Enhanced scheme using frame coherence: Modelling with the suggested relaxation algorithm might be used efficiently to confirm reliable correspondences in a frame. Since multi-view correspondences have a frame coherence, we use 3D marker tracking information to enhance our algorithm. First, our procedure uses Kalman filtering [2] to obtain predicted 3D pose data from the previous frame. We then project predicted 3D pose data into every 2D image plane. Next, we search any marker at close range from the projected 2D position. If the detected markers have the same correspondence as the relationships between markers of the previous frame, then we confirm this relation as a reliable correspondence in the current frame. Otherwise, we align the epipolar line, apply it in our relaxation algorithm, and extract their correspondences with undetected markers. Finally, we place together all correspondences, calculate the 3D pose and repeat the above procedure for every frame.

We can speed up our previous algorithm using time coherence, because the relationships of cliques are not suddenly broken during the short interval of frames. Since we use previously confirmed clique information, we just use the relaxation algorithm for only a few 2-D data.
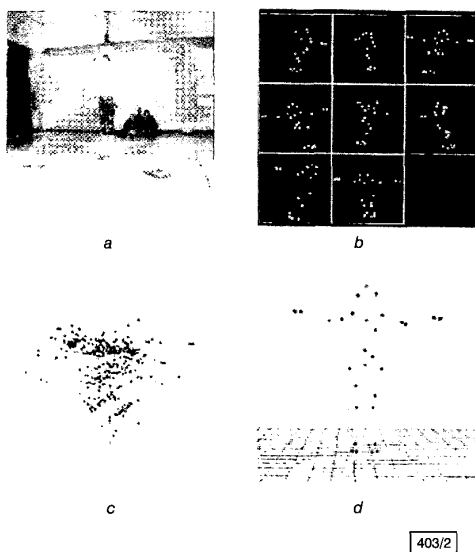

a Motion captured environment
b Captured 2D image data in any frame
c 3D data extraction result using epipolar constraints alone
d 3D data extraction result using proposed algorithm

Fig. 2 *Process of extracting 3D point in motion capture system*

Experimental results: We implemented the algorithm and a test environment on a Pentium II PC which had a 350MHz CPU and 64Mbyte memory. The simulated experimental setting consisted of 33 markers within the restricted space, and the initial positions of four virtual cameras. We used Tsai's calibration method [3] to calibrate the camera system. We evaluated our algorithm using synthetic and real data. First, 100 random data sets were used for synthetic data with the Monte Carlo method, which leads to errors in the position of the 3D data due to camera noise and occlusion phenomena. The experimental result shows that 99.2% of 2D markers have reliable correspondences. It takes 0.22s/frame, and 0.11 s/frame when using 3D tracking information. We performed the same experiments with real captured 2D data set (2633 frames) acquired from an EVa Hires Motion capture system, which is a high-performance commercial system. The results show 0.33 ghost/frame (compared with 0.583 for the EVa system), 0.0175 missing marker/frame (compared with 0.1525 for the EVa system). Here, ghosts (false alarm) and missing markers were the main causes of the need to carry out time-consuming post-processing in the optical motion capture system. Fig. 2 depicts the process of 3D data extraction in the second experiment.

Conclusions: An efficient ghost removal method for acquiring reliable correspondence information for 2D data matching using relaxation labelling has been presented. We have also presented an enhanced mechanism exploiting 3D marker tracking information in order to achieve substantial savings in computation time.

Experimental results illustrate that the proposed method is likely to be very useful for acquiring accurate and stable multiview correspondence information and obviating the need for very time-consuming post-processing in optical motion capture system.

Dong Hoon Lee and Soon Ki Jung (*Department of Computer Engineeinrg, Kyungpook National University, 1370 Sankyuk-dong, Puk-gu, Taegu, 702–701, Korea*)

E-mail: skjung@knu.ac.kr

Kwangyun Wohn (*Department of Computer Dcience, Korea Advanced Institute of Science and Technology, 373-1 Kusong-dong, Yusong-gu, Taejon, 305-701, Korea*)

References

1 HONG, SHEN, LI, K., PAN, Y., YOUNG, G.H., and ZHENG, S.Q.: 'Performance analysis for dynamic tree embedding in *k*-partite networks by a random walk', *J. Parallel Distributed Comput.*, 1998, **50**, pp. 144–156

2 WELCH, G., and BISHOP, G.: 'An introduction to the Kalman filter'. TR 95-041, University of Carolina at Chapel Hill, 1995

3 TSAI, R.Y.: 'An efficient and accurate camera calibration technique for 3-D machine vision'. Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition, June 1986, pp. 364–374

4 ZHANG, Z., and FAUGERAS, O.: '3D dynamic scene analysis' (SpringerVerlag, Berlin, Heidelberg, NY, 1992)

Shadow generation from stereo images

Jen-Hui Chuang, Lou-Wei Kuo, Hao-Jui Kuo and Jain-Shing Liu

A novel approach to generating object shadows using two dimensional (2D) information is presented. Existing approaches to shadow generation are generally based on 3D geometry. The proposed method is based on deriving accurate shadows of a 3D object using only stereo images of the object taken from two different viewpoints. The approach can be extended easily to solve more general problems which involve multiple objects as well as multiple planar surfaces on which object shadows are to be cast.

Introduction: For computer generated 3D scenes, shadows are crucial to the realism of synthetic 2D images. Existing approaches to shadow generation are generally based on 3D geometry. Relevant 3D data required in the shadow computation include the object model, the location of light sources and the surfaces on which object shadows are to be cast. A survey of shadow algorithms can be found in [1]. For the scan-line shadow algorithm presented in [2], polygonal boundaries are projected onto the scanned object and then projected onto the viewing screen during the display phase. Preprocessing is performed to determine pairs of objects that can interact to produce shadows. In [3], volumes which enclose shadowed regions of space are generated. Their boundaries are then used to determine if a visible surface is in shadow. In [4], the visibility of object surfaces viewed from the light source are determined. Those invisible object surfaces correspond to the shadowed area. Apart from the above, other approaches such as the z-buffer [5] and ray tracing [6] techniques have also been developed to generate shadows.

Shadow region derivation: In the following, we present a method of using cross-ratios to derive accurate shadows of a 3D object which only requires stereo images of the object. In the approach it is assumed that at least five reference points on the base plane, upon which the object shadow is to be cast, can be identified in the

images. For example, Fig. 1a and b show stereo images of an object, and a base pentagon formed by the five reference points, obtained from viewpoints V_1 and V_2, respectively. The base plane region blocked by the object in Fig. 1a, the border of which is shown with solid line segments in Fig. 1c, corresponds to the shadow cast on the base plane due to a point light source located at V_1. Fig. 1d shows the object and its shadow obtained for Fig. 1b with the shadow generation process discussed next.
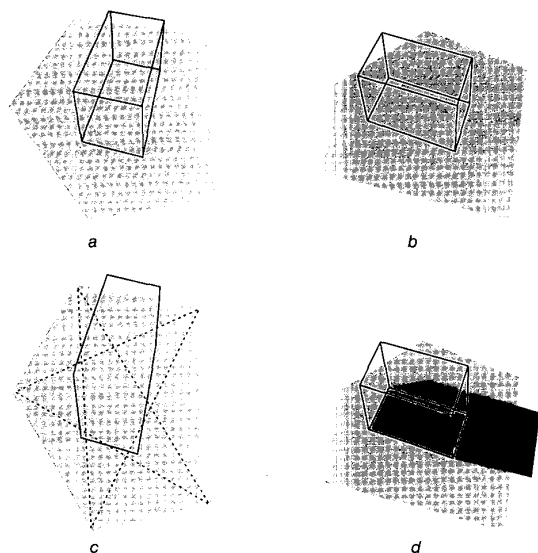

Fig. 1 *Object images from different viewpoints and generated shadow*

a Object image obtained from viewpoint V_1
b Object image obtained from viewpoint V_2
c Base plane region blocked by object as seen from V_1
d Shadow generated from light source located at V_1

In Fig. 1c, if the five reference points can be located precisely, we can also identify four reference points along each diagonal of the pentagon: two end points of the diagonal, and the intersections of that diagonal with two other diagonals. To locate the extended line of a border segment of the blocked base region, at least two of its intersections with the extended lines of the diagonals of the pentagon need to be found.
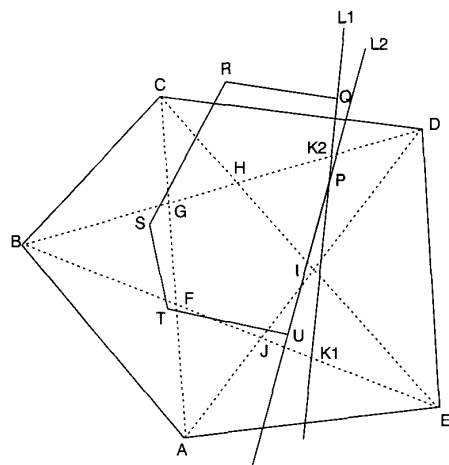
Fig. 2 *Object shadow generated from light source located at V_1*

Consider for example point K_1, which is the intersection of $\overline{BE}$ and L_1 in Fig. 2 and the corresponding $k_1 = (k_{1x}, k_{1y})$, which is the intersection of $\overline{be}$ and l_1 in Fig. 3. According to the view-invariant property of the cross-ratio

$$\frac{BK_1 \cdot FE}{FK_1 \cdot BE} = R \qquad (1)$$

(here, point J can be used in place of point F if a nearer zero denominator in eqn. 1 will not be generated) we have

$$k_{1x} = \frac{Rf_x(b_x - e_x) - b_x(f_x - e_x)}{R(b_x - e_x) - f_x + e_x} \qquad (2)$$

Similarly, k_{1y} can be determined. In exactly the same way, the position of $k_2 = (k_{2x}, k_{2y})$ which is the intersection of $\overline{bd}$ and l_1 can be obtained. With k_1 and k_2, l_1 can be determined. Thus, the shadow region shown in Fig. 3 can be obtained from the extended lines of its border segments. Figs. 4 and 5 show additional shadow generation results for simulated and real data, respectively.
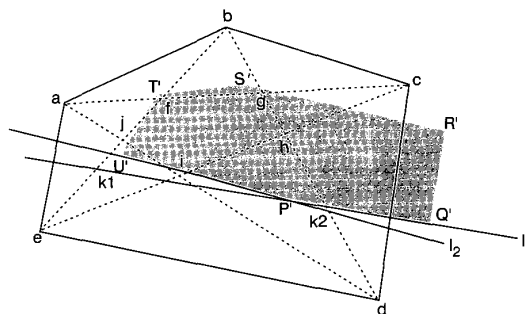
Fig. 3 *Perspective view obtained from V_2 for shadow region shown in Fig. 2*
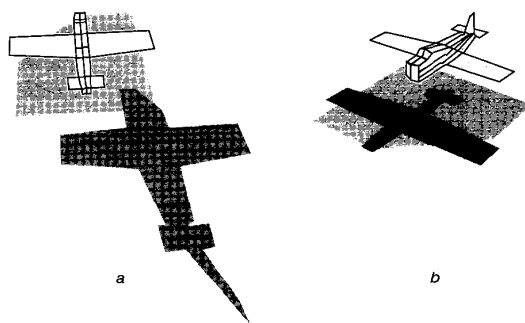
Fig. 4 *Shadows of airplane generated from light sources located at two different viewpoints*
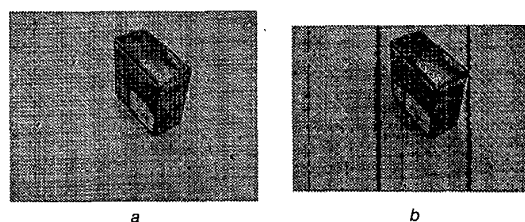
Fig. 5 *Real image and object and additional shadow generated using second image*

a Real image of box
b Object and additional shadow generated using second image (not shown)

Conclusions: A novel approach to generating shadows is presented in this Letter. Unlike existing shadow algorithms which require 3D object models, the proposed approach works with stereo images of polyhedral objects placed on a base plane from which five reference points can be identified. The approach can be extended easily to take into account more general situations where shadows of multiple objects are cast onto polyhedral surfaces.