

광학식 동작 포착에서 동적 모델을 이용한 신뢰성있는 3-D 좌표 추정 및 추적

(Robust 3-D Points Estimation and Tracking with Dynamic Model for Optical Motion Capture)

이 동 훈 * 추 창 우 ** 김 성 진 ** 정 순 기 ***
(Dong Hoon Lee)(Chang Woo Chu)(Sung Jin Kim)(Soon Ki Jung)

요 약 광학식 동작 포착(optical motion capture) 시스템에서 신뢰할만한 3차원 좌표의 획득과 마커의 궤적 추적은 매우 중요한 문제이다. 3차원 좌표를 획득하기 위해서는 다중의 카메라로부터 2차원의 데이터 간의 대응관계를 구해야 한다. 본 논문에서는 각 카메라에서의 2차원 마커들 간의 대응관계를 k partite graph로 모델링하고, 릴랙세이션 알고리즘을 사용하여 고스트가 제거된 신뢰성있는 클릭을 추출한다. 이를 통해 정확하고 안정적인 3차원의 좌표를 생성할 수 있다. 또한 추출된 3차원 마커의 궤적의 추적을 위해 칼만 필터를 사용한 마커의 예측과 데이터 연계 문제의 해결을 위한 전략을 제안하고, 사라진 마커의 궤적을 유지시키기 위해 다이내믹 모델을 사용한 추적 알고리즘을 제시한다.

Abstract It is a very important issue to extract reliable 3-D marker data and track them robustly in optical motion capture system. In this paper, we propose effective ghost removal algorithms to extract reliable 3-D marker data. We model relations of point matching using weighted k -partite graph and make a globally convergent iterative algorithm to extract reliable maximum cliques from the graph using probabilistic relaxation labeling. Also this paper proposes a tracking algorithm, which constructs a dynamic model of subject and uses the information to track 3D markers robustly. Experimental results show the effectiveness of the proposed methods.

1. 서 론

동작 포착(motion capture)이란 분석 또는 재생을 목적으로 인체 또는 다른 관체체의 움직임을 기록하는 기술이다[4]. 동작 포착은 기존의 키 프레임을 사용하여 동작을 생성하는 애니메이션과 비교하였을 때, 많은 수작업을 덜어주고, 실제 다관체체의 움직임을 그대로 포착하기 때문에 자연스러운 동작의 애니메이션이 가능하다는 이점을 가지고 있다. 동작 포착 시스템은 센서의 종류에 따라 보철식(prosthetic), 음파식(acoustic), 자기

식(magnetic), 광학식(optical) 동작 포착 시스템으로 나눌 수 있으며, 이 중 광학식 동작 포착 시스템은 연기가 반사 마커를 몸에 부착하고, 여러 대의 카메라로 움직임을 포착하여, 3차원의 데이터를 추출한다. 많은 상업화된 동작 포착 시스템이 개발되었고, 컴퓨터 캐릭터 애니메이션에 좋은 성능을 보이고 있지만, 대부분의 시스템에서 많은 시간을 소모하는 후처리 단계를 거쳐야만 한다. 이는 정확하고 안정적인 3차원 좌표의 획득과 획득된 3차원 좌표의 신뢰성있는 궤적을 추적하기가 어렵다는 점에 기인한다.

3차원 좌표는 둘 이상의 다중 카메라에서 포착한 2차원 데이터들 사이의 다중 시각 대응관계(multiview correspondence)로부터 얻을 수 있다. 본 논문에서는 다중 시각 대응관계를 획득하기 위해 에피폴라 제약조건(epipolar constraints)[8]을 사용하였고, 이를 적용하였을 때, 잘못된 대응관계인 고스트가 발생할 수 있다(그림 1 참조). 또한 카메라로부터 은닉된 점들로 인해 모든 카메라에서 대응되는 점이 존재하는지의 여부를

* 학생회원 : 경북대학교 컴퓨터공학과
dhlee@vr.knu.ac.kr

** 비 회 원 : 경북대학교 컴퓨터공학과
thrash@vr.knu.ac.kr
saintjin@vr.knu.ac.kr

*** 종 신 회 원 : 경북대학교 컴퓨터공학과 교수
skjung@knu.ac.kr

논문접수 : 2000년 3월 22일

심사완료 : 2000년 9월 16일

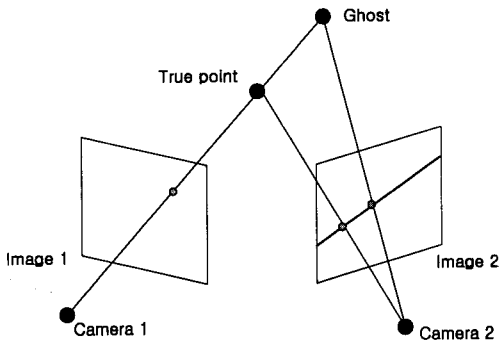


그림 1 다중 시각 대응관계에서의 고스트의 발생 예

보장할 수 없다. 이러한 복잡한 환경에서 다중의 카메라로부터 신뢰할 만한 2차원의 대응관계를 찾는 것은 어려운 문제이다. 본 논문에서는 마커(marker)들 간의 관계들을 k -partite 그래프로 보고, 그래프의 관점에서 릴렉세이션 레이블링 알고리즘을 이용하여 고스트를 제거하는 효과적인 방법을 기술한다.

시간의 흐름에 따른 3차원 좌표의 추적을 위해서 가장 간단하게는 현재의 프레임에서 이전 프레임의 특징좌표의 위치와 가장 가까운 곳의 특징점을 추적하는 방법이 있다(nearest neighborhood method)[2]. 그러나, 이 방법은 포착 대상의 움직임이 많을 때에는 적당하지 않으며, 포착된 동작에 센서 잡음(noise)이 포함되므로, 잡음이 존재하는 시스템에서 안정적으로 추적할 수 있는 방법이 필요하다. 따라서, 본 논문에서는 잡음이 존재하는 물리적 시스템의 상태를 예측, 추정, 추적, 제어하는데 널리 사용되는 칼만 필터(Kalman filter)[4,7,9]를 이용한다. 칼만 필터는 실시간에서의 동작의 예측치와 측정치의 비교를 통해 추적을 제어한다. 그러나, 실제 측정치가 없는 경우 즉, 예측값에 대응되는 특징점이 없는 경우에는 현재 상태의 올바른 갱신이 이루어지지 않는다. 본 논문에서는 포착 대상의 모델을 자동으로 생성, 유지함으로써 사라진 특징점의 측정치를 추정하여, 단시간 동안 사라진 특징점에 대해서도 궤적이 끊기지 않게 하는 방법을 제안한다. 또한, 급격하게 이동 방향을 바꾸는 마커의 궤적을 유지하는 방법에 대해서 기술한다.

본 논문의 구성은 다음과 같다. 2장에서는 k -partite 그래프를 사용한 문제의 모델링과 고스트 제거를 위한 릴렉세이션 알고리즘을 기술하고, 3장에서는 2장에서 제시한 알고리즘의 가속화를 위하여 시간에 따른 2차원 추적 정보를 이용한 복합적인 접근 방법을 기술한다. 4

장에서는 칼만 필터링을 사용한 전반적인 궤적 추적 과정과 이동 방향을 급격하게 바꾸는 특징점의 추적에 대해서 설명한다. 5장에서는 모델을 구성하는 방법과 모델을 이용하여 사라진 특징점의 위치 추정에 대해서 설명한다. 6장에서는 제시한 방법을 사용하여 실험한 결과를 보인다. 끝으로 7장에서 결론을 맺는다.

2. 릴렉세이션에 의한 3차원 좌표 추출

3차원 좌표의 추출은 다중의 카메라에서 포착된 2차원 마커간의 대응관계를 통해 얻어진다. 본 장에서는 신뢰할만한 마커간의 대응관계를 구하기 위해 다중 카메라의 관계를 k -partite 그래프로 모델링하고, 확률적 릴렉세이션 알고리즘을 이용하여 고스트를 제거하는 방법을 기술한다.

2.1 k -partite 그래프를 이용한 모델링

그림 1과 같이 임의의 카메라 1의 특정한 점에 대해 다른 카메라 2에서 대응되는 점은 영상 2의 평면 상에 존재하는 직선상의 한 점으로 나타나게 되며, 이 직선을 에피폴라 직선(epipolar line)이라고 한다[8]. 이 때 대응되는 두 점은 그래프의 관점에서 연결 관계가 있는 에지로 볼 수 있다. k -partite 그래프는 같은 집합 내에서 어떠한 노드도 연결되어 있지 않은 k 개의 비접합(disjoint) 집합으로 이루어진 그래프를 의미한다[3]. 우리의 문제는 k -partite 그래프로 모델링 될 수 있다. 왜냐하면, 동일한 카메라 내의 마커들 사이에는 어떠한 연결 관계도 없기 때문이다. 각각의 노드들은 임의의 카메라 s 로부터 획득한 마커에 해당되며, 각각의 카메라는 k -partite 그래프에서의 비접합 집합들을 의미한다. 따라서 카메라의 총 수 S 는 비접합 집합의 수 k 이다.

2.2 릴렉세이션 레이블링

카메라로부터 측정된 2차원 마커 정보는 카메라의 잡음에 의해 에피폴라 직선에 정확히 매칭되지 않는다. 이를 수용하기 위해서 에피폴라 적용시 임의의 임계치를 설정하며, 이로 인해 실제 대응관계로 추출되어야 할 마커의 집합을 확보할 수는 있으나, 동시에 다양한 대응관계들로 인해 많은 고스트가 발생한다. 임의의 카메라에 대하여 에피폴라 직선에 가장 근접한 점을 가장 신뢰할만한 대응관계에 있다고 확신할 수 없으므로, 전역적인 탐색을 통해 마커들 간의 관계가 일관성을 유지할 수 있는 전략이 필요하다. 전역적으로 만족해야 하는 일관성은 임의의 마커에 대해 대응관계를 가지는 관계는 유일하게 하나만 존재해야 한다는 것이다.

릴렉세이션(relaxation) 알고리즘은 각각의 객체들이 가질 수 있는 특성을 레이블들로 부과하여 국부적인 관

계의 제약을 전역적으로 전파함으로써 전체적으로 일관성을 지닌 레이블을 추출하는 알고리즘이다. 이러한 문제는 그래프, 오토마타, 그래프 컬러링, 영상 이해 등 여러 분야에서 사용되어져 왔다[1]. 모든 레이블링 문제는 객체들(objects)의 집합, 객체들 간의 관계(relations)에 대한 유한 집합, 레이블들(labels)의 유한 집합, 제약 조건들(constraints)로 표현된다. 본 논문에서는 레이블링 문제를 k -partite 그래프에서의 최대 클릭(maximum clique)을 구하는 문제로 다음과 같이 모델링한다.

(a) 객체들(objects)의 집합 : 존재하는 모든 객체들은 모든 카메라에서의 마커들의 집합이다. 즉, 하나의 객체는 임의의 카메라 s 로부터 획득한 임의의 마커를 의미한다.

(b) 객체들 간의 관계(relations)에 대한 유한 집합 : 하나의 객체에 대해 인접한 객체들의 배치는 에피폴라 제약조건을 만족하는 마커들 간의 관계로 표현된다.

(c) 레이블들(labels)의 유한 집합, 또는 의미를 지닌 신발들의 집합 : 레이블은 임의의 카메라의 특징점에 대해 에피폴라 제약조건[8]을 만족하는 특징점의 집합인 튜플, 즉 최대 클릭의 집합이다. 하나의 마커에서 각 카메라에 에피폴라 제약조건을 만족하는 마커들로부터 클릭을 형성하고 그 중 최대 클릭들의 집합이 그 마커에 대한 레이블의 후보들이 된다.

(d) 제약 조건들(constraints) : 모델링을 위한 제약 조건들은 존재하지 않는다. 왜냐하면, 각각의 마커는 반드시 하나의 레이블로 수렴한다는 조건 자체가 릴랙세이션 구조에 내포한 성질이기 때문에, 특별한 제약 조건들을 두지 않고도 문제는 해결 가능하다.

(e) 연결 강도의 유한 집합 : 레이블을 이루는 클릭은 에지의 집합으로 구성되어 있고, 에지는 앞에서 언급한 바와 같이 에피폴라 제약조건을 만족하는 관계성을 지니고 있다. 연결강도는 에피폴라 직선과 떨어진 거리인 에러를 역사상한 값으로 모델링된다.

가. 초기 가중치의 부과

확률적 릴랙세이션(probabilistic relaxation)에서 하나의 객체에 대한 가중치의 부과는 0과 1사이의 값이며, 그 합은 1이다. 따라서, 초기 가중치는 다음과 같이 부과된다.

$$\forall i \sum_{k=1}^K P(\theta_i = w_k) = 1, \quad \forall i \forall k P(\theta_i = w_k) = \frac{1}{R_i}. \quad (1)$$

여기에서 θ_i 는 마커 m_i 에 레이블링 되어질 하나의 레이블을 의미하고, w_k 는 레이블의 후보이다.

는 w_k 가 마커 m_i 에 레이블링 될 확률을 의미하고, R_i 는 객체 θ_i 의 후보 레이블의 갯수를 나타낸다.

나. 레이블 갱신 전략

s 번째 반복에서의 임의의 마커에 부과된 임의의 레이블 θ_i 와 인접한 임의의 마커에 후보로 존재하는 모든 레이블 w_k 사이의 관계를 통한 신뢰도 q_j^i 는 θ_i 의 후보 레이블 w_k 와 동일한 레이블을 가진 인접 마커의 후보 레이블의 확률을 취한다.

$$q_j^i(\theta_i = w_k) = P^s(\theta_j = w_k) \quad (2)$$

레이블링을 위한 식 (2)에 내포된 또 다른 성질은 클릭의 에지의 개수가 크면 클수록 가중치가 커진다는 사실이다. 이것은 레이블이 하나의 에지로 구성된 것이 아니라, 에지의 집합 즉, 클릭이 하나의 레이블을 이루므로 이웃 객체에 대한 레이블 탐색 시, 튜플의 수만큼 방문하여 가중치가 전달되기 때문이다.

하나의 객체의 임의의 레이블에 대해 인접한 모든 객체들의 후보 레이블들과의 관계를 통해 w_k 가 레이블이 될 확률은 다음과 같다.

$$Q^s(\theta_i = w_k) = \sum_{j=1}^N c_{ij} q_j^i(\theta_j = w_k) \quad (3)$$

여기에서 c_{ij} 는 객체와 객체간의 연결강도로 에피폴라 직선과 특징점과의 거리가 작을수록 실제 마커일 가능성이 크다는 성질을 이용하여 식 (4)와 같이 에러값 d 를 0과 1사이의 값으로 역사상하여 모델링한다. 이를 통해 에피폴라 직선에 근접한 마커일수록 더 큰 가중치를 갖는다.

$$c_{ij} = 1 - \frac{d}{threshold}. \quad (4)$$

반복 횟수 s 에서 임의의 마커가 w_k 를 레이블로 가질 확률은 인접한 이웃들과의 관계를 통해 결정된 확률 $Q^s(\theta_i = w_k)$ 와 현재 단계 s 에서 결정된 후보 레이블 w_k 가 레이블이 될 확률 $P^s(\theta_i = w_k)$ 를 통해 다음 단계 $s+1$ 에서의 확률이 결정된다.

$$P^{s+1}(\theta_i = w_k) = \frac{1}{K} P^s(\theta_i = w_k) Q^s(\theta_i = w_k), \quad (5)$$

여기에서 K 는 정규화 상수이며 다음 식과 같이 정의된다.

$$K = \sum_{k=1}^N P^s(\theta_i = w_k) Q^s(\theta_i = w_k). \quad (6)$$

3. 3차원 추적을 이용한 릴랙세이션 알고리즘의 개선

릴렉세이션을 통한 모델링은 하나의 장면에 대해 올바른 마커를 판별하는데 유용하게 쓰일 수 있다. 또한 연속된 프레임에 있어서 시간의 간격은 매우 적기 때문에, 2차원의 마커들 간의 관계 또한 프레임 간 큰 변화 없이 유사한 구조를 유지할 것이다. 본 논문에서는 알고리즘의 속도를 가속화하는 방안으로 3차원 위치정보의 예측값을 2차원으로 역투영하여 추적함으로써 이전 프레임과 동일한 연결관계를 지닌 구조에 대해 정보를 유지하는 방법을 사용하였다.

먼저 첫 번째 프레임에 대하여 릴렉세이션 알고리즘을 수행한다. 릴렉세이션 알고리즘의 수행 결과로 구해진 대응관계를 가진 2 차원 위치정보를 이용하여 3 차원 위치정보를 구할 수 있다. i 번째 프레임에서 임의의 3 차원 위치정보 P_a^i 는 k 대의 카메라에서 대응관계를 이루는 2 차원 위치정보의 집합으로 표현할 수 있다.

$$P_a^i = \{m^1, m^2, \dots, m^k\}. \tag{7}$$

2 차원 추적을 위한 기술로 4장에서 기술할 칼만 필터링(Kalman filtering)을 사용하여 첫 번째 프레임에 대한 3 차원 추적의 예측값을 구하고, 예측값을 2 차원 영상으로 역투영한다. 그림 2와 같이 $i+1$ 번째 프레임에서, 이전 프레임에서의 3 차원 위치정보 P_a^i 가 예측(prediction)된 3 차원 위치정보를 Q_a^{i+1} 라 두었을 때, 이를 영상면으로 역투영시켜 2 차원 위치정보의 예측값을 구할 수 있다.

$$projection(Q_a^{i+1}) = \{m'_1, m'_2, \dots, m'_k\}. \tag{8}$$

이를 통해 이전 프레임의 확정된 클릭의 추적이 가능하다. 만약 확정된 클릭과 동일한 정보, 즉 확정된 클릭의 다음 프레임의 예측값에 마커가 모두 존재할 경우,

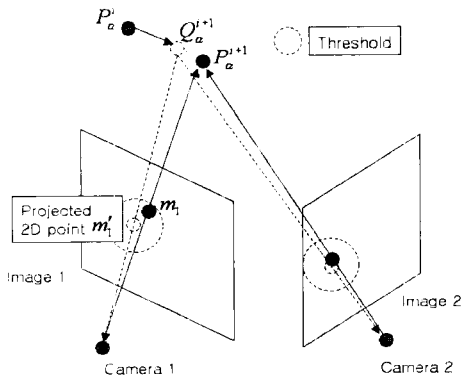


그림 2 시간에 따른 정보를 이용한 신뢰성있는 3 차원 위치정보의 복구

포착된 마커를 현재 프레임에서 3 차원으로 복원될 2 차원 위치정보로 사용한다. 예측값의 위치에 측정값이 존재하지 않는 경우는, 2장에서 기술한 방법에 따라 에피폴라 제약조건을 수행하고 클릭을 추출하여 릴렉세이션 알고리즘을 수행하여 3 차원 위치정보를 추출한다.

만약 프레임 간의 위치 변화가 크다면, 이전 프레임의 2 차원 정보의 관계성이 유지될 수 없고, 매 프레임마다 마커 간의 관계를 새로이 구해야만 한다. 그러나, 실제로 프레임 간의 2 차원 정보의 관계성 그래프에서 큰 변화는 나타나지 않았다. 따라서 극소의 변화된 클릭의 구조에 대해서만 새로이 릴렉세이션을 수행하고, 대부분은 이전 프레임에서 구한 클릭을 그대로 유지함으로 3 차원 좌표 추출을 위한 시간을 단축하였다.

4. 마커의 궤적 추적

마커의 궤적 추적은 이전 프레임의 마커들이 현재 프레임에서는 어떤 마커인지를 결정하는 과정이다. 본 장에서는 마커의 궤적을 추적하기 위해 칼만 필터를 모델링한 방법과 칼만 필터를 통해 예측된 마커의 위치와 측정된 마커간의 연관 관계를 짓는 방법에 대해 기술한다.

4.1 칼만 필터(Kalman Filter) 모델링

포착된 동작은 시간에 따라 인체 운동의 역학적 모델(dynamic model)을 변화시키는 시스템으로 볼 수 있다. 이러한 역학적 시스템의 행태는 그 시스템의 성질을 나타내는 일군의 상태 변수(state variable)로 표현할 수 있다. 상태 변수를 모아 하나의 벡터로 만든 상태 벡터(state vector)의 시간에 따른 변화는 대상 시스템의 고유한 성질을 반영하는 선형 또는 비선형의 수식으로 모델링되는데, 이를 역학적 모델이라 한다. 외부에서 직접 측정 또는 제어 가능하며 상태 벡터로부터 계산할 수 있는 성질을 시스템의 관측값이라 하고, 상태 벡터와 관측값의 관계를 표현한 것을 시스템의 관측 모델(measurement model)이라 한다.

각 마커의 위치 이동을 나타내는 상태 벡터는 다음 식 (9)과 같이 표현한다.

$$s = [p^T v^T]^T. \tag{9}$$

즉, 각 마커의 위치 이동을 등속 운동으로 가정할때, 상태 벡터는 위치(p)와 속도(v)로 나타낼 수 있다. 여기서 등속 운동은 데이터가 처리되는 시간 t 의 부근에서 국부적으로 등속도를 유지함을 의미한다. 이후의 상태 벡터의 갱신 과정을 거친 후의 결과는 등속 모델에 의한 예측치와 입력으로 사용되는 실제 데이터와 합쳐진 것으로, 입력 데이터가 전체적으로 등속 운동으로 정합(fitting)된 것과는 다르다.

시간 t 에서의 상태 벡터를 s_t 라 하고 상태 전이 함수를 f 라 할 때, 역학적 모델은 식 (10)과 같이 나타낼수 있다.

$$s_t = f(s_{t-1}) + n_{t-1}, \quad (10)$$

이때 n_{t-1} 은 상태 전이 과정의 잡음으로 정규 확률 분포(normal probability distribution)를 가지는 백색 잡음으로 모델링한다. 등속 운동 모델에서 상태 전이 행렬 F 는 식(11)와 같이 간단하게 표시할 수 있다.

$$F = \begin{pmatrix} I_3 & I_3 \\ 0_3 & I_3 \end{pmatrix} \quad (11)$$

여기에서 I_3 은 3×3 단위 행렬을, 0_3 은 3×3 영행렬을 의미한다.

또한 관측 벡터를 z_t , 관측 함수를 h 라 할 때 관측모델은 식(12)과 같다.

$$z_t = h(s_t) + v_t, \quad (12)$$

이때 v_t 는 관측에 포함된 랜덤 노이즈이다. 관측값의 예측은 관측 행렬 $[I_3 \ 0_3]$ 를 이용해서 구하고, 포착된 마커 데이터를 실제 관측값으로 사용한다.

이렇게 모델링된 시스템은 그림 3과 같이 시간적 갱신(time update)와 측정치 갱신(measurement update)의 두 과정을 반복한다[7].

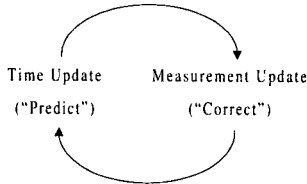


그림 3 칼만 필터 순환(kalman filter cycle)[7]

시간적 갱신에서는 식(10)과 식(13)을 이용하여, 현재 상태의 예측치와 선형적(priori) 오차 공분산을 구한다.

$$P_{t|t-1} = F_{t-1} P_{t-1} F_{t-1}^T + Q_{t-1}. \quad (13)$$

측정치 갱신에서는 오차공분산, 관측 행렬 H_t , 관측 모델 잡음의 공분산 R_t 로 Kalman gain K_t 를 식 (14)에 의해 구한다. 식 (15)에서는 예측된 관측치 \hat{z} 과 실제 관측치 z 의 차이(measurement residual)에 Kalman gain을 가중치로 두어 예측된 상태 벡터에 더함으로써 현재 상태 벡터를 갱신한다. 식 (16)에서는 posteriori 오차 공분산 P_t 를 갱신한다.

$$K_t = P_{t|t-1} H_t^T (H_t P_{t|t-1} H_t^T + R_t)^{-1} \quad (14)$$

$$s_{t|t} = \hat{s}_{t|t-1} + K_t (z_t - H_t \hat{s}_{t|t-1}) \quad (15)$$

$$P_t = (I - K_t H_t) P_{t|t-1} \quad (16)$$

위 과정을 통해 상태 벡터는 대상 시스템의 현재 상태의 근사치를 계속 유지하게 된다.

4.2 마커의 궤적 추적

마커의 궤적 추적은 다음의 3단계로 이루어진다.

- ① 현재 프레임에서의 마커의 위치 예측,
- ② 예측값(prediction)과 측정값(measurement) 사이의 대응관계 설정,
- ③ 마커의 상태 및 공분산 조정.

첫 번째와 세 번째 단계는 각각 칼만 필터 순환(kalman filter cycle)의 시간적 갱신(time update), 측정치 갱신(measurement update)[7]에 해당한다. 궤적 추적의 핵심이 되는 두 번째 단계는 3차원 마커의 예측치로부터 실제 측정치들과의 대응 관계를 찾는 것으로 전통적인 통계적 데이터 연계 기법[2]들이 적용 가능하다. 하지만 본 논문에서는 카메라로부터 얻은 측정치의 높은 신뢰성을 바탕으로 이분할 그래프(bipartite graph)에서 최대흐름(maximum flow)을 구하는 문제로 모델링한다.

마커들의 예측된 위치를 갖는 집합 P와 측정 위치를 갖는 집합 M이 있을 때, 예측값에 가장 가까운 측정값을 사상(mapping)시키는 것이므로 두 집합은 비집합(disjoint)관계를 가지게 된다. 두 집합의 원소간 거리가 최대 속도로 표현되는 임계치(threshold) 범위를 만족할 때 두 원소는 대응관계를 가진다. 이때, 그림 4와 같이 애매모호성(ambiguity)이 발생한다. 이를 그래프로

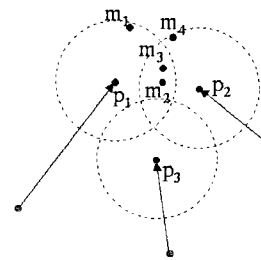


그림 4 예측치와 측정치의 대응관계

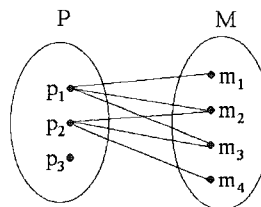


그림 5 대응 그래프

나타내면 그림 5와 같고, 이 그래프에서 최대수의 대응되는 쌍을 찾는 것을 목표로 한다.

전통적인 이분할 그래프(bipartite graph)에서 최대흐름(maximum flow)을 찾는 문제에서는 모든 에지(edge)에 같은 가중치(weight)를 주어서 해결하는데, 본 논문에서는 예측치와 측정치의 거리 차이가 작은 쌍에 더 비중을 두어서, 최대한 많은 쌍의 대응관계를 찾되, 선택된 연결은 선택 가능한 연결 중 가능한 가장 짧은 거리를 가진 연결을 선택하도록 한다. 이러한 조건을 만족시키기 위해서 본 논문에서는 선택 가능한 모든 에지들을 거리값으로 정렬시켜서, 정렬된 순서대로 P와 M의 원소 중 어느 하나도 중복이 되지 않도록 에지를 선택한다. 그러나, 그림 5의 p_3 와 같이 대응되는 측정치가 없는 마커가 생긴다. 이유는 두가지로 설명된다. 첫번째는 카메라의 측정 오차 혹은 모든 카메라에 대해 마커가 가려져서 실제 측정값이 현재 프레임에서는 사라지는 경우이고, 두번째는 각 마커의 상태를 등속도 운동으로 모델링한 역학적 모델이 추적에 실패하여, 현재 프레임의 탐색 범위 밖에 측정치가 있는 경우이다. 이러한 경우에 궤적을 유지시키기 위해서 5장에서 설명할 모델을 이용한 이분할 그래프(bipartite graph)의 최대흐름(maximum flow) 문제로의 재모델링 및 마커의 위치 예측을 통해 해결할 수 있다.

5. 동적 모델을 이용한 마커의 위치 예측

칼만 필터 순환의 "측정치 갱신" 과정에서는 예측값과 측정값을 이용해 현재 상태와 공분산을 조정한다. 그러나, 이 단계에서 측정값이 없으면, 현재 상태의 조정이 이루어지지 않게 된다. 이렇게 마커의 상태가 조정되지 않고 그대로 유지된다면, 마커는 시간의 흐름에 따라 계속 같은 방향으로 움직이는 것으로 예측될 것이므로, 몇 프레임동안 위치 정보를 가지지 못한 마커는 원래의 궤적에 다시 포함시킬 수 없게 된다. 본 장에서는 이러한 문제점을 해결하기 위해, 캡처 대상의 모델을 구성하고 연결된 마커의 정보를 이용하여, 사라진 마커의 위치를 예측하는 방법에 대해 설명한다.

5.1 동적 모델의 생성 및 갱신

모델을 생성하기 위해 장면의 초기 자세를 그림 6(a)와 같이 양팔을 벌린 자세로 가정한다. 첫 번째 프레임에서 마커들의 3차원 좌표를 이용하여 그림 6(b)와 같이 초기의 모델을 구성한다. 그러나, 이렇게 구성된 모델은 데이터의 잡음과 캡처 대상마다 각 세그먼트의 길이가 다르기 때문에, 정확하게 고정(rigid) 관계를 가지는 연결만을 가지지는 않는다. 따라서, 대상체의 움직

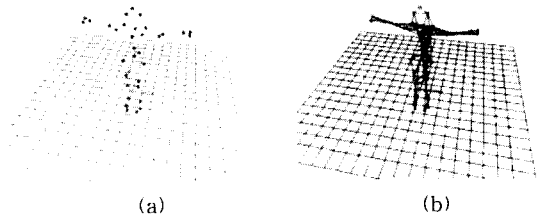


그림 6 (a)초기 자세 (b)초기 모델

임에 따라 바뀌는 마커간 거리를 이용하여 고정(rigid) 세그먼트만을 가질 수 있도록 각 프레임마다 모델 정보를 갱신한다.

모델의 정보로는 마커간 거리, 평균, 분산을 이용한다. 궤적이 정해진 현재 프레임의 마커들은 이전 프레임까지의 모델 정보에 현재 프레임에서의 거리를 포함하여 정보를 갱신한다. 마커가 사라진 경우의 모델 정보는 주위의 마커의 상관 관계를 이용하여 예측된 마커의 위치를 모델 정보로 활용한다. 이에 대해서는 5.2절에서 상세히 기술할 것이다.

현재 프레임의 임의의 두 마커가 이전 n-1 프레임까지 연결을 유지하여 왔을때, n번째 프레임에도 연결을 유지할 수 있는지 가능하기 위해서 식(17)과 식(18)을 이용하여 거리의 평균과 분산을 갱신한다.

$$E_n = \{E_{n-1} \times (n-1) + x_n\} / n \tag{17}$$

$$V_n = \frac{\sum_{i=1}^n x_i^2 / n - E_n^2}{n-1} = \left(\frac{\sum_{i=1}^{n-1} x_i^2 + x_n^2}{n} \right) / n - E_n^2 \tag{18}$$

(단, x_n : 현재 프레임의 두 마커간 거리,

E_n : n 번째 프레임까지의 평균, V_n : 분산)

이때, 거리와 평균의 차가 $|x_n - E_n| > threshold_1$ 이거나 분산이 $V_n > threshold_2$ 이면, 이 연결은 연결성이 없으므로, 연결을 끊어준다. 그림 7은 연결이 끊어지는 예를 나타낸다. 그림 7에서 마커 m_1, m_2, m_3 가 다음 프레임에서 각각 m_1', m_2', m_3' 으로 이동할 때, $\overline{m_2 m_3}$ 에 비해 $\overline{m_2' m_3'}$ 가 현저하게 커졌으므로, 연결이 끊어지게 된다.

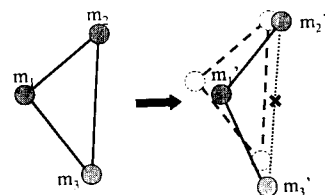


그림 7 마커간의 연결이 끊어지는 경우

이렇게 함으로써 시간에 따른 각 마커들의 거리의 고정성(rigidity)을 유지하는 세그먼트만을 갖는 동적 모델을 자동적으로 생성, 유지 가능하다.

5.2 사라진 마커의 위치 예측

4장에서 언급한 예측값과 측정값의 대응관계를 찾지 못하는 두 가지 경우를 모델을 이용하여 단계적으로 해결할 수 있다. 임계치를 이용한 국부적인 탐색을 통한 대응 관계 설정에 실패한 예측값과 측정값들을 전역적으로 탐색하여 이분할 그래프(bipartite graph)를 재구성한다. 이 그래프에서 최대 흐름(maximum flow)은 이전 프레임의 고정(rigid) 세그먼트 조건에 얼마나 잘 부합하는지에 가중치를 두어 구한다. 이 과정을 통해서도 대응 관계를 찾지 못한 예측치는 실제값이 없는 경우이므로, 다음의 방법으로 마커의 위치를 추정한다.

구성된 모델을 이용하여 사라진 마커의 위치를 추정한다. 그림 8처럼 이전 프레임까지 연결이 유지되어온 마커 m, m_0, m_1 중 m_0, m_1 은 레져 추적 결과 현재 프레임에서 측정값으로 p_0, p_1 을 가지고, m 은 예측값 p' 만을 가진다고 가정한다. 마커 m 과 연결된 m_0, m_1 이 각각 거리가 d_0, d_1 이므로, 현재 프레임에서도 같은 거리 상에 측정값 p 가 존재해야 한다. 이 조건을 만족시키기 위해 p_0 를 중심으로 하고 반지름이 d_0 인 구와 p_1 을 중심으로 반지름이 d_1 인 구의 방정식을 만든다.

두 개의 구의 방정식으로부터 두 개의 구가 만나는 원을 포함하는 평면 방정식을 구할 수 있다. 예측된 위치 p' 에서 부터 p_2 를 중심으로 하는 원상의 점들 중 가장 가까운 점 p 를 구하여 m 의 측정값으로 한다.

각 프레임마다 마커간의 연결 정보를 갱신하지만, 임의의 프레임에서 임의의 마커에 연결되어 있는 마커들이 모두 고정(rigid) 관계를 나타내는 것은 아니다. 따라서, 위와 같은 방법을 적용하기 위해서는 현재 프레임에서 사라진 마커와 이전 프레임에서 연결된 여러개의 마

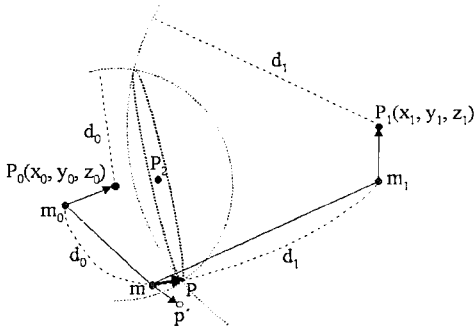


그림 8 모델을 이용한 마커 위치 추정

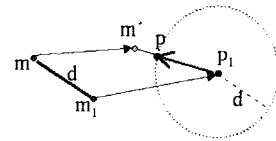


그림 9 연결된 마커가 한 개인 경우

커 중 어떤 마커를 선택하는지가 문제가 된다. 본 논문에서는 신뢰성있는 마커의 선택을 위해 가장 오랫동안 연결이 유지되었고, 마커간 거리의 분산이 가장 적은 최대 두 개의 마커를 선택하도록 하였다. 포착 대상의 움직임이 많을수록 서로 고정되지 않은(non rigid) 관계를 가진 마커들은 빨리 연결성을 잃을 것이다. 만약, 선택된 마커가 한 개인 때는 그림 9와 같이, 알고 있는 측정치 p_1 을 중심으로 이전 프레임에서 두 마커 m_1, m_2 의 거리 d 를 반지름으로 하는 구와 p_1 에서 m_2 의 예측값 m_2' 로 향하는 벡터가 만나는 점 p_2 를 측정값으로 선택하였다.

6. 실험 결과

6.1 릴렉세이션(Relaxation)을 이용한 3차원 좌표 추출

제한한 알고리즘에 대한 실험으로 가상의 시뮬레이션 실험 환경을 조성하였다. 먼저 그림 10과 같이 제한된 공간(497.843 mm * 788.865 mm * 748.485 mm) 내에 33개의 3차원 데이터를 무작위로 추출하였다. 총 4대의 가상 카메라를 대칭적으로 위치시키고, 무작위로 구한 3차원 좌표를 각 카메라에 투영시켜서 2차원 마커정보를 얻는다. 이때 보다 현실감 있는 실험 환경을 위하여 $\sigma = 2$ 화소(pixel)인 가우시안(Gaussian) 잡음을 추가하였다. 또한 임의의 마커에 대해 평균 한대의 카메라의 2차원 데이터의 투영을 금지함으로써 실제환경에서의 은닉 효과를 주었다. 그림 10은 실험에서 사용된 3차원 마커의 예이다.

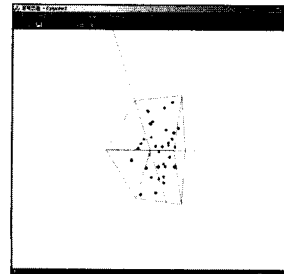
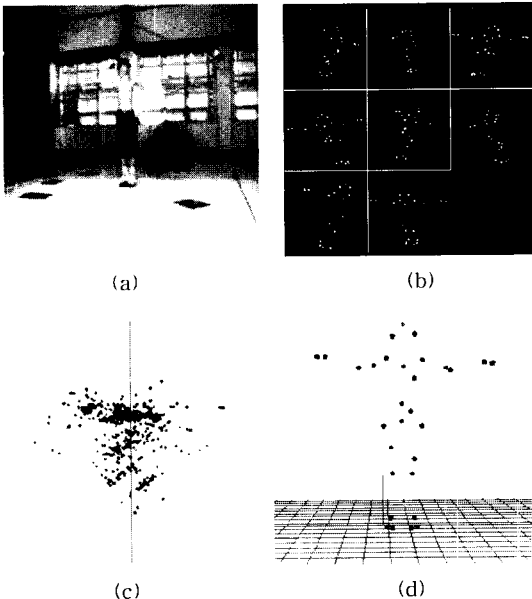


그림 10 임의의 범위 내에서 무작위로 추출한 3차원 정보

구현환경은 펜티엄II, 350Hz의 CPU와 64MB의 메모리를 가진 PC에서 Visual C++ 6.0을 사용하였다. 100개의 data set에 대한 실험 결과 약 99.2% 정확한 대응관계를 추출하였다. 실험에서 에러로 판명된 경우는 대부분 임의의 2차원 평면에 대해 은닉된 특징점의 위치에 다른 특징점이 임계값을 만족하여 클릭으로 생성된 경우이다. 프레임 당 속도는 0.22초로 개선이 요구되어진다.

실제 환경의 실험을 위해 Motion Analysis사의 상용 시스템인 Eva Motion Capture system[10]에서 동작 포착한 총 2633 프레임의 이차원 데이터를 사용하여 실험하였다. 동작 포착을 위해 사용된 카메라의 개수는 8대이며, 마커 수는 33개이다. 동일한 2차원 데이터에 대하여 Eva 시스템과의 비교 결과, 본 논문에서 제시한 알고리즘은 프레임당 0.33개의 고스트가 발생하였고, Eva 시스템은 프레임당 0.583개의 고스트가 발생하였다. 추출되어야 할 마커가 사라지는 현상의 경우, 본 알고리즘은 프레임 당 0.0175개의 은닉이 생겨났고, Eva 시스템은 0.1525개의 은닉이 발생하였다.



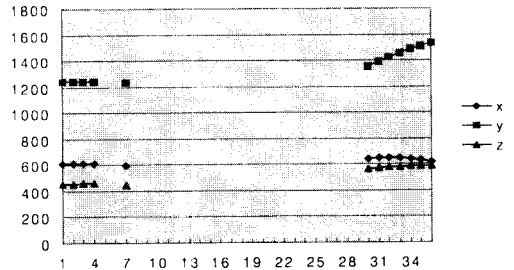
(a) 동작 포착 환경
 (b) 임의의 프레임에서 포착된 2차원 영상 데이터
 (c) 에피폴라 제약조건만을 사용하여 3차원 좌표 추출 결과
 (d) 제안한 알고리즘을 사용한 경우

그림 11 동작 포착 시스템에서 3차원 좌표 추출의 과정

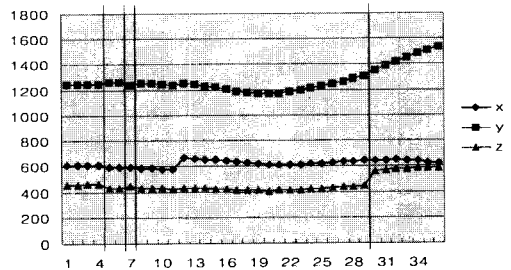
6.2 3차원 궤적 추적

본 실험의 실험 데이터로는 마찬가지로 Motion Analysis사[10]의 EVa를 통해 댄서의 움직임을 포착한 3차원 위치 데이터를 사용하였다.

그림 12는 단시간 동안 실제 마커가 사라진 왼쪽 이두근(biceps)에 부착한 마커를 추적한 결과를 타임라인으로 나타낸 것으로, 전체 2633개 프레임 중 265번째 프레임에서 300번째 프레임까지를 샘플링한 것이다. 그림 12(a)와 같이 실제 샘플링된 데이터의 5, 6, 8~29번째 프레임에서 실제마커가 없으며, 이 구간 내에서는 제안한 방법으로 유추한 마커의 위치는 그림 12(b)와 같은 타임라인으로 나타난다.



(a)



(b)

그림 12 왼쪽 이두근에 부착한 마커의 움직임

그림 13은 왼발 바깥쪽에 부착한 마커를 추적한 결과를 나타낸 것이다. 발을 뒤쪽으로 들다가 앞으로 내딛는 동작을 두 프레임 간격으로 샘플링하여 나타내었다. 이 경우 마커의 이동 방향이 급격하게 다른 방향으로 바뀌어서 EVa System에서 추적에 실패하였지만, 제시한 알고리즘을 사용하였을 경우 궤적이 유지됨을 알 수 있다.

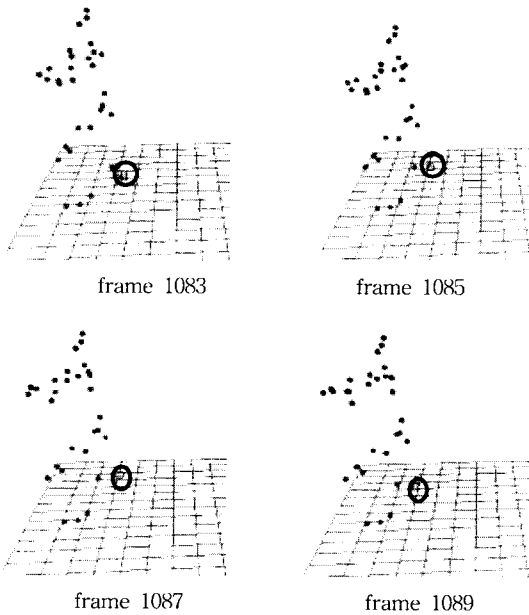


그림 13 왼발 바깥쪽에 부착한 마커의 움직임

실험에 사용한 EVa의 데이터는 33개의 실제 마커를 부착하여 포착한 동작이다. 전체 2633 프레임에 대해 184개의 궤적을 가졌으나, 본 논문에서 제시한 방법을 적용해서 궤적을 125개로 줄일 수 있었다.

7. 결 론

본 논문에서는 여러 대의 카메라를 통해 포착된 2차원 데이터들 간의 신뢰할 만한 대응관계를 구하기 위한 효율적인 고스트 제거 알고리즘을 제시하였다. 문제를 그래프의 관점에서 정의하여 전역적 탐색을 통해 확률적으로 수렴하도록 릴렉세이션 레이블링 알고리즘을 문제에 맞게 모델링하였고, 속도 문제의 개선을 위해, 3차원 추적 정보를 이용하여 복합적인 방안을 또한 제시하였다. 본 논문에서 제시한 알고리즘을 통한 실험을 통해 약 99.2 %의 신뢰할 만한 대응관계를 얻었다. 이를 통해 보다 정확하고 안정적인 3차원의 좌표가 생성될 것이다.

또한 각 마커의 상태를 등속도 운동으로 모델링한 칼만 필터를 이용하여 궤적 추적을 수행하였다. 궤적 추적시 핵심이 되는 예측치와 측정치의 대응 관계 설정은 이분할 그래프(bipartite graph)로 모델링하여 해결하였다. 포착 대상의 모델을 동적으로 구성하여 마커의 사라

짐, 급격한 움직임으로 인한 궤적의 끊김을 방지할 수 있었다. 이러한 결과를 통해 동작 포착 시스템의 후처리 수작업에 대해 상당량의 작업을 줄일 수 있다.

감사의 글

본 논문은 2000년 한국전자 통신연구원(ETRI)의 “PC 기반 광학식 동작 포착 시스템 및 소프트웨어 개발” 연구 과제의 지원에 의한 것이다

참 고 문 헌

- [1] Dana H. Ballard and Christopher M. Brown, *Computer Vision*, Prentice Hall, pp.408-430, 1982.
- [2] I. J. Cox, "A Review of Statistical Data Association Techniques for Motion Correspondence," *International Journal of Computer Vision*, vol. 10, no. 1, pp.53-65, 1993.
- [3] Shen Hong, K. Li, Y. Pan, G. H. Young, and S. Q. Zheng, "Performance Analysis for Dynamic Tree Embedding in k-Partite Networks by a Random Walk," *Journal of Parallel And Distributed Computing* 50, pp.144-156, 1998.
- [4] SoonKi Jung, "Motion Analysis of Articulated Objects For Optical Motion Capture," PhD thesis, KAIST 1997.
- [5] K.R. Pattipati, S. Deb, Y. Bar-Shalom, and R.B. Washburn, *Multitarget multisensor Tracking : Advanced Applications*, pp. 219-246, Norwood, MA: Artech House, 1990.
- [6] R. Y. Tsai, "An Efficient and Accurate Camera Calibration Technique for 3-d Machine Vision," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, (Miami Beach, FL), pp.364-374, June 1986.
- [7] Greg Welch and Gary Bishop, "An Introduction to the Kalman Filter," TR 95-041, University of North Carolina at Chapel Hill, Department of Computer Science.
- [8] Z. Zhang and O. Faugeras, *3D Dynamic Scene Analysis*, Berlin Heidelberg, NY : Springer-Verlag, 1992.
- [9] 정순기, 설창환, 원광연, "동작 포착을 이용한 인체 동작의 생성", *Journal of the Korea Computer Graphics Society*, pp. 21-29, vol. 4, no. 1, 1998.
- [10] "http://www.motionanalysis.com," Motion Analysis Corporation.



이 동 훈

1999년 동서대학교 컴퓨터공학과 졸업 (학사). 1999년 ~ 현재 경북대학교 컴퓨터공학과 석사과정 재학중. 관심분야는 Image-Based Modeling and Rendering, Character Animation, Virtual Reality.



추 창 우

1999년 경북대학교 컴퓨터공학과 졸업 (학사). 1999년 ~ 현재 경북대학교 컴퓨터공학과 석사과정 재학중. 관심분야는 Image-Based Modeling and Rendering, Motion Capture, Web3D, Virtual Reality.



김 성 진

1999년 경북대학교 컴퓨터공학과 졸업 (학사). 1999년 ~ 현재 경북대학교 컴퓨터공학과 석사과정 재학중. 관심분야는 Image-Based Modeling and Rendering, Motion Capture, Multimedia Streaming, Computer Graphics.



정 순 기

1990년 경북대학교 컴퓨터공학과 졸업 (학사). 1992년 한국과학기술원 전산학과 졸업(석사). 1997년 한국과학기술원 전산학과 졸업(박사). 1997년 ~ 1998년 University of Maryland, Visiting Researcher. 1998년 ~ 현재 경북대학교 컴퓨터공학과 조교수. 관심분야는 Virtual Reality, Artificial Intelligence, Computer Vision, Image Processing, Computer Graphics.