

다관절체 운동의 분석을 위한 실루엣 상의 특징점 추출

이동훈⁰ 정순기

경북대학교 컴퓨터공학과

dhlee@vr.knu.ac.kr, skjung@knu.ac.kr

Detection of Feature on the Silhouette for Articulated Body Motion

Dong Hoon Lee⁰ Soon Ki Jung

Dept. of Computer Engineering, Kyungpook National University

요 약

인체와 같은 다관절체의 운동은 세그먼트 단위로 강체의 특성을 포함하고 있으나, 움직임에 따라 옷의 주름과 같은 잡음과 관절의 변화 등의 비강체의 성질을 가진다. 따라서 잡음과 관절의 움직임에 따라 다양한 특징점이 추출되며, 영상 간 특징점의 변화가 심하므로, 대상의 특징을 비교적 정확히 반영하는 대응관계를 가진 특징점을 추출하는 것은 어려운 문제이다. 본 논문에서는 인체와 같은 다관절체의 움직임으로부터 신뢰성 있는 대응관계를 가진 특징점을 추출하기 위한 기법을 제안한다. 또한 쌍 대응(pairwise correspondence)의 문제를 확장하여 특정 동작을 이루는 다수의 프레임에서의 일관성 있는 특징점 대응을 위한 간단한 전략을 제안한다. 이를 통해 정확하고 안정적인 대응관계를 가진 특징점을 추출할 수 있다.

1. 서론

영상에서의 특징점 대응문제(correspondence problem)는 컴퓨터 비전에서 고전적인 분야로서 스테레오 시차(stereo disparity), 물체의 인식(object recognition) 및 운동 측정(motion estimation) 등의 많은 비전 응용분야에서 사용된다. 많은 연구가 광학 흐름(optical flow)과 같이 화소(pixel) 단위의 대응점 추출의 문제에 기여하는 반면, 물체를 대표할 수 있는 특징점(feature point)을 사용한 대응 문제 또한 많은 연구가 진행되고 있다. 특히 MPEG-4 기반의 비디오 통신 등과 같은 응용분야에서는 최소한의 필요한 정보를 이용하여 물체를 표현하고, 보간 등의 기법으로 복원하는 연구가 활발하게 진행 중이다. 이러한 특징점을 이용하여 물체를 표현하여 사용한 연구로는 자동화된 립싱크(automatic lip synching)[2], 뷰 기반의 모델 포착(view-based model capture)[1] 그리고 자동화된 영상 모핑(automatic image morphing)[3] 등을 들 수 있다.

특징점 간의 대응관계는 물체의 성질 및 운동에 따라 강체 대응(rigid correspondence) 과 비강체 대응(nonrigid correspondence)으로 분류할 수 있다. 본 연구에서 사용한 인체와 같은 다관절체의 운동은 세그먼트 단위로 강체의 특성을 포함하고 있으나, 움직임에 따라 옷의 주름과 같은 잡음과 관절의 변화 등의 비강체의 성질을 가진다. 또한 이전 연구가 주로 쌍(pair)을 이루는 물체 간의 대응관계를 다루어 온 반면, 본 연구는 특정 동작의 정보를 포함한 다수의 프레임에서 일관성 있는 대응관계를 추출하는데 관심이 있다. 본 논문에서는 인체와 같은 다관절체의 동작 포착을 위해 실루엣(silhouette)으로부터 특징점을 추출하고, 특징점이 다수의 프레임에서 대응관계의 일관성을 유지할 수 있는 기법을 제안한다.

논문의 구성은 다음과 같다. 2장에서는 쌍을 이루는 영상에서 다관절체의 특징점을 추출하고 대응관계를 구하기 위한 전략을 소개하고, 3장에서는 다중 프레임에서 일관성 있는 대응관계를 유지하기 위한 알고리즘을 제안한다. 그리고 4장에서는 제안한 알고리즘을 사용한 실험결과를 보이며, 끝으로 5장에서는 결론을 맺는다.

2. 비강체 물체의 특징점 대응 자동화 알고리즘

쌍 간의 대응관계(pairwise correspondence) 문제는 컴퓨터 비전에서 많이 다루어져 왔던 문제이다. 일반적으로 물체의 경계 간의 대응되는 변환함수를 구하는 문제는 특히 강체(rigid) 변환에 초점을 두어 연구되어 왔다. 그러나, 본 논문에서는 비강체의 물체 간의 대응관계를 다룬다. 비강체 물체의 특징점의 대응관계를 자동적으로 추출하기 위한 방법으로는 Hill[4]이 제안한 기법을 기초로 문제를 설계하였다. 이 알고리즘은 greedy 알고리즘을 사용하여 평가함수를 최적화하는 기법을 통해 희소 다각형 근사 기법(sparse polygonal approximation)을 통해 구해진 각 물체의 특징점의 대응관계를 구한다. 평가함수는 Hill[4]이 제안한 방법과는 달리 물체의 모양에서 이웃한 특징점과 이루는 곡률과 상대적인 길이의 비를 이용하여 모델링한다.

구하고자 하는 대응 관계를 이를 대상을 $A = \{A_i, 1 \leq i \leq n_a\}$, $B = \{B_i, 1 \leq i \leq n_b\}$ 라 두었을 때, 대상의 쌍 간의 대응관계는 순서쌍의 집합 $\Phi = \{\phi_i = (\alpha_i, \beta_i), 1 \leq i \leq n_a\}$ 로 표현할 수 있다. 여기에서 A, B 는 대응 관계를 이를 대상 실루엣의 화소값의 집합이며, $\{\alpha_i\}, \{\beta_i\}$ 는 각각 A, B 사이의 대응관계를 이루는 임의의 화소에 대한 인덱스(index) 정수값이다. 따라서 $\{\alpha_i\}, \{\beta_i\}$ 에 대한 $A' = \{A'_i = A_{\alpha_i}, 1 \leq i \leq n_a\}$, $B' = \{B'_i = B_{\beta_i}, 1 \leq i \leq n_b\}$ 는 A, B 쌍 간의 대응관계를 가지는 희소 부분 다각형으로 정의된다. 위와 같은 쌍 간의 대응관계를 가지는 부분 다각형을 구하기 위한 알고리즘은 다음과 같이 3개의 부분으로 구성된다 [4].

1. 대상 A, B 각각에 대한 희소 다각형 A', B' 의 생성 - 이 단계는 희소 다각형 근사 기법을 사용하여 대상의 특징점을 추출하므로, 두 대상 간의 대응관계가 설정되지 않는다. 따라서 일반적으로 A, B 간의 특징점은 다른 크기를 가진다 ($n_a \neq n_b$)
2. 대응되는 희소 다각형 A', B' 의 초기화 - A', B' 를 사용하여 A, B 외각선의 경로 길이를 기반으로 대응관계를 구한다. 여

기에서 순서쌍 n_o 의 크기는 $n_o = (n_a + n_b)/2$ 로 고정한다.

3. Greedy 알고리즘을 사용한 초기 대응관계의 최적화 - 이웃한 특징점과의 길이와 곡률의 관계를 모델링한 평가함수를 사용하여 대상 A, B 간의 유사성을 최적화한다. 최적화 전략은 두번째 단계에서 구한 최소 다각형 중 하나를 참조 대상으로 고정하여, α_i 또는 β_i 의 값을 변화시켜 수행된다.

2.1. 특징점 추출 알고리즘

A에 대한 최소 다각형 A'를 구하기 위해 본 논문에서는 Zhu와 Chirlian[5]에 의해 고안된 특징점 포착(critical point detection; CPD) 알고리즘을 사용하였다. CPD 알고리즘은 두개의 인접한 특징점과 이루는 삼각형의 면적을 특징값(critical value)으로 두고, 반복적으로 낮은 특징값을 제거하여 원하는 수준의 특징점을 추출하는 알고리즘이다. 특징값에 사용된 삼각형에서 높이는 임의의 특징점이 인접한 두개의 특징점과 이루는 곡률의 크기를 내포하고 있고, 밑변의 길이는 인접한 특징점 간의 거리를 내포하고 있다. 즉 곡률이 상대적으로 작더라도 인접 특징점과의 거리가 크다면 특징점으로서 추출되어야 하고, 곡률이 크더라도 주위 인접 삼각형이 가까이 있다면 잡음으로 처리 가능하다는 성질이 삼각형의 면적을 통해 모델링된다. 대상의 모양의 정보는 사각 좌표계(rectangular coordinates)를 극좌표계(polar coordinates)로 변환하여 사용한다. 2-D의 모양 정보(x,y)를 두개의 1-D 정보($\rho(i), \theta(i)$)로 분리함으로써, 회전(rotation)과 크기 변환(scaling)을 쉽게 다룰 수 있다. 특징값의 추출을 위한 초기의 특징점 획득은 특징점이 될 가능성이 있는 후보들(pseudo critical point)로서, 식 (1)(2)의 조건을 만족하는 극좌표의 두 성분이 이루는 두개의 1-D 곡선 각각의 zero-crossing 지점을 사용한다.

$$\text{sign}(\rho(p_i) - \rho(p_{i-1})) \neq \text{sign}(\rho(p_{i+1}) - \rho(p_i)) \quad (1)$$

$$\text{sign}(\theta(p_i) - \theta(p_{i-1})) \neq \text{sign}(\theta(p_{i+1}) - \theta(p_i)) \quad (2)$$

2.2. 경로 기반 대응점 추출 기법

본 절에서는 대응되는 특징점의 집합 A'와 B'의 초기 측정치를 구하기 위한 방법을 기술한다. 경로 기반 대응점 추출을 위한 알고리즘은 대상 A와 B가 유사한 형상을 지닌다고 가정한다. 즉 임의의 특징점 A_{α_i} 가 물체 A에서 차지하는 인접한 특징점까지 공간의 크기는 B_{β_i} 가 물체 B에서 차지하는 인접한 특징점까지의 공간의 크기와 유사하다. 경로 기반 대응점 추출 알고리즘은 이러한 성질에 기반하여, 대상 물체에 자신의 특징점을 투영하여 좌표상의 위치의 유사도를 평가함수로 측정하여 대응점을 결정한다. 투영을 위한 기준점은 대상 물체의 외각을 이루는 모든 화소의 집합에 대해 반복적으로 방문하여 유사도를 평가함수를 통해 확인하고 평가함수의 값이 최소가 되는 지점을 기준으로 결정한다(식 3참조).

$$\text{Min } E_i^2 = \sum_{j=1}^{n_a+n_b} \|A_{\alpha_i} - B_{\beta_j}\|^2 \quad (3)$$

물체 간에 2-D 공간 상의 회전이나 이동이 가능하나, 본 논문에서는 인체와 같은 다관절체 물체를 실험 대상으로 함으로 2-D 회전변환은 발생하지 않는다고 가정하며, 2-D 이동 변환에 대해서는 물체의 중심으로부터의 상대좌표를 사용하기 때문에 이동변화에 불변(translation invariant)한 결과를 얻을 수 있다. 양 방향으로의 대상체에 대한 투영의 결과로 각각의 물체는 $\Phi = \{\phi_i, 1 \leq i \leq (n_a + n_b)\}$ 의 대응관계를 가지게 된다. 이러한 결과는 유사한 지점에 중복되는 점들이 과잉으로 추출되므로, 앞 절에서 설명한 특징점을 추출하는 것과 유사한 제거과정을 사용하여 식 (4)와 같이 특징값을 추출하여 대응되는 특징점의 수가 $n_o = (n_a + n_b)/2$ 가 될 때까지 잉여점을 줄여나간다.

$$T_i = \text{Max}(\text{Area}(A_{\alpha_i}, A_{\alpha_{i-1}}, A_{\alpha_{i+1}}), \text{Area}(B_{\beta_i}, B_{\beta_{i-1}}, B_{\beta_{i+1}})) \quad (4)$$

2.3. 최적화 기법

최적화를 위한 Greedy 알고리즘은 그림 1과 같다.

```

For j=N, N-1, ..., 1
  set f = 1/2
  do
    for all indices  $\beta_i$ 
      set  $\delta = \text{backwards}(\beta_{i-1}, \beta_{i+1}, f)$ 
      set  $\beta'_i = \beta_i - \delta$  for  $k=i-j+1...i$ 
      evaluate E using  $\beta'_i$  in place of  $\beta_i$ 
      set  $\delta = \text{forwards}(\beta_i, \beta_{i+1}, f)$ 
      set  $\beta'_i = \beta_i + \delta$  for  $k=i-j+1...i$ 
      evaluate E using  $\beta'_i$  in place of  $\beta_i$ 
      accept the best (if any) improvement in E for i
    while improvement in E continues
    set f = f/2
  while maximum possible movement of any  $\beta_i$  is  $\geq 1$  pixel
    
```

그림 1. Greedy algorithm used to minimize E in (5) for fixed A' [4].

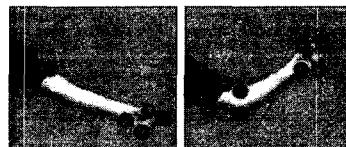
여기에서 E는 평가함수로서 대응되는 두 물체의 특징점 사이의 곡률의 차이와 인접한 다음 특징점 간의 거리를 사용한다(식 5 참조).

$$E = \lambda E_c + (1-\lambda) E_L \quad (5)$$

δ 는 forward와 backward 방향으로 움직일 최소 단위의 거리로서 인접한 화소의 실제 거리에 대한 거리의 비 f에 의해 조절된다. δ 에 의해 조절된 거리의 양만큼 현재의 특징점에 부과되어 평가함수에 의해 적절성이 판단되며 현재의 특징점이 가지는 에러보다 낮은 에러를 가진 지점으로 판단될 경우 비교 대상체의 특징점과 보다 유사한 지점으로 특징점이 고정된다. 이러한 반복과정은 조절될 거리가 최소화소가 될 때까지 반복되어 알고리즘을 수행한다. N은 한꺼번에 움직일 특징점의 개수를 의미하며 실험적으로 조절하여 적당한 값을 결정한다. 앞 절에서 설명한 초기화를 위한 경로 기반 대응점 추출 기법이 전역적으로 가장 유사한 특징점의 선택인데 반해 최적화를 위한 greedy 알고리즘은 국부적으로 특징점의 위치를 변환하여 보다 정밀하게 특징점을 결정하는 알고리즘이다.

3. 다중 프레임에서의 특징점 대응 알고리즘

본 연구의 최종적인 목표는 두개의 영상 간의 신뢰성 있는 대응관계를 추출하는 것을 확장하여 다수의 영상열에서 정확한 대응관계를 구하는 것이다. 2장에서 기술한 특징점 대응 알고리즘은 쌍을 이루는 두 개의 물체 사이의 대응관계를 구하는 것으로 다수의 영상으로 이를 확장시킬 수 없다. 또한 그림 2와 같이 임의의 프레임에서는 존재하지 않았던 중요한 특징점이 물체의 움직임에 따라 새로이 생성될 경우 이를 포함시킬 수 없다. 따라서 다수의 영상에서의 신뢰성 있는 대응점을 추



● Critical Point

그림 2. 물체의 움직임에 따라 새로이 생성된 특징점의 예.

출하기 위한 새로운 방안을 구축해야만 한다. 본 장에서는 이전 프레임의 영상에서 추출된 최적화된 특징점의 정보를 연계할 수 있는 방안으로 앞 장에서 기술한 특징점의 가능성을 내포하는 특징값을 사용하여 다중 프레임에서의 신뢰성 있는 특징점 추출을 위한 전략을 제안한다. 다중 프레임에서의 일관성 있는 특징점의 대응을 구하기 위한 절차는 그림 3과 같다.

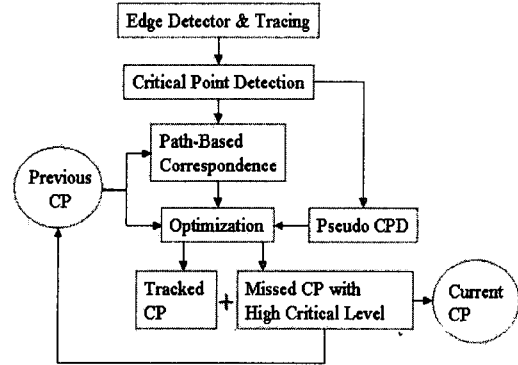


그림 3. 다중 프레임에서의 일관성 있는 특징점 추출.

먼저 이전 프레임의 최적화된 특징점의 집합을 이용하여 경로 기반 대응점 추출 기법을 현재의 영상에서 추출한 특징점에 적용한다. 초기에 결정된 특징점의 개수를 유지하기 위한 전략으로 앞 장에서의 경로 기반 대응점 기법에서 양방향으로 경로의 길이 정보를 투영한 것과는 달리 이전 프레임의 경로 정보만을 현재 프레임에 투영한다. 따라서 잉여정보 제거를 위한 별도의 작업이 필요하다. 다음으로 greedy 알고리즘을 이용하여 현재 프레임의 특징점을 최적화한다. 여기에서 고정되는 정보는 이전 프레임의 특징점의 집합이다. 위 두 번의 단계는 현재 프레임에서 물체의 어떠한 정보도 사용하지 않고 이전 프레임에서 결정된 정보만을 연계하여 이용하여 현재 프레임에서의 물체의 변화에 대해서 직접적인 반영이 이루어지지 않은 상태이다. 세 번째 단계에서는 현재 프레임에서의 특징점의 변화를 포착하고, 이를 이전 프레임에 반영하는 단계로 현재 프레임에서의 의사(pseudo) 특징점을 이용하여 전략을 수행한다. 의사 특징점은 물체가 가질 수 있는 특징점의 모든 후보를 포함하고 있으므로 짧은 시간 간격 사이의 모션의 변화에 대해 거의 변화가 없다고 간주할 수 있다. 따라서 대부분의 특징점은 위치 정보만을 가지고 추적이 가능하다. 본 단계에서는 추적을 위해 단순한 nearest neighboring 방법을 이용하고, 적절한 임계치를 만족하는 점에 대해서는 매칭된 특징점으로 선택한다. 대응관계의 선택은 데이터 연계(data association) 문제로 볼 수 있으며, 단지 가까이 인접한 특징점들이 존재할 경우 거리만을 가지고 선택이 어려운 상황이 발생할 수 있다. 따라서 특징점으로 선택될 가능성의 확률을 내포하는 각 후보들의 특징값이 높을수록, 이전 프레임과의 곡률의 변화량이 적을수록, 인접한 특징점과의 거리의 차가 적을수록 대응관계로 선택될 가능성이 증가하며, 이를 식 6과 같은 평가함수로 모델링한다.

$$E = \frac{CriticalValue}{\lambda E_C + (1-\lambda)E_L} \quad (6)$$

선택되지 않은 의사특징점 중 특징값이 높은점은 그림 2에서와 같이 새로이 발생한 특징점으로 간주할 수 있다. 이러한 특징점은 이웃과의 경로 거리를 이용하여 이전 프레임에 적용한다.

4. 실험

제안한 알고리즘에 대한 실험으로 한 대의 카메라로부터 얻은

몇 개의 영상열을 사용하였다. 영상의 실루엣을 추출하기 위해 수작업을 통한 크로마키를 생성하였다. 배경은 빨간색을 사용하였으며, 이러한 크로마키의 생성은 배경 모델링(background modeling) 등의 기법으로 대체하여 자동화할 수 있다. 그림 4는 본 논문에서 제안한 알고리즘 모듈별 수행 결과이다(임계치 ; 특징점 10).

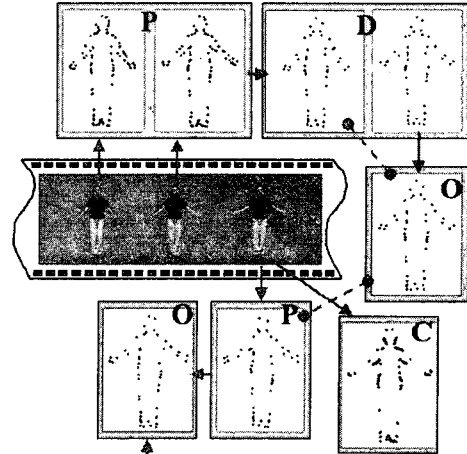


그림 4. P; 경로기반 대응 결과, D; 잉여점 제거 결과, O; greedy 알고리즘 기반 최적화 결과, C; 의사특징점 추출 결과

5. 결론

본 논문에서는 다수의 프레임에서 인체와 같은 다관절체의 움직임의 대응점을 일관성 있게 추출할 수 있는 효과적인 알고리즘을 제안하였다. 적절한 평가분석을 위해 특징점의 일관성에 크게 의존하는 외평의 수행 결과 대체적으로 만족할 만한 결과를 얻을 수 있었다. 제시한 알고리즘의 정확성을 높이기 위해 차후 영상 내부의 컬러 정보를 최적화를 위한 평가함수의 요소로 추가할 계획이다. 또한 본 논문에서 제시한 자동화된 특징점 추출 알고리즘을 사용하여 다관절체의 2-D 외평에 관한 연구가 진행 중이다.

감사의 글

이 논문은 1999년도 한국 학술진흥재단 선도과제(KRF-99-E00294)와 한국과학기술원 가상현실연구센터의 지원을 받았음

참고문헌

[1] D. Beymer, T. Poggio, "Face Recognition from One Example View," MIT AI Memo 1536, 1995.
 [2] C. Bregler, S. Omohundro, M. Covell, M. Slaney, S. Ahmad, D. Forsyth and J. Feldman, "Probabilistic Models of Verbal and Body Gestures," Computer Vision in Man-Machine Interfaces, Cambridge University Press, 1996.
 [3] M. Covell, M. Withgott, "Spanning the gap between motion estimation and morphing," Proc IEEE International Conference on Acoustics, Speech and Signal Processing, 1994.
 [4] Andrew Hill, Chris J. Taylor, and Alan D. Brett, "A Framework for Automatic Landmark Identification Using a New Method of Nonrigid Correspondence," IEEE Transaction on Pattern Analysis and Machine Intelligence, vol.22, no.3, March 2000.
 [5] Pengfei Zhu and Paul M. Chirlian, "On Critical Point Detection of Digital Shapes," IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 1.17, no.8, August 1995.