

# 개선된 직교분해기법을 사용한 빠른 구조 복원 및 융합

박종승<sup>+</sup>, 윤중현<sup>\*\*</sup>

## 요 약

본 논문에서는 비디오에서의 특징점 추적을 통해 얻은 2차원 좌표를 이용하여 3차원 구조를 계산하고 부분적으로 복원된 형상들을 점진적으로 융합하여 전체 형상을 생성하는 기법을 제안한다. 영상의 각 프레임에서 공통적으로 추적된 특징점들을 이용하여 형상을 추정한다. 3차원 좌표 추정 방법으로 개선된 직교분해기법을 사용하였다. 개선된 직교분해기법에서는 3차원 좌표를 복원함과 동시에 카메라의 위치와 방향을 계산할 수 있다. 복원된 부분 형상의 융합을 통해 입체적인 전체 형상을 만든다. 복원된 부분 데이터들의 서로 다른 좌표계를 기준 좌표계로 변환하여 하나의 전체 형상으로 융합한다. 형상 추정 과정과 융합 과정이 통합적으로 수행되며 반복적 최적화 작업을 수행하지 않고 선형적으로 이루어진다. 이는 기존 융합 방법인 ICP(Iterative Closest Point) 방법보다 융합 속도를 향상시켜 빠른 형상 복원이 가능하다. 융합 시간은 평균 0.01초 이내의 수행 속도를 보이며 융합의 오차는 평균 1.0mm 이하의 오차를 보였다.

## Fast Structure Recovery and Integration using Improved Scaled Orthographic Factorization

Jong-Seung Park<sup>+</sup>, Jong-Hyun Yoon<sup>\*\*</sup>

## ABSTRACT

This paper proposes a 3D structure recovery and registration method that uses four or more common points. For each frame of a given video, a partial structure is recovered using tracked points. The 3D coordinates, camera positions and camera directions are computed at once by our improved scaled orthographic factorization method. The partially recovered point sets are parts of a whole model. A registration of point sets makes the complete shape. The recovered subsets are integrated by transforming each coordinate system of the local point subset into a common basis coordinate system. The process of shape recovery and integration is performed uniformly and linearly without any nonlinear iterative process and without loss of accuracy. The execution time for the integration is significantly reduced relative to the conventional ICP method. Due to the fast recovery and registration framework, our shape recovery scheme is applicable to various interactive video applications. The processing time per frame is under 0.01 seconds in most cases and the integration error is under 0.1mm on average.

**Key words:** Improved Acaled Orthographic Factorization(개선된 직교분해기법), 3D Structure Registration(3차원 형상 융합, 정합), Camera Pose Estimation(카메라 위치 추정)

## 1. 서 론

가상세계의 효과적인 구현을 위해서는 실세계 물

체의 정확한 기하 데이터를 3차원 가상세계에서 표현할 수 있는 물체의 3차원 모델링 절차가 필요하다. 실물의 3차원 모델을 획득하는 방법으로는 일반적인

※ 교신저자(Corresponding Author) : 박종승, 주소 : 인천시 남구 도화동 177번지(402-749), 전화 : 032)770-8428, FAX : 032)766-6894, E-mail : jong@incheon.ac.kr

접수일 : 2006년 6월 2일, 완료일 : 2007년 1월 15일

<sup>+</sup> 종신회원, 인천대학교 컴퓨터공학과 조교수

<sup>\*\*</sup> 준회원, 인천대학교 컴퓨터공학과 석사과정  
(E-mail : jhyoon@incheon.ac.kr)

※ 이 논문은 인천대학교 2006년도 자체연구비 지원에 의하여 연구되었음.

범용 카메라를 사용하여 모델링하는 방법과 3차원 깊이 카메라를 사용하는 방법이 있다. 레이저 스캐너와 같은 깊이 카메라를 통해서 정확도가 높은 형상을 얻을 수 있다[1,2]. 이 방법의 가장 큰 장점은 복원의 정확도가 우수하며 대상의 미세한 부분까지 복원을 할 수 있다는 것이다. 그러나 깊이 카메라는 이동성이 떨어지고 촬영 범위에 제약이 클 뿐만 아니라 고가의 부가장치가 필요하므로 실제 사용에 많은 제약이 있다. 또한 기계적인 장치의 동작시간으로 인해 데이터 획득 시간이 매우 오래 걸린다. 대화형 멀티미디어 응용 시스템이 요구하는 형상 복원을 위해서는 이동이 용이한 범용 카메라를 사용하는 빠른 복원 및 정합방법이 적합하다[3].

범용 카메라를 통해 얻은 영상의 2차원 정보를 사용해서 3차원 형상을 추정하는 방법은 크게 보정(calibrated) 방법과 비보정(uncalibrated) 방법의 두 가지로 분류할 수 있다[4]. 보정 방법은 획득한 영상으로부터 Euclidean 좌표까지 형상 복원이 가능하나 카메라 보정 파라미터 정보가 없으면 복원을 수행할 수 없다[5]. 반면에, 비보정 방법은 카메라의 보정 파라미터 정보 없이도 3차원 모델링이 가능한 방법이다[6,7]. 본 연구에서는 카메라 파라미터가 불변이라는 가정 하에 카메라 파라미터를 고정시키고 촬영한 이미지를 사용하여 카메라 파라미터를 구하지 않고도 복원 및 융합이 가능한 고속 형상 추정 방법의 개발을 목표로 한다.

빠른 복원이 가능한 기법으로는 분해(factorization) 기법이 있다[8]. 대표적인 분해기법에는 직교분해(orthographic factorization), 크기변환된 직교분해(scaled orthographic factorization), 근접 원근분해(para-perspective factorization) 기법들이 있다[8,9]. 직교분해기법에서는 카메라의 모션정보 추정에서 깊이값 계산이 불가능하고 복원된 모습이 실제 형상과 다르게 계산되는 단점이 있다. 근접 원근분해기법은 가장 정확한 형상 추정결과를 보이나 노이즈에 민감하여 형상 추정이 불안정하며 계산 과정이 어렵다는 단점이 있다. 이런 단점 때문에 근접 원근분해기법의 적용이 실질적으로 곤란한 경우가 많다[10,11].

복원 과정을 거쳐 추정된 데이터는 복원하려는 대상 객체의 일부분이므로 전체를 포함하는 형상 추정을 위해서는 부분 데이터를 융합하는 과정이 필요하다.

대표적인 형상 융합 방법으로 ICP(Iterative Closest Point) 알고리즘을 들 수 있다. ICP 알고리즘은 일치 부분을 갖는 두 부분 데이터를 융합하는 방법으로, 반복적 최적화 단계를 통해 오차를 줄이는 융합 방법이다. 이는 더 정확한 융합을 할 수 있으나 처리 시간이 매우 길다는 단점이 있다[12]. 이들 방법은 대화형 응용(interactive application) 시스템과 같은 실시간성을 요구하는 경우에는 처리 속도의 문제 때문에 부적합하여 사용되지 않고 있다. ICP 알고리즘의 단점인 처리 속도를 개선하기 위한 많은 연구들이 진행되고 있다. 융합을 위한 데이터를 미리 처리하여 융합시간을 줄이는 등의 방법들이 소개되고 있고 많은 연구를 통해 처리 시간을 단축하였으나 아직 실시간 처리와는 거리가 멀다[13].

이와 같은 일반적인 방법들에 대한 단점들을 보완하는 방법으로 본 논문에서는 특수한 깊이 카메라가 아닌 범용 CCD 카메라를 통해 얻은 2차원 영상을 통해 복원 및 융합의 오차와 속도가 향상된 3차원 형상 추정 방법을 제시한다. 빠르고 안정된 형상의 복원을 위해서 계산 시간이 빠른 직교분해기법의 개선된 방법을 제시한다. 이 방법을 통해 물체의 형상의 복원과 카메라의 모션의 추정이 동시에 이루어진다. 개선된 직교분해기법은 비디오 프레임의 2차원 정보를 이용하여 3차원 형상을 추정하는 방법이다. 특징점 좌표를 이용한 분해기법을 통해 특징점에 대응되는 3차원 모델의 형상과 카메라의 이동경로를 추정한다[14].

획득된 부분 데이터들은 서로 다른 좌표계를 갖기 때문에 모든 부분 형상 데이터를 기준이 되는 하나의 좌표계로 변환한다. 좌표계의 변환 관계는 두 부분 데이터의 공유 가시 영역에서 공통점 지정을 통해 계산할 수 있다. 융합과 복원의 통합화를 위해서 자동으로 검출되는 특징점들 중에서 공유되는 특징점들을 선별하는 절차를 포함시켜서 복원과 동시에 복원된 부분 데이터가 즉시 전역 형상 데이터에 등록되도록 한다. 모든 과정은 선형적으로 처리되기 때문에 처리 속도의 향상 결과를 얻을 수 있고 점진적인 복원 및 융합을 통해 카메라로부터 실시간으로 들어오는 영상 프레임을 처리하는 실시간 대화형 응용시스템으로의 적용이 가능하다.

본 논문의 구성은 다음과 같다. 2장에서 개선된 직교분해기법을 이용한 복원과 복원데이터의 융합

방법에 대해 기술하며, 3장에서 제안한 방법에 대한 실험 결과를 제시하고, 마지막으로 4장에서 결론을 맺는다.

## 2. 3차원 형상 추정 및 융합 방법

본 논문에서는 대화형 증강 현실 시스템에서 응용 가능한 3차원 형상의 빠르고 안정적인 복원 및 융합 방법을 제안한다. 개선된 직교분해 기법은 직교분해 기법에 기반한 복원 방법에 크기변환 인자를 고려하여 기준이 되는 참조위치에 대한 깊이정보도 정확히 추정하도록 개선한 방법이다. 모든 미지의 형상 인자와 카메라 인자가 선형시스템으로 표현되도록 하여 초기치가 필요하지 않고 물체 추적이 실패하더라도 즉시 재시작될 수 있도록 한다. 각 프레임에 대해서 장면의 공간 정보 및 카메라 자세가 복원되며 그 처리시간은 직교분해 기법에 기반하므로 매우 빠르게 수행된다.

복원된 3차원 형상 융합의 목적은 부분적으로 복원된 형상 데이터로부터 3차원 모델을 형성하는 것이다. 복원된 형상들은 융합과정을 거쳐 완전한 한 객체를 이룰 수 있다. 대표적인 융합 방법으로써

Iterative Closest Point (ICP) 알고리즘을 들 수 있다 [16]. ICP 알고리즘은 두 개의 점집합을 융합하는 일반적인 방법으로 많은 응용에 널리 적용되고 있다. 다른 좌표계로 표현된 두 점 집합 간의 최적의 변환을 계산하여 하나의 좌표계로 표현되도록 한다. 이는 최적화의 반복적 수행을 통해서 융합의 오차를 최소화 한다. 본 논문에서는 실시간 응용 시스템에서의 활용이 적합한 실시간적 복원과 융합 방법에 대하여 연구하였다. 이런 관점에서 반복적 수행을 하는 ICP 알고리즘은 실시간 융합을 위해 적합하지 않은 방법이라 할 수 있다. ICP 알고리즘은 단지 두 점 집합을 반복적인 오차 최소화 전략으로 융합하는 방법이며 실시간 복원에서는 각각의 부분적으로 복원된 형상에 대한 점진적인 융합이 필요하다. 공통 영역을 가지는 두 복원된 형상들은 점진적으로 융합과정을 수행한다. 카메라로부터 들어오는 연속적인 이미지 프레임을 통해서 점진적으로 복원하고 융합한다.

그림 1은 객체의 형상을 복원하고 카메라의 모션을 계산하여 부분 형상을 복원하고 복원된 부분 형상을 전체 형상에 융합하는 과정으로 구성된 통합된 복원 단계를 설명하고 있다.

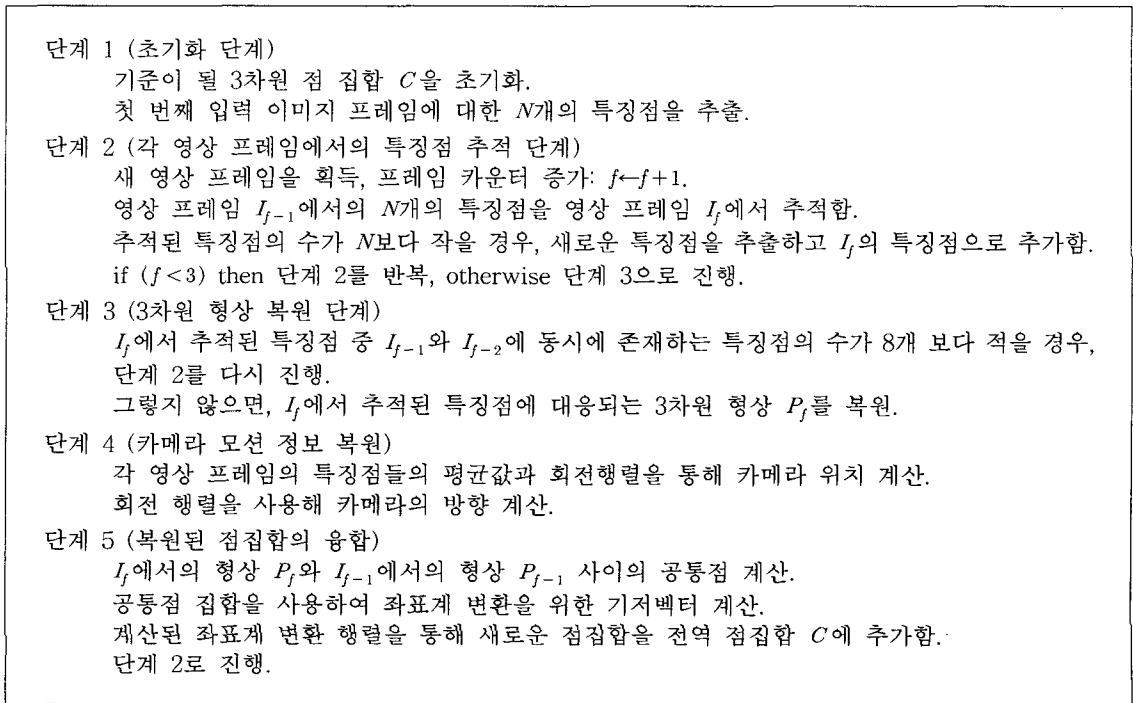


그림 1. 형상 복원, 카메라 모션 추정 및 융합 과정

2.1 형상 및 카메라 모션 복원을 위한 개선된 분해기법

범용 디지털 카메라를 통해 획득되는 세 개의 영상 프레임에 대하여  $N$ 개의 특징점이 주어져 있다고 가정한다. 영상 프레임 인덱스  $f$ 와 특징점 인덱스  $p$ 를 사용하여 특징점들을 식 (1)과 같이 표기한다.

$$\{(x_{fp}, y_{fp}) | f = 1, 2, 3, p = 1, \dots, N\} \quad (1)$$

이들 특징점들로부터 두 행렬  $X$ 와  $Y$ 를 구성한다.  $X$ 는 세 프레임에서의 특징점들의 x-좌표의 열벡터로 구성된  $3 \times N$  행렬이고,  $Y$ 는 세 프레임에서의 특징점들의 y-좌표의 열벡터로 구성된  $3 \times N$  행렬이다. 개선된 분해기법의 입력으로 사용될  $6 \times N$  행렬  $W$ 을 식 (2)와 같이 구성한다.

$$W = [X^T Y^T]^T \quad (2)$$

분해기법에 사용되기 위한 형식으로  $X'$ 와  $Y'$ 를  $X' = [x_{fp}']$ 와  $Y' = [y_{fp}']$ 으로 구성하여  $W' = [X'^T Y'^T]^T$  형태로 변형시킨다. 여기서,  $x_{fp}'$ 와  $y_{fp}'$ 는 식 (3)과 같이 계산된다.

$$x_{fp}' = x_{fp} - \frac{1}{N} \sum_{p=1}^N x_{fp}, \quad y_{fp}' = y_{fp} - \frac{1}{N} \sum_{p=1}^N y_{fp} \quad (3)$$

변형된  $W'$ 은 Singular Value Decomposition (SVD) 기법을 사용하여  $W' = UDV$ 로 분해한다[15].  $U$ 의 처음 세 개의 열로부터  $U'$ 을 구성하고,  $D$ 의 처음 세 개의 열중 처음 세 개의 행으로부터  $3 \times 3$  행렬인  $D'$ 을 구성하고,  $V$ 의 처음 세 개의 행으로부터  $V'$ 을 구성한다. 이렇게 구성된  $U'$ ,  $D'$ ,  $V'$ 로  $W'$ 를 재구성한다.

$W'$ 를 재구성된  $U'$ ,  $D'$ ,  $V'$ 를 사용하여 식 (4)와 같이 두 부분으로 나누어 정의하면  $W' = R'S'$ 과 같이 만들 수 있다.

$$R' = U' [D']^{1/2}, \quad S' = [D']^{1/2} V' \quad (4)$$

따라서  $R'$ 과  $S'$ 이 각각 카메라 회전 행렬과 3차원 형상 행렬이 될 수 있다. 실제로  $R'$ 과  $S'$ 은 정확한 회전과 형상을 표현하는 것이 아니라 미지의 선형변환에 좌우되는 표현이다.  $R'$ 은  $6 \times 3$ 의 크기를 갖는 실제 회전행렬  $R$ 의 선형변환이고,  $S'$ 은  $3 \times N$ 의 크기를 갖는 실제 형상행렬  $S$ 의 선형변환이다. 회전행렬의 각 두개씩의 열은 각 프레임에서의 x-축과 y-축

에 대한 기저벡터이다. 형상행렬의 각 행은 각 특징점의 형상좌표에 해당한다.  $R'$ 과  $S'$ 로부터 식 (5)를 통해 실제의 회전행렬  $R$ 과 실제의 형상행렬  $S$ 를 구할 수 있다[14].

$$W = R'S' = (R'Q)(Q^{-1}S') = RS \quad (5)$$

이제 미지의 선형변환  $Q$ 를 결정해야 유일한 회전행렬  $R$ 과 실제의 형상행렬  $S$ 가 정해진다. 선형변환  $Q$ 을 결정하는 방법은 각 영상 프레임에서의 두 기저벡터의 크기가 동일하다는 가정과 두 기저벡터의 내적이 0이라는 가정을 사용한다. 또한 두 기저벡터의 크기가 각각 1이라는 가정도 추가한다. 이 세 가정을 가장 잘 만족하는 선형변환  $Q$ 를 결정한다.

각 프레임에 대해서 이전 프레임에 상대적인 카메라의 이동 위치와 회전 방향을 계산한다. 카메라의 이동 위치는 각 영상 프레임의 특징점들의 평균값으로 표현될 수 있다. 카메라 위치의 x-좌표와 y-좌표는 추적에 성공한 모든 특징점의 x-좌표와 y-좌표의 평균을 취하여 계산되며 식 (6)과 같이 나타낼 수 있다.

$$c_x = \frac{1}{N} \sum_{p=1}^N x_{fp}, \quad c_y = \frac{1}{N} \sum_{p=1}^N y_{fp} \quad (6)$$

카메라의 방향은 분해기법으로 얻어진 회전행렬  $R$ 로부터 구해진다.  $R$ 은 식 (7)과 같이 표현할 수 있다. 여기에서 모양의 복원 시 계산된  $R$ 을 통해 카메라의 방향을 계산할 수 있다.  $R$ 은 식 (7)과 같은 형태이다.

$$R = [i_1 \ i_2 \ i_3 \ j_1 \ j_2 \ j_3]^T \quad (7)$$

여기서  $i_f$ 와  $j_f$ 는  $f$ 번째 프레임의 x-축과 y-축에 대한 기저벡터이다. z-축에 대한 기저벡터  $k_f$ 는 두 기저벡터  $i_f$ 와  $j_f$ 는 외적하여 얻을 수 있다.  $k_f$ 는 카메라가 향하는 방향과 일치한다.

크기변환된 직교분해기법에서는 카메라의 위치의 깊이 정보를 계산할 수 있다. 그러나 계산된 z-좌표값은 단위가 정규화(normalize)되어 있으므로 카메라 참조위치의 x-좌표와 y-좌표도 단위를 적절히 조절해야 한다. 크기변환된 직교분해기법을 개선하면 직교분해기법에 기반한 복원방법에 크기변환 인자를 고려한 카메라의 참조위치에 대한 깊이정보의

추출이 가능하다. 카메라의 참조위치를  $C_f$ 라고 하면,  $m_f = 1/|i_f|$ 이고 사용한 영상의 해상도가  $v \times h$ 일 때  $C_f$ 는 식 (8)과 같이 계산할 수 있다.

$$C_f = ((c_x/m_f - v/2)/v, (c_y/m_f - h/2)/h, m_f) \quad (8)$$

또한 각 프레임에서 특징점의 평균으로 계산된 카메라의 이동 경로는 서로 방향(orientation)이 다르다.  $R_f = [i_f \ j_f \ i_f \times j_f]^T$ 를  $R$ 로부터 구성하고 이를 통해 식 (9)와 같은 과정을 거쳐서 방위가 맞춰진 올바른 카메라의 참조위치를 계산할 수 있다.

$$C'_f = R_f C_f \quad (9)$$

방향은 프레임 별 이동 경로에 각각의 회전 행렬을 적용시킴으로써 일치시킬 수 있다. 단위를 조절한 x-좌표 및 y-좌표와 계산된 z-좌표를 함께 각 프레임 별 회전행렬에 적용시킴으로써 올바른 카메라의 모션 정보를 추정할 수 있다.

### 2.2 복원된 3차원 형상 데이터의 점진적 융합 기법

점집합의 융합의 선행 작업으로서 부분적으로 복원된 형상의 3차원 점집합  $P_i$ 와  $P_j$ 가 복원되어 있음을 가정한다. 두 점 집합  $P_i$ 와  $P_j$ 는 서로 공통된 점들을 포함하고 있다고 가정한다. 각 점 집합은 서로 다른 좌표계로 표현됨으로 두 부분 형상의 융합을 위해서 좌표계의 변환 과정이 필요하다. 하나의 점집합  $P_i$ 의 각 점  $X_i$ 를 다른 점집합  $P_j$ 에서 해당하는 점  $X_j$ 로의 변환을 위한 변환행렬을  $T_{ij}$ 라 하면 변환은 식 (10)과 같이 표현할 수 있다.

$$X_i = T_{ij} X_j \quad (10)$$

여기서

$$X_i = \begin{bmatrix} x_i \\ y_i \\ z_i \\ 1 \end{bmatrix}, X_j = \begin{bmatrix} x_j \\ y_j \\ z_j \\ 1 \end{bmatrix}, T_{ij} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

이다. 식 (10)은 다음의 식 (11)과 같이 나타낼 수 있고,

$$\begin{aligned} r_{11}x_j + r_{12}y_j + r_{13}z_j + t_1 - x_i &= 0 \\ r_{21}x_j + r_{22}y_j + r_{23}z_j + t_2 - y_i &= 0 \\ r_{31}x_j + r_{32}y_j + r_{33}z_j + t_3 - z_i &= 0 \end{aligned} \quad (11)$$

이를  $Mq=0$  형태인 동차방정식으로 변형한다. 벡터  $q = [r_{11}, \dots, r_{33}, t_1, t_2, t_3, 1]^T$ 는 구하고자 하는 12개의 미지수를 포함하는  $13 \times 1$  벡터이다.  $3 \times 13$  행렬  $M$ 은  $X_i$ 와  $X_j$ 의 3차원 좌표로 구성되어 있는 이미 알고 있는 정보이다. SVD를 사용하여 벡터  $q$ 를 계산할 수 있다.  $q$ 를 구하기 위하여 공통점의 쌍이 5개 이상 필요하다. 해를 구하기 위한 최소 점의 수인 5쌍으로 표현된 동차방정식은 식 (12)와 같은 선형식으로 표현할 수 있다.

$$Aq=0, A = \begin{bmatrix} M_1 \\ \vdots \\ M_5 \end{bmatrix} \quad (12)$$

식 (12)에서의  $15 \times 13$  행렬  $A$ 는 부분적으로 복원된 형상의 점집합의 좌표들로부터 계산되므로 벡터  $q$ 를 선형식으로 유일하게 계산할 수 있다. 구해진 벡터  $q$ 로부터 변환행렬  $T_{ij}$ 을 구성하고 두 부분 점집합  $P_i$ 와  $P_j$ 를 융합한다.

기준 점집합  $C$ 가 여러 융합 과정을 거쳐  $C = X_{p=1:k}$ 와 같이  $k$ 개의 점들로 이루어져 있고,  $P_i$ 를 마지막으로 융합한 점집합이라고 할 때,  $P_j$ 는  $P'_j = T_{ij}P_i$  과정을 거쳐  $C$ 와 같은 좌표계를 동일한 점집합으로 다시 구성된다. 그리고  $n$ 개의 점들로 구성된  $P'_j$ 은  $C$ 로 융합되어  $C = X_{p=1:(k+n)}$ 의 새로운 점집합으로 확장된다. 그림 2는 부분적으로 복원된 점집합의 점진적인 융합 과정을 보여준다.

### 3. 실험 결과

영상 프레임 정보를 통해 3차원 부분 형상 데이터를 복원하는 방법으로 제안된 개선된 직교분해기법

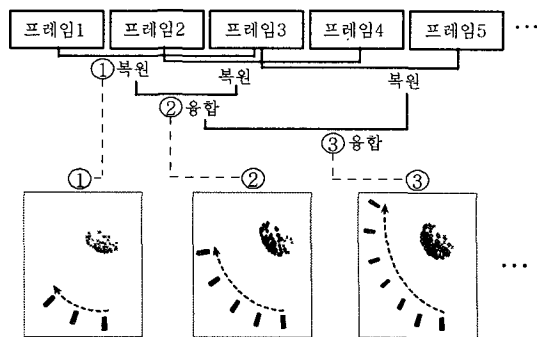


그림 2. 점진적인 형상 복원을 위한 복원 및 융합 과정

의 유용성을 보이기 위해 실제 영상 데이터로부터 복원 실험을 수행하였다. 일반적인 범용 디지털 카메라로 물체를 향하면서 상이한 위치에서 촬영한 640 × 480 해상도의 다수 영상들을 실험에 사용하였다. 실험 환경으로 복원 및 융합을 위하여 Intel Pentium 4 2.8GHz PC를 사용하였고, 영상 촬영을 위한 범용 디지털 카메라로써 Nikon COOLPIX 4500을 사용하였다. 각 영상들은 카메라의 이동에 따른 획득 순서와 일치하도록 입력되며 순차적으로 입력된다. 입력된 각 영상 프레임에 대해서 일정 수의 특징점을 추출하고 이를 지속적으로 추적한다. 세 프레임 이상에서 추적된 특징점들에 대해서는 매 프레임마다 복원을 수행하고 계산된 3차원 좌표는 전체 형상에 융합된다. 각 프레임에 대해서 카메라의 모션 정보도 동시에 계산된다.

### 3.1 패턴 상자 영상을 사용한 형상 복원

표면에 격자무늬 패턴을 붙인 정육면체에 대한 영

상 프레임에 대해 제안된 형상 복원 기법을 적용해 보았다. 상자의 각 면에 규칙적인 격자무늬를 부착하여 복원의 정확성을 측정할 수 있도록 하였다. 각 영상에 대해서 모서리 특성을 잘 표현할 수 있는 특징점들을 추출하고 이를 다음 프레임에서 추적한다. 패턴 상자 영상과 추출된 특징점들이 그림 3에 있다. 이전 프레임에서 추출된 특징점이 다음 프레임에서 추적되지 않을 수 있다. 이 경우에는 추적에 실패한 개수만큼 새로운 특징점을 프레임에 추가해서 새로 나타나는 영역에 대한 형상의 복원이 가능하도록 하였다.

상자는 여러 영상을 통하여 다양한 각도에서 보여진다. 먼저 세 개의 부분 형상을 복원하였다. 세 부분 복원에 사용된 영상은 각각 14개, 10개, 7개이다. 각 복원에서의 특징점의 개수는 각각 43개, 61개, 61개이다. 특징점의 개수는 영상의 선명도와 모서리점의 포함 정도에 따라서 자동으로 결정된다. 복원된 세 개의 부분 형상이 그림 4에 있다.

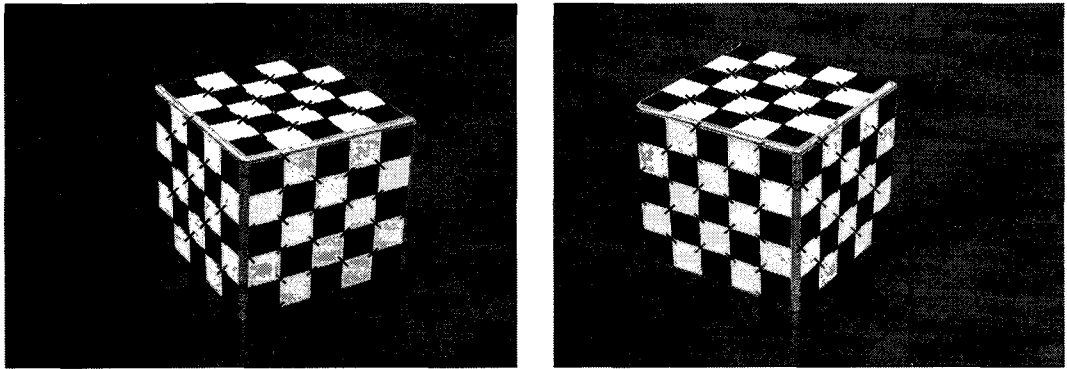


그림 3. 패턴 상자 데이터의 영상 프레임과 추출된 특징점들

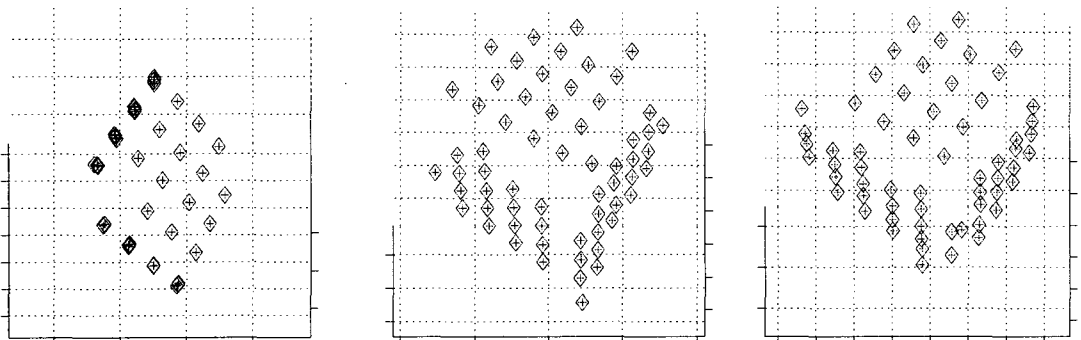


그림 4. 패턴 상자의 복원된 세 부분 형상: 14개 프레임에서 41개 특징점 사용 (왼쪽), 10개 프레임에서 61개 특징점 사용 (중간), 7개 프레임에서 61개 특징점 사용 (오른쪽)

복원된 세 개의 부분 영상 데이터를 융합하여 정면, 위면, 두 측면의 네 면을 모두 포함하는 전체 영상을 생성하였다. 첫 부분 영상 데이터를 기준으로 하여 다음 두 부분 영상 데이터의 좌표계를 변환하는 식으로 융합하였다. 복원된 부분 영상 데이터의 융합 시에 카메라의 이동경로와 방향을 기준이 되는 부분 영상 데이터 좌표계로 일치시켜서 전체적인 카메라 움직임의 궤적을 계산한다. 카메라의 이동경로와 방향을 영상 변환과 동일한 변환행렬을 적용하여 변환하였다. 그림 5에서 패턴상자의 융합된 결과를 볼 수 있다. 그림은 복원된 세 개의 부분 영상 데이터를 사용하여 융합한 결과로 각 부분 데이터를 다른 밝기로 표시하였다. 그림 5의 왼쪽 그림은 복원된 패턴상자를 위에서 본 모습이며, 중간 그림은 측면에서 본 모습이다. 상자 표면의 격자무늬의 각 모서리점이 올바르게 융합되었음을 알 수 있다.

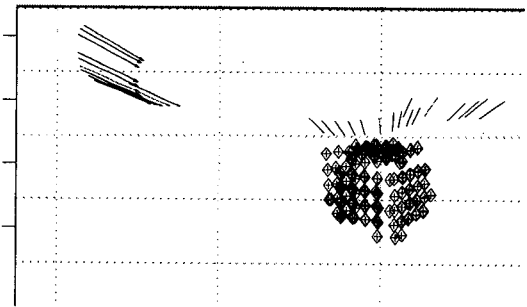
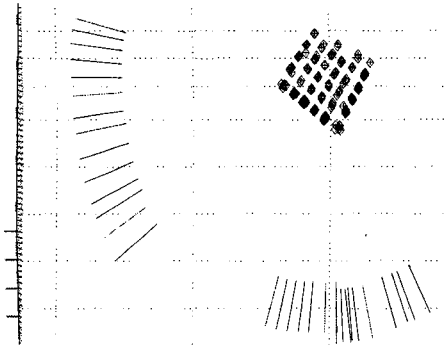


그림 5. 패턴 상자 영상 프레임으로부터 복원된 부분 영상 데이터들의 융합 결과: 위에서 본 모습 (왼쪽), 측면에서 본 모습 (오른쪽)

### 3.2 양(sheep) 인형 영상을 사용한 영상 복원

물체의 모양이 불규칙한 경우에 대해서도 영상의 복원실험을 수행하였다. 복원 대상으로 특징점의 위치를 인위적으로 지정하기가 어려운 인형 모양을 사용하였다. 그림 6은 실험에 사용된 양 인형 영상들과 추적되는 특징점의 모습을 보여주고 있다. 패턴 상자 실험에서와 마찬가지로 세 개의 부분적인 영상 데이터를 복원하였다. 생성된 부분 데이터는 인형의 오른쪽, 앞쪽, 왼쪽에 해당하는 부분 데이터이다. 그림 7에서 각 부분 데이터의 복원된 모습과 카메라의 모션 정보를 확인할 수 있다. 각 부분 복원에 대해서 사용된 프레임의 수는 각각 8개, 7개 11개의 프레임이 사용되었다. 각 부분 복원에서의 특징점은 각각 155개, 143개, 146개가 사용되었다.

각 부분적으로 복원된 데이터는 물체의 촬영에 대

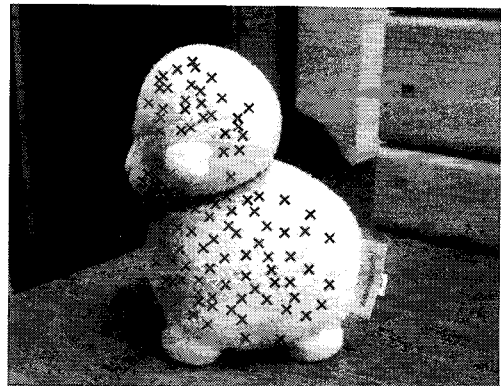
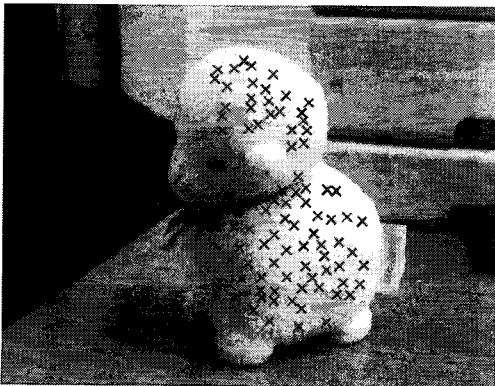


그림 6. 양 인형 영상 프레임과 추출된 특징점들

한 카메라 이동 경로가 부드럽게 변하지 않더라도 융합이 가능하다. 그림 8은 복원된 세 개의 부분 데이터를 융합한 결과이다. 형상 복원과 더불어 구해지는 각 영상 프레임의 카메라 위치와 방향을 선분으로 표시하였다.

카메라의 촬영 경로가 거의 비슷한 간격으로 부드

럽게 움직이는 더 많은 영상 프레임이 주어지는 경우에는 더 정교한 융합 결과를 얻을 수 있다. 그림 9는 연속되는 부분 형상 데이터를 융합한 3차원 구조를 나타내고 있다. 융합된 모습은 인형의 오른쪽부터 정면까지의 모습을 나타내고 있다. 카메라 이동 궤적도 선분으로 표시하였다.

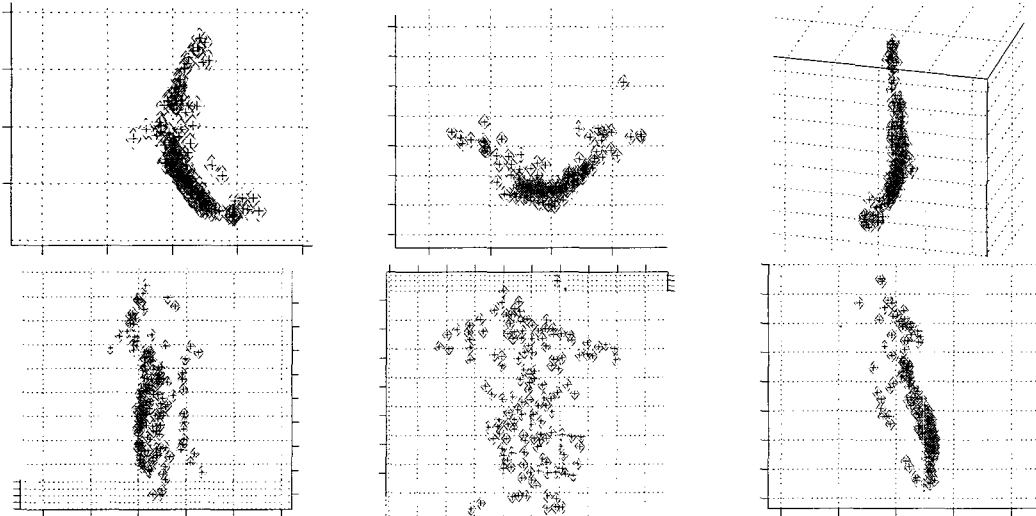


그림 7. 인형 데이터의 세 개의 부분 복원된 형상 (왼쪽, 중간, 오른쪽): 위에서 본 각 부분 형상의 모습 (위), 앞에서 본 각 부분 형상의 모습 (아래)

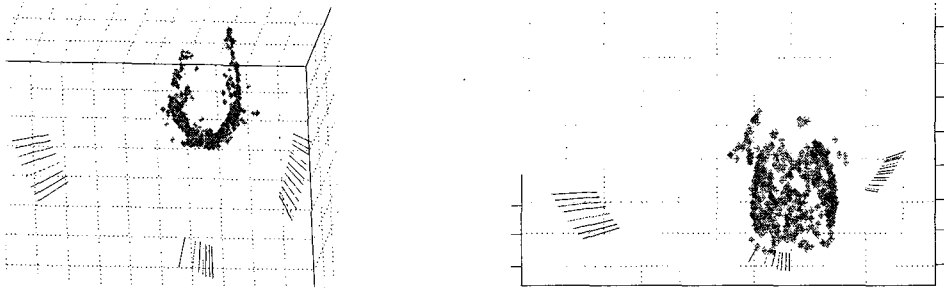


그림 8. 인형 데이터의 불연속적인 세 형상의 융합 결과: 위에서 본 모습 (왼쪽), 정면에서 본 모습 (오른쪽)

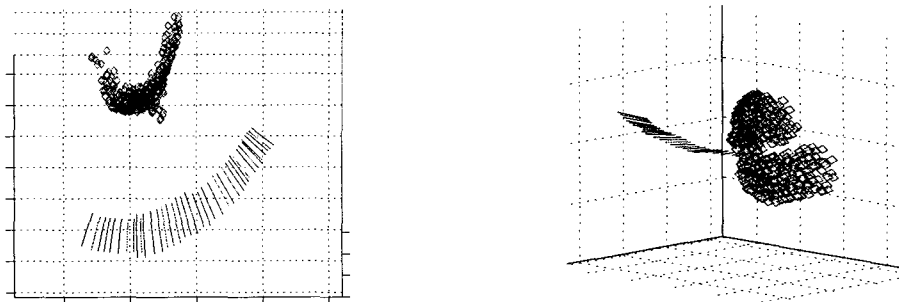


그림 9. 인형 데이터의 연속적인 부분 복원 형상의 융합 결과: 위에서 본 모습 (왼쪽), 측면에서 본 모습 (오른쪽)



### 3.3 축구공 영상을 사용한 형상 복원

곡면을 가진 객체에 대한 형상 복원 실험을 하였다. 그림 10은 실험에 사용된 축구공의 영상의 모습을 보여주고 있다. 이 실험에서는 물체의 완전한 형상 및 카메라 움직임을 복원하기 위해서 카메라를 물체 주위로 한 바퀴 회전하면서 물체를 촬영하였다.

복원에는 총 23 프레임을 사용하였다. 매 세 프레임마다 부분 데이터의 형상 복원을 하였으며, 각 복원 시마다 융합을 실시하였다. 그림 11을 통해서 공의 복원된 모습을 확인할 수 있다. 왼쪽 모습은 복원된 형상의 옆면 모습이고, 오른쪽 그림은 복원된 형상을 위에서 본 모습을 나타낸다. 복원과 더불어 계산한 카메라 움직임도 함께 표시하였다.

### 3.4 형상 복원의 정확도 분석

특징점을 사용한 형상의 부분적인 복원과 각 부분

형상의 융합에 대한 정확도를 분석하였다. 부분적인 형상 복원에 대한 정확성과 전체 형상의 생성을 위한 융합 절차에 대한 정확성을 각각 구분하여 분석하였다. 부분적인 형상 복원에 대한 정확성의 분석을 위해서는 실측(ground truth) 정보가 필요하다. 패턴 상자 영상에 대해서 각 특징점의 실측 거리를 측정하고 복원의 정확도를 계산된 형상의 좌표값과 실측 좌표값의 차이로 계산하였다.

먼저 패턴 상자 데이터에 대해 복원된 형상의 정확도를 분석하였다. 그림 3에 나타난 패턴 상자의 격자무늬의 간격은 가로와 세로 모두 20mm으로 일정하게 제작되었다. 두 개의 부분적으로 복원된 형상에 대한 정확도를 분석한 결과가 그림 12에 있다. 두 부분 복원 형상 데이터에 대하여 실측 형상과 비교한 결과가 그림 12에 있다. 오차가 평균 1mm 정도로 매우 정확한 형상이 복원됨을 알 수 있었다. 복원된 특징점들에 대응되는 패턴 상자의 48개의 격자의 길

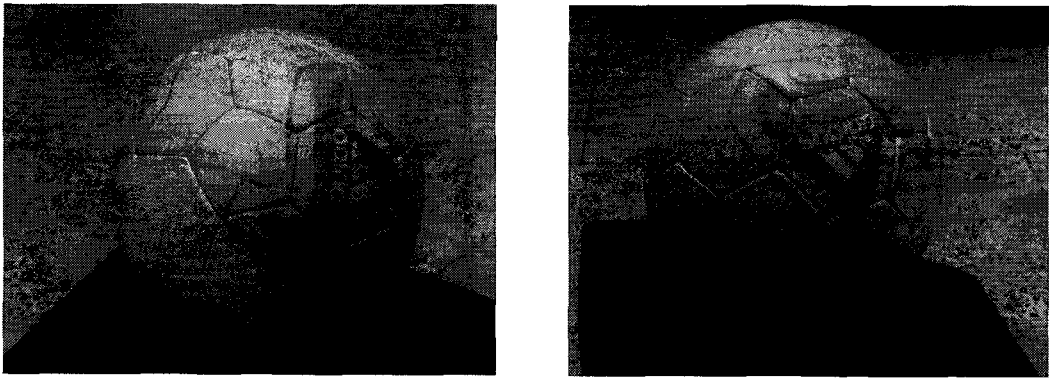


그림 10. 축구공 입력 영상 프레임

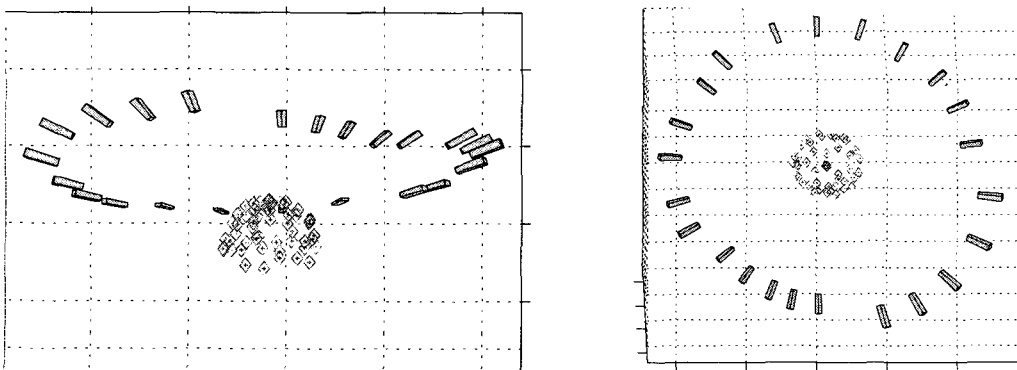


그림 11. 축구공 데이터의 연속적인 부분 복원 형상의 융합 결과: 옆에서 본 모습 (왼쪽), 위에서 본 모습 (오른쪽)

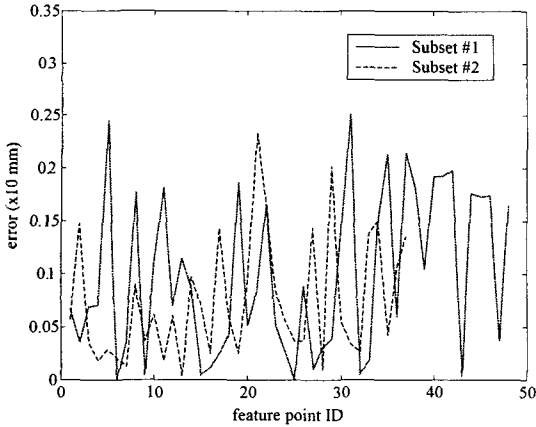


그림 12. 복원된 두 부분 형상 데이터의 각 점의 실측 거리와의 오차

이를 실측 길이와 비교한 결과, 최대 오차는 2.5mm 정도였고 평균 오차는 0.986mm로 나타났다. 최대 오차와 평균오차가 특징점 추적의 정확도에 의해서 약간의 차이가 있으나, 평균 1.0mm 이하의 오차이므로 형상 복원의 정확도가 매우 높은 복원이 이루어졌다고 할 수 있다.

부분적으로 복원된 형상의 융합에 대한 오차는 최종적으로 융합된 형상과 융합 전의 각 부분 형상을 비교하여 분석할 수 있다. 즉 융합 오차는 최종적인 형상과 각 부분 형상의 상대적인 기하구조와의 일관성 정도로 측정할 수 있다. 그림 13과 그림 14는 복원된 부분 데이터의 융합오차에 대한 실험 결과이다. 두 그림에서 x-축은 융합에 사용된 공유점의 번호를 나타낸다. 그림 13은 패턴 상자의 부분 데이터 두 개를 융합할 때 지정한 공유점들의 변환 오차이다. 이는 그림 5의 실험에서의 융합 오차로써, 두 부분 데이터 사이의 변환 행렬을 구할 때에 사용된 20개의 공통점에 대한 변환 오차를 계산한 것이다. 오차의 계산 시에 복원된 형상은 실제 물체의 모양뿐만 아니라 크기도 일치해야 하므로 절대적인 좌표의 차이로 오차를 계산한다. 따라서 두 좌표 값의 차이를 그 점에서의 오차로 사용하는 부호가 있는 오차(signed error)를 고려하였다. x-좌표와 y-좌표의 경우는 오차의 크기가 약 1.5mm 이내로 제한되는 결과를 보였으나 z-좌표의 경우 최대 5.0mm에 가까운 오차가 발생하였다. 이는 가시적인 물체의 형상보다 물체까지의 깊이 계산이 영상에서의 특징점 좌표의 변화에 더

민감하기 때문이다.

그림 14는 양 인형 실험에서 네 개의 부분 복원 형상을 점진적으로 융합하는 경우에서의 융합 오차를 나타낸다. 그림은 각 두 부분 복원 형상의 공통점들인 30개의 형상 점에 대한 오차를 표시하였다. 형상점의 x-좌표와 y-좌표의 경우는 오차의 크기가 약 2.1mm 이내로 제한되는 결과를 보였고 z-좌표의 경우 최대 6.5mm 정도의 오차를 나타내고 있다. 패턴 상자의 복원이 인형의 복원보다 약간 더 정확한 결과를 얻을 수 있었다. 약 0.6mm 정도의 차이를 보이고 있고, 이는 패턴 상자의 경우 특징점을 더 정확하게 추적할 수 있음에 의존하는 결과임을 알 수 있다. 그러나 형상의 복잡도는 상자보다 인형의 경우가 더 복잡하나 두 경우에서의 오차가 그다지 큰 차이를 보이지는

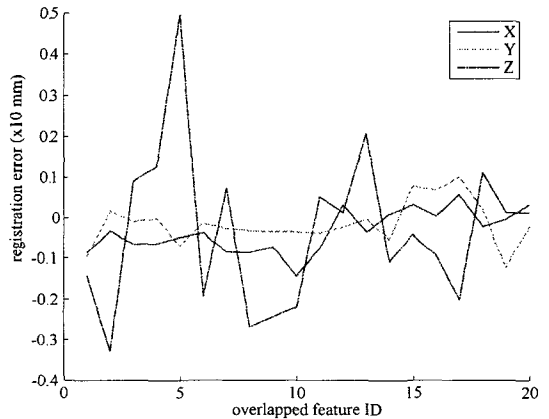


그림 13. 패턴 상자의 두 부분 데이터의 융합오차

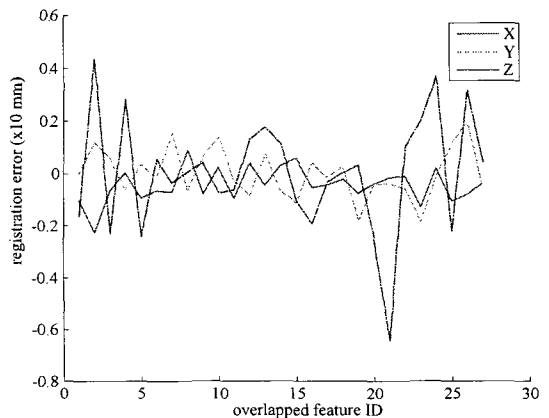


그림 14. 인형의 부분 데이터 사이의 융합 오차

않는다. 이는 형상의 복원이 패턴의 복잡도가 아니라 특징점 추적의 정확도에만 의존되기 때문이다.

그림 15는 축구공 복원 실험에 대한 오차를 나타내고 있다. 축구공의 실측 좌표와 복원된 형상의 좌표를 비교하였다. 실험을 통해서 복원된 결과가 기존 방법보다 우수함을 확인할 수 있다. 그림 16은 복원된 축구공의 융합오차를 나타내고 있다. 논문에서 제안한 융합 방법과 고전적인 ICP 방법을 비교하였다. 제안한 방법과 ICP 방법이 근사한 오차를 나타냄을 알 수 있다. 논문에서 제안한 선형적인 융합 방법은 반복적 최적화 방법인 ICP 보다 융합시간이 빠른 뿐만 아니라 융합의 정확도도 비슷하게 나오는 결과를 확인할 수 있다.

### 3.5 형상 복원의 처리 시간 분석

형상의 복원 및 융합 시간은 실시간 응용 시스템에서의 활용 가능성을 결정하는 중요한 요소이다. 실제 환경에서의 영상 데이터에 대해서 제안된 형상

복원 방법의 처리 시간을 분석하였다. 그림 3의 패턴 상자 실험과 그림 6의 양 인형 실험에 대한 부분 형상 복원 및 융합의 통합 처리 시간이 그림 17에 있다. 표시된 처리 시간은 각 입력 영상에 대해서 부분 형상을 복원하고 융합하는 시간을 모두 합친 시간으로 프레임 당 복원 및 융합 시간으로 간주할 수 있다.

두 경우 모두 각 프레임 당 처리시간이 최대 0.12초의 성능을 보였다. 평균 처리시간은 패턴 상자의 경우 약 0.07초이고 양 인형의 경우 약 0.09초이다. 즉 평균적으로 초당 약 10 프레임의 처리 속도를 유지함을 확인할 수 있다. 이러한 수행속도는 실시간 응용에서 필요한 수행시간에 미치지 못하지만 특징점 처리 알고리즘에 있어서 가속화의 여지가 있고 고성능 하드웨어의 사용으로 인한 처리시간 단축이 가능하므로 충분히 대화형 응용 시스템에 사용될 수 있는 성능이라고 볼 수 있다.

## 4. 결 론

본 논문에서는 직교분해기법을 기반으로 빠르고 안정적으로 세 개의 영상 프레임으로부터 형상을 복원하고 점진적으로 융합하는 방법을 제안하였다. 제안한 방법은 각 프레임마다 국부적으로 복원되는 부분 형상들은 선형적으로 빠르게 전체 형상에 융합할 수 있다. 형상의 복원 및 융합 과정을 요약하면 다음과 같다. 각 영상 프레임에서의 특징점을 자동 추출 및 추적하고 이를 통해 3차원 형상을 복원한다. 영상을 촬영할 때의 카메라의 이동 및 회전 정보도 형상의

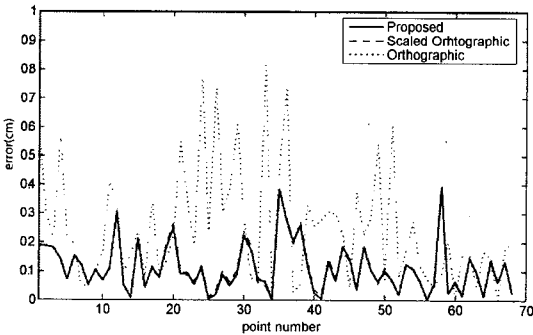


그림 15. 축구공 부분 복원 형상의 복원 오차

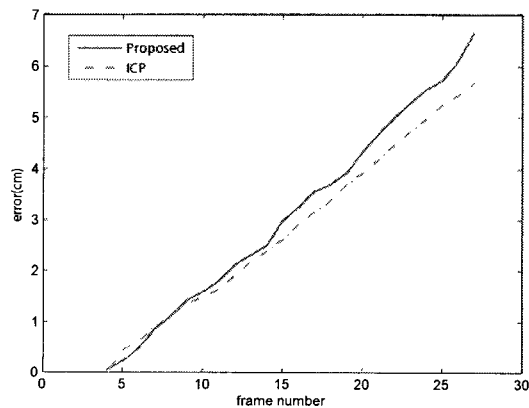


그림 16. 축구공 복원 형상들의 융합 오차

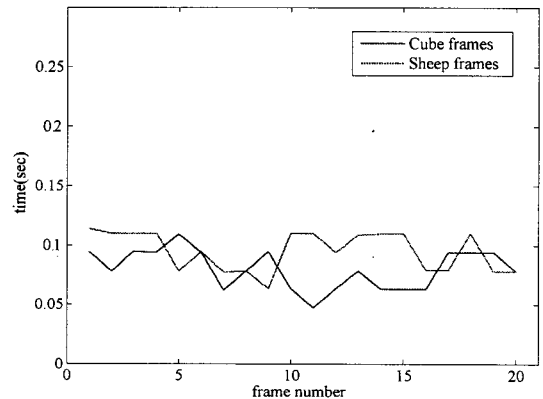


그림 17. 각 패턴상자 프레임과 인형 프레임에 대한 복원과 융합의 통합 수행 시간

복원 시에 동시에 계산된다. 부분 형상의 융합 시에 각 부분 점집합에서 공유되는 5개 이상의 공통점을 사용하여 한 전역 기준 좌표계로 변환하는 변환행렬을 계산한다. 변환행렬을 현재의 부분 형상 데이터에 적용하여 전체 형상 데이터의 기준 좌표계로 변환하여 점진적으로 융합된 형상 데이터를 획득할 수 있다.

논문에서 제안한 방법의 정확도 측정을 위하여 다양한 실제 영상 프레임으로 실험을 수행하였다. 촬영된 형상이 패턴 상자와 같이 균일한 특징점을 가지는 경우에 대해서는 특징점 추출 및 추적이 용이하므로 비교적 정확한 복원을 할 수 있었고 융합된 모습도 실제의 형상과 유사한 결과를 얻을 수 있었다. 추정된 카메라의 모션 정보도 실제 측정된 카메라의 위치 및 방향과 오차가 매우 작음을 알 수 있었다. 인형과 같이 형상의 모습이 매우 불균일한 형상의 경우 복원 형상의 에러가 균일 형상의 복원 오차보다는 크지만 형상 복원과 카메라 추정 및 융합에서 실물과 유사한 결과를 얻을 수 있었다. 융합 시에 필요한 공유점을 탐색할 때에 5개의 점이 필요하지만 더 많은 공유점이 사용 가능할 경우에 이를 모두 사용하였다. 이 경우 공유점에 대한 오차를 조금 커지나 전체적인 점에 대한 평균 오차가 줄어드는 것을 확인할 수 있었다.

제시한 형상 복원 방법은 세 영상으로부터의 선형적으로 3차원 형상을 복원하고 각 영상으로부터의 부분 형상을 점진적으로 융합한다. 모든 과정을 선형적으로 처리하여 반복적인 최적화 융합 방법의 단점인 처리 시간의 문제를 해결할 수 있었다. 평균적으로 초당 약 10 프레임의 처리 속도를 유지함을 확인할 수 있다. 이러한 수행 속도는 특징점 처리 알고리즘에 있어서 가속화의 여지가 있고 고성능 하드웨어의 사용으로 인한 처리시간 단축이 가능하므로 충분히 대화형 응용 시스템에 사용될 수 있는 성능이라고 볼 수 있다. 또한 구현에 있어서 성능 개선의 여지가 많이 남아 있으므로 다양한 대화형 응용 시스템에 적용이 가능할 것이다.

제시된 방법은 객체의 전체적 형상의 복원이 가능하나 부분 형상의 복원과 각 부분 형상의 융합 시에 실제의 형상과는 약간의 차이가 있다. 이러한 에러는 특징점 추출 및 추적에 대한 정확성 개선을 통해서 크게 줄일 수 있다. 복원된 전체 형상 데이터에 대해서는 형상의 표면을 표현할 수 있는 메시 구조를 계산하도록 하여 입력 영상 프레임에 기반한 텍스처

매핑을 통해 사실적인 묘사를 위한 실시간 렌더링 시스템에 응용될 수 있다.

## 참 고 문 헌

- [1] J. Lee, B. Moghaddam, H. Pfister, and R. Machiraju, "Silhouette-Based 3D Face Shape Recovery," *Proc. of Graphics Interface*, pp. 21-30, 2003.
- [2] L. Zhang, Z. Zhang, C. L. Tan, and T. Xia, "3D Geometric and Optical Modeling of Warped Document Images from Scanners," *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, Vol. 1, pp. 337-342, 2005.
- [3] P. Sturm, "Multi-View Geometry for General Camera Models," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 1, pp. 206-212, 2005.
- [4] M. Han and T. Kanade, "Creating 3D Models with Uncalibrated Cameras," *Proceeding of IEEE Computer Society Workshop on the Application of Computer Vision*, pp. 178-185, 2000.
- [5] J.-M. Frahm and R. Koch, "Camera Calibration and 3D Scene Reconstruction from Image Sequence and Rotation Sensor Data," *Proceedings of the Vision, Modeling, and Visualization*, pp. 79-86, 2003.
- [6] Y. F. Li and R. S. Lu, "Uncalibrated Euclidean 3-D Reconstruction Using an Active Vision System," *IEEE Transactions on Robotics and Automation*, Vol. 20, No. 1, pp. 15-25, 2004.
- [7] B. Micusik, D. Martinec, and T. Pajdla, "3D Metric Reconstruction from Uncalibrated Omnidirectional Images," *Proc. Asian Conf. Computer Vision*, Vol. 1, pp. 545-550, 2004.
- [8] C. Tomasi and T. Kanade, "Shape and Motion from Image Streams: a Factorization Method," *Technical Report CMU-CS-91-172*, Carnegie Mellon University, Pittsburgh, PA, September 1991.

[9] C. J. Poelman and T. Kanade, "A Paraperspective Factorization Method for Shape and Motion Recovery," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, pp. 206-218, 1997.

[10] L. Matthies, T. Kanade, and R. Szeliski, "Kalman Filter-based Algorithms for Estimating Depth from Image Sequences," *International Journal of Computer Vision*, Vol. 3, pp. 209-236, 1989.

[11] C. Tomasi and T. Kanade, "Shape and Motion without Depth," *Third International Conference on Computer Vision*, pp. 91-95, 1990.

[12] S. Rusinkiewicz and M. Levoy, "Efficient Variants of the ICP Algorithm," *Third International Conference on 3D Digital Imaging and Modeling*, pp. 145-152, 2001.

[13] P. Yan and K. W. Bowyer, "A Fast Algorithm for ICP-Based 3D Shape Biometrics," *Fourth IEEE Workshop on Automatic Identification Advanced Technologies*, pp. 213-218, 2005.

[14] C. Bregler, A. Hertzmann, and H. Biermann, "Recovering Non-Rigid 3D Shape from Image Streams," *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, Hilton Head, South Carolina, Vol. 2, pp. 690-696, 2000.

[15] G. H. Golub and C. Reinisch, "Singular Value Decomposition and Least Squares Solutions," In *Handbook for Automatic Computation*, Vol. 2, pp. 134-151, 1971.

[16] A. Johnson and S. B. Kang, "Registration and Integration of Textured 3-D Data," *Proceedings of 3DIM*, pp. 234-241, 1997.



박 종 승

1992년 경북대학교 전자계산학과 (이학사)  
 1994년 포항공과대학교 컴퓨터공학과 (공학석사)  
 1999년 포항공과대학교 컴퓨터공학과 (공학박사)  
 1999년~2001년 한국전자통신연구원 가상현실연구부 선임연구원  
 2001년~2003년 디지털아리아 시각정보처리연구소 책임연구원  
 2004년~현재 인천대학교 컴퓨터공학과 조교수  
 관심분야: 영상처리, 3D 비전, 증강현실, 게임공학



윤 종 현

2006년 인천대학교 컴퓨터공학과 (공학사)  
 2006년~현재 인천대학교 컴퓨터공학과 석사과정  
 관심분야: 동작기반 인터페이스, 증강현실 게임, 카메라 추적