*MVA2002* IAPR Workshop on Machine Vision Applications, Dec. 11-13, 2002, Nara-ken New Public Hall, Nara, Japan

# 3—30 Recovering and Analyzing Wide Range Human Motions Based on Mobile Cameras

Tomonori Tabusa[1]
Department of Computer Science
Yuge National Collage of Technology

Joo Kooi Tan[2]
Department of Mechanical and Control
Engineering Kyushu Institute of Technology

Seiji Ishikawa[3]
Department of Mechanical and Control
Engineering Kyushu Institute of Technology

## Abstract

This paper proposes a technique for recovering 3-D wide range human motions employing a set of uncaliblated mobile cameras. This paper also claims availability of such 3-D recovery in analyzing human motions. Three video cameras mounted on a mobile frame track and capture a human motion in a large field from three different directions. Feature points on the subject are registered in a measurement matrix. Since the frame motion contains rotation, the obtained measurement matrix is not a full matrix. The matrix is transformed into a full matrix by taking account of relative relation between the subject and the mobile frame. The 3-D coordinates of feature points on the subject are calculated by factorization with every sample time. Using the proposed technique, several wide range human motions recover in a 3-D way. Analysis of the motions shows significance of the technique.

## 1 Introduction

A motion capture technique is becoming more and more important, since it has large application fields mainly related to modeling human motions such as 3-D human characters producing in video games or in a virtual reality space, and to analyzing human motions such as form analysis in various sports.

A motion capture system with magnetic sensors realizes real time motion recovery. It restricts motion of the subject. however, since one should remain in a limited magnetic field. This technique is therefore not applicable to recovery of a wide range motion. Stereo vision [1] is another established technique for performing such motion recovery. It always necessitates calibration of the employed fixed cameras before taking video streams. This makes the technique difficult in recovering wide range motions, since there is no way of calibrating arbitrarily moving stereo cameras. Tan at el. [2,3] propose an optical 3-D motion recovery technique based on more than three uncalibrated fixed cameras. This technique is employable in any place, only if one can take at least three image streams of the subject from different views. Having escaped from camera calibration, it still suffers from fixed cameras when one wants apply it to the recovery of wide range motions.

In the present paper, we propose a technique for achieving 3-D recovery of wide range human motions employing a mobile set of cameras. The similar recovery system has not yet been proposed to date. The technique will surely be introduced in broader areas having interests in 3-D human motion modeling. Using the proposed technique, 8 human wide range motions recovered their 3-D motions successfully in the performed experiment.

## 2 Recovering Wide Range Human Motions

A subject $O$ in a 3-D space is taken images from $F$ ($\geq 3$) orientations by the same number of cameras fixed on a mobile frame. Feature points $p$ ($p$=1,2,..., $P_t$) on a subject must be commonly visible by the $F$ cameras at time $t$ and are projected onto the points ($x_{fp}(t)$, $y_{fp}(t)$) ($t$=1,2,...,$T$) on the image plane of the $f$ th ($f$ =1,2, ..., $F$) camera. Correspondence of those projected feature points are examined among the $F$ images at each sample time $t$ and it is written into a measurement matrix $W(t)$ of the form

$$W(t) = \begin{bmatrix} x_{11}(t) & x_{11}(t) & \cdots & x_{1P_t}(t) \\ \vdots & \vdots & \ddots & \vdots \\ x_{F1}(t) & x_{F1}(t) & \cdots & x_{1P_t}(t) \\ y_{21}(t) & y_{21}(t) & \cdots & y_{2P_t}(t) \\ \vdots & \vdots & \ddots & \vdots \\ y_{F1}(t) & y_{F1}(t) & \cdots & y_{FP_t}(t) \end{bmatrix}, \quad (1)$$

where the rows represent camera $f$($f$=1,2,...,$F$), and the columns correspond to the feature points.

The matrices $W(t)$ ($t$=1,2,...,$T$) corresponding to respective camera (or mobile frame) orientations are merged into a single matrix of the form

$$W''' = \begin{bmatrix} W(1) & & & \\ & \cdots & & \cdots \\ & & & W(T) \\ & & W(t) & \end{bmatrix}. \quad (2)$$

Here the $f$ 'th row of sub-matrix $W(t)$ corresponds to the $f$ 'th camera's observation on the mobile frame at time $t$. Sub-matrices $W(t)$ ($t$=1,2,...,$T$) contain the coordinates of the projected feature points, whereas the other entries of matrix $W'''$ remain vacant. (Static feature points, if any, are regarded here, for simplicity, as different feature points at different sample times.) Since the frame motion generally contains rotation, the obtained matrix $W'''$ is in this way not a full matrix and therefore it cannot be factorized. The matrix $W'''$ is transformed into a full matrix by considering the subject's relative motion with respect to the mobile frame.

[1] Address: 1000 Shimoyuge, Yuge-cho, Ochi-gun, Ehime 794-2593, Japan. E-mail: tabusa@info.yuge.ac.jp
[2] Address: 1-1 Sensui-cho, Tobata, Kitakyushu, Fukuoka 804-8550, Japan. E-mail: etheltan@is.cntl.kyutech.ac.jp
[3] Address: 1-1 Sensui-cho, Tobata, Kitakyushu, Fukuoka 804-8550, Japan. E-mail: ishikawa@is.cntl.kyutech.ac.jp

The idea of the proposed technique is to shift sub-matrices $W(t)$ to an arbitrary camera location at time $t$ in matrix $W'''$. This yields an extended measurement matrix $W$ of the form

$$W = \left( W(1) \mid W(2) \mid ...... \mid W(T) \right) \tag{3}$$

Obviously this matrix is a full matrix and, after having been transformed into matrix $\widetilde{W}$ whose entries represent deviation from the average point of the projected feature points, it is applied singular value decomposition and other linear computation algorithms [4]. The matrix $\widetilde{W}$ decomposes into camera orientation matrix $M$ and shape matrix $S$ as follows;

$$\widetilde{W} = M \cdot S . \tag{4}$$

## 3 Experimental Results

In order to show performance of the proposed technique, an experiment is performed in which 8 wide range motions are captured images by a mobile frame (See **Figure 1**). Average speed of subjects and recovery errors are shown in **Table 1**. These subjects were sampled on video images with the interval of 1/15 second, and had 15 feature points. Employed three video images with respect to Exp.8 are shown in **Figure 2**. The motion successfully recovered in a 3-D way. With respect to the recovered motion corresponding to image frame 1 to 27, every six or seven successive frames were superposed and displayed as illustrated in **Figure 3**, from which a periodical movement of the subject can be observed. This is clearer from the top view of the subject's motion as given in **Figure 4**. In the figure, only the subject's head and shoulders lines are drawn for simplicity. One can easily see that the lines stay at some locations repeatedly while moving forward. Thus we can discover periodical motion of the subject from 3-D recovered images of the motion and, if necessary, the velocity, the period, etc., can be further computed from the 3-D data.

## 4 Conclusions

A technique was proposed for recovering 3-D wide range human motions employing mobile cameras. Advantages of the technique were shown in analyzing human motions as well as the fact that camera calibration is not necessary in the technique. Instead, a certain amount of recovery errors cannot be avoided because of the approximation of the imaging by orthographic projection and the disturbance on the captured images caused by the movement of the image-taking frame. Broader application areas are expected by the present technique.

## References


[1] Jain, A. K., Fundamentals of Digital Image Processing, Prentice Hall(1989).

[2] Tan J. K., Kawabata, S. Ishikawa, S. "An efficient technique for motion recovery based on multiple views", Proc. IAPR Workshop on Machine Vision Applications. 270-273 (1998).

[3] 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).

[4] Tomasi, C., Kanade, T. "Shape and motion from image streams under orthography: A factorization method", Int. J. of Computer Vision, 9, 2, 132-154 (1992).


**Tabel 1.** Average speed and recovery errors

| Exp No. | Average speed(m/s) | Errors(%) |
|---|---|---|
| 1 | 1.04 | 3.09 |
| 2 | 0.97 | 2.65 |
| 3 | 0.85 | 2.18 |
| 4 | 0.76 | 2.48 |
| 5 | 0.54 | 2.19 |
| 6 | 0.46 | 2.96 |
| 7 | 0.37 | 2.26 |
| 8 | 0.31 | 3.12 |


**Figure 1.** Mobile frame with three cameras

**Figure 2.** Conveying a heavy load with a rope (Exp.8)

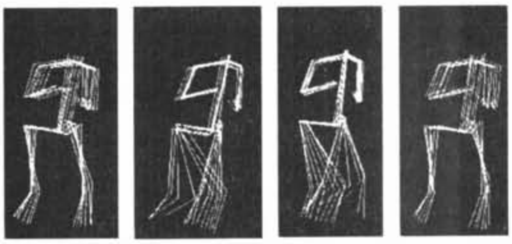

**Figure 3.** Superposition of recovered images

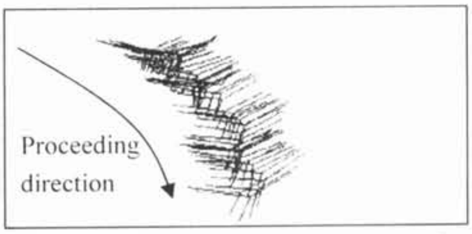


**Figure 4.** Top view of the recovered motion