Handbook Volume 1..6," ALTERA Corporation



(a) Left image

(b) Center image

(c) Right image

Fig. 8 Input images