A Study on Image Editing Technology for Synthetic Sensation

Michitaka HIROSE, Kazuhisa TAKAHASHI, Tomoki KOSHIZUKA
Faculty of Engineering, University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113, JAPAN
Tel:+81-3-3812-2111 (ext.6369), Fax:+81-3-3818-0835
e-mail:hirose@ihl.t.u-tokyo.ac.jp
kazz@sanpo.t.u-tokyo.ac.jp
koshi@hilbert nasl.t.u-tokyo.ac.jp
and
Yoichi WATANABE
Engineering Department 2, Kinden Corporation
5-25-12 Higashi Gotanda, Shinagawa-ku, Tokyo 141, JAPAN
Tel:+81-3-3447-3167, Fax:+81-3-3447-7383
e-mail:watanabe@ihl.t.u-tokyo.ac.jp

Abstract

Several systems have been developed for synthesizing virtual environments from real photographic images. This type of system can be considered as an alternative way of generating virtual worlds as opposed to using computer graphics with 3D CAD models, This system can also easily provide very realistic sensations of being present 3D space. For example, the authors have developed a system called the "Virtual Dome". However, this system only allows the user to look around at previously recorded images from fixed viewpoints. This means that the user is not allowed to move around freely while changing viewpoints continuously in the virtual environment. Without any special algorithms to provide the user with images from all possible viewpoints would require enormous amounts of recorded image data. This paper proposes a new system which can generate an image from various viewpoints by interpolating already recorded images. In addition, the capability and limitations of the first prototype are discussed.

Keywords: Virtual Reality, Virtual Environment, Digital Video Images, Image Editing

Category: Artificial Reality / Virtual Reality

1.Introduction

Most of the virtual worlds are generated by computer graphics with 3D models by using very powerful 3D graphics computers. Recently, however, an alternative way to generate virtual worlds is becoming possible. 3D virtual worlds can be generated from 2D realistic photographic images. This methodology becomes possible because the real image data can be easily captured and synthesized by low cost computers. Until now, several systems have been developed for synthesizing scenes from real images. For instance, the authors have developed a system called the "Virtual Dome" [1], [2]. The limitation of this system is that the user is only allowed to look around from a fixed view point. It is also impossible to walk around in the virtual environment. If we try to support the translational movement of viewpoints in three dimension space, it is necessary to prepare huge image data arrays corresponding to the location in three dimension space. The authors do not believe that this is a practical solution.

The purpose of this study is to propose an alternative solution for synthesizing various images from arbitrary viewpoints, (ie.using the images from various viewpoints taken by the video camera with positional sensors [3]).

2.Concept

As an image recording system, we can use a Video Cassette Recorder (VCR). However, the experience of seeing VCR images does not have any interactivity. The audience has to follow the fixed sequence of images or given camera work as recorded. However, if we can generate new sequences of camera movement by using the previously recorded image sequences, we will be able to experience this image world interactively. Based on this idea, the system shown in Figure 1 can be considered. This system includes three constituent subsystems which are the image recording subsystem, the image database subsystem, and image synthesis subsystem.

The image recording subsystem consists of the video camera and the position sensor. This subsystem is for recording image data and camera position simultaneously, then image data I (x, y, z, θ, ϕ, ω, t) can be obtained and recorded as a sequential data set simultaneously. The image database subsystem is for rearranging the recorded time sequential data into the 3D data space indexed by (x, y,

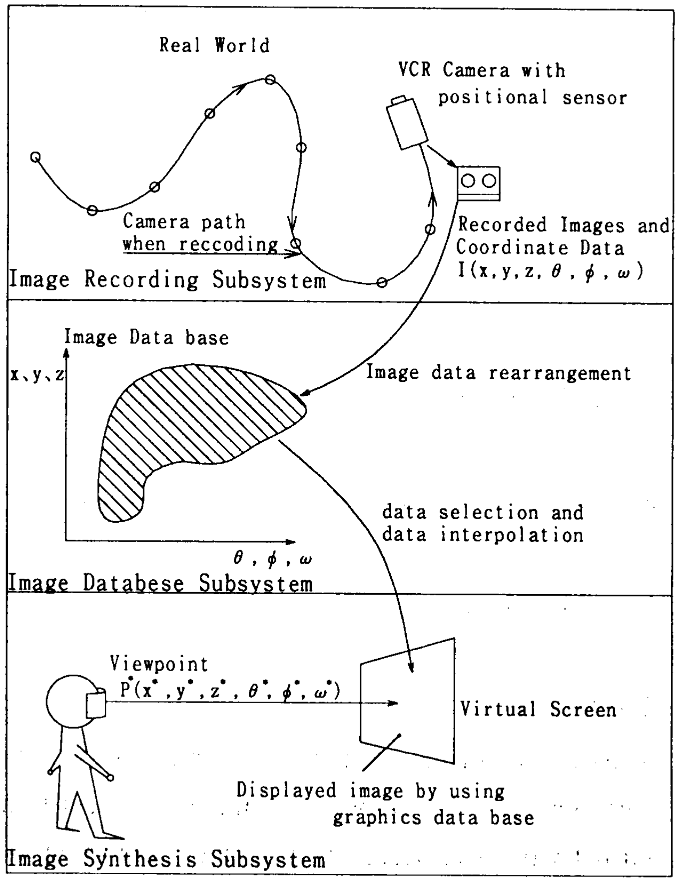

Figure 1: System Concepts

z, θ , ϕ , ω) coordinates. The image synthesis subsystem is for synthesizing and displaying the images corresponding to the new user's viewpoint, by referring to and interpolating the limited image database.

Using this system configuration, the user can experience continuous changes of scene depending on his movement. In other words, the user can walk around a large 3D virtual environment.

3. Basic Implementation

3.1 Image Recording System

The schematic implementation of the Image Recording Subsystem is shown in Figure 2. Three rotary encoders are used as a position sensor in this prototype system. For simplification, degrees of freedom are reduced to three.

The layout of this subsystem is shown in Figure 3. This subsystem includes a cart unit and

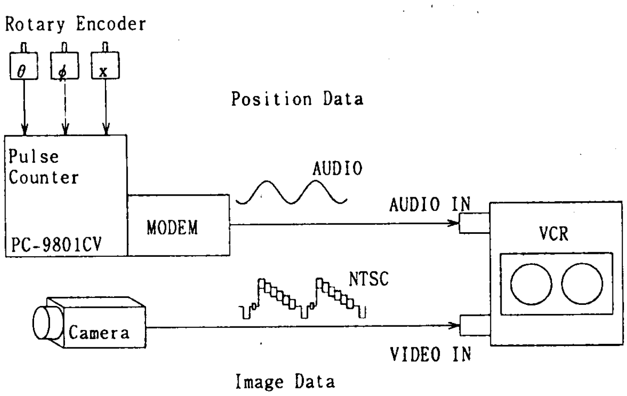

Figure 2: Concept of Image Recording Subsystem

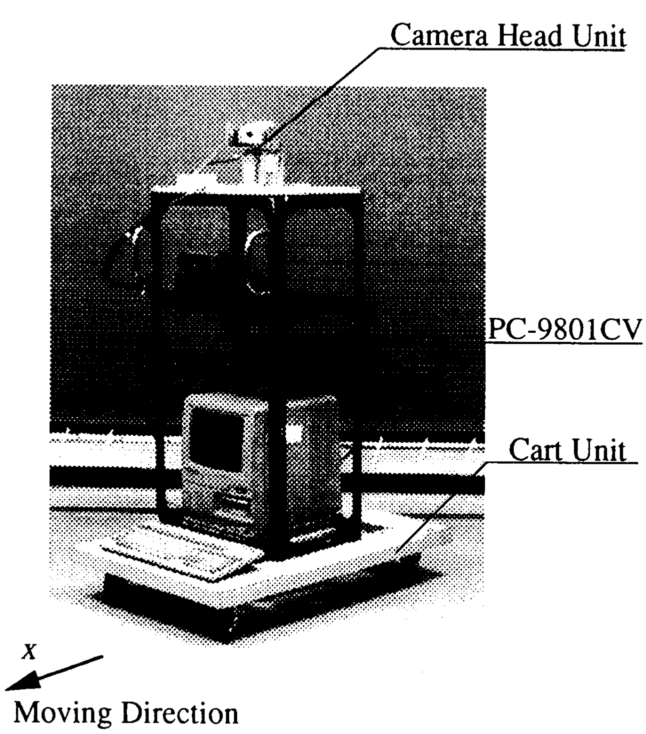

Figure 3: Image Recording Subsystem

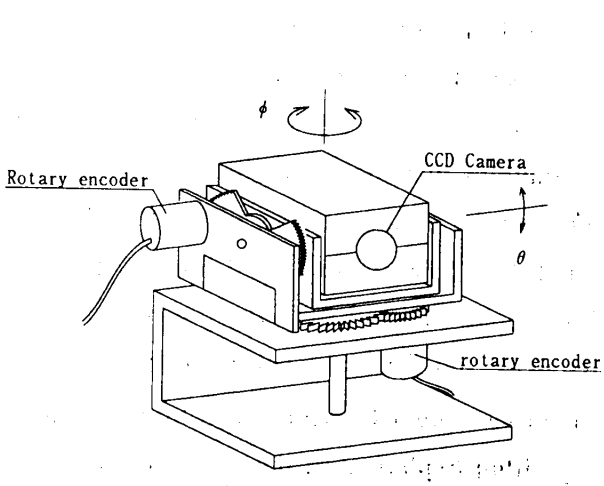

Figure 4: Camera Head Unit

a camera head unit. The cart unit measures the camera's translational movements(x) from the rotation of the wheels. The camera head unit has 2 degrees of freedom as shown in Figure 4. The horizontal angle (ϕ) and vertical angle (θ) of the camera are measured by the rotary encoders. The computer (PC-9801 CV) is used for counting the output pulses from the encoders.

All the positional data sets (x, θ , ϕ) for camera movement are recorded into the sound track of the VCR via a 2400bps modem.

Thus, the output of this subsystem is a video tape which records sequential images and corresponding positional data sets (x, θ , ϕ) .

3.2 Image Database Subsystem

The basic implementation of the Image Database subsystem is shown in Figure 5. The purpose of this subsystem is to capture the recorded images and the positional data into the graphics workstation, and to arrange the time series data into an image database indexed by the positional data.

The recorded image is captured by an up-converter (RGB Spectrum; VIEW 2050) and a Frame-Grabber (Silicon Graphics; Video-Lab) and transferred from the VCR to the graphics workstation (Silicon Graphics; IRIS 4D 210VGX). The resolution of the digitized image is 140×110 pixels and the depth of each pixel is 24 bits (8bits $\times$3(RGB)). On the other hand, positional information recorded on the sound track is transferred to the workstation via modem and serial port. Thus, in the graphics workstation, an array of image data indexed by position data is constructed as $\{I(x, \theta, \phi)\}$.

3.3 Image Synthesis Subsystem

The role of the Image Synthesis Subsystem is to select appropriate images from the database, and then to synthesize a new appropriate image. For this purpose, we used a Graphics workstation (Silicon Graphics: Skywriter) because it was more suitable for quick handling of bitmapped images. Several images are selected from the graphics database and the scene corresponding to a given viewpoint is synthesized by using the algorithm explained in the next section.

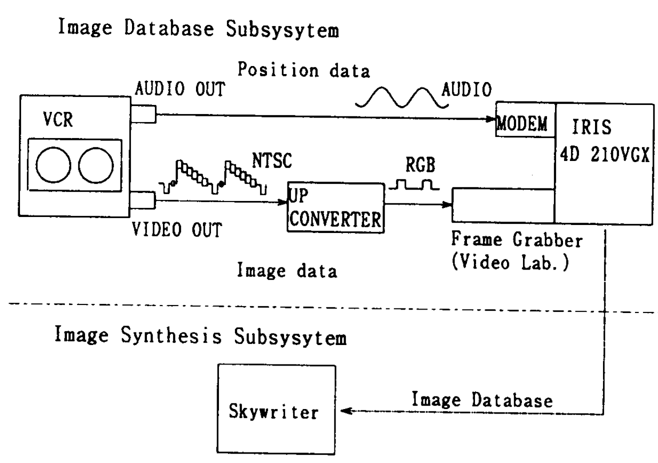

Figure 5: Concept of the Image Database Subsystem and Image Synthesis Subsystem

4. Image Synthesis Method

Figure 6 shows an example of camera movement during image recording. I_1~I_7 are the images taken from camera position P_1~P_7 respectively. By using the Image Database Subsystem, 7 images (I_1~I_7) can be referred to by using P_1~P_7 as a key.

From this Image database, we can generate other sequences of viewpoints P^*_1,P^*_2,P^*_3,etc.. . By applying operation F to I, what we have to do is estimate and synthesize an image I^*_i for P^*_i. Namely,

$$I^* = F (I)$$

Figure 7 shows examples of operation F for a very simple operation.

In case (a); movement is parallel to the original Image, F is just a image shifts of Image I.

In case (b); movement is normal to the original Image, F is a image magnification of Image I.

Figure 8 shows a more complex case in which the user looks around from the location Q*. According to the rotation of viewpoint (P^*_1~P^*_5), images I^*_1~I^*_5 are generated as illustrated. The algorithm used in this prototype system is a combination of such operations mentioned above. Another problem is to select images I for I*. In this prototype, the image nearest to P* is simply selected.

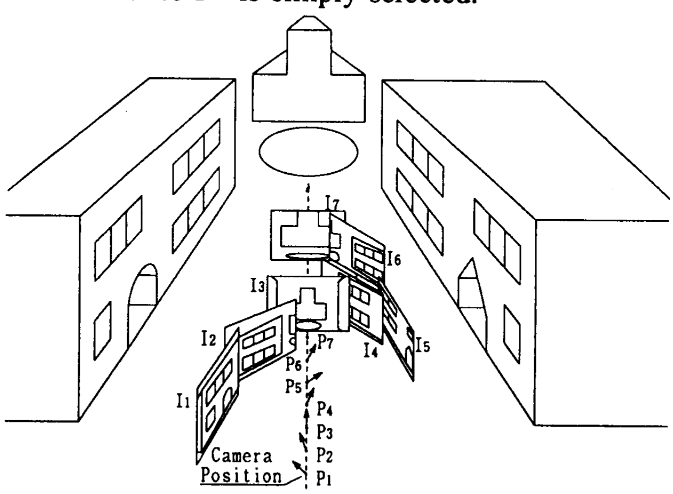

Figure 6: Example of Taking Images

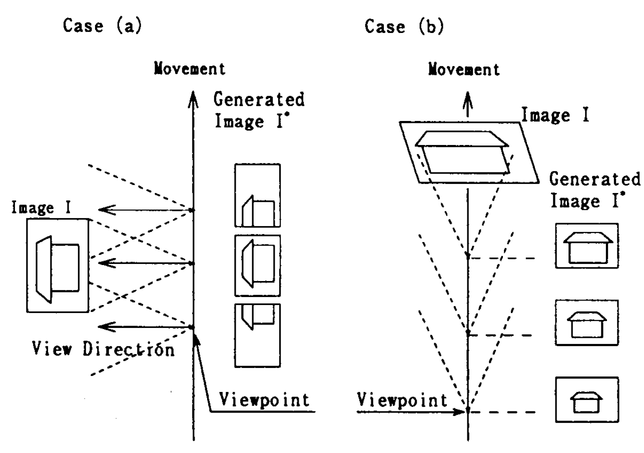

Figure 7: Example of Operation F

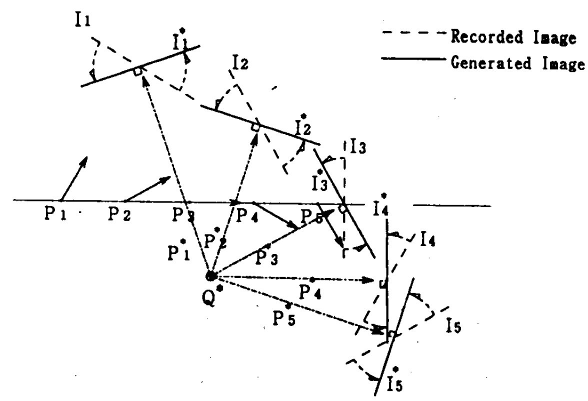

Figure 8: Example of Operation F
(Looking around from the location Q*)

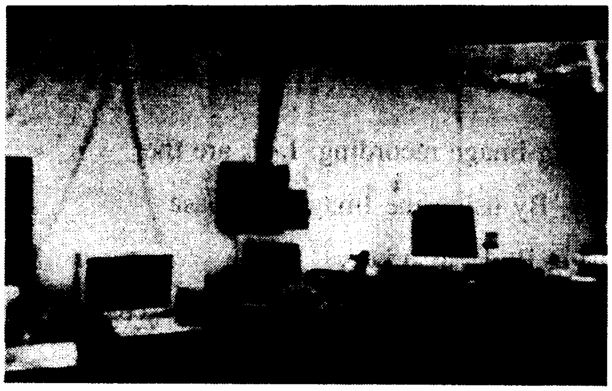

Figure 9-1: Example of Synthesized Image
(Looking in the right handed direction -1)

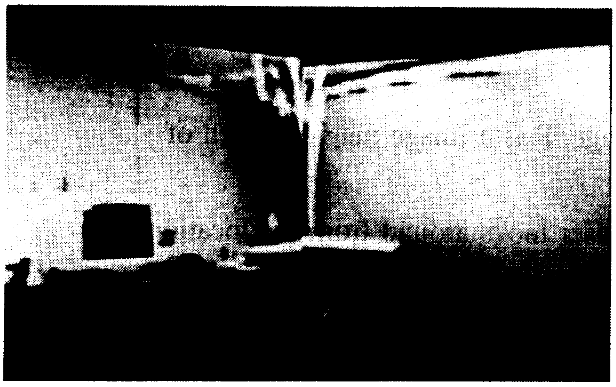

Figure 9-2: Example of Synthesized Image
(Looking in the right handed direction -2)

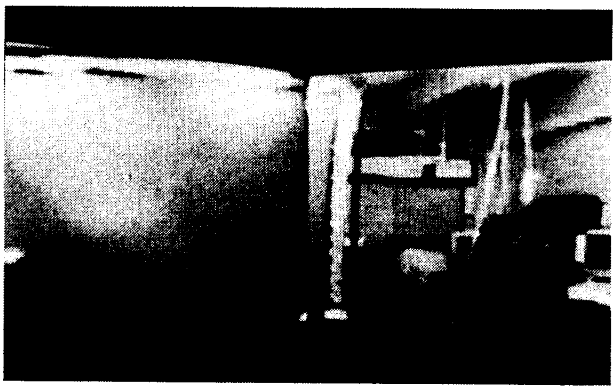

Figure 9-3: Example of Synthesized Image
(Looking in the right handed direction -3)

Figure 9-4: Example of Synthesized Image
(Looking in the right handed direction -4)

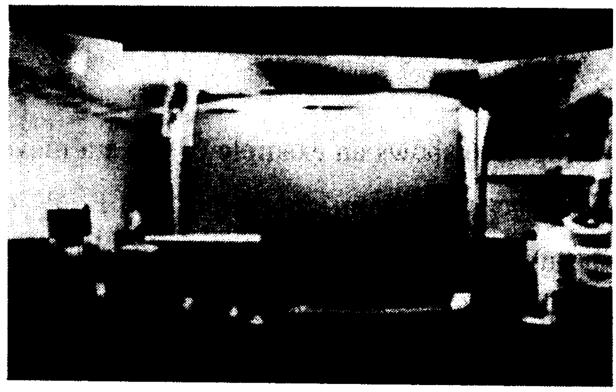

Figure 9-5: Example of Synthesized Image
(Moving Forward -1)

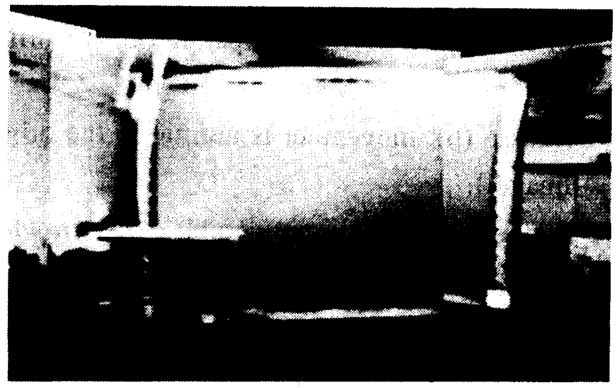

Figure 9-6: Example of Synthesized Image
(Moving Forward -2)

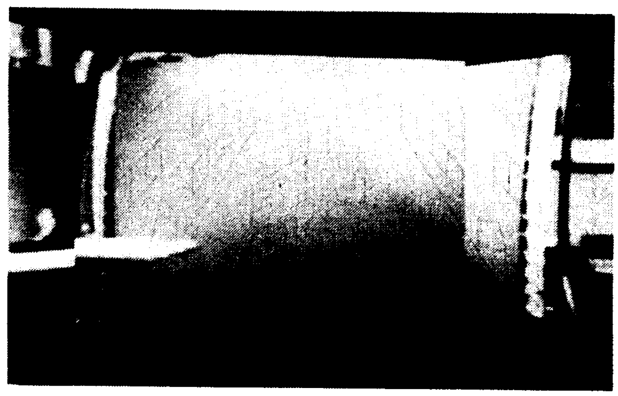

Figure 9-7: Example of Synthesized Image
(Moving Forward -3)

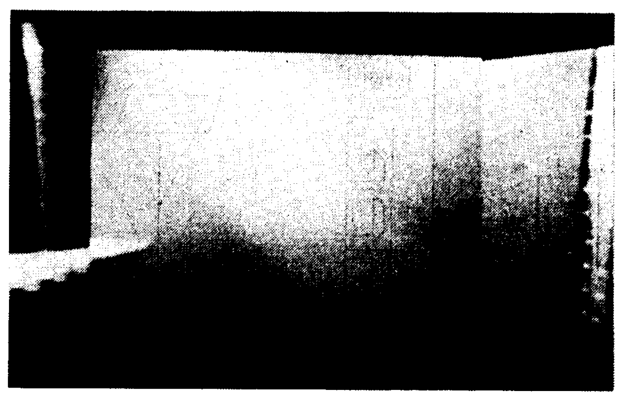

Figure 9-8: Example of Synthesized Image
(Moving Forward -4)

5 Experimental Results and Discussion

An example of the result of image synthesis by the system introduced here is shown from Figure 9-1 to 9-8. First the user is looking around clockwise (9-1~9-4), and then moving forward (9-5~9-8) . When seeing such a experimental scene, most of the users could get a very good sense of presence, even though the prototype system was very primitive. Strictly speaking, there should be large errors in synthesizing motion parallax by using this prototype system. However, interestingly enough, the errors were not so serious because images are very quickly changing due to user movement.

6. Improvement of the Prototype System

Future direction of improvement of this prototype system in future is as follows;

6.1 Use of Advanced Image Interpolation

The image synthesis system used for this prototype is still in a very primitive stage. Operation F was a very simple combination of simple algorithms. As a next step, more sophisticated operation should be considered. If we have two images from different view-points, and the difference between two image is small, we should be able to interpolate the image of the third viewpoint from the two original images. By using this kind of sophisticated frame interpolation, movement of image will become much smoother.

6.2 Use of Digital Video Technology

Currently, captured images are recorded in a hard disk without any image compression. This means the image data base still has a large volume. The authors are considering employing an image compression system. Recently, this kind of digital image processing

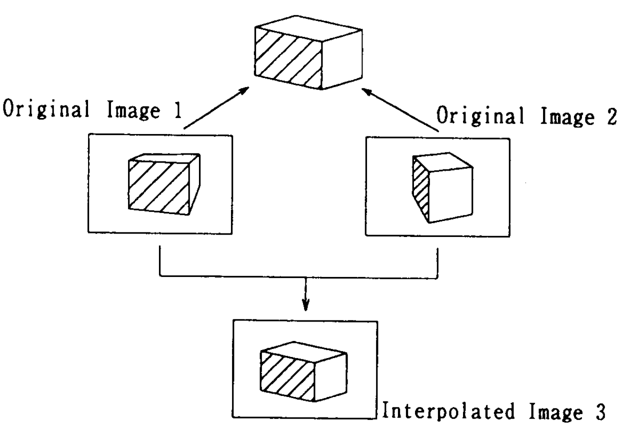

Figure 10: Concept of Using Advanced Image Interpolation

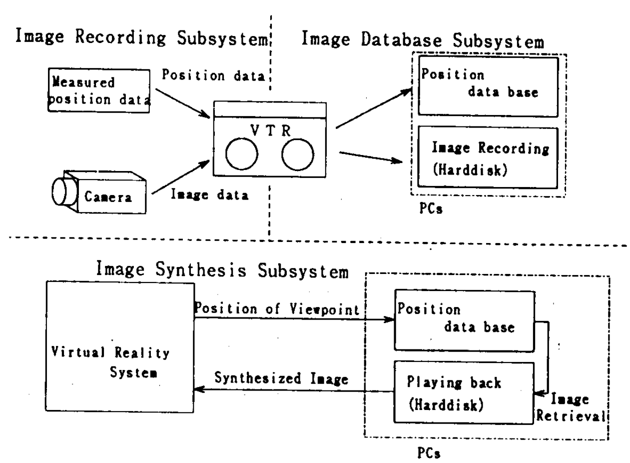

Figure 11: Concept of Using Digital Video Technology

system has archived very rapid progress. Even by using personal computers, processing and editing motion pictures will be possible in a near future.

6.3 Advanced Positional Data Measurement

Currently, only 3 encoders are used for measuring viewpoints; 1 for translation and 2 for rotations. More sophisticated methodology should be prepared. For example, spatial sensors which have more degrees of freedom, wider sensing area and higher resolutions are preferable. The GPS sensor will be one of the candidates for this approach.

7. Conclusions

Ideas to convert video tape image which are non-interactive media into interactive media were proposed. For this purpose, a prototype VCR camera which can record its positional information as well as surrounding images was developed. By using these 2D images with 3D positional information, virtual 3D space could be generated. Conventional ways to generate virtual 3D space involve using computer graphics considering 3D models and rendering technology. However, this methodology is too expensive or too primitive to be used in realistic applications. The authors believe that this paper is proposing a practical alternative. This prototype system, although very primitive, worked very well. Users could feel like that they were walking around in the virtual recorded space. Several ideas for sophistication of this prototype system are proposed. For example, improvement of the image interpolation algorithm or positional sensing methodology should be included in future research.

References

[1] Michitaka HIROSE, Kensuke YOKOYAMA: *VR Application for Transmission of Synthetic Sensation*, The Second International Conference on Artificial Reality and Tele-existence, pp145-154, (1992)

[2] Michitaka HIROSE, Sin'ichi SATO: *Synthesis and Transmission of Realistic Sensation Through Virtual Reality Technology*, Proceedings of 9th Symposium on Human Interfac e, pp111-116, (1993)

[3] Yoichi WATANABE, Michitaka HIROSE, Kazuhisa TAKAHASHI and Tomoki KOSHIZUKA :"*A Study on Image Editing Technology for Synthetic Sensation.*", Human Interface News and Report, Vol.9, No.3, pp273-278, (1994)