A Study on data input of natural human motion for virtual reality system

Michitaka HIROSE
Guy DEFFAUX and Yoshiyuki NAKAGAKI

Department of Mechano-Informatics, University of Tokyo,
7-3-1 Hongo, Bunkyo-ku, Tokyo 113, Japan
E-mail : {hirose, guy, nakagaki}@ihl.t.u-tokyo.ac.jp

Abstract

This paper describes the authors' motion capture system by which the motion of a virtual body, identical to the real human body, can be directly controlled by data supplied from sensors positioned on the user's limbs (head, arm, leg, foot, etc...). The main point of this study concerns the optimal arrangements of the sensors. We focused on the treatment of the data supplied from the sensors to obtain a realistic (performed with an inverse kinematics filter) and natural (performed with a data smoothing filter) body motion. As an application of this motion capture system, automatic work analysis is also introduced in this paper, showing how body posture categories based on the OWAS method can be automatically classified.

1. Introduction

The purpose of a motion tracking technique is to generate virtual human motion by duplicating real human motion. In our laboratory, we have developed a motion capture system based on a magnetic spatial sensor. The main part of this research concerns the method of human body modeling and the sensor environment. The minimum number and arrangement of the sensors must be obtained to give realistic and natural movement to the virtual body (created in the virtual space) as a copy of the working person in the real world.

One of the important applications of this technology lies in automatic work analysis. Using virtual reality technology, one method of analysis is presented in this paper. Using this method, OWAS codes could be easily generated automatically by measuring the virtual body.

2. Motion Tracking System

2.1 Space Sensors

Various technologies can be considered to complete the motion capture operation. Basically, however, two different approaches can be considered.

The first approach is to capture the subject's motion by using a video camera and to obtain 3D information from the 2D pictures by using image processing technology. The advantage of this method is that the real human body is completely free from measuring or monitoring devices. However, since data from this method has much noise, real time analysis is very difficult.

The second approach is to attach 3D space sensors to the subject's body (ie. direct measurement). This methodology is more suitable to real time measurement.

There are currently four main technologies which exist for the measurement of position in 3D space : link, magnetic, ultra-sonic, and optical sensors. For the purposes of this study, in order to trace all subject movements, sensors are required to :

1) Allow movement without restriction. This means that the actor's activity should not be limited by the measurement system (sensors, emitters, receptors, cables, etc...)

2) Have enough precision (ie. less than 5cm error on any measurement).

3) Allow real time measurement

4) Require only marginal processing for data input

5) Be reasonable in cost

A sensor called Fastrak, developed by Polhemus, satisfied the above requirements and was used for implementation in the authors' system. Each sensor has 6 degrees of freedom (DOF) that are represented by position (x,y,z) and orientation (pitch, yaw, roll). Other technical characteristics of this sensor are given in Table 1.

Measurement precision : position : 0.8mm, angle : 0.15° Simultaneous measurement of 16 receivers Sampling : one data set 120 times per second Domain with standard precision : 75cm radius sphere Largest measurement domain : 300 cm radius sphere Magnetic source amplification possible

Table 1. Specifications of Polhemus magnetic sensors

2.2 *Body Model and Algorithm*

It is said that the human body consists of more than 200 bones and has very high degrees of freedom. Therefore, it would be impossible to measure the exact position of all these parts with a restricted number of sensors. Thus, a human body model must be simplified by means of a hypothesis concerning body limbs and articulations.

In this paper, the presented model is skeletal. This leads to a model like the one shown in Figure 1. If we select the chest first, as a reference point, it has 6 DOF. Linked to it, the neck has an additional 3 DOF, followed by the shoulder (3 DOF), elbow (1 DOF), and forearm (1 DOF due to rotation). For the lower half body, the pelvis has 3 DOF, the hip 3 DOF, the knee 1 DOF, and 2 DOF for the ankle. Consequently, this basic model has 34 DOF in total.

Regarding to the model definition, the following 3 major questions have to be answered :

What should be the degree of complexity for the body model?

Of course, the answer to this question depends on the type of application such as movies, games, industrial work, etc. However, the first goal of this study is aimed at posture analysis. In the next chapter, the necessary data items for our posture analysis will be listed. The important body parts are the chest, wrist, hips, knees and feet. The center of gravity must also be processed. To obtain this information with a minimum number of sensors, articulations were simplified in a low DOF link. After simplification, the obtained model is the basic model which was previously shown in Figure 1.

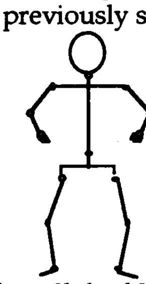


Fig. 1 Skeletal Model

What is the minimum number of sensors from a theoretical/practical perspective ? Where should they be attached ?

Each Polhemus sensor has 6 DOF. Therefore, 6 sensors should be enough because 6 sensors x 6 DOF = 36 DOF, which is larger than the 34 DOF of our basic model which defines all the necessary main parts. Due to the complicated architecture of all the links (called the kinematic chain), for any combination sensor attachment, the entire body posture can not be completely deduced just from the sensors. This is only possible with 7 sensors or more. The cumbersome demonstration is not given here. Simply considering all the attachment possibilities leads to this conclusion.

With 7 sensors, 42 DOF can be obtained. However, this leaves us with 8 DOF which are redundant when considering the case of our model. Generally speaking, when the DOF of the sensors exceeds the DOF of the real body which is being modelled, data is excessive and measurement error is synonymous with a violation of the problem hypothesis. Thus, a method of data correction is necessary to modify the mismatched data in concordant data. On the other hand, when the DOF of the sensors is less than the defined DOF of the subject's body, all parts of the body can not be completely defined.

How should sensor data be treated ?

In all our experiments, Polhemus sensor measurement error was not considered a crucial issue. It was usually more than 0.8mm (cf. Table 1) due to magnetic perturbation, but never exceeded 5 cm. In conclusion, serious data correction was not required. A simple inverse kinematic algorithm was effectively employed.

As a first hypothesis, the position given by the chest sensor can be considered to be sufficiently accurate [step A]. The main reason is because the chest sensor is nearest to the magnetic source. The magnetic field at this point is almost never affected by outside elements with electromagnetic properties. The sensor data from other portions (head, forearms, etc...) are then received by the computer [step B]. With the hypothesis from the body model, the position of the undetermined parts can be calculated. For example, in Figure 2, elbow DOF gives upper arm orientation for longitudinal axis [step C]. Finally, any aberration is just corrected by a simple translation of the not-yet-linked parts to the chest [step D].

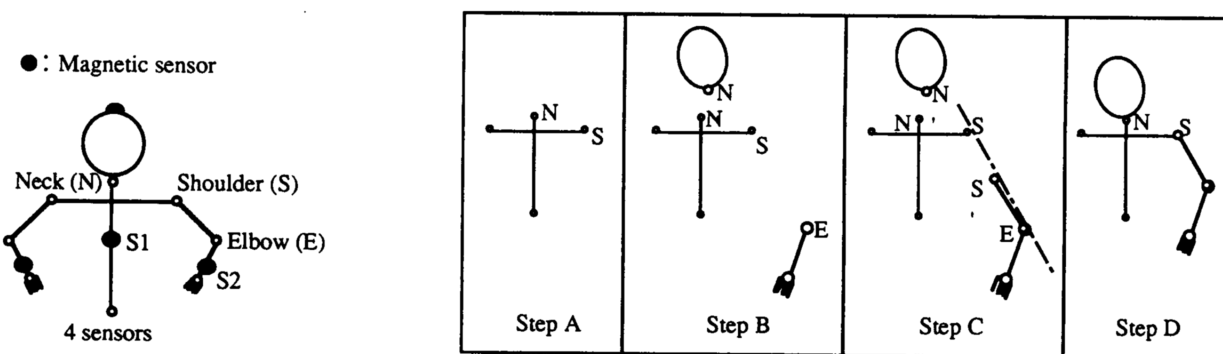
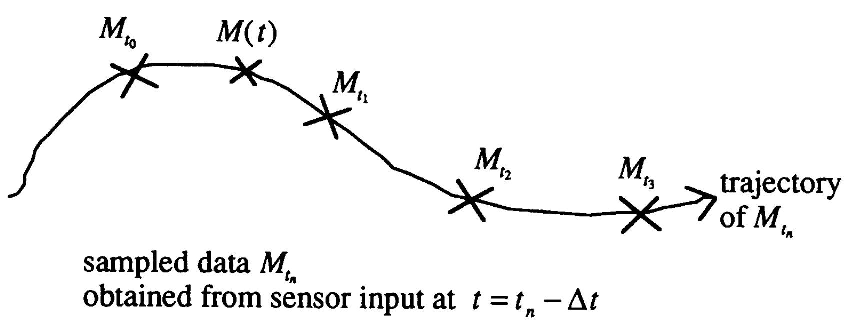


Fig. 2 Correction of measurement aberration

Experimentally, due to the small errors of the sensors as mentioned above, the motion of the virtual body looked considerably "natural" in most cases. One drawback of this algorithm, however, was that articulations looked broken in a few cases, especially for elbows and knees. The fact is that all articulations have a movement range and our real arm cannot stretch outwards for example. These limits are not taken into consideration in the current version of our model algorithm. Future work in our laboratory will look at taking this problem into consideration.

Another problem is data interpolation. In the case of the Polhemus sensors, they are quite slow. From the Polhemus deck, the data import for each one sensor takes a fixed duration (1/120 sec.). This means that the data refresh rate takes n times longer when n sensors are used. For instance, in case of 8 sensors, data sampling rate Δt is 1/15 seconds. Under such a low data sampling rate, fast animation can not be directly created. Between the given data, real-time interpolation is necessary for smooth animation of the virtual body. Hence, the authors employed the hypothesis that position and speed vector must be continuous.



The trajectory of M is a polynomial of degree 2 because this is a dynamical system.

$\vec{M}(t) = \vec{M}_{t_n} + (t - t_n) \cdot \vec{V}_{t_n} + (t - t_n)^2 \cdot \vec{A}_{t_n}$ *for* $t \in [t_n, t_{n+1}]$ where M_{t_n} is the position data at $t = t_n$ obtained from the sensor input at $t = t_{n-1}$, $\vec{V}_n$, $\vec{A}_n$ are the coefficients related to velocity and acceleration. $\vec{V}_n$, $\vec{A}_n$ can be calculated from the condition that $\vec{M}(t)$ and $\frac{d\vec{M}(t)}{dt}$ are continuous, namely :

$$
\begin{cases} \vec{M}(t_{n+1}) = \vec{M}_{t_{n+1}} = \vec{M}_{t_n} + \Delta t \cdot \vec{V}_{t_n} + \Delta t^2 \cdot \vec{A}_{t_n} \\ \dfrac{d\vec{M}(t)}{dt}(t_{n+1}) = \vec{V}_{t_{n+1}} = \vec{V}_{t_n} + 2\Delta t \cdot \vec{A}_{t_n} \end{cases} \quad \text{Eq. (1)}
$$

3. Experimental Design

The first algorithm mentioned above was examined experimentally. In particular, the relation between the realism of the movement and the number and position of the sensors was examined. The schematic diagram of the experimental system is shown in Figure 4, and the geometrical arrangement of the motion capture system is shown in Figure 5.

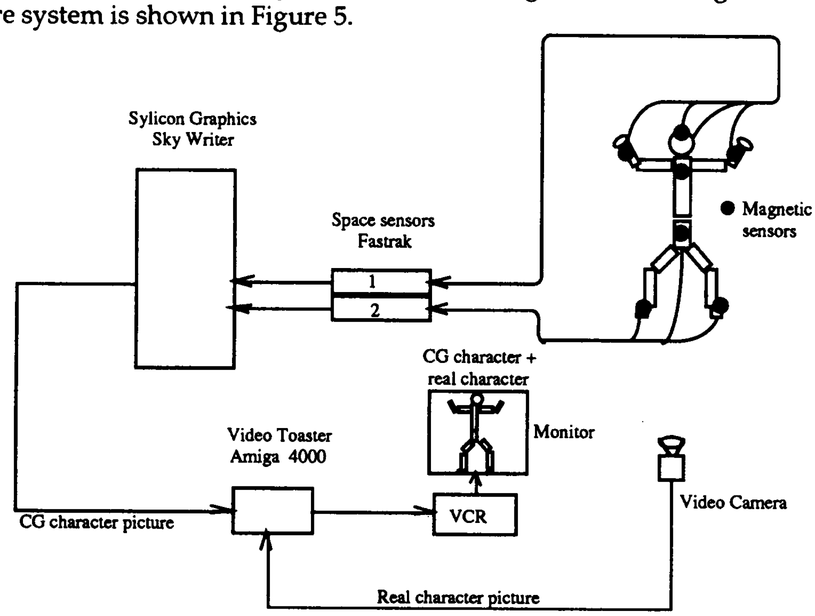


Fig 4. System layout

While the real human body is in action, his/her motion is taken by CCD camera, positional data is measured by the sensors, and the CG character is drawn in the workstation. Both images of the camera image graphics (real body) and the CG character (virtual body) are then superimposed, as shown in Figure 6. By the method of superposition, it was possible to evaluate the CG character resemblance with the real character. Of course, this superimposed image could be recorded by a VCR.

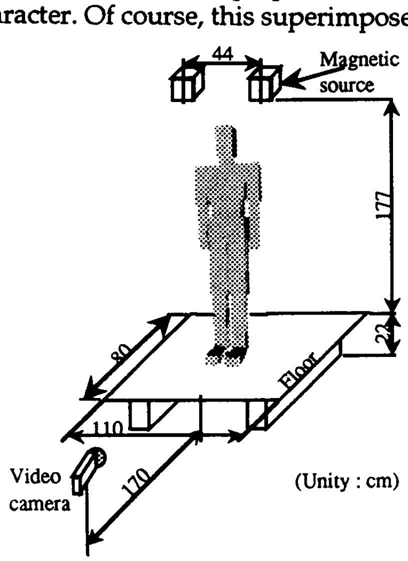


Fig. 5 Position of magnetic source and body

Fig. 6 Image superimposed by CG character and real character

There were 3 series of experiments : half upper body, half lower body and entire body. For each series, several sensor arrangements were provided.

3.1 Half Upper Body Movement Input

To measure the arm movement, two types of sensor arrangements have been tried, as shown in Figure 7. One way is with 6 sensors (head, chest, upper, and forearms) and the other one way is with 4 sensors (head, chest, and forearms).

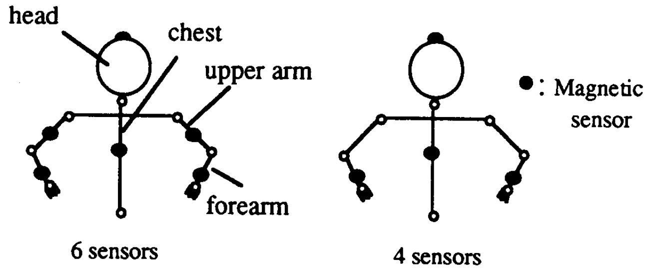


Fig. 7 Sensors positions for the upper half body

3.2 Half Lower Body Movement Input

Four kinds of sensor arrangements were tried, as shown in Figure 8. The body parts concerned are the hips, thighs, legs, and feet. The case with 3 sensors is the one used also for all body attachments.

3.3 Entire Body Movement Input

These experiments have been performed after the two previous sets of experiments. Using 7 sensors, located at the head, forearms, chest, hips, and feet. The arrangement is shown in Figure 9.

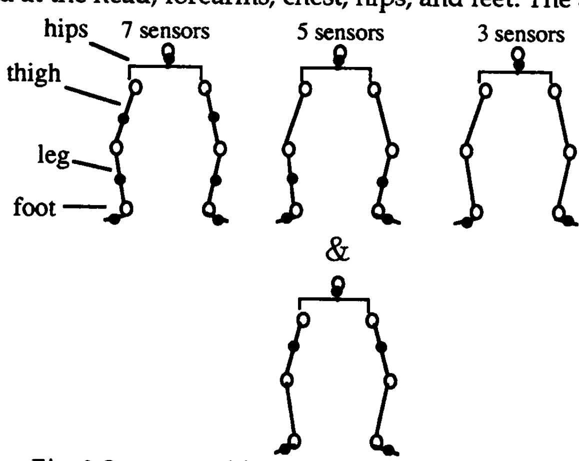


Fig. 8 Sensors positions for the lower half body

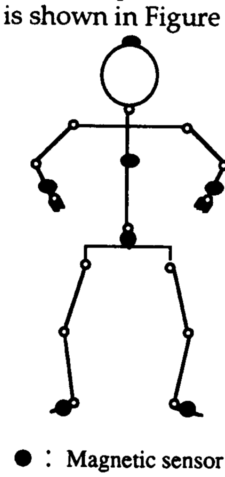


Fig. 9 Sensors positions for the entire body

4. Results

4.1 Half Upper Body

Both results of 6 and 4 sensor arrangements looked almost the same. In the case of 4 sensors, the movement input is enough to play back a natural and precise movement. Arms, chest and the head of the virtual body were correctly located by this 4 sensor tracking system.

4.2 Half Lower Body

Four different arrangements were tried. In the case of 7 and 5 sensors, the animation was less natural than for the case of 3 sensors. This result was somewhat surprising, but could be explained by the following 2 reasons :

- The first reason is that the Polhemus sensors are not efficient in the case of a high number of sensors. For instance, the data refresh rate for 7 sensors takes more than twice the time for 3 sensors. Even using interpolation algorithms, the movement lacks some realism due to data insufficiencies.

- The second reason is that in the case of 3 sensors, there are 18 DOF, exactly equal to the DOF of the model we are using (6 for hips, (3+1+2) for legs, 18 in total). There is no redundancy of information because all the data is used once to draw a leg conforming to its DOF. In case of 7 and 5 sensors, the redundancy is not so helpful because of the simplicity of the model. For example, the knee has only 1 DOF for rotation (real knee has more DOF), but the information coming from the thigh and leg sensors always incorrectly shows a small rotation on the two other rotation axis.

In conclusion, 3 sensors are enough. The virtual body's leg posture can be measured with fewer sensors and with good realism and precision.

4.3 Entire Body

The author grouped the results of 4-1 and 4-2. With a minimum of 7 sensors, the animation looks relatively natural. The only drawback is a measurement delay, as explained in section 4-2 : only a few data sets per second were available. There was a slight delay between real and virtual body motion.

5. Work Analysis

As an application of the system developed here, the authors considered automatic work analysis. The structure is shown in Figure 10.

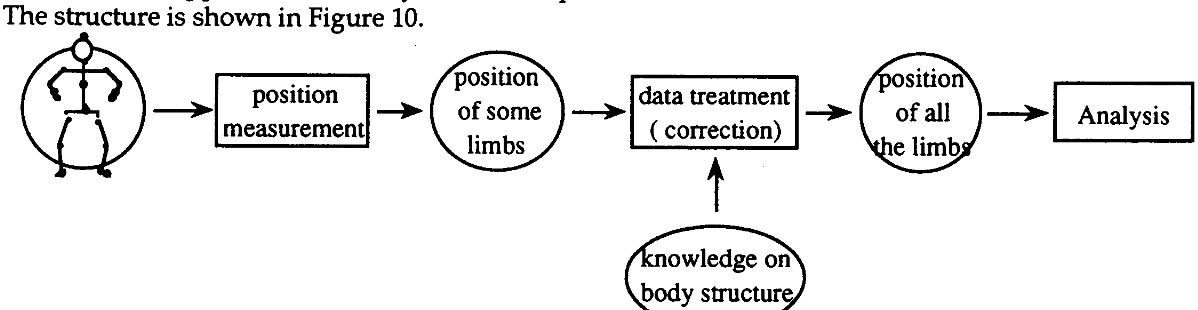


Fig. 10 Graph of the OWAS system

As an example of the work analysis method, OWAS (Ovako Working Posture Analysis System) was selected. OWAS was developed in Sweden more than 10 years ago. At that time, measurements were conducted by using a simple movie camera. This analysis is based on the idea that posture importance factors are considered to be of hand position (up or down), trunk orientation (left/right, bending), and stability on the feet (place of the center of gravity and knees respectively to the feet).

For each factor, respectively 3, 4, and 6 categories are defined. That makes 72 different postures. The last factor is a number from 1 to 4 to quantify the weight of a burden carried by the worker. For each set of 4 factors (3 for posture and 1 for burden), called a OWAS code, a coefficient from 1 to 4 is attributed to express the amount of difficulty to maintain the posture for a certain period of time. In Figure 11, examples of body posture and corresponding OWAS codes are illustrated.

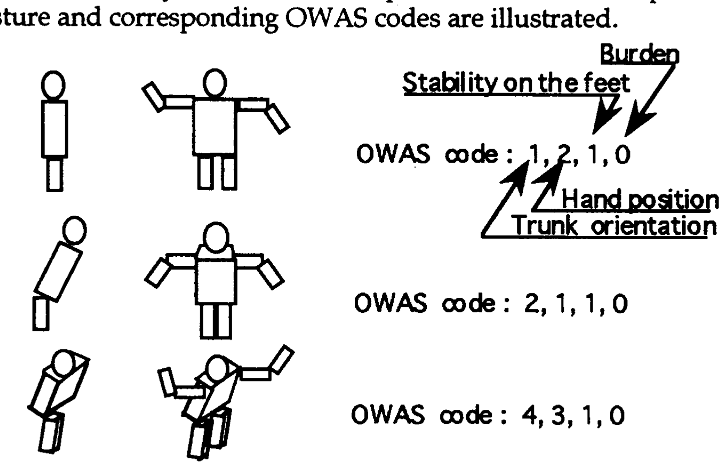


Fig. 12 3D Character prototype

In this study, the authors attempted to generate the correct OWAS code automatically. By doing so, manufacturing activities can also be optimized via computer. The authors concentrated their work on the first 2 digits of the OWAS code. The definitions of the 2 first digits are shown in Table 2.

	First digit	Second digit
1	The worker's trunk is straight	both hands are under the neck
2	The worker is bent down	one hand is over the neck, one is under
3	The worker is turned left or right	both hands are over the neck
4	The worker is turned left or right and is bent down.	

Table 2. The first 2 digits of OWAS code

Two types of experimental work were selected. One type chose each different OWAS code with 6 different postures and calculated the accuracy rate. The second type dealt with manufacturing work which simulates loading/unloading tasks. In both cases, with 3 different subjects, an accuracy rate of approximately 85% was obtained. The failures are due to the error caused by the dislocation of sensors; especially the one put on the chest because it was attached on the body harness fitted on the shoulders. When the arms were moving, there was a slight dislocation of the chest sensor.

6. Summary and Future Work

In this paper, the authors have presented a motion capture system using virtual reality technology. In addition, this system was applied to automatic posture analysis. By using a simple but complete body model, and a minimum number of magnetic sensors, the system could clone a real human body into a virtual body. In its first version, the program takes advantage of the DOF of the articulations for drawing the virtual human body. For automatic posture analysis, with 7 sensors, an accuracy rate of approximately 85% was obtained.

The next step of this study will be the amelioration of several points :

- the inverse kinematics function is not yet strong enough. More constraint of the human body model should be taken into consideration.
- the smoothing function is strong, but further improvement is required.

As for future applications, the authors are planning to add sophistication to the current skeletal model. Figure 12 shows the 3D character prototype. This character will be a virtual actor in the authors' virtual environment, which is being developed as a virtual shared space where 2 people located in remote places can meet each other through the network.

Acknowledgements

Part of this research was performed in collaboration with Mr. Tohru Hirano from the CAE Center of DAIKIN Industries.

References

(Books)
1) Alan and Mark WATT, "Advanced Animation and Rendering Techniques, theory and practice" ACM Press, Addison-Wesley, pp.345-368

(Conference proceedings)
2) Kazuaki Iwata and Toshimichi Moriwaki, "Modelling and Computerized Analysis of 3D-Kinematics of Human Motion in Operation", Proceedings of the 1st International Conference on Human Factors in Manufacturing, pp.69-78, 1984

3) Jiyun Lee and Kwangyeon Wohn, "Fuzzy Aggregation of Motion Factors for Human Motion generation", Virtual Reality Software & technology, Proceedings of the VRST '94 Conference, pp. 55-69, 1994

4) Wes Trager, "A practical Approach to Motion Capture : Acclaim's optical motion capture system", Proceedings of Imagina, pp. 58-70, 1995