
3차원 발 자세 추정을 위한 새로운 형상 기술자

송호근* · 강기현** · 정다운** · 윤용인**

Shape Descriptor for 3D Foot Pose Estimation

Ho-Geun Song* · Ki-Hyun Kang** · Da-Woon Jung** · Yong-In Yoon**

이 논문은 2008년도 한서대학교 교내 연구과제 지원사업의 연구비지원에 의해 연구되었음.

요 약

본 논문은 3차원 발 자세를 추정하기 위한 효과적 형상 기술자를 제안하였다. 처리 시간을 단축시키기 위하여 특수 제작된 3차원 발 모형을 2차원 투영하여 발 형상 데이터베이스를 구축하고, 3차원 자세 요약정보를 메타 정보로 추가한 2.5차원 영상 데이터베이스를 구성하였다. 그리고 특징 공간 크기가 작고 다른 형상 기술자에 비하여 자세 추정 성능이 뛰어난 수정된 Centroid Contour Distance를 제안하였다. 제안된 기술자의 성능을 분석하기 위하여, 검색 정확도와 시공간 복잡도를 계산하고 기존의 방식들과 비교하였다. 실험 결과를 통하여 제안된 기술자는 특징 추출 시간과 자세 추정 정확도면에서 기존의 방식들보다 효과적인 것으로 나타났다.

ABSTRACT

This paper proposes the effective shape descriptor for 3D foot pose estimation. To reduce processing time, silhouette-based foot image database is built and meta information which involves the 3D pose of the foot is appended to the database. And we proposed a modified Centroid Contour Distance whose size of the feature space is small and performance of pose estimation is better than the others. In order to analyze performance of the descriptor, we evaluate time and spatial complexity with retrieval accuracy, and then compare with the previous methods. Experimental results show that the proposed descriptor is more effective than the previous methods on feature extraction time and pose estimation accuracy.

키워드

형상기술자, 내용기반색인, 자세추정

Key word

Shape description, Contents-based Retrieval, Pose estimation.

* 한서대학교 컴퓨터정보공학과

** 중앙대학교 첨단영상대학원

접수일자 : 2009. 11. 03

심사완료일자 : 2009. 12. 03

I. 서 론

인체의 움직임을 기반으로 한 응용 시스템을 구현하기 위해서는 인체 움직임 분석 시스템이 필수적이며, 감시 카메라나 특수효과, 게임, 애니메이션, 지능형 인터페이스(Perceptual Interface)등에 주로 응용되고 있다[1].

이때 인체 움직임 추적 기술은 어떠한 환경에서도 강건하게 추적이 가능하며 실시간이어야 한다. 그러나 카메라의 위치와 자세 그리고 추적할 인체 부분에 따라 추적 방법이 달라진다. 따라서 인체 움직임을 추적하는 기술에 대하여 보다 정확하고 신속한 결과를 보장하기 위한 알고리즘이 다양하게 연구되고 있다[2].

한편 발과 관련된 연구로는 신발 디자인을 위한 가상 제작 환경 툴을 비롯하여[3], 가상 착용[4] 분야 위주로 발 관련 연구들이 주로 진행되고 있다. 그리고 카메라 트래킹과 증강 현실 기술을 응용하여 가상으로 신발 착용하는 Virtual Mirror [5]가 대표적이다. 그 밖에 J. Wang[6] 등은 다중 카메라로 취득된 영상으로부터 발 형상을 3차원적으로 재구성하는 연구를 수행하였다. 그러나 위 연구 결과들은 정확한 발 특징의 기술과 자세 추정 및 실시간 처리가 해결 과제로 남아있다.

따라서 본 논문에서는 카메라로부터 취득된 발 실루엣 영상의 3차원 자세를 효과적으로 추정하는 것을 목표로, 영상 검색에 사용되는 다양한 형상 기술자(Descriptor)들의 성능을 검토하고 새로운 기술자를 적용한 발 자세 추정 연구를 수행하였다. 이를 위해서 특수 제작된 3차원 발 모형을 2차원으로 투영하여 발 형상 영상 데이터베이스를 생성하되, 그 3차원 자세 요약 정보를 메타 정보로 추가한 2.5차원 영상 데이터베이스를 구성하였다. 또한 발의 자세를 효율적으로 추정하기 위하여 기존의 다양한 형상 기술자들의 성능을 비교 평가하고, 새로운 기술자를 제안하여 효율적인 발 동작 이해와 자세 추정에 기여하고자 하였다.

본 연구는 인체 검출(body detection), 행동 분석(behavior analysis)에 적용할 수 있는 원천 기술로서 활용이 가능하며, 나아가 인간의 생활 형태에 근접한 디지털 시뮬레이션 시스템이나 아바타(Avata), 디지털 맞춤형 같은 인체 움직임 분석을 기반으로 하는 응용 연구 개발이나 증강 현실 시스템 등에 활용될 수 있다.

II. 발의 특징 및 유사 연구

발은 많은 뼈들로 구성된다. 이 뼈들을 둘러싼 무수한 인대와 근육들은 발에 가해지는 압력과 균형에 대하여 효율적으로 변화하며 다양한 상황에 맞게 미세한 조절 활동을 통해 형상이 변화한다. 하지만 발가락 관절의 경우 손가락과 달리 모두 조밀하게 뭉쳐있으며, 마디가 매우 짧고 그 움직임의 변화 또한 매우 미세하다. 그리고 발의 기능적 특성 상 양말이나 신발을 신는 경우가 빈번하기 때문에 외형적인 특징 정보가 하나의 덩어리로 나타나며, 바닥을 지탱하기 때문에 지면과 발간격 사이에 발생하는 그림자는 정확한 인식 및 추출을 방해한다[7].

한편 인체 검출 및 분석을 하는 기존의 유사 연구들은 다음과 같다. Christopher[8] 등은 의학 영상으로부터 취득된 3차원 척추골의 투영된 2차원 형상 정보를 이용하여 회전된 뼈의 자세를 추정하였고, 이를 3차원 객체 정합(Registration)에 활용하였다. 그리고 객체의 2차원 형상 정보를 형상 기술자로 정의하고, 자세 정보를 포함한 메타 데이터와 결합시켜 데이터베이스에 인덱싱한 뒤, 영상 검색 기법을 통하여 주어진 형상을 데이터베이스로부터 찾아 객체의 자세를 추정하는 새로운 접근의 연구들이 활발히 진행되고 있다[9]. 이는 2차원 형상 기술자를 통하여 3차원 자세를 추정하기 위한 효율적 대안이기 때문이며, 본 논문에서도 위와같은 방식을 취하되 특히 발 영상 검색 및 자세 추정에 적합한 기술자를 연구하였다.

하지만 3차원 자세 추정을 위한 형상 기술자는 인지적으로 유사한 형상들의 회전이나 이동, 크기변화에 강건해야 하며, 높은 인식률을 바탕으로 적은 데이터 양, 짧은 추출 및 비교 검색 시간을 충족시켜야 한다[10].

따라서 본 연구에서는 인체 검출과 분석들에 자주 사용되는 형상 기술자들 중에서 Chain Code, Hue Invariant Moment, Binary Sequence, Axis Projection, Centroid Contour Distance, Shape Context 방법을 선택하고, 이를 발 형상 검출 및 3차원 자세 추정을 위한 검증 도구로 활용하였다. 그리고 발 자세 추정을 위한 인식률을 만족하는 시공간 복잡도, 추출 신뢰도 등을 비교하여 제안하는 기술자의 성능이 우수함을 보였다.

III. 실험 환경과 구성

3.1 데이터베이스 생성

다양한 발 자세에 대한 형상 기반 기술(Shape-based Description)을 위하여, 그림 1과 같이 기준 발 객체를 자체 제작하고, 3D 모델링 과정을 거쳐 총 1197장의 실루엣(Silhouette) 영상을 생성하였다. 축 변환은 각각 10°씩, X축 -40°부터 +40°, Y축 0°부터 180°, Z축 -30°이다. 표 1은 이러한 실험 영상 생성 기준을 보인다.

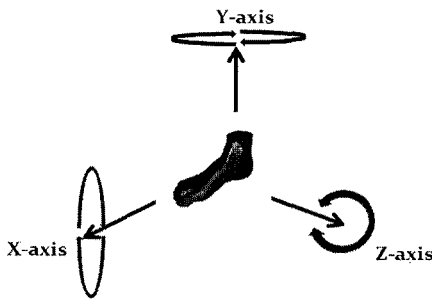


그림 1. 3D 발 모델객체와 3차원 축
Fig. 1 3D foot model and its 3D axis

표 1. 발 DB영상을 위한 축(Axis) 회전각도
Table 1. Axis rotation angle for foot DB images.

축(Axis)	분류수	축(Axis)별 취득각도
X	9	-40° ~ 0° ~ +40°, (10°씩)
Y	19	0° ~ 180° (10°씩)
Z	7	-30° ~ 0° ~ +30° (10°씩)

그림 2는 각각의 기준에 의하여 생성된 영상의 예를 보인다. 생성된 영상들은 각 축의 회전 정보 및 형상 특징(Shape Feature)과 함께 인덱싱 과정을 거쳐 데이터베이스를 생성한다. 이는 실시간 모델 검색을 위한 것으로 특정 모델의 형상 특징을 미리 저장하는 것은 질의 영상과 모델 영상 간의 비교 처리 시간을 단축시키기 위함이며, 회전 정보 즉 특정 모델의 자세 정보를 저장하는 것은 최종적으로 3차원 자세 추정을 위한 계산 시간 단축 및 계산 오차를 최소화하기 위한 것이다.

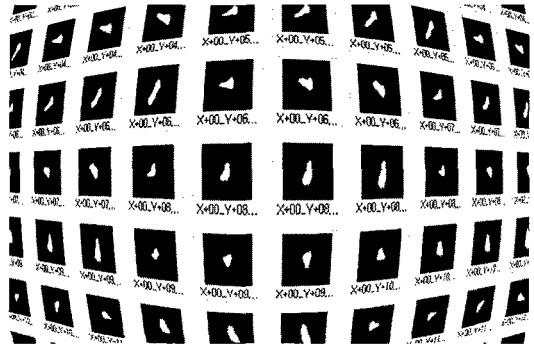
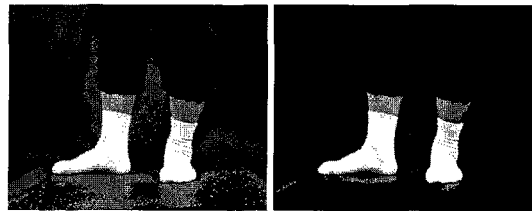


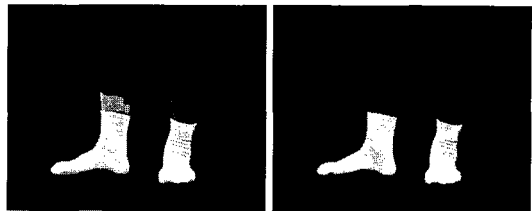
그림 2. 발 DB 영상들
Fig. 2 Foot DB images

3.2 질의 영상(Query Image)의 생성

카메라로부터 취득된 실제 영상으로부터 질의 영상을 생성한다. 질의 영상은 실제 영상으로부터 발부분만을 이진 객체로 표현한다. 이를 위하여 먼저 Blue Screen 기법을 이용하여 배경 영상을 제거하고, 발 객체의 색상 정보를 이용하여 객체를 분할하였으며, 조명, 그림자, 반사 등의 외부 영향에 의한 잡음을 보완하기 위해 형태학적 기법(Morphology)을 사용하였다.



(a) (b)



(c) (d)

그림 3. 질의 영상의 생성
Fig. 3 Creation of query image

3.3 발 형태 정의를 위한 기존 형상 기술자들

본 연구에서는 2차원으로 투영된 발의 모형을 기술하고 비교 평가를 수행하기 위하여, Chain Code[11], Hue Invariant Moment[9], Binary Sequence [12], Axis Projection, Centroid Contour Distance[13], Shape Context[14]의 총 6가지의 형상 기술자들의 성능을 비교하고, 이를 기반으로 새로운 기술자를 제안하였다. 이때 형상 기술자들은 비교 실험을 위하여 크기 변화에 강건하도록 각각 정규화를 수행하였으며, 실험을 위하여 일부 특징은 간소화 또는 수정되었다. 체인 코드(Chain Code)는 형상 객체를 구성하는 외곽선에 대하여 연속하는 외곽점들의 방향 나열을 통해 객체를 정의하는 방법이다. 본 실험에서는 4근방 탐색 방식으로 외곽선을 정의하였다. 한편 코드의 추출은 형상 모델의 좌측 상단을 기준으로 수행되며, 크기 변화에 불변하도록 총 360개로 정규화를 수행하였다. 휴 불변 모멘트(Invariant Moment)는 영상의 기하학적, 통계적 정의를 이용하며, 이동, 회전, 크기 변화에 안정적인 기술이 가능하다. 또한 추출된 특징의 양이 매우 적은 장점이 있다. 휴 불변 모멘트를 구성하는 각각의 특징들은 다음과 같다.

$$\phi_1 = \mu_{20} + \mu_{02} \tag{1}$$

$$\phi_2 = (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2 \tag{2}$$

$$\phi_3 = (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2 \tag{3}$$

$$\phi_4 = (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2 \tag{4}$$

$$\phi_5 = (\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] + (3\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] \tag{5}$$

$$\phi_6 = (\mu_{20} - \mu_{02})[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] + 4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03}) \tag{6}$$

$$\phi_7 = (3\mu_{21} - \mu_{03})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] \tag{7}$$

2진 문자 순서(Binary Sequence)는 객체 영역에 속하는 픽셀은 1로, 그렇지 않은 경우 0으로 표현된 2진 영상을 대상으로 수행된다. 가로와 세로 $m \times n$ 인 i 의 크기를 갖는 임의의 영상 $I(x, y)$ 에 대한 2진 문자 순서는 다음과 같이 정의 가능하다.

$$Binary\ Sequence [I_i] = \begin{cases} 0, & \text{if } I(x, y) = 0 \\ 1, & \text{otherwise} \end{cases} \tag{8}$$

이 기법은 해상도에 종속적인 크기를 갖는다. 따라서 256개의 특징 크기를 갖도록 영상을 16×16 으로 변환하여 특징을 추출하였다.

축 투영(Axis Projection)은 x축에 대한 열의 값 누적, y축에 대한 행의 값 누적을 통하여 해당 축에서의 객체 형상 및 공간 분포 특성을 표현한다. 크기가 $m \times n$ 인 영상 $I(x, y)$ 가 주어졌을 때, 각각의 x축과 y축으로 투영된 특징 값 X, Y 는 다음과 같이 정의된다.

$$X[x] = \sum_{y=0}^{m-1} I(x, y) \tag{9}$$

$$Y[y] = \sum_{x=0}^{n-1} I(x, y) \tag{10}$$

위 방법은 해상도의 가로 길이 m 과 세로 길이 n 에 종속적이므로 각 축당 128개, 총 256개의 크기를 갖도록 크기 정규화를 128×128 로 수행하였다. 웨임 콘텍스트(Shape Context)는 외곽선에 위치한 기준 픽셀을 중심으로, 섹터(Sector)로 분할된 동심원에 포함되는 정보의 양을 기술한다. 형상을 구성하는 n 개의 포인트들 p_n 에 대하여, 각도 θ 별, 섹터 k 에 대한 웨임 콘텍스트는 그림 4의 예에서와 같이 외곽선에 위치한 픽셀들의 관계적 형상 분포를 자세히 표현할 수 있다.

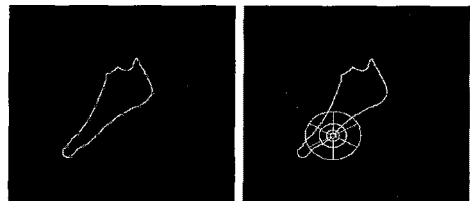


그림 4. 발 영상에 대한 웨임 콘텍스트 추출
Fig. 4 Shape context extraction from foot image

형상에 대한 외곽선 중 총 16개를 샘플 포인트로 정하였으며, 섹터의 수 k 는 3, 각도 θ 는 45° 씩 8로 분할 설정하였다. 따라서 최종적으로 한 형상에 대한 총 384개의 웨임 콘텍스트 특징이 추출된다.

3.4 제안된 형상 기술자

제안된 형상 기술자는 기존의 CCD 기술자를 2차 가공한 특징을 사용하며 다음과 같다.

a. CCD

객체의 외곽 포인트들로 구성된 외곽선을 $P(p_1, p_2, \dots, p_n)$ 라고 할 때, 무게중심 $C(X_c, Y_c)$ 는

$$\begin{aligned} X_c &= \frac{\sum_x \sum_y f(x,y)x}{\sum_x \sum_y f(x,y)}, \\ Y_c &= \frac{\sum_x \sum_y f(x,y)y}{\sum_x \sum_y f(x,y)} \end{aligned} \quad (11)$$

와 같이 구한다. 이때 무게 중심 C 로부터 객체 임의의 외곽선 포인트 p' 까지의 거리는 r 이라하면, 각도 α 에 따른 C 로부터의 거리 집합 $R(r_0, r_1, r_2, \dots, r_{360})$ 을 구할 수 있다. 이렇게 계산된 무게 중심으로부터의 각도별 거리 R 을 CCD(Centroid Contour Distance)라 한다.(그림 5-(a))

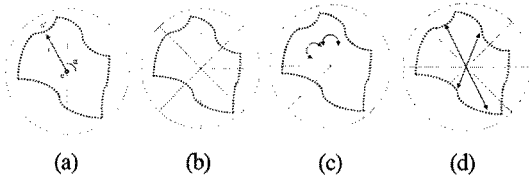


그림 5. 제안된 형상기술자
(a) CCD (b) 영역분할 (c) AAD (d) SAD
Fig. 5 Proposed shape descriptor
(a) CCD (b) Area division (c) AAD (d) SAD

b. 수정된 CCD

위와같은 CCD를 각도 별로 기술하면 총 360개의 특징이 기술된다. 하지만 형상 검출의 정확도를 유지하면서 특징을 개수를 줄이면 실시간 처리에 도움이 된다. 따라서 본 논문에서는 보다 효과적인 형상의 구분 가능한 특징 기술자를 위하여 다음과 같이 정보량은 줄이고 성능을 향상시킬 수 있는 새로운 CCD 기술 방법을 제안하였다. 먼저 CCD를 기술하되 전체 360도의 중심 각도를 임의의 각도 α' 로 분할 영역을 구성한다. 이때 생성되는 영역(area) A^i 의 수는 $i = 360^\circ / \alpha'$ 에 의해서 결정된다. 따라서 $\alpha' = 45^\circ$ 이면 A^i 는 45개(45bin)의 거리

를 포함하는 영역이 되며, 전체 영역 A 는 다음과 같이 표현된다.(그림 5-(b)).

$$A = \{A_n^1, A_n^2, A_n^3, A_n^4, A_n^5, A_n^6, A_n^7, A_n^8\} \quad (12)$$

이제 분할된 영역들을 대상으로 인접 영역간의 거리 (Adjacent Area Distance : AAD)를 구한다. 이때 AAD 의 크기는 분할 영역 i 와 같으며, 다음과 같이 계산된다.

$$AAD_i = \sum_{i=0}^p \sum_{j=0}^q \|A_j^i - A_j^{i+1}\| \quad (13)$$

여기서, p 는 분할된 영역의 개수, q 는 영역의 크기 (α' 와 같은 크기)이다.(그림 5-(c))

위 AAD 는 분할된 CCD 영역에서 이전 영역에 대한 다음 영역의 상대 변이값(relative variance)에 해당되며, 기존 CCD 방법에 대한 2차적인 통계기술법이다. 주어진 형상이 규칙적이거나 상대적 변화가 적은 경우에는 적은 값을 나타내고, 불규칙적이거나 상대적 변화가 큰 경우에는 큰 값을 나타내는 성질을 가진다.

다음으로 대칭 영역간의 거리(Symmetric Area Distance: SAD)를 구한다. 이때 SAD 는 대칭되는 영역간의 거리 차 누적에 해당되고, SAD 의 크기는 $i/2$ 와 같으며, 다음과 같이 계산된다.

$$SAD_i = \sum_{i=0}^p \sum_{j=0}^q \frac{\|A_j^i - A_j^{symmetric}\|}{1 + A_j^i - A_j^{symmetric}} \quad (14)$$

여기서 p 는 $i/2$, q 는 영역의 전체 개수이다. 그리고 위 식에서 분모에 1을 더한 것은 0으로 나뉘는 것을 피하기 위한 최소 상수값이다.

한편 위 SAD 는 AAD 가 상대 변이값을 나타내기 때문에 절대 차이값을 반영하지 못하는 단점을 보완하기 위해 도입되었다. 즉, 작은 값 간의 차와 큰 값 간의 차이가 구별되게 하고, 누적 거리 차에 반영되어 형상 변별력을 높여게 된다.(그림 5-(d))

특히 발 모델의 특성상 좌우 주변 영역간의 차는 발목과 발끝을 제외하고 거의 비슷하기 때문에 큰 차이가 없으나 대칭일 경우 거리차가 크게 발생하는 경우가 많았다. 따라서 AAD 는 일반적인 거리 차 누적을 하였고

SAD는 관계적인 거리 차 누적을 하여 발 모델의 형상 변별력을 높일 수 있게 된다.

마지막으로 자기 영역 누적값(Self Area Accumulation : SAA)을 구한다. 이때 SAA는 해당 구간 자체에서의 누적값에 해당되고,

$$SAA_i = \sum_{i=0}^p \sum_{j=0}^q A_j^i \quad (15)$$

와 같이 계산된다. 여기서 p 는 분할된 영역의 수, q 는 영역의 크기이다.

본 논문에서는 분할 각도를 $\alpha' = 20^\circ$ 으로 하였다. 이 때, 분할 영역의 개수 $i = 18$ 이다. 따라서 AAD의 경우 18개, SAD의 경우에는 9개, SAA는 36개의 특징값을 가지게 된다. 결국 제안된 형상 기술자는 특정 모델에 대하여 총 63개의 특징값으로 기술된다. 이를 통하여 기존의 CCD의 360개 특징값에 대하여 데이터량을 약 60% 수준으로 줄임으로서 실시간 검색을 도모하고, 더불어 인식 정확도를 향상시키는 장점을 얻을 수 있었다.

제안된 형상 기술자를 통한 발 자세 추정은 비교 검색을 위한 데이터량이 감소되어 검색 속도가 향상되고, CCD를 통한 3가지 통계 특성들을 결합하여 검색 정확도를 향상시키는 결과를 얻게 되었다.

IV. 실험

본 장에서는 발의 3차원 자세 추정에 효율적인 제안된 형상 기술자를 검증하기 위하여 기존의 기술자들과 제안된 기술자에 대한 시간 복잡도와 공간 복잡도 및 검색 정확도를 다음과 같이 비교 평가하였다.

4.1 시스템 구성

본 논문의 실험은 Pentium 4 Dual Core 3.2 Ghz, 2GB 메모리 사양의 컴퓨터에서 수행되었고, Visual Studio 2005 환경에서 MFC를 이용하여 구현하였다. 발 형상 데이터베이스는 3장에 기술된 것같이 3차원 발 모델에 대하여 1197장의 실루엣 영상을 제작하고, 각각의 형상 특징값과 해당 3차원 자세 정보를 메타 정보로서 결합하여 구성하였다. 질의 영상은 실제 취득된 발 영상에 대하여 배경과 잡음을 제거한 발 영상으로 하였다.

구현된 시스템의 전체 흐름은 다음과 같다. 질의 영상이 입력되면, 선택된 특징에 해당하는 데이터베이스와 특징간 유사도 비교 연산을 수행하여, 유사도가 가장 높은 모델 영상을 선택한 후, 3차원 자세 정보를 포함하는 메타 데이터(Meta Data)를 참조하여 최종적으로 질의 영상내 발의 3차원 자세를 추정한다.

4.2 공간 복잡도와 시간 복잡도

앞에서 살펴본 총 6가지 형상 기술자와 제안된 형상 기술자의 인덱싱에 필요한 공간 복잡도를 정리하면 표 2와 같다.

표 2. 공간 복잡도 비교
Table 2. Spatial complexity comparison

기술자(Descriptor)	특징(Feature) 수
Chain Code	360
Hu Moment	7
Binary Sequence	256
Axis Projection	256
Shape Context	384
Proposed method	63

여기서 공간 복잡도는 형상 기술자가 포함하는 특징값의 개수이다. 이때 공간 복잡도는 영상 데이터베이스를 검색할 때 질의 영상과 모델 영상간의 유사도 연산 횟수와 비례하므로 영상 데이터베이스의 크기가 커질수록 실시간 처리에 직접적인 요인이 된다. 표 2를 보면, 휴모멘트와 제안된 기술자를 제외하면 평균 323개의 특징값을 가진다. 따라서 모델 영상 데이터베이스를 검색할 때 휴볼변 모멘트가 가장 유리하지만, 제안된 기술자는 평균 복잡도에 대하여 약 20% 수준으로 매우 우수한 조건임을 알 수 있었다.

한편 각각의 기술자에 대한 평균 특징 추출 시간을 표 3과 같이 정리하였다. 모든 특징 추출 시간은 기술자에 특성에 맞게 적용된 전처리 시간이 포함된 전체적 시간을 의미한다. 일반적으로 평균 처리 시간은 연산 복잡도와 비례한다. 따라서 입력 질의 영상에 대한 특징 추출 시간은 전체 검색 시간에 있어 기본적으로 소요되는 상수 시간의 의미를 가진다.

표 3. 시간 복잡도 비교
Table 3. Time complexity comparison

기술자(Descriptor)	평균추출 시간
Chain Code	0.000861
Hu Moment	0.008119
Binary Sequence	0.001504
Axis Projection	0.001815
Shape Context	0.036798
Proposed method	0.00296

실시간 영상 검색을 위해 적은 시간일수록 유리하나, 위 공간 복잡도에 의한 처리 시간이 증가될수록 위 시간 복잡도의 의미는 줄어든다고 본다. 표 3을 보면 체인 코드가 가장 빨랐고, 가장 느린 shape context를 제외한 평균 시간 복잡도는 약 0.00346였다. 제안된 형상 기술자의 시간 복잡도는 평균 시간 복잡도에 비하여 약 86%로 우수한 성능을 확인할 수 있었다. 따라서 제안된 형상 기술자는 기존의 기술자들에 비해 평균 복잡도를 기준으로 볼 때 공간 복잡도와 시간 복잡도가 모두 유리한 조건임을 알 수 있었다.

4.3 검색 정확도

검색 정확도는 다음과 같이 추정하였다. 100장의 무작위 선택된 질의 영상을 이용하여 총 1197장의 모델 영상에 대한 유사도 거리값을 계산하였다. 유사도 크기 순으로 10장의 후보 모델 영상을 선택하였다. 그리고 질의 영상의 x,y,z 축으로의 자세 정보와 검색 결과 영상의 x,y,z 축으로의 자세 정보를 비교하여 그 오차를 계산하였다. 이때 검색 순위별로 10회 평균 오차를 각 축별, 순위 구간에 대하여 계산하였다. 기존의 기술자들과 제안된 기술자의 검색 정확도는 표 4에 보였다. 표 4에서 Rank(2)는 검색 결과 1, 2위의 순위 구간에서 평균 오차, Rank(5)는 검색 결과 1~5위의 평균 오차, 그리고 Rank(10)는 검색 결과 1~10위의 평균 오차를 x,y,z 축별로 각각 나타내고, 마지막으로 x,y,z 3축 오차의 전체 평균값을 표시하였다.

한편 n 개의 길이로 구성된 입력된 발 형상에 대한 특징 F_n 과 총 k 개의 영상으로 구성된 데이터베이스 영상들의 특징 F_n^k 간의 유사도는

$$Dist(F_n, F_n^k) = \sum_{i=0}^n |F_i - F_i^k| \quad (16)$$

와 같이 계산하였다. 이때 거리차가 최소인 F^k 이미지가 입력된 영상 F 에 대한 3차원 자세 정보를 가진 검색 결과가 된다.

표 4. 검색 정확도
Table 4. Retrieval accuracy

ChainCode	x-axis	y-axis	z-axis	avg.
Rank(2)	11.5	25.5	5	14
Rank(5)	17	27.8	6.4	17.1
Rank(10)	20.5	29.6	7.9	19.3

Axis Pjt	x-axis	y-axis	z-axis	avg.
Rank(2)	3	3.5	0	2.2
Rank(5)	8	8.6	1	5.9
Rank(10)	12.1	11.8	2.4	8.8

Hu Mmt	x-axis	y-axis	z-axis	avg.
Rank(2)	5.5	15	2.5	7.7
Rank(5)	10.8	29.2	5.6	15.2
Rank(10)	12.7	36.8	6.7	18.7

Binary Str	x-axis	y-axis	z-axis	avg.
Rank(2)	3	3.5	0	2.2
Rank(5)	6.6	7.6	0.8	5
Rank(10)	9.5	10.1	1.4	7

ShapeCtxt	x-axis	y-axis	z-axis	avg.
Rank(2)	3	4	0.5	2.5
Rank(5)	6.2	8.6	1.2	5.3
Rank(10)	8.4	11.6	2.4	7.5

Proposed	x-axis	y-axis	z-axis	avg.
Rank(2)	2.5	2.5	0	1.7
Rank(5)	5.6	5.8	0.2	3.9
Rank(10)	7.9	8.3	0.9	5.7

표 4의 결과는 각각의 기술자에 대하여, 3차원 축 정보 및 순위별 평균 오차를 나타낸다. 위 결과를 통하여 제안된 방법이 추정된 각각의 축에 대한 모든 순위별 정확도에서 비교 우위를 나타내는 것을 확인할 수 있었다. 제안된 방식은 수정된 CCD로서 발 형상과 같이 심한 굴곡의 변화 및 관절과 같이 접히는 골격구조가 없는 형상 객체의 인식에 상대적으로 우수한 성능을 나타내기 때

문이었다. 한편 웨입 콘텍스트(Shape Context)의 경우 콘센트릭 서클을 구성하는 θ 와 섹터의 수를 결정하는 r 의 수를 증가시켜 특징을 추출하면 세밀한 형상 정보의 기술(Description)과 향상된 인식 성능을 얻을 수 있다. 하지만, 모든 기술자들에 대하여 가장 느린 추출시간과 가장 큰 저장 공간을 차지하는 기술자의 특성상, 보다 정교한 기술을 위한 실험 설정치의 변화는 더 많은 시간 복잡도와 공간복잡도를 유발하기 때문에 실시간 처리에 적합하지 않다. 체인 코드의 경우 가장 낮은 성능의 인식을 나타냈는데, 이는 실험에 적용된 4방향 기술 방법이 가지는 제한된 특성과 크기변화와 비교연산을 위해 수행한 샘플링 과정에서 기인하며, 8방향 기술을 하여도 유사한 결과를 얻었다. 그 외의 기술자들은 대체로 안정적인 성능을 보였다. 한편 검색 정확도를 검색 평균 오차의 역수로 정의하여 환산했을 때, 제안된 방식의 검색 정확도가 가장 정확했으며 기존 방식들의 평균 정확도보다 약 36% 정확한 결과를 얻는 것으로 확인되었다. 그림 7는 유사도 순서에 따른 각각의 알고리즘에 대한 검색 결과의 예를 보인다.

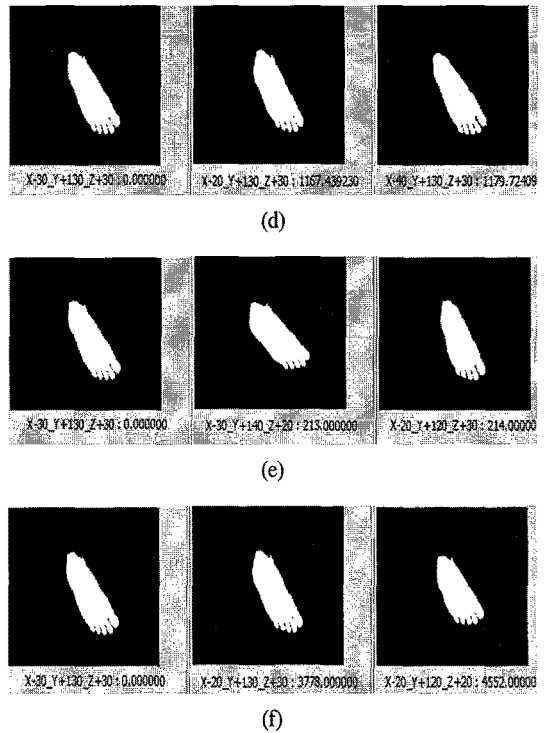
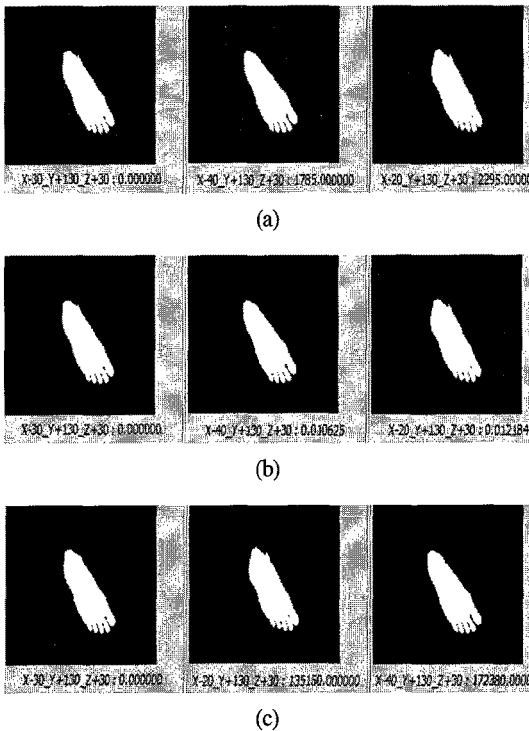


그림 7. 검색결과
 (a) 휴 모멘트 (b) 이진문자순서 (c) 축투영
 (d) 체인코드 (e) 웨입 콘텍스트 (f) 제안방식
 Fig. 7 Retrieval result
 (a) Hu moment (b) Binary sequence (c) Axis projection (d) Chain code (e) Shape context (f) Proposed method

V. 결론

본 논문에서는 발 실루엣 영상의 3차원 자세를 효과적으로 추정하는 것을 목표로 영상 검색 기법을 위한 다양한 형상 기술자들의 성능을 검토하고 새로운 기술자를 적용한 발 자세 추정 연구를 수행하였다. 자세 추정 처리 시간을 단축시키기 위하여 먼저 특수 제작된 3차원 발모형을 2차원 투영하여 발 형상 데이터베이스를 구축하고, 3차원 자세 요약 정보를 메타 정보로 추가한 2.5차원 영상 데이터베이스를 구성하였다. 그리고 3차원 발 자세 변화에 강건하며, 높은 인식률을 바탕으로 적은 데이터량, 짧은 추출 및 비교 검색 시간 조건을 충족시키기

위한 형상 기술자를 제안하였다. 제안된 방식의 성능을 검증하기 위하여 기존의 형상 기술자들의 공간 및 시간 복잡도와 자세 추정 정확도를 비교하는 실험을 수행하였다. 제안된 방식은 공간 복잡도에 있어서는 기존 방식들의 평균값보다 약 20%, 시간 복잡도는 그 평균값보다 약 86% 수준으로 우수하였다. 한편 자세 추정 정확도는 그 평균값보다 약 36% 더 정확하였다. 위 실험 결과를 통하여 제안된 방식은 기본 방식들보다 적은 데이터량, 빠른 검색 시간 그리고 높은 인식율을 나타냄으로서 효율적인 3차원 발 자세 추정 기법으로 활용될 수 있음을 확인하였다. 본 연구는 인체 검출, 행동분석 분야와 기타 인체 움직임 분석기반 응용 연구 그리고 증강 현실 시스템 등에 응용될 수 있다. 앞으로 추가적인 처리 시간 단축을 위해 공간 유사도뿐만 아니라 시간 유사도등을 활용하는 알고리즘에 대한 연구가 필요하다.

참고문헌

- [1] A. Agarwal and B. Triggs, "Recovering 3D Human Pose from Monocular Images," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 28, No. 1, 2006.
- [2] A. O. Balan, L. Sigal, M. J. Black, J. E. Davis and H. W. Haussecker, "Detailed Human Shape and Pose from Images," *International Conference on Computer Vision and Pattern Recognition*, pp. 73-80, 2007.
- [3] P. Azariadis, V. Moulianitis, S. Alemany, J. C. Conzalez, P. D. Jong, M. V. D. Zande and D. Brands, "Virtual Shoe Test Bed: A Computer-Aided Engineering Tool for Supporting Shoe Design," *Computer-Aided Design and Applications*, Vol. 4, No. 6, pp. 741-750, 2007.
- [4] S. G. Papageorgiou, V. C. Moulianitis, and N. A. Aspragathos, "Transfer VM Concepts to Cloth Design and Manufacturing," *Eurasia-Tex Conference on 3D Body Scanning and Virtual Try-On Systems*, pp. 41-49, 2003.
- [5] P. Eisert, J. Rurainsky and P. Fechteler, "VIRTUAL MIRROR : Real-Time Tracking of Shoes in Augmented Reality Environments," *International Conference on Image Processing*, pp. 557-560, 2007.
- [6] J. Wang, H. Saito, M. Kimura, M. Mochimaru and T. Kanade, "Human Foot Reconstruction from Multiple Camera Images with Foot Shape Database," *IEICE Transactions on Information and Systems*, Vol. 89, No. 5, 2006.
- [7] J. T-M. Cheung and M. Zhang, "Finite Element Modeling of the Human Foot and Foot Wear," *ABAQUS User's Conference*, pp. 145-159, 2006.
- [8] C. M. Cyr, A. F. Kamal, T. B. Sebastian, and B. B. Kimmia, "2D-3D Registration based on Shape Matching," *Proceedings of Mathematical Method in Biomedical Image Analysis*, pp. 198-203, 2000.
- [9] P. Tresadern and I. Reid, "An Evaluation of Shape Descriptors for Image Retrieval in Human Pose Estimation," *The 18th British Machine Vision Conference*, Vol. 2, pp. 800-809, 2007.
- [10] D. Zhang and G. Lu, "Review of Shape Representation and Description Techniques," *Pattern Recognition*, Vol. 37, pp. 1-19, 2004.
- [11] Herbert Freeman, "Computer processing of line-drawing images," *Computing Surveys*, Vol. 6, No. 1, pp. 57-97, 1974.
- [12] B. G. Prasad, K. K. Biswas and S.K. Gupta, "Region based image retrieval using integrated color, shape & location index." *Computer Vision and Image Understanding*, Vol. 94, pp. 193-233, . 2004
- [13] Z. Wang, Z. Chi and D. Feng, "Shape based leaf image retrieval," *IEE Proceedings of Image and Signal Processing*, Vol. 150, No.1, pp. 34-43, 2003.
- [14] G. Mori, S. Belongie and J. Malik, "Efficient Shape Matching using Shape Context," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 27, No. 11, pp. 1832-1837, 2005.

저자소개



송호근(Ho-Geun Song)

1991년 중앙대학교 전자공학과 학사
1993년 중앙대학교 전자공학과
정보제어 석사
1997년 중앙대학교 전자공학과
정보공학 박사

1996-현재 한서대학교 컴퓨터정보공학과 교수
※관심분야: 신호처리, 컴퓨터비전, 3차원영상처리,
증강현실



윤용인(Yong-In Yoon)

1986년 중앙대학교 전자공학과 학사
1988년 중앙대학교 전자공학과 석사
2003년 중앙대학교 첨단영상대학원
영상공학과 박사.

1991-1993년 대우고등기술연구원.
2003~2005년 중앙대학교 첨단영상대학원 Post-Doc.
2006~2006년 중앙대학교 문화컨텐츠연구원
선임연구원
2006~현재 중앙대학교 첨단영상대학원 연구교수
※관심분야: 증강현실, 신호처리, 컴퓨터비전, 영상
처리



강기현(Ki-Hyun Kang)

2006년 한서대학교
컴퓨터정보학과 학사
2006-2008년 중앙대학교 첨단영상
대학원 영상학과 석사

2008-2009 중앙대학교 첨단영상대학원 박사과정
2009-현재 한미반도체 제2연구소 비전연구소
※관심분야: 영상처리, 패턴인식, 내용기반영상검색



정다운(Da-Woon Jung)

2001년 덕성여자대학교 전산학과
학사
2007년 중앙대학교 첨단영상
대학원 영상공학과 석사

2007년~현재 중앙대학교 첨단영상대학원 첨단영상
학과 박사과정
※관심분야: 컴퓨터비전, 휴먼모션, 컴퓨터 그래픽스,
증강 현실