

3차원 기하정보 및 특징점 추적을 이용한 다시점 거리영상의 온라인 정합

백 재 원[†] · 문 재 경^{**} · 박 순 용^{***}

요 약

본 논문에서는 물체의 3차원 모델을 복원하기 위하여 거리영상 카메라에서 획득한 다시점 3차원 거리영상을 온라인으로 정합(registration)하는 기술을 제안한다. 3차원 모델 복원을 위하여 거리영상 카메라를 복원하고자하는 물체 주위로 이동하여 연속된 다시점 거리영상과 사진영상을 획득하고 물체와 배경을 분리한다. 분리된 다시점 거리영상의 정합을 위하여 이미 등록된 거리영상의 변환정보 그리고 두 거리영상 사이의 기하정보를 이용하여 정합을 초기화한다. 위 과정을 통해 서로 인접한 거리영상에서 영상 특징점을 선택하고 특징점에 해당하는 거리영상의 3차원 점군을 이용하여 투영 기반(projection-based) 정합을 실시한다. 기하정합이 완료되면 사진영상 간의 대응점을 추적하여 정합을 정제(refinement)하는 과정을 거치는데 KLT (Kanade-Lucas-Tomasi) 추적기를 수정하여 대응점 탐색의 속도와 성공률을 증가시켰다. 영상 특징점과 추적된 대응점에 해당하는 3차원 점군을 이용하여 거리영상을 정제하였다. 정합과 정제의 결과를 통해 추정된 변환 행렬과 정합된 대응점들 사이의 거리를 계산하여 정합 결과를 검증하고 거리영상의 사용 여부를 결정한다. 만약 정합이 실패하더라도 경우에도 거리영상을 실시간으로 계속 획득하고 정합을 다시 시도한다. 위와 같은 과정을 반복하여 충분한 거리 영상을 획득하고 정합이 완료되면 오프라인에서 3차원 모델을 합성하였다. 실험 결과들을 통해 제안한 방법이 3차원 모델을 성공적으로 복원할 수 있음을 확인 할 수 있었고 오차 분석을 통해 모델 복원의 정확도를 검증하였다.

키워드 : 다시점 거리영상, KLT 정합, 스테레오, 온라인정합

Online Multi-view Range Image Registration using Geometric and Photometric Feature Tracking

Baek, Jaewon[†] · Moon, Jaekyoung^{**} · Park, Soon-Yong^{***}

ABSTRACT

An on-line registration technique is presented to register multi-view range images for the 3D reconstruction of real objects. Using a range camera, we first acquire range images and photometric images continuously. In the range images, we divide object and background regions using a predefined threshold value. For the coarse registration of the range images, the centroid of the images are used. After refining the registration of range images using a projection-based technique, we use a modified KLT(Kanade-Lucas-Tomasi) tracker to match photometric features in the object images. Using the modified KLT tracker, we can track image features fast and accurately. If a range image fails to register, we acquire new range images and try to register them continuously until the registration process resumes. After enough range images are registered, they are integrated into a 3D model in offline step. Experimental results and error analysis show that the proposed method can be used to reconstruct 3D model very fast and accurately.

Key Words : Multi-View Range Image, Klt Registration, Stereo, Online Registration

1. 서 론

3차원 컴퓨터비전기술은 카메라로 획득한 영상 정보로부터 실환경(real scene)의 3차원 정보를 정확하고 빠르게 추

출하려는 노력에서 비롯되었다. 국내외의 3차원 컴퓨터비전 기술은 주로 로봇, 디스플레이, 그리고 디지털 콘텐츠(digital contents) 등의 응용분야에 적용되고 있다. 3차원 정보를 빠르고 정확하게 추출함으로써 로봇이나 자동차가 카메라를 이용하여 자율적인 주행이나 위험감지, 자동주차 등에 사용할 수 있다. 또한 역사적인 유물 또는 유적을 3차원으로 디지털화 하여 문화원형을 보존 하고 특수효과 및 가상현실을 구현하는 등의 다양한 분야에 활용할 수 있다. 일반적으로

※ 본 연구는 2006년 정부(교육인적자원부)의 재원으로 한국학술진흥재단의 지원을 받아 수행된 연구임. (KRF-2006-331-D00618)

† 준 회원 : 경북대학교 컴퓨터공학과 석사과정

** 정 회원 : 경북대학교 전기전자컴퓨터학부 BK계약교수

*** 정 회원 : 경북대학교 컴퓨터공학과 조교수 (교신저자)

논문접수: 2007년 3월 22일, 심사완료: 2007년 7월 24일

실환경의 3차원 정보를 빠르게 추출하기 위해서는 거리영상 카메라(range imaging camera)를 사용한다. 하지만 사람의 눈과 마찬가지로 물체의 가려진 부분을 볼 수 없기 때문에 물체의 여러 방향에서 거리영상을 획득하여야 물체의 완전한 3차원 정보를 획득할 수 있다. 다시점(multi-view)에서 획득된 거리영상들은 서로 독립적인 좌표계를 가지기 때문에 이들을 하나의 좌표계로 이동시켜야 하는데 이를 3차원 정합(3-dimensional registration)이라 한다.

3차원 정합에 대한 초기 연구는 3차원 점군(point clouds) 또는 곡면(surface)을 정확하게 정합하고자 하였는데 Besl[1] 등에 의해 제안된 ICP(iterative closest point) 알고리즘은 가장 많이 사용되는 방법의 하나이다. Johnson[2] 등은 ICP에 색상 정보를 함께 사용한 Color ICP를 제안하였다. 그리고 Levoy[3] 등은 복셀(voxel)에 기반을 둔 방법으로 정합을 수행하였으며 카네기 멜론 대학의 Huber[4] 등은 그래프 이론을 응용하여 어떤 변환정보도 가지지 않은 다시점 거리영상을 정합하는 연구를 수행하였다. Urfalioglu[5] 등은 전역 최적화를 이용해 크기와 초기 위치에 관계없이 정합 성능이 좋은 연구를 하였다. 스탠포드 대학의 컴퓨터 그래픽스 연구실에서는 다비드상 등의 역사적인 문화유적을 디지털화한 디지털 미켈란젤로 프로젝트를 예로 들 수 있다[3]. 컬럼비아 대학의 로봇공학 연구실에서는 이동로봇에 거리영상획득 장치를 장착하고 뉴욕시의 문화유적에 대한 디지털화를 진행하고 있다[6]. Akbarzadeh[7] 등은 차량에 센서를 장착하고 INS와 GPS등의 장비를 사용해 도시를 모델링 하는 연구를 하였다. 하지만 최근에는 고속으로 거리영상을 정합하려는 시도가 많이 진행이 되고 있고 실시간으로 3차원 정합을 구현하려는 기술이 연구되고 있다.

다시점 거리영상을 실시간 혹은 온라인으로 연구한 대표적인 연구사례는 다음과 같다. Liu와 Heidrich[8]는 그래픽 하드웨어를 이용하여 스테레오 기반의 다시점 거리영상을 실시간으로 정합하는 연구를 수행하였다. Jaeggli와 Koninckx[9]는 패턴 프로젝터를 이용하여 다시점 거리영상을 획득하고 이를 온라인으로 정합하고 오프라인에서 정합 결과를 정제하는 연구를 수행하고 있다. 또한 Popescu[10] 등은 레이저 포인트 빛 비디오카메라를 이용한 실시간 3차원 거리영상 센서를 이용하여 물체 및 실내 환경 복원에 대한 연구를 수행하고 있다. 프린스턴 대학의 Rusinkiewicz[11]는 패턴 프로젝터와 투영기반 ICP를 사용하여 실시간으로 거리영상을 정합하고 3차원 모델을 재구성하는 연구를 수행하였다.

현재까지 실시간 정합에 대한 연구는 오프라인 정합에 비하여 그 수가 적은 편이고 실시간으로 거리영상을 획득할 수 있는 패턴프로젝터 또는 레이저 스캔 장치 등의 추가적인 센서를 필요로 하는 것이 대부분이다. 레이저 센서의 경우에는 거리영상의 정밀도는 높지만 텍스처를 사용하기 위해서는 추가로 카메라가 필요하고 2차원 센서와 3차원 센서의 정보를 일치시키기 위한 작업이 필요하다는 단점이 있다. 패턴프로젝터의 경우에는 2차원 영상과 3차원 영상이

동시에 획득되기 어려운 단점이 있다. 또한 거리영상 카메라와 그래픽 하드웨어를 이용하여 실시간으로 모델링 한 성과가 있지만 거리영상과 사진영상(photometric image)을 동시에 이용하여 온라인 정합을 시도한 연구는 보고된 바가 없다.

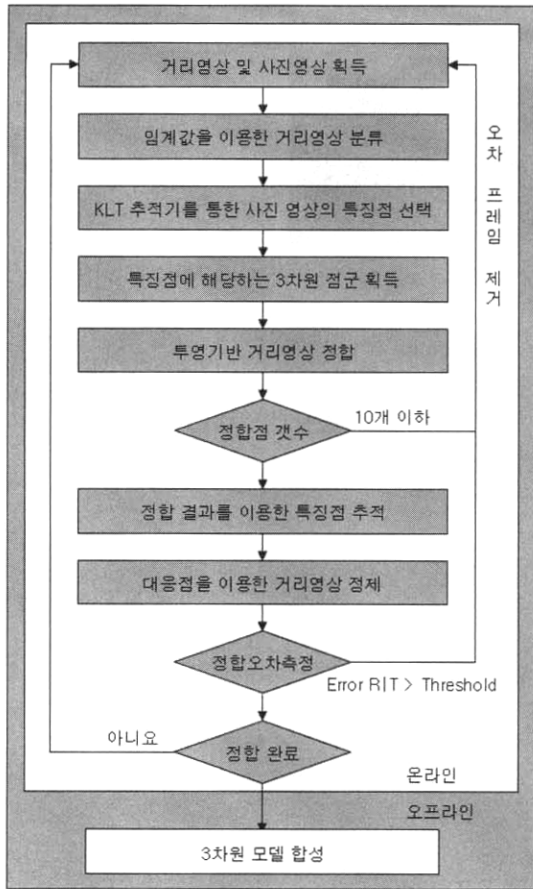
본 논문에서는 휴대 가능한 스테레오 기반의 거리영상 카메라를 이용하여 연속적으로 거리영상을 획득한다. 이들을 투영기반 정합기술로 1차적으로 정합하고 사진영상의 특징을 이용하여 2차적으로 정제함으로써 고속으로 더욱 정확하게 정합할 수 있다. 또한 사용자가 온라인으로 정합결과를 모니터링 하면서 카메라의 시점을 원하는 곳으로 이동하여 물체의 완전한 3차원 모델을 생성할 수 있도록 인터랙티브(interactive) 기능을 가지는 3차원 모델링 장치를 구현하였다.

본 논문에서 제안하는 방법을 요약하면 다음과 같다. 우선 첫 번째 거리 영상이 획득되면 3차원 점군을 샘플링하여 투영변환행렬(perspective projection matrix)을 추정하고 획득된 점군의 거리정보를 이용하여 배경과 물체를 분리한다. 투영기반 정합이 성공적으로 실시되려면 두 거리영상간의 초기 오차가 적을수록 좋은데 이를 위해 본 논문에서는 두 거리영상의 중점간의 거리차이를 이용하여 고속으로 두 거리영상의 이동오차를 줄인다. 가까워진 두 거리영상을 투영기반 정합 기술을 사용하여 정합하였다. 또한 KLT 추적기를 사용하여 영상 특징점들의 대응점을 찾아 투영기반 정합 기술의 결과를 정제 하였다[12,13]. 이때 초기정합의 결과를 이용하여 사진영상의 특징점을 탐색하는 범위를 줄일 수 있어 정합을 빠르게 구현할 수 있다. 위 과정을 통해 거리영상이 정합되면 두 거리영상의 회전오차와 이동오차를 측정하여 자동적으로 정합 결과를 검증하고 거리영상의 사용 여부를 판단한다. 만약 거리영상이 정합에 실패하면 사용자는 카메라의 시점을 변경하면서 정합이 다시 성공하도록 하여 정합과정을 계속한다. 충분한 거리영상이 정합되면 오프라인에서 3차원 점군을 하나의 3차원 모델로 합성(integration)한다. 제안된 시스템을 사용하여 실험실 환경과 외부 환경에서 물체에 대하여 실험을 수행하였다.

본 논문의 구성은 다음과 같다. 2절에서는 온라인 정합의 전체적인 구조를 다루고 투영기반 거리영상 정합 기술은 3절에서 다룬다. 4절에서는 영상기반 정합기술을 보이며 5절은 정합 결과의 오차를 계산하고 사용여부를 결정하는 방법을 설명한다. 6절에서는 정합 과정을 실시간으로 사용자에게 보여주는 온라인 프리뷰와 실험결과 및 오차분석을 다룬다. 7절에서는 본 연구의 결론을 기술한다.

2. 온라인 정합 과정

본 논문에서 다시점 거리영상의 정합은 (그림 1)과 같은 과정으로 이루어진다. 먼저 거리영상 카메라를 통해 거리영상 및 사진영상을 획득한다. 첫 번째 획득된 거리영상을 선형 샘플링하여 약 100개의 3차원 점군을 추출하여 투영변환행렬을 추정한다. 투영변환행렬의 추정에는 선형 최소자승



(그림 1) 다시점 거리영상의 정합 과정

법(linear least square method)을 사용하였다. 또한 획득된 거리영상은 배경과 물체의 구분이 없는 상태인데 일반 환경에서는 스테레오 알고리즘의 특성상 특징이 없는 벽 등에는 잡음이 많아서 정합에 사용하기에 적합하지 않다. 거리영상에서 배경에 대한 부분을 제거하기 위해 임계값을 기준으로 배경이라고 판단되는 부분을 제거한다.

두 개 이상의 거리영상을 획득하면 거리영상의 기하적인 특성에 기반을 둔 투영기반 거리영상 정합을 수행한다. 하지만 다시점에서 획득된 거리영상은 서로 초기 오차가 큰 경우가 많기 때문에 투영기반 거리영상 정합을 수행할 때 정합점을 탐색하지 못하는 문제가 발생하는데 이러한 문제를 해결하기 위해 정합할 두 거리영상의 중점(centroid)차를 계산하여 초기 위치를 결정함으로써 이동 오차(translation error)를 최소화 한다. 기하정합을 수행하기 위해 KLT 추적기를 사용하여 n 번째 영상의 특징점을 선택한다. 기존의 KLT 추적기는 특징점을 선택하는데 사진영상만의 정보를 사용하지만 본 논문에서는 특징점에 해당하는 3차원 좌표가 없는 경우에는 선택하지 않았다. 이를 통해 투영기반 거리영상 정합과 영상기반 정합에 사용되는 특징점은 배경을 제외하고 물체에 대해서만 선택된다. 선택된 특징점들에 해당하는 3차원 좌표를 투영기반 거리영상 정합의 입력 좌표로 사용하여 $[n-1, n]$ 의 한 쌍으로 정합을 수행

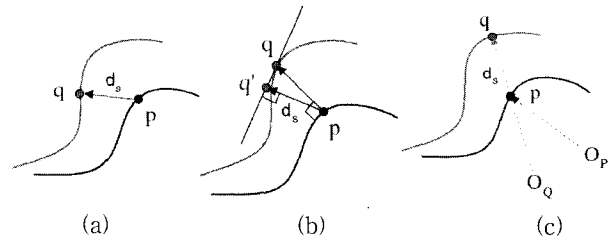
한다. 투영기반 거리영상 정합에 대해서는 3절에서 자세히 설명한다.

투영기반 거리영상 정합의 결과를 정제하기 위해 수정된 KLT 추적기를 사용한다. n 번째 영상의 특징점 p 에 해당하는 3차원 좌표 P 가 투영기반 거리영상 정합을 통해 p 의 대응점 q 에 해당하는 3차원 좌표 Q 와 가까워 졌다고 가정하고 정합된 P 점을 $n-1$ 번째 영상에 투영하여 KLT 추적기의 시작 좌표로 사용한다. 수정된 KLT 추적기를 사용하여 특징점을 $n-1$ 번째 영상에서 추적하여 대응점을 찾는다. 특징점과 대응점에 해당하는 3차원 점군을 획득하여 투영기반 정합의 결과를 정제한다.

위와 같이 정합을 수행하고 나면 두 거리영상이 정합이 성공적인가를 검증하는 과정이 필요하다. 본 논문에서는 정합의 결과인 회전변환행렬의 대각 성분의 RMS(root mean square) 오차를 측정하는 방법으로 회전 오차(rotation error)를 측정하고 기하정합에서 사용된 정합점들의 거리를 평균하여 이동 오차를 측정하였다. 측정된 오차가 임계값 이상일 경우에는 정합이 실패하였다고 판단하고 해당 프레임은 삭제한다. 위 과정을 반복하면서 거리영상 카메라의 시점을 이동하고 충분한 데이터를 얻으면 정합을 종료하고 3차원 합성 과정을 통해 하나의 모델을 구성한다.

3. 투영기반 정합

거리영상의 정합은 (그림 2)에서 보는 것과 같이 크게 점대점(point to point), 점대면(point to plane), 그리고 점대투영점(point to projection)의 세 가지로 나눌 수 있다[14,15]. 일반적으로 알려져 있는 ICP(iterative closest point) 방법은 점대점 또는 점대면 방법을 말하며 정확성이 높지만 속도가 느린 단점이 있다. 점대투영점 방식은 2.5차원의 거리영상과 3차원 곡면사이의 투영관계를 이용한다. ICP와는 달리 속도가 빠르지만 두 곡면의 초기 자세가 큰 오차를 가지고 있을 때는 ICP에 비해 정확성이 떨어지는 단점이 있다.



(a)점대점 방법 (b)점대면 방법 (c)점대투영점 방법
(그림 2) 대표적인 거리영상 정합 기술

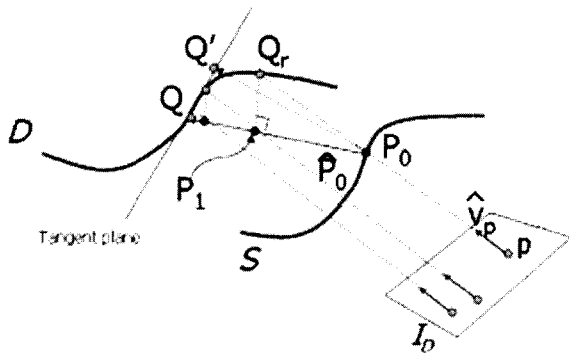
본 논문에서는 이러한 문제를 해결하기 위해 점대면과 점대투영점 방식의 장점을 결합한 정합 기술인 IPP(iterative-projection-point) 방식을 사용하였다[16]. 이 방식은 점대투영점 기술을 변형하여 점 P 와 Q 를 각각 영상면과 P 의 법선(normal)에 연속적으로 투영함으로써 점대면

방식을 고속으로 구현할 수 있다.

3.1 IPP(Iterative Projection Point) 알고리즘

(그림 3)에서 IPP 방식을 간단히 설명하고자 한다. S 곡면상의 점 P_0 의 정합점을 곡면 D에서 찾는 것이 정합의 주목적이다. 우선 P_0 를 D 곡면의 2차원 영상으로 투영하여 좌표 p를 구한다. 좌표 p에 해당하는 D 곡면상의 점 Q는 거리영상을 이용하여 구할 수 있다. 이 점을 다시 P_0 의 법선에 투영하여 P_1 을 구하고 위의 과정을 P_k 가 수렴할 때까지 반복하여 Q를 구한다. 이러한 방법으로 충분한 수의 대응점의 집합 {P}와 {Q}를 구하였을 때 두 점집합 사이의 정합 오차 ϵ 를 최소화하는 3차원 변환 행렬(transformation matrix) T를 다음 선형 방정식을 이용하여 구한다. 본 논문에서는 수식 (1)을 풀기 위해 SVD(singular value decomposition) 방법을 사용하였다.

$$\epsilon = \sum_i \| P_i - TQ_i \| \quad (1)$$



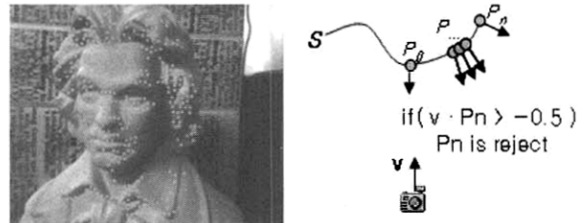
(그림 3) IPP 기술을 이용한 정합점 탐색

위 과정을 반복하여 최종적인 변환 행렬을 계산한다. 본 논문의 실험에서 사용된 두 개의 거리영상에 대하여 IPP 알고리즘을 적용한 결과 일반적인 경우 20회 이상의 정합을 수행하여도 정합 결과가 변하지 않았으며, 따라서 본 논문에서는 투영기반 거리영상 정합은 20회만 수행하였다.

3.2 IPP 입력 좌표 선택

본 논문에서는 IPP의 입력 좌표 집합 {P}를 선택하기 위해 다음과 같은 방법을 사용한다. 먼저 사진영상에서 KLT 추적기를 사용하여 특징점을 선택한다. 선택한 특징점들에 해당하는 3차원 좌표가 있는지를 확인하여 특징점이 대상 물체 위에 존재하는가를 판단하고 배경에 해당하는 특징점은 사용하지 않는다.

위 과정을 통해 (그림 4)에서 보는 것과 같이 물체에 해당하는 특징점만을 추출 할 수 있는데 이것은 영상기반 정제 과정에서 사용된다. 또한 카메라의 주시 방향(viewing direction)의 벡터와 입력된 3차원 점 P의 법선의 내적이 큰 경우에는 점 P가 곡면의 경계에 위치할 가능성이 크다. 스



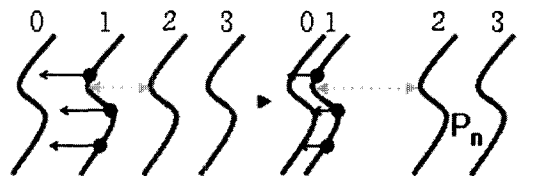
(그림 4) (a)유효한 특징점과 (b)IPP에 적합한 3차원 점 선택

테레오 알고리즘의 특성상 물체의 경계선에 위치한 경우 잡음이 발생할 가능성이 크기 때문에 실험에서는 내적이 -0.5 이상인 3차원 점들은 정합에 사용하지 않았다.

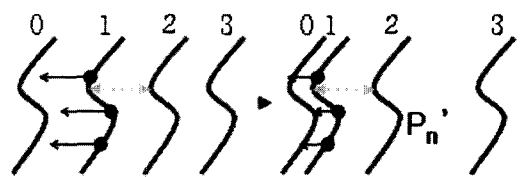
3.3 쌍정합과 초기 오차 최소화

본 논문에서 투영기반 정합은 한 쌍(pair-wise)의 거리영상 사이에서 이루어진다. 다수의 연속된 영상에 대한 정합을 실시할 때 쌍정합의 문제점은 (그림 5(a))를 통해 알 수 있다. 다수의 영상에서 0번과 1번 곡면을 정합하면 1번과 2번 곡면은 실제로 획득된 거리보다 멀어지게 되므로 정확한 정합점을 찾기 힘들다. 이러한 문제점을 해결하기 위해 수식 (2)와 같이 1번과 2번 곡면을 정합하기 전에 0번과 1번 곡면 사이의 변환 행렬을 2번 곡면에 곱해주면 1번과 2번 곡면은 다시 원래의 변환 관계로 돌아간다.

$$P_n' = T_{n-1}P_n \quad (2)$$



(a)



(b)

(그림 5) (a)쌍정합의 문제점과 (b)해결방법

또한 본 논문에서 사용하는 투영기반 거리영상 정합 방법은 두 거리영상의 초기 이동 오차가 큰 경우에는 정합점의 탐색 성능이 떨어진다. 이러한 문제점을 해결하기 위해 수식 (2)를 통해 원래의 변환관계로 돌아간 두 곡면의 중점을 계산하고 그 차이를 이용하여 두 곡면의 초기 이동 오차를 최소화 한다.

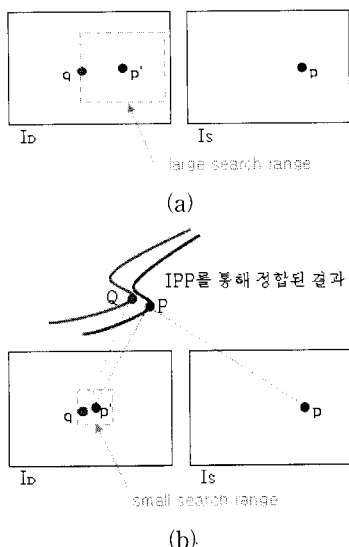
$$C_D = \frac{\sum_{i=0}^k (P_i \in D)}{k} \quad (3)$$

$$P' = P + (C_D - C_S) \tag{4}$$

곡면 S 를 D 곡면으로 이동시켜 초기 이동 오차를 최소화 하기 위해서는 우선 수식 (3)을 통해 곡면의 중점을 계산해야 한다. D 곡면의 중점 C_D 를 계산할 때는 10% 정도의 3차원 점을 선형 샘플링하여 계산하는데 수식에서 k 는 샘플링된 3차원 점의 수를 말한다. 정합할 두 곡면의 중점 C_D 와 C_S 가 계산되면 수식 (4)를 통해 두 중점 사이의 차이가 두 곡면 사이의 이동 오차라 가정하고 S 곡면에서 두 중점 사이의 차를 더해 주어 S 곡면을 D 곡면으로 이동시킨다. 제안된 방법을 통해 정합할 두 곡면은 이동 오차가 최소화 되어 투영기반 거리영상 정합을 수행하기에 적합한 것을 알 수 있었다.

4. 영상기반 정합

기존의 영상기반 정합기술은 거리영상이 존재할 때 사진 영상간의 대응점을 추적하여 해당하는 3차원 점군을 정합에 이용하였다. 하지만 좌표변환이 큰 거리영상 사이의 대응점을 추적 할 경우에는 탐색 범위의 크기에 따라 추적의 성공률과 추적에 걸리는 시간이 상충(trade off)한다. 실험에서는 보정(rectification)된 사진영상과 이것을 기반으로 생성된 거리영상을 사용하였다. (그림 6(a))는 기존 KLT 추적기가 사진영상 I_S 상의 특징점 p 의 대응점 q 를 I_D 에서 추적하는 과정을 나타낸다. 기존 KLT 추적기는 p 와 동일한 좌표를 가지는 p' 를 탐색의 시작 좌표로 사용하기 때문에 넓은 탐색범위를 필요로 한다. 하지만 제안된 방법은 (그림 6(b))와 같이 IPP를 통해 정합된 결과를 사용하여 탐색 시작점 p 를 결정하는 방법으로 KLT 추적기의 입력 좌표를 수정함으로써 상대적으로 좁은 탐색 범위로 대응점 추적 성공률을 높였고 추적 속도를 개선하였다.

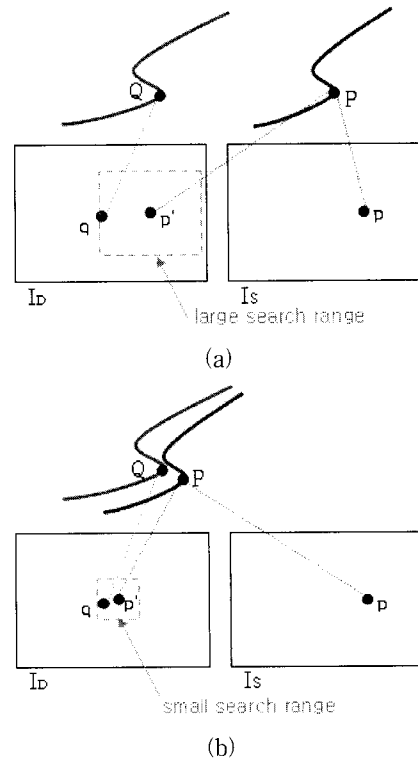


(그림 6) (a)기존 KLT와 (b)수정된 KLT의 탐색 시작점 비교

(그림 7)은 본 논문에서 제안하는 방법을 나타낸다. 특징점 p 에 대응되는 3차원 거리를 P 라 할 때 수식 (5)와 같이 I_D 에 투영하여 p' 을 구한다. 이 때 모든 3차원 점군 데이터는 초기정합 과정을 통해 0번 거리영상에 정합되어 있으므로 I_D 거리영상의 역변환행렬 T_{n-1}^{-1} 를 I_S 거리영상에 곱해주었고 투영변환행렬 M 을 곱하여 I_D 사진영상에 투영하였다.

$$p' \cong M(T_{n-1}^{-1}(P)) \tag{5}$$

정합 전에는 투영된 P 가 기존 KLT의 탐색 시작 좌표와 동일하지만 정합 후에는 추적해야 할 대응점 q 에 근접하는 것을 (그림 7(b))를 통해 알 수 있다. 이를 통해 기존 KLT에 비해 작은 탐색 범위를 사용하여 대응점을 추적하고 p 와 q 의 집합을 구할 수 있다. p 와 q 에 해당하는 3차원 점들의 집합을 수식 (1)에 대입하여 기하정합 결과를 정제하였다.



(그림 7) (a)정합 전과 (b)정합 후의 정합점 탐색 범위 비교

5. 정합 성공의 판정

본 논문에서는 거리영상을 정합할 때 항상 두 장의 거리영상을 사용한다. 다수의 거리영상을 정합할 때 쌍정합의 정합 오차가 높은 경우가 발생하면 이후의 거리영상은 잘못 정합된 거리영상을 기준으로 정합되게 된다. 본 논문에서는 쌍정합을 수행하고 변환행렬을 이용하여 정합의 오차를 계

산하였다. 정합의 오차는 다음 수식 (6)을 통해 계산한다.

$$T_{n-1,n} = T_{n-1} T_n^{-1} \quad (6)$$

첫 번째로 거리영상에 대한 변환행렬 T 는 0번째 거리영상을 중심으로 계산되어 있으므로 쌍정합의 정합 오차를 계산하기 위해서는 $n-1$ 번째 거리영상에 대한 변환행렬 T_{n-1} 와 T_n^{-1} 를 곱하여 $n-1$ 번째 거리영상과 n 번째 거리영상 사이의 변환행렬을 계산한다. 계산된 변환행렬 $T_{n-1,n}$ 에서 회전 변환행렬을 이용해 수식 (7)을 통해 회전 오차(e_R)를 계산한다.

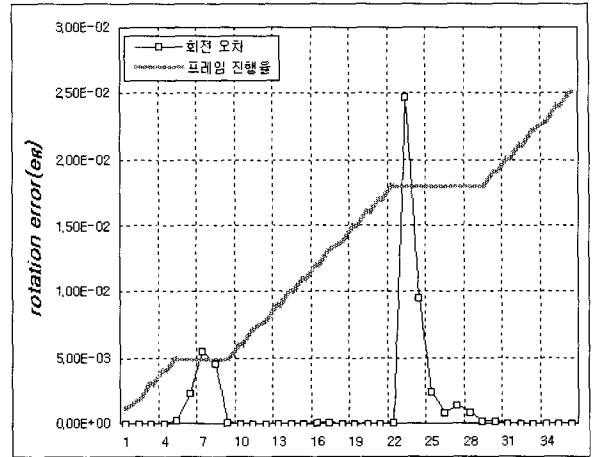
$$e_R = \frac{\sqrt{\sum_{i=0}^2 (I_{ii} - R_{ii})^2}}{3} \quad (7)$$

$$e_T = \frac{\sum_{i=0}^k |Q_i - P_i|}{k} \quad (8)$$

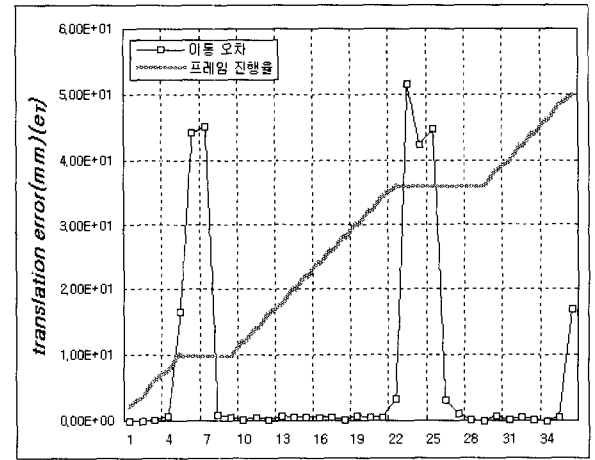
정합 결과의 이동 오차(e_T)는 수식 (8)을 통해 계산할 수 있다. 회전변환행렬은 정합이 성공적인 경우 단위행렬에 가까워진다고 가정한다. 단위행렬과 회전변환행렬의 대각 성분의 차에 대한 RMS 값을 회전 오차로 사용한다. 하지만 이동 오차는 사용자가 카메라의 시점을 얼마나 이동 하는가에 따라 차이가 나고 정합 결과가 성공적이라도 이동 변환행렬의 이동 오차가 낮다고 가정할 수 없다.

본 논문에서는 정합에 사용한 정합점 P_i 와 탐색된 대응점 D_i 의 평균 거리를 측정하여 두 거리영상의 이동 오차를 계산한다. 본 논문에서 사용한 베토벤상에 대한 실험을 통하여 정합이 성공적인 경우와 실패하는 경우의 회전 오차와 이동 오차를 조사하고 적절한 임계값을 설정하였다. 거리영상이 정합이 실패한 경우, 우리는 거리영상을 계속 획득하고 정합을 다시 시도한다. 거리영상 카메라를 사용하는 사용자는 모니터를 통하여 정합의 과정을 관측하고 정합위치를 추정할 수 있다. 정합이 다시 시작될 때까지 사용자는 카메라의 위치를 수정하여 영상을 획득한다.

(그림 8)은 실험에 사용된 물체에서 24프레임의 거리영상을 온라인으로 획득하며 시점을 이동시켜 정합한 경우의 정합 결과 및 오차 측정 결과이다. (그림 8)을 보면 총 36프레임의 거리영상이 획득되었으나 오차가 크다고 판단되어 정합에 사용되지 않은 프레임이 12개 존재하는 것을 알 수 있다. 그림에서 짙은 회색 선이 실제 정합에 사용된 거리영상의 수를 나타낸다. 정합에 사용된 거리영상의 수가 증가하다가 정합이 실패한 경우, 즉 오차가 큰 경우, 증가하지 않고 다시 정합이 시작되면 거리영상의 수가 증가함을 알 수 있다. (그림 8)에서는 회전 오차가 0.0008 이하이고 거리 오차가 2mm 이하인 경우에 정합이 성공적이라고 판단하였다.



(a)



(a)

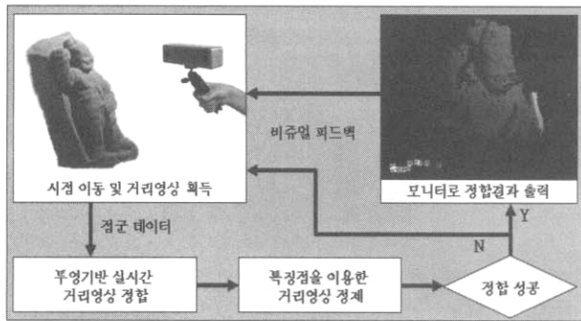
(그림 8) 정합이 실패한 경우 프레임 진행에 따른 (a) 회전오차와 (b) 이동오차의 변화

6. 실험 결과

본 논문에서는 캐나다 PointGrey에서 개발된 BumbleBee 스테레오 카메라를 사용하여 320x240 해상도의 사진영상과 거리영상을 획득하였다. 카메라는 사용자가 쉽게 휴대가 가능하고 실외 실험의 경우 노트북을 이용하였다. 실험은 실험실 환경과 일반 환경에서 서로 다른 물체를 사용하여 진행되었다. 정합에 사용된 컴퓨터는 실험실에서는 Pentium 3.4 GHz, 일반환경에서는 Pentium 1.83GHz의 노트북을 사용하였다. 제안된 방법은 위 시스템에서 두 장의 거리영상을 정합하는데 색상 정보를 포함할 경우 약 800ms, 포함하지 않을 경우 약 600ms의 정합 속도를 보였다.

6.1 온라인 정합 프리뷰

제안된 시스템에서는 다시점에서 획득된 거리영상을 실시간으로 정합하는 과정을 사용자에게 보여준다. 사용자는 이를 통해 카메라의 다음 시점을 결정할 수 있도록 시각적 피드백(visual feedback)을 받을 수 있다. 사용자가 카메라를



(그림 9) 인터랙티브 모델링 시스템

손에 들고 거리영상을 획득하면서 화면상에 3차원 점군을 표시해준다.

사용자는 획득된 거리영상에 빈 영역(hole)이 존재하는지 프리뷰를 통해 확인하고 빈 영역에 대한 거리영상을 더 획득할 수 있다. 또한 물체의 완전한 3차원 모델을 복원하기 위해 어떤 부분에 대한 거리영상을 획득하여 정합해야 하는지 알 수 있다. 본 논문에서 온라인 정합에 대한 프리뷰는 OpenGL을 사용하여 구현하였다.

(그림 9)에서 보는 것과 같이 모니터를 통해 거리영상의 정합 과정을 볼 수 있고 또한 정합 결과로 추정된 변환 행렬을 통해 거리영상 카메라의 위치를 추정하여 화면에 표시해준다. 사용자는 이를 통해 거리영상을 획득하여 정합 후 오차 때문에 정합이 중단된 경우에는 카메라가 위치했던 마지막 위치를 파악하여 정합을 재개할 위치를 알 수 있다.

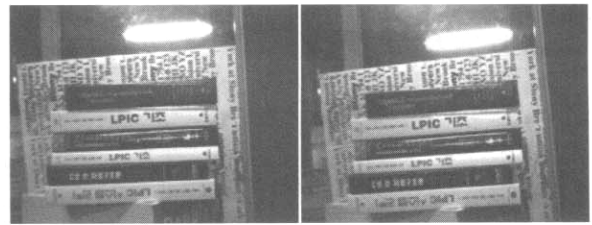
6.2 기존 방법과의 성능 비교

실험에서는 동일한 KLT 탐색 범위를 사용하여 제안한 방법과 기존의 KLT를 사용했을 경우에 추적 성공률 및 정합 결과에 대한 비교를 수행하였다. 실험은 (그림 10(a))의 두 영상을 사용하였고 KLT의 search range와 window size는 5로 설정하였다. (그림 10)의 (b),(c)는 특징점에 대한 추적 결과를 운동 벡터로 표현한 것이다. 두 결과 모두 잘못된 추적 결과는 포함되어 있으나 기존의 KLT를 사용한 경우에 더 많은 수의 잘못된 추적 결과가 존재함을 알 수 있다.

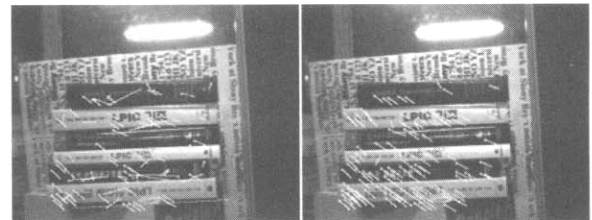
(그림 11)은 수정된 KLT와 기존 KLT를 사용하여 영상기반 정합을 수행한 결과를 나타낸 것이다. 기존 KLT의 경우 (그림 10(b))와 같이 추적에 실패한 경우가 영상기반 정합의 입력으로 사용되므로 (그림 11(a))와 같이 좋지 않은 정합 결과를 보였다. 정합 결과에 대해 정량적인 평가를 하기 위해 거리영상의 약 10%를 선형 샘플링하여 영상기반 정합에 사용하였다. <표 1>에서 보는 것과 같이 제안된 방법의 정합 오차가 더 낮은 것을 알 수 있다.

<표 1> 거리영상 정합 및 모델합성의 속도 및 오차

	특징점	추적된 특징점	정합 전 오차(mm)	정합 후 오차(mm)
기존 KLT	266	84	25.59	18.88
제안된 방법	266	128	3.61	2.23



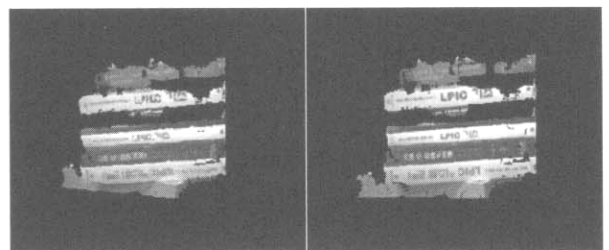
(a)



(b)

(c)

(그림 10) KLT 입력 영상(a)과 (b)기존 KLT와 (c)제안된 방법의 추적 결과



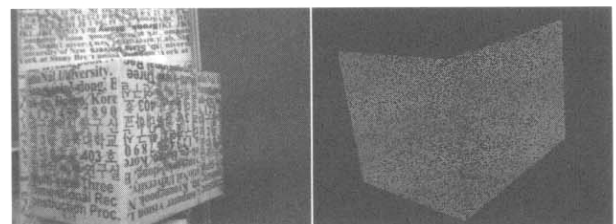
(a)

(b)

(그림 11) color point cloud의 정합 결과 비교 (a)기존 KLT와 (b)제안된 방법

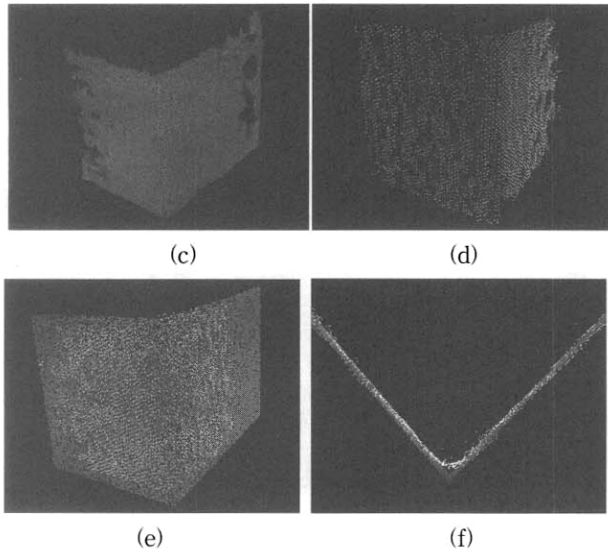
6.3 모델 복원의 오차 분석

이 절에서는 3차원 모델 복원의 정확도를 분석하기 위하여 가상의 3차원 모델과 복원된 모델을 비교하였다. 실험에서는 (그림 12(a))와 같은 박스를 실측하여 얻어진 수치로 (그림 12(b))와 같은 215*305*253(mm)의 ground truth 모델을 생성하였다. 그리고 (그림 12(a))에 대한 40장의 거리영상을 획득하여 (그림 12(c))와 같은 메쉬 모델을 합성하였다. 생성된 메쉬 모델은 폐곡면을 가정하고 생성된 것이므로 ground truth 모델과의 정합을 통해 모델 복원의 정확도를 분석하려면 표면에 해당하는 점군만을 추출해야 하는데 적절한 임계값을 사용하여 (그림 12(d))와 같은 점군을 추출할 수 있었다.



(a)

(b)

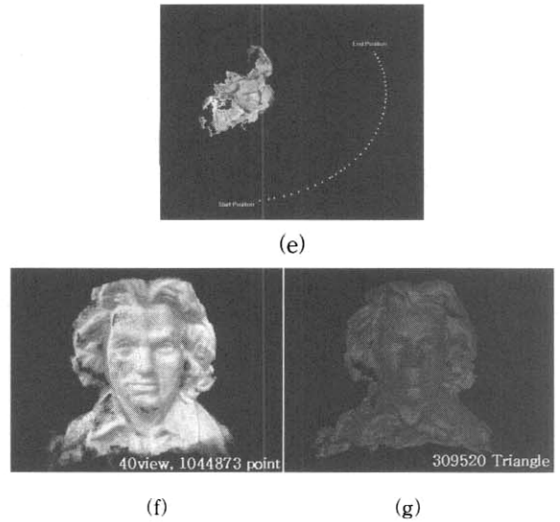


(그림 12) 실험에 사용된 (a)물체와 (b)가상의 거리영상, (c)합성된 모델, (d)실험을 위해 모델에서 추출된 점군, (e)정합 결과, (f)모서리를 확대한 결과

Ground truth 모델과 복원된 모델에서 추출된 점군을 ICP 방법을 사용하여 정합한 결과를 (그림 12(e))에서 볼 수 있는데 육안으로 보기에 정확히 정합된 것을 알 수 있으며 정합 결과는 <표 2>에서 정리하였다. 본 논문에서 거리영상의 획득하는 스테레오 센서의 특성 때문에 추출된 점군의 모서리 부분은 물체와 다르게 나온 것을 (그림 12(f))에서 확인할 수 있다. 위 실험을 통해 본 논문에서 제안하는 시스템의 결과로 나온 3차원 모델은 물체에 가깝게 복원되었음을 알 수 있다.

<표 2> 합성된 모델의 정확도 분석 결과 (mm)

ICP 정합점 수	평균 오차	RMS 오차
2956	1.220462	1.416079



(a)보정된 사진영상, (b)거리영상, (c) 특징점 선택, (d)추적된 특징점의 운동 벡터, (e)추정된 카메라 위치, (f)정합된 점군, (g)합성된 모델

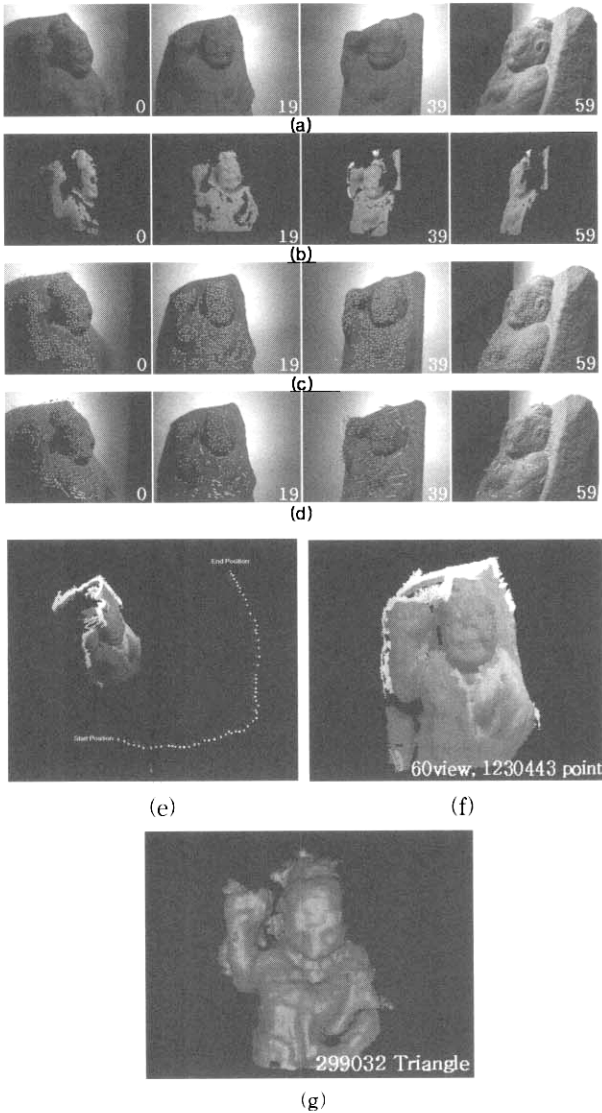
(그림 13) 실험 1. 베토벤상에 대한 실험

6.4 실험 결과

(그림 13)은 실험실 환경에서 베토벤상에 대해 40장의 연속된 사진영상 및 거리영상을 획득하여 정합을 수행한 결과이다. 데이터를 획득 할 때 배경과 카메라는 고정시키고 물체만 회전관위에서 회전시키는 방법을 사용하였다. 획득된 모든 거리영상을 정합하고 메쉬 구조로 3차원 모델이 생성된 것을 확인할 수 있다. 메쉬 모델의 생성은 Marching Cubes 알고리즘을 사용하였다[17]. 메쉬 모델을 합성할 때는 전체 거리영상을 사용한 경우와 20장의 영상의 사용한 경우로 실험하였다. (그림 13(g))의 메쉬 모델을 합성할 때의 복셀 크기는 3mm로 지정하였다.

(그림 14)는 일반 환경의 박물관에 소장중인 금강역사 부조에 대해 60장의 연속된 사진영상 및 거리영상을 획득하여 정합을 수행한 결과이다. 데이터를 획득 할 때 물체는 고정되어 있고 카메라를 손에 휴대하는 방법으로 촬영하였다. 실험 환경에 비해 카메라의 움직임이 불안정하고 3차원 점군에 약간의 잡음도 포함되어 있다. 또한 박물관의 조명 특성상 물체에 그늘진 곳이 생겨 다시점에서 데이터를 받았지만 3차원 점군의 데이터를 획득할 수 없는 곳이 있었다. 결과적으로 금강역사 부조의 경우 정합 속도가 약간 느린 것을 알 수 있었다. (그림 14(g))의 금강역사 부조 모델의 합성에는 복셀 크기를 2mm로 지정하였다.

(그림 15)는 바위와 나무가 함께 있는 조경을 약 5m 정도 거리영상 카메라를 손에 이동하며 연속된 40장의 거리영상을 정합한 결과이다. 입력되는 거리영상에 많은 잡음이 존재하고 식물 등 정합하기 어려운 부분이 존재한다. 하지만 영상 윗부분의 바닥과 같은 잡음이 심한 부분을 임계값을 통해 제거하고 정합에 사용하여 성공적인 결과를 보였다.



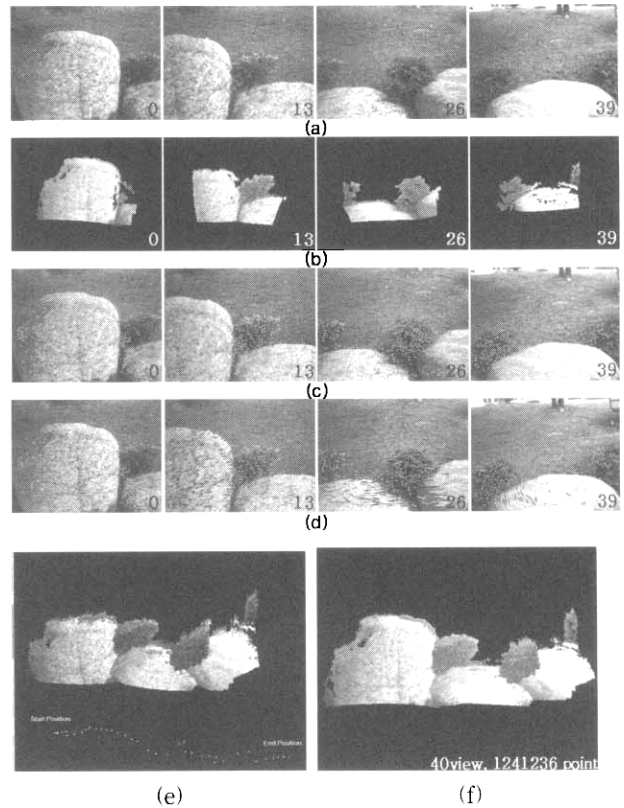
(a)보정된 사진영상, (b)거리영상, (c) 특징점 선택, (d)추적된 특징점의 운동 벡터, (e)추정된 카메라 위치, (f)정합된 점군 (g)합성된 모델

(그림 14) 실험 2. 사천왕상에 대한 실험

<표 3> 거리영상 정합 및 모델합성의 속도 및 오차

실험 물체	평균정합속도 (sec)	정합오차		복셀 크기 (mm)	모델합성속도 (sec)		
		이동 오차 (mm)	회전 오차		사용프레임 수		
					20	40	60
베토벤상	0.71/frame	0.47	0.00052	3	350	720	N.A.
금강역사 부조	0.85/frame	0.62	0.00028	2	220	450	660
조경	0.93/frame	4.18	0.00032	10	400	1119	N.A.

<표 3>에서는 두 물체 대한 모델링 실험의 정합 오차 및 속도를 보여준다. 정합 속도는 쌍정합의 속도를 전체 정합 횟수로 나누어 평균을 계산하였고 합성속도는 전체 영상을 사용해서 합성했을 경우, 일부 영상을 사용했을 경우로 나



(a)보정된 사진영상, (b)거리영상, (c) 특징점 선택, (d)추적된 특징점의 운동 벡터, (e)추정된 카메라 위치, (f)정합된 점군, (g)합성된 모델 (그림 15) 실험 3. 실외 조경에 대한 실험

누었다. 한 프레임의 정합 속도가 1초 이하로 비교적 빠르게 정합을 할 수 있었다. 사용자가 거리영상 카메라를 휴대하고 영상을 획득하는 경우 1초 이하의 속도로 각 프레임 정합할 수 있기 때문에 온라인 모델링이 가능하다.

7. 결론 및 향후 연구

본 논문에서는 다시점 거리영상을 온라인으로 정합하기 위한 기술을 제안하였다. 제안된 방법을 사용함으로써 온라인으로 거리영상 카메라의 시점을 이동하며 거리영상을 획득하고, 정확하게 다시점 거리영상을 정합할 수 있었다. 초기 오차를 최소화하여 시점의 변화가 비교적 큰 거리영상이 발생하더라도 성공적으로 정합이 될 수 있도록 하였으며 사용자가 거리영상 카메라를 손에 들고 사용한 경우에도 정합 성능이 크게 떨어지지 않음을 알 수 있었다. 그리고 획득된 거리영상이 성공적으로 정합되었는지 여부를 자동으로 판단하여 정합오차가 큰 거리영상은 3차원 모델링에 사용하지 않음으로써 더 정확한 3차원 모델을 생성할 수 있다. 또한 온라인 프리뷰를 통해 정합이 되는 과정을 실시간으로 확인할 수 있었고 데이터가 부족하여 홀이 생긴 부분이나 카메라의 추정된 위치 등을 파악할 수 있으므로 인터랙티브하게 3차원 모델을 생성할 수 있었다.

추후에는 그래픽 하드웨어를 사용하여 정합이 수행되는

소요시간을 높여 좀 더 실시간(real-time)에 가깝게 모델링을 할 수 있는 방법과 사진영상을 이용하여 3차원 모델의 텍스처를 실사에 가깝도록 하는 연구를 진행할 예정이다.

참 고 문 헌

[1] P. J. Besl and N. D. McKay, "A Method for Registration of 3-D Shapes", IEEE Trans. on Pattern Recognition and Machine Intelligence, Vol. 14, No. 2, pp. 239-256, 1992.

[2] A. E. Johnson and S. B. Kang, "Registration and integration of textured 3D data", Image and Vision Computing, Vol. 17, No. 2, pp. 135-147, 1999.

[3] M. Levoy, K. Pulli, B. Curless, S. Rusinkiewicz, D. Koller, L. Pereira, M. Ginzton, S. Anderson, J. Davis, J. Ginsberg, J. Shade, and D. Fulk. "The digital michelangelo project : 3D scanning of large statues", In SIGGRAPH, pp. 131-144, 2000.

[4] Daniel Huber and Martial Hebert. "3-D Modeling Using a Statistical Sensor Model and Stochastic Search", In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 858-865, 2003

[5] Onay Urfalioglu, Patrick Mikulastik, Ivo Stegmann, "Scale Invariant Robust Registration of 3D-Point Data and a Triangle Mesh by Global Optimization", Proceedings of the 8th International Conference on Advanced Concepts for Intelligent Vision Systems (ACIVS 2006), LNCS Vol. 127, pp. 1059-1070 Antwerp, 2006

[6] P. Allen, S. Feiner, A. Troccoli, H. Benko, E. Ishak, B. Smith, "Seeing into the Past: Creating a 3D Modeling Pipeline for Archaeological Visualization", Second International Symposium on 3D Data Processing, Visualization and Transmission 3DPVT04, pp. 751-758, 2004.

[7] A. Akbarzadeh, J.-M. Frahm, P. Mordohai, B. Clipp, C. Engels, D. Gallup, P. Merrell, M. Phelps, S. Sinha, B. Talton, L. Wang, Q. Yang, H. Stewenius, R. Yang, G. Welch, H. Towles, D. Nister and M. Pollefeys, "Towards Urban 3D Reconstruction From Video", Proc. 3DPVT06, 2006.

[8] Yushuang Liu, Heidrich, W. "Interactive 3D model acquisition and registration", Computer Graphics and Applications, 2003. Proceedings. 11th Pacific Conference, pp 115-122, 2003

[9] T. Jaeggli, T.P. Koninckx, L. Van Gool, "Online 3D Acquisition and Model Integration", IEEE International Workshop on Projector-Camera Systems - ICCV03, cdrom proc, 2003, France.

[10] V. Popescu, E. Sacks, and G. Bahmutov, "The Model Camera: a Hand-Held Device for Interactive Modeling," Proceedings of 3DIM03, pp. 285-292, 2003.

[11] S. Rusinkiewicz, O. Hall-Holt, and M. Levoy. "Real-time 3d model acquisition", In Proceedings of Siggraph, Vol. 21, No. 3, pp. 438-446, 2002.

[12] Bruce D. Lucas and Takeo Kanade. "An Iterative Image Registration Technique with an Application to Stereo Vision", International Joint Conference on Artificial Intelligence, pp. 674-679, 1981.

[13] Carlo Tomasi and Takeo Kanade. "Detection and Tracking

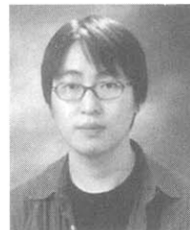
of Point Features", Carnegie Mellon University Technical Report CMU-CS-91-132, April 1991.

[14] R. Benjema and F. Schmitt, "Fast global registration of 3D sampled surface using a multi-z-buffer technique", First International Conference on Recent Advances in 3-D Digital Imaging and Modeling, pp. 113-123, 1997.

[15] R. Bergevin, M. Soucy, H. Gagnon, and D. Laurendeau, "Toward a general multi-view registration technique", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 18, No. 5, pp. 540-547, 1996.

[16] Soon-Yong Park and Murali Subbarao, "An Accurate and Fast Point-to-Plane Registration Technique", Pattern Recognition Letter, Vol. 24, No. 16, pp. 2967-2976, 2003.

[17] William E. Lorensen, Harvey E. Cline, "Marching cubes: A high resolution 3D surface construction algorithm", ACM SIGGRAPH Computer Graphics, Vol. 21, No. 4, pp.163-169, 1987.



백 재 원

e-mail : depia@vision.knu.ac.kr

2006년 동서대학교 멀티미디어공학과 (학사)

2006년~현재 경북대학교 컴퓨터공학과 석사과정

관심분야: 3차원 컴퓨터비전, 컴퓨터그래픽스, 가상현실



문 재 경

e-mail : jkmoon@ee.knu.ac.kr

1991년 경북대학교 전자공학과(학사)

1993년 경북대학교 전자공학과(공학석사)

1999년 경북대학교 전자공학과(공학박사)

1993년~2001년 한국전자통신연구원 선임연구원

2000년~2004년 뉴욕주립대 스토니브룩 방문연구원

2005년~2006년 경북대학교 센서기술연구소 연구원

2006년~현재 경북대학교 전기전자컴퓨터학부 BK계약교수

관심분야: Image Processing, Depth-from-Defocus



박 순 용

e-mail : sypark@knu.ac.kr

1991년 경북대학교 전자공학과(학사)

1993년 경북대학교 전자공학과(공학석사)

2003년 뉴욕주립대 스토니브룩(공학박사)

1993년~1999년 한국원자력연구소

선임연구원

2003년~2004년 뉴욕주립대 스토니브룩 박사후연구원

2004년~2005년 한국전자통신연구원 디지털컨텐츠연구원

선임연구원

2005년~현재 경북대학교 컴퓨터공학과 조교수

관심분야: 3D Computer Vision, Image Processing, Digital

Video Effect, Stereo Vision