J-65

Free Viewpoint Television

Department of Information Electronics, Graduate School of Engineering, Nagoya University
Purim Na Bangchang Toshiaki Fujii Masayuki Tanimoto

1. Abstract

In this paper, we propose the new television system named "Free Viewpoint Television" which allows users to freely change viewpoint in realtime. A new view image is generated using a collection of input images captured by array of cameras. The missing ray information between cameras can be obtained by interpolation technique. We describe the system implementation for both hardware and algorithm in detail. The result of the new system is good in term of image quality and refresh rate.

2. Introduction

Even though television is considered to be a good visual receiver tool in visual communication system, it still has some limitations. Many researches are aimed at improving and developing a new television system which will overcome those present limitations.

We considered the new television system named the "Free Viewpoint Television". The aim of this system is to overcome the problem that the television's viewpoint is fixed and can not be changed by users. With the new system, the user is allowed to freely change the viewpoint in realtime.

The basic idea of this system is based on the ray-space method in which the arbitrary photo-realistic view can be generated from a collection of real view images. Theoretically, in order to generate the arbitrary view, the dense ray data is needed. They can be acquired using a very large number of cameras. But it is practically impossible to do so. Thus, captured ray-space always constitutes only a sub-sampled version of the whole information. The missing ray information between each camera can be obtained by interpolation.

Such an idea as the user should be allowed to freely control the viewpoint in realtime was first proposed as the "Bird's Eye View System for ITS"[3]. A hemisphere dome with 16 CCD cameras is constructed for view image capturing. Because the limited number of cameras is used and the distance between each camera is far, the arbitrary view image is generated using model-based interpolation technique which needs a simple 3D object model. In that experiment, a miniature car is used as the object and the arbitrary view image of the moving car was generated in realtime. But because that system used object dependent interpolation technique, its application is limited to only a case that the object can be modeled.

The interpolation technique we use here is based on object independence, ray-space data interpolation[2]. Thus, our new television system does not have shape limitation and can be applied to any object.

3. Proposed System

The system outline is explained as below.

3.1 Hardware

Images are acquired by 16 CCD cameras placed parallel in one dimensional array. The distance between each camera is close, 2cm, in order to obtain easy interpolation. Each camera is connected to one client processor which is responsible for image capturing and image preprocessing (e.g. image rectification). Images captured by clients are sent to server via Ethernet connection. The server's main duty is to generate new view image based on received image information and user's input view position. This configuration depicts as in Fig.1.

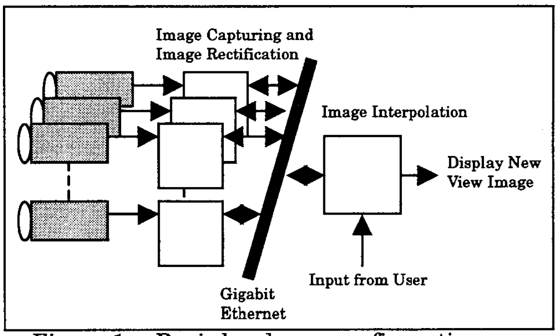

Figure 1: Basic hardware configuration.

3.2 Algorithm

The main algorithm can be separated into 2 parts, Rectification and Interpolation.

3.2.1 Rectification

In our system, set of input images is captured by using the array of cameras. Even though we carefully set camera's position, some small position errors still remain. In this case, the interpolation of multi-camera images is impossible to achieve directly. In order to do interpolation, the rectification must be applied.

Traditional rectification which applies to stereo camera case is not suitable as it doesn't allow multi image interpolation. Special rectification using calibration pattern shown in Fig2 is adopted. The rectification determines a transformation of each image of image array into one plane such that the same set of conjugate epipolar lines become collinear and parallel. The projection matrixes are calculated by assume the same point in world coordinate shall be mapped to same y position of all images.

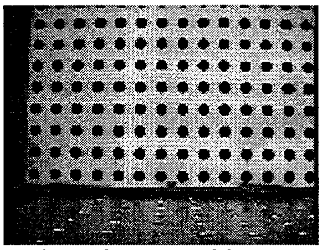

Figure 2: Rectification calibration pattern.

3.2.2 Interpolation

The interpolation is one of the most important parts in our system as it allows us to obtain missing ray data. The technique we use is based on the ray-space data interpolation[2]. The technique has many advantages. First, the interpolated image has better quality compare to the conventional method. Second, the system parameter, such as the number of interpolation input image, can be control flexibly. More importantly, this technique is object independent which allows the system to be applied to any type of object.

The interpolation is pixel-based but the detection of correlated pixel is block-based. The block matching is used to detect the oblique straight line for interpolation. First, we take two blocks on the original horizontal lines and calculate the matching degree between these blocks. Then, the straight line that has the highest matching degree is select and the interpolated pixel value is calculated using the values of the center of these two blocks.

4. Results

We implemented the complete system described above.

The specifications of the computers in this experiment are as follows:

- Server(OS:Linux)
  - CPU : Pentium3 850MHZ
  - Memory : 256MB
- Client(OS:Linux)x16
  - CPU : Pentium3 850MHZ
  - Memory : 128MB
  - Video Capture : IO Data GV-VCP2/PCI
- Network
  - Gigabit Network

As mentioned early, this system needs to operate in realtime. So algorithm parameters are chosen to satisfy the realtime constraint.

The input image is captured at 30 frames/sec, 160x120 pixel resolution. The rectification and interpolation are carried on with the same image resolution. The rectification is done in client processor in order to share some processing task from the server. The interpolation is done using only 2 original images to obtain fast interpolation.

Between each pair of cameras, 15 view images are interpolated. Totally 226 different viewpoints can be selected which allows smoothly change of viewpoint.

The results are shown as in Figure3. Fig.3(a) show the rectified images received from one of the client computers. Fig.3(b) shows the interpolated image.

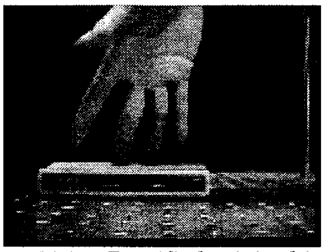

Figure 3(a): Rectified original image.

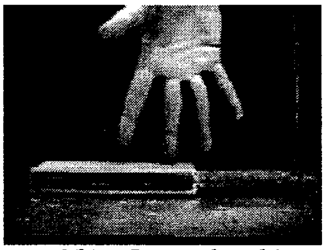

Figure 3(b): Interpolated image.

The quality of the interpolated image is good especially in the zone of interesting (hand image). Some errors still occur at small background object. By using simple interpolation, a small and narrow background object is difficult to interpolate correctly. More complicated interpolation will reduce the errors but it will affect processing time. On the other hand, if we allow little more errors, faster refresh rate can be obtained as we can obtain up to 30 frames/sec. In experiment, the optimum performance of the system, now, is 20 frames/sec.

5. Conclusions and Future Direction

In this paper we proposed the experimental system of Free Viewpoint Television. Using this system, the arbitrary view image of various objects can be generated in realtime. The interpolated image quality is good and the generation rate is quite sufficient for the normal use.

Although the result is satisfied, many points still need to be improved. The image quality must be improved. It can be obtained by developing a new interpolation method. The display rate also needs to be increased by sharing the processing task between the server and the client computer in the network.

6.Reference

[1] T.Fujii, T.Kimoto, M.Tanimoto, "Ray Space Coding for 3D Visual Communication", PCS'96, pp.447-451.

[2] A.Nakanishi, T.Fujii, T.Kimoto, M.Tanimoto,"Ray-Space Data Interpolation for View Image Generation", IDW'01, pp.1385-1388.

[3] M.Sekitoh, T.Kutsuna, K. Toyota, T.Fujii, T.Kimoto, M.Tanimoto,"Bird's-Eye View System for ITS", IEEE, Intelligent Vehicle Symposium, pp.119-123, May 2001.