Realizing a Mixed Reality Space Guided by a Virtual Human ¾ Creating a Virtual Human from Incomplete 3-D Motion Data

Shinsuke Abe*, Iku Yamaguti*, Joo Kooi Tan*, Seiji Ishikawa*

*Department of Mechanical and Control Engineering
Kyushu Institute of Technology
Sensuicho 1-1, Tobata, Kitakyushu 804-8550, Japan
email: abe@is.cntl.kyutech.ac.jp

Abstract: Recently the VR technique has evolved into a mixed reality (MR) technique, in which a user can observe a real world in front of him/her as well as virtual objects displayed. This has been realized by the employment of a see-through type HMD (S-HMD). We have been developing a mixed reality space employing the MR technique. The objective of our study is to realize a virtual human that acts as a man-machine interface in the real space. It is important in the study to create a virtual human acting naturally in front of a user. In order to give natural motions to the virtual human, we employ a developed motion capture technique. We have already created various 3-D human motion models by the motion capture technique. In this paper, we present a technique for creating a virtual human using a human model provided by a computer graphics software, 3D Studio Max©. The main difficulty of this issue is that 3D Studio Max© claims 28 feature points for describing a human motion, but the used motion capture system assumes less number of feature points. Therefore a technique is proposed in the paper for producing motion data of 28 feature points from the motion data of less number of feature points or from incomplete motion data. Performance of the proposed technique was examined by observing visually the demonstration of some motions of a created virtual human and overall natural motions were realized.

Keywords: Mixed reality, See-through HMD, Motion capture, Interpolation

1. Introduction

Recently, with the evolution of the Hyper Transport (HT) technology, Mixed Reality (MR) research has been developed actively and rapidly. MR is a technique that merges the real world and a virtual world seamlessly in real time so that it becomes an equal part of our natural environment. The technique is applicable to medical treatments, welfare, disaster prevention, city planning, amusement, *etc*.

The 3-D data obtained from the proposed motion capture technique [1] was expressed employing points or wire frames in the former version of our study. However this expression lacked reality as a virtual object. In this paper we propose a technique for transforming the 3-D data into the form that are acceptable by the CG software (3D Studio Max©) to generate a virtual object whose reality is higher, *i.e.*, more natural than points or wire-frame expression. The main difficulty of this issue is that 3D Studio Max© claims 28 feature points for describing a human motion, whereas the used motion capture system assumes less number of feature points. Moreover occlusion of some feature points is often the case with 3-D human motion recovery; *i.e.*, feature points attached to some part of a human body are often hidden by another part of the body by motion. To solve these difficulties, a technique is proposed in this paper for producing motion data of 28 feature points from the motion data of less number of feature points or from incomplete motion data.

We then represent the obtained virtual object *i.e.* virtual human in the developed MR space. The S-HMD which we used for displaying a virtual human in the real world has two half-mirrors; one in front of the right eye and the other in front of the left eye: User can see the real world in front through the half-mirrors, and, at the same time, the user observes virtual objects displayed on two CRTs through the half-mirrors.

In the following part of the paper, interpolation and unification techniques are presented in section 2 and section3, respectively. Experimental results are shown in section 4 and discussion and conclusions are finally given in section 5.

2. Interpolation Method of Feature Points

According to the literature [1], a novel 3-D shape recovering technique based on uncalibrated multiple video cameras is outlined in the following.

In the proposed motion capture system, more than three cameras are placed around a non-rigid objects *i.e.* a human in motion. The cameras take the motion images, producing three or more image streams. Some markers are attached on the non-rigid objects as feature points. Correspondence of the feature points among the images is taken at each observation time to yield a single extended measurement matrix $\tilde{W}$. Factorization is applied to the matrix $\tilde{W}$ to yield a shape matrix S that provides the 3-D coordinates of all the feature points on the object at every sample time.

Although the 3-D coordinates of 15-17 feature points a re acquired at every sample time by the above procedure, it is not enough for expressing a natural virtual human by that amount of 3-D data. One may think that increasi ng the number of cameras may solve this problem, but t his may result in new problems such as extraction of lar ge number of feature points, integration of all the 3-D fe ature points, *etc.*

We present an interpolation method to create a natural human model provided by a computer graphics software, 3D Studio Max©. However, the software claims at least 28 feature points for describing a natural human model. Therefore, interpolation technique needs to be developed for creating 28 feature points from the obtained 15 to 17 feature points. In addition to this interpolation issue caused by the difference of the used basic number of feature points, those missing feature points caused by occlusion must be interpolated as well. The occlusion of feature points occurs when some part of human body hides feature points on the body by motion.

In this paper, we propose two techniques for interpolating feature points from motion capture data: (i) interpolation by parallel shift, and (ii) interpolation by vectors. In **Fig. 1**, light gray circles show positions of the 15 feature points used in the motion capture experiment and the black circles indicate the position of the interpolated feature points. In the case of 17 points, feature points of tiptoes of feet are added.

The interpolation that uses parallel shift calculates the direction of the coordinates and the width (or depth) of the interpolation from given feature points' locations. **Figure 2** shows the interpolation by parallel shift. The point of a right waist is transformed parallel to the origin of the coordinate system in the first place. The angle the straight line of the waist and the *x* axis make is then calculated and the straight line receives rotation around the axes to coincide with the *x* axis by using the 3-D affine transform. The point $P(x, y, z)$ in the 3-D space is transformed into the point $P'(x', y', z')$ by 3-D affine transform defined by the following equation;

$$[x' \quad y' \quad z' \quad 1] = [x \quad y \quad z \quad 1]\begin{bmatrix} a & b & c & 0 \\ d & e & f & 0 \\ g & h & i & 0 \\ t_x & t_y & t_z & 1 \end{bmatrix} \qquad (1)$$

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} \neq 0$$

Actually the rotation in Eq.(1) only contains the rotation *q* around the *z* axis, since relative locations have simply to be determined with respect to the feature points to be interpolated. Thus Eq. (1) is written in the following way;

$$[x' \quad y' \quad z' \quad 1] = [x \quad y \quad z \quad 1]\begin{bmatrix} \cos q & \sin q & 0 & 0 \\ -\sin q & \cos q & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad (2)$$

And the parallel shift is used and interpolated for the black feature points in Fig. 1. The width/depth of the interpolation here is calculated from the distance of two points composing the waist. After the interpolation, all the feature points are described in the original coordinate system by inverse application of the rotation by *q* and the parallel translation.

The interpolation method using vectors computes the outer product of vectors defined on the feature points. In this interpolation method, judgment is indispensable for determining if the feature point exists in front or at the rear of a human model. If we denote two vectors by

$$\mathbf{a} = (a_x \quad a_y \quad a_z), \quad \mathbf{b} = (b_x \quad b_y \quad b_z)$$

calculation of the outer product is given by

$$\begin{cases} (\mathbf{a}\times\mathbf{b})_x = a_y * b_z - a_z * b_y \\ (\mathbf{a}\times\mathbf{b})_y = a_z * b_x - a_x * b_z \\ (\mathbf{a}\times\mathbf{b})_z = a_x * b_y - a_y * b_x \end{cases} \cdot \qquad (3)$$

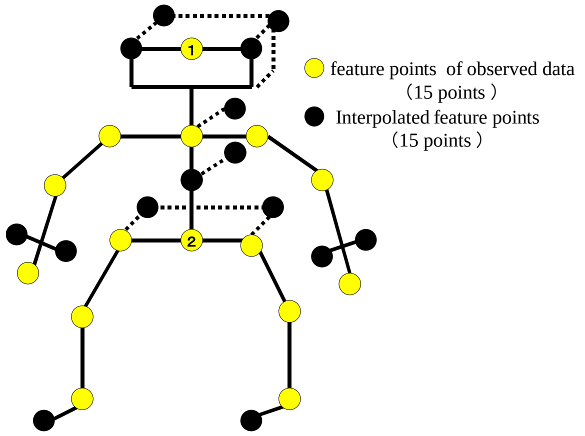
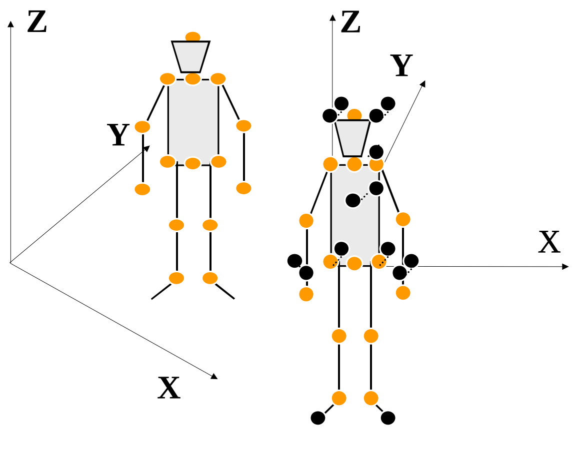


Fig.1 Positions of the observed feature points and interpolated feature points.



(a) (b)

Fig.2 Interpolation by parallel shift: (a) Experimental coordinate system, and (b) parallel shift and rotation around the axes.

Figure 3 explains the interpolation using vectors. First of all, the interpolated width is calculated from the outer product of two vectors (normal vectors) employing two points of both shoulders and points at waist or two points of both waists and points of chest. Then the following vectors are calculated; vectors from chest feature points to shoulder feature points, vectors from elbow feature points to right and left hand feature points, and vectors from waist feature points to chest feature points. Using these vectors, the interpolation is done. The point in front of the head is interpolated using the vector from the chest feature point to the shoulder feature point. The point at the back of the head is interpolated using the outer product vector of the chest. The middle point in the body is interpolated using the vector from the waist feature point to the chest feature point. Two points of the back are interpolated using the outer product vector of the chest. The point at the back of the waist is interpolated using the outer product vector

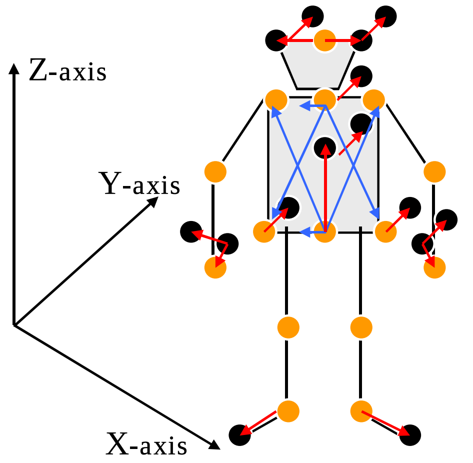


Fig.3 Interpolation by vectors.

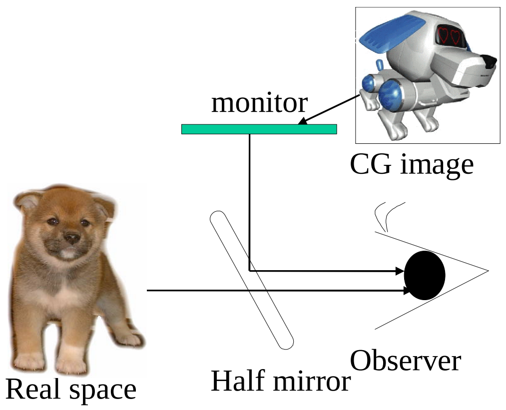


Fig.4 Unifying principle of virtual and real spaces employing see-through HMD.

of the waist. The point at the arm is interpolated by the vector from elbow feature point to hand feature point and the vector of the shoulder. The point at the foot ahead is interpolated using the outer product vector of the waist.

3. Principle of Unifying and Displaying Real Space and Virtual Space [2]

The position and the direction of the head in the real space are sensed with the magnetic sensor. A virtual character with high level of reality is unified and displayed in the real space by using HMD employing the above information. In order to realize a MR space, a display is important which does not have sense of incongruity in displaying real images and virtual images. The display used by this study is an optical see-through HMD of the transparent type. **Figure 4** shows the principle of the optical see-through HMD.

The optical see-through HMD is realized by putting the half mirror in front of a user. The spectacle a user is facing in the real world penetrates directly through the half mirror to the eyes. The user can see simultaneously with the real world a virtual world by reflecting virtual objects to the half mirror. By this system, the real world and the virtual world are unified on the retina. Though the bright part of the displayed virtual object is visible through the half mirror, its dark part is not visible because of no reaching light. Therefore, if the background beyond the virtual object displayed is made dark in the virtual space, the reflected light from the real space in sight penetrates the low portion of brightness, and the real space is observed by the user.

4. Experimental Results

Figure 5 shows part of the data ‘walking’ and **Fig.7** shows part of the data ‘standing up and sitting down’. They show original images of some sample times. 3-D recovered models of using the proposed interpolation by vectors are shown in **Fig.6** and **Fig.8**, respectively.

Moreover, in the experiment, the two cases of moving virtual human were displayed in a see-through HMD to realize a MR space. **Figure 9** shows part of the images observed through the see-through HMD, when the simulation is executed.

5. Conclusions

A technique was presented for producing a virtual character whose reality is higher than points set display or wire frame display employed in the former version of the study. For this purpose, a 3-D computer graphics software was employed for making a virtual human model. In order to use incomplete motion capture data, two methods were proposed that interpolate locations of missing feature points. The created virtual character was then unified to the real space through a see-through HMD and its effectiveness was confirmed. In the future, more real virtual character expression is scheduled to be examined by adding improvements to the developed MR system and the interpolation methods.

References

[1] Tan, J. K., Ishikawa, S.: “Human motion recovery by the factorization based on a spatio-temporal measurement matrix”, *Computer Vision and Image Understanding*, **82**, 2, 101-109, 2001.

[2] Tan, J. K., Yamaguchi, I., Ishikawa, S.:“Unifying a Virtual Human from Motion Capture Data into Real Space”, Int. Conf. on Robotics, Vision, Information, and Signal Processing, ROVISP2003.

Fig.5 Some of original images of data 'walking'.

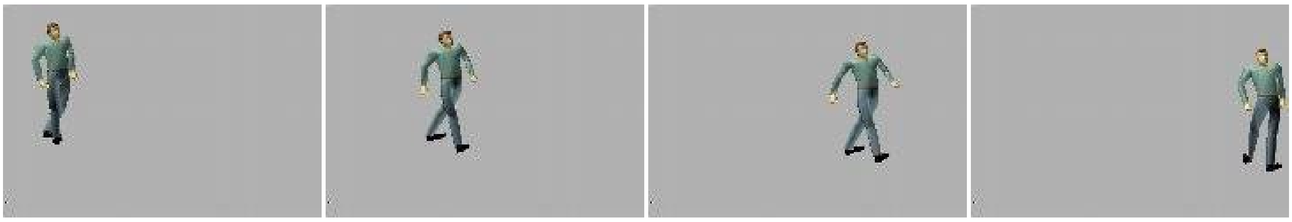


Fig.6 A 3-D model of data 'walking'.

Fig.7 Some original images of data 'standing up & sitting down'.

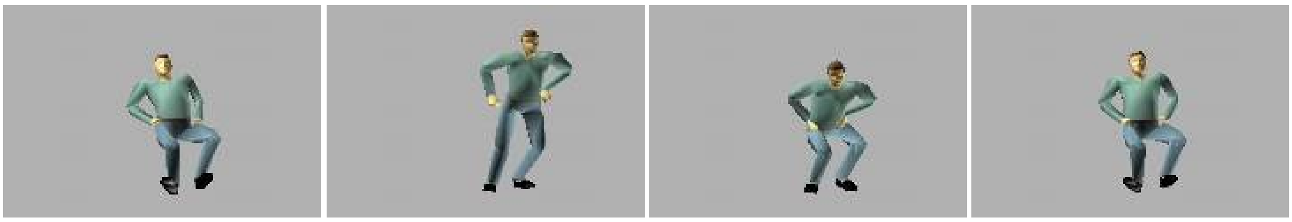


Fig.8 A 3-D model of data 'standing up & sitting down'.

Fig.9 Part of the images in the MR space observed through the monitor of HMD, when the simulation is executed