

논문 2007-44SP-6-4

# 발의 움직임 추적에 의한 3차원 신발모델 정합 시스템

## (Registration System of 3D Footwear data by Foot Movements)

정다운\*, 서용호\*, 최종수\*\*

(Da Un Jung, Yung-Ho Seo, and Jong-Soo Choi)

### 요약

정보화 성장과 함께 인간의 생활도 발전하면서, 정보의 접근이 보다 간편한 시스템들이 개발되고 있다. 본 논문에서는 한 대의 카메라를 사용하여 3차원 신발 모델을 발에 정합하는 시스템을 제안한다. 인체 움직임 분석에서 전신 움직임에 대한 연구가 대부분인 것과 달리, 우리는 발의 움직임을 기반으로 한 새로운 움직임 분석 시스템을 제안한다. 본 논문은 시스템이 구현되는 과정과 결과를 설명한다. 3차원 신발모델을 이미지의 발에 투영하기 위해 발 추적, 투영, 자세 추정 과정으로 구성했다. 이 시스템은 2차원 영상 분석과 3차원 자세추정으로 나뉜다. 먼저 발 추적을 위해 발의 형태학적 특성에 따라 특징점을 찾는 방식을 제안한다. 그리고 별도의 영상 교정 없이 한 대의 카메라로 2차원 좌표와 3차원 좌표의 관계를 설정하는 기하학적 수식을 제안한다. 제안한 방법에 따라 응용 시스템을 구현하고 거리 오차를 측정한 결과 거의 유사한 위치로 정합 되는 것을 확인할 수 있었다.

### Abstract

Application systems that easy to access a information have been developed by IT growth and a human life variation. In this paper, we propose a application system to register a 3D footwear model using a monocular camera. In General, a human motion analysis research to body movement. However, this system research a new method to use a foot movement. This paper present a system process and show experiment results. For projection to 2D foot plane from 3D shoe model data, we construct processes that a foot tracking, a projection expression and pose estimation process. This system divide from a 2D image analysis and a 3D pose estimation. First, for a foot tracking, we propose a method that find fixing point by a foot characteristic, and propose a geometric expression to relate 2D coordinate and 3D coordinate to use a monocular camera without a camera calibration. We make a application system, and measure distance error. Then, we confirmed a registration very well.

**Keywords :** Human Motion Analysis, Silhouette, Pose Estimation, HCI

## I. 서론

보편화 된 네트워크 시스템과 더불어, 노트북, PDA, DMB 모바일 폰 등 개인의 개성을 최대한 반영할 수 있는 다양한 디자인과 기능을 갖춘 개인용 통신 기구들은 소형화, 경량화를 통한 이동성뿐만 아니라, 지역적

한계를 넘어선 정보 접근능력을 사용자에게 제공하였고, 나아가 정보 통신 기술 환경을 보다 더 개별화된 양상으로 발달시키는데 기여하였다. 그리고 이러한 현상은 정보 통신 기술을 통하여 접근하는 생활 전반의 정보를 대하는 태도와도 긴밀한 연관을 맺고 있다.

최근 네트워크 시스템의 발전에 따라 인터넷 쇼핑물 관련 사업이 눈부시게 성장하고 있다. 비교적 저렴한 가격도 크게 작용했지만, 다양한 상품 선택과 시공간 제약이 없는 쇼핑 정보의 접근성이 보다 더 큰 이유라고 할 수 있다. 그러나 그 대가로 소비자들이 감수하고 있는 것은, 그 자리에서 직접 착용해 볼 수 없다는 물리적 제약과 이로 인한 맞지 않는 사이즈나 사진과의 차이로 인한 상품에 대한 불만과 불편함이다. 이것은 표

\* 학생회원, \*\* 평생회원, 중앙대학교 첨단영상대학원 영상공학과

(Dept. of Image Engineering, GSAISM, Chung-Ang University)

※ 본 연구는 교육인적자원부, 산업자원부, 노동부 출연금으로 수행한 최우수실험실지원사업, 서울시산학 협력사업, 2단계 BK21 지원과제임.

접수일자: 2007년3월27일, 수정완료일: 2007년10월30일

준화된 규격의 한계라기보다는 상품정보 디스플레이 방식이 여타 정보 통신 환경의 발달에 비해 상대적으로 낙후된 인터페이스를 가지고 있다는 점에 기인한다<sup>[1]</sup>.

이러한 정보통신 기술 양상에 따르는 환경 제약에 대응하기 위해서는 개인화, 개별화 된 인터페이스의 개발이 이루어져야 한다. 따라서 한층 더 인간의 생활 형태에 가까워 지기위해 디지털 시뮬레이션 시스템이나 아바타, 디지털 맞춤형 같은 인체 움직임 분석(Human Motion Analysis)을 기반으로 한 연구, 개발이 진행되고 있다<sup>[2-3]</sup>.

인체 움직임 분석(Human Motion Analysis)을 기반으로 한 시스템을 구현하기 위해서는 시스템 환경에 적합한 영상 이론, 제한되는 환경(Constraint), 움직임 분석의 대상이 될 객체분석의 세 요소에 초점을 두고 흐름을 결정해야 한다.

첫째로 영상 분석에서의 인체 움직임 분석은 일반적으로 인체 검출(Human Detection), 인체 움직임 추적(Human Motion Tracking), 행동 이해(Behavior Understanding)의 세 가지 형태로 분류한다. 인체 검출은 사람이 포함된 영상에서 인체만을 분류하는 것이고, 매 프레임마다 인체의 위치, 모양 등 특징정보를 찾아 추적하는 것이 인체 움직임 추적이다. 그리고 추적된 인체의 동작을 분석하고 인식하는 것이 행동 이해 과정이다. 이 세 분류는 구현하고자 하는 시스템의 최종 목표에 따라 서로 적절히 조합되거나 독립적으로 적용할 수 있다<sup>[4]</sup>.

두 번째로 환경의 제한이다. 강건한 추적과 시스템 구현을 위해서 실시간(Real-time) 여부와 카메라의 개수, 영상 교정(Camera Calibration), 마커의 사용여부 등을 미리 정하여 구현해야 하므로 환경에 제한이 따른다. 실시간 여부를 결정하려면, 발을 추적하고 정합(Registration)을 하는 과정까지 데이터 처리량과 시간 대비 하드웨어 성능에 대한 고려가 중요하다. 3차원 데이터를 취득하는 방법과 가려지는 영역(Occlusion)이나 시차(Depth)정보 등의 문제점에 대한 보완책에 따라 카메라 개수를 결정한다. 그리고 마커의 사용 여부는 시스템 구현 목적에 따라 결정된다.

마지막으로 대상이 되는 객체의 특징이 무엇인가에 대한 분석이다. 추적할 대상 객체가 무엇이고 형태학, 기하학적으로 어떤 모양을 하고 있으며, 고유 특징은 무엇인지 미리 분석하여 동적 환경 변화에도 온건하게 추적할 수 있도록 해야 한다.

본 논문에서는 실생활에 사용될 수 있는 응용 시스템

으로, 한 대의 카메라를 사용하여 발을 추적하고 3차원 정보 취득하여 발 위에 신발을 투영하는 시스템을 제안한다.

본 논문의 구성은 다음과 같다. II장에서는 전체적인 시스템 환경과 흐름에 대해 간략하게 설명한다. III장에서는 발 객체 추적과 3차원 좌표값 추출에 대해 설명한다. IV장에서는 실제 발에 신발을 투영하기 위한 3차원 자세 추정에 대해 설명하며, V장에서는 제안하는 시스템의 실험 결과를 보여준다. 마지막으로 VI장에서는 실험에 대한 평가를 하고 결론을 맺는다.

## II. System Overview

### 1. 시스템 환경

인체 움직임을 기반으로 한 응용 시스템을 구현하기 위해서 환경 설정과 인체에서 추적해 낼 움직임이 어떤 것인지에 대한 가정이 필요하다. 본 논문의 시스템은 구매자가 상점에서 신발을 구입할 때 무릎 아래로만 보이도록 한 대의 카메라를 정면에 설치하고 별도의 디스플레이를 통해 구매자의 발에 직접 3차원 신발을 증강(Augmentation)하여 마치 신발을 직접 신은 듯한 효과를 내고자 한다. 전체적인 시스템 구성 환경은 그림 1과 같다.

일반적으로 배경에서 인체를 검출할 때에는 인체의 고유 색상인 피부 색상(Skin color)을 사용하여 영역을 구분하지만, 본 논문에서는 피부 색상이 조명이나 그림

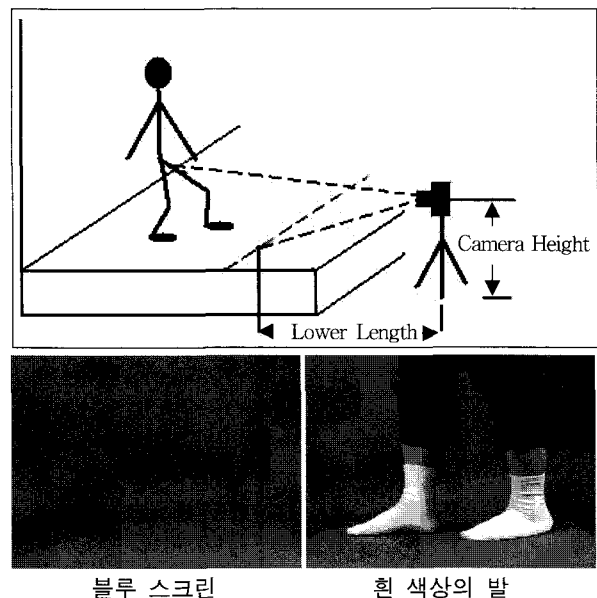


그림 1. 시스템 구성 환경  
Fig. 1. System Environment.

자 등의 여러 변화요소에 의한 방해에 민감하다는 문제점을 최소화하기 위해 어떤 변화에도 강건하도록 배경을 파란색(Blue Screen) 그리고 발을 흰색(White Object)으로 색상을 지정하였다<sup>[5]</sup>.

움직이는 발에 매 프레임마다 3차원 신발을 증강하기 위해서는 2차원 영상에서 3차원 데이터를 취득해야 하는데, 이를 위해 영상 교정(Camera Calibration)이 필요하다. 본 논문의 시스템은 한 대의 카메라를 가지고 별도의 영상 교정 없이 2차원 좌표와 3차원 좌표를 서로 대응시키는 기하학적 수식을 사용한다. 기하학적 수식을 구성하기 위해서는 최소 개수의 실제측정 파라미터를 알고 있어야 한다. 본 논문의 시스템은 Lower Length와 Camera Height의 실제 측정 파라미터가 필요하다. 이 파라미터들은 영상에 투영되는 실제계영역에 포함되는 것으로 Lower Length는 카메라에서부터 영상에 투영되는 지면의 거리를, Camera Height는 지면을 기준으로 한 카메라의 높이를 측정한다.

2. 시스템 흐름도

본 논문의 시스템은 발의 움직임을 찾는 2차원 영상 분석과 발의 자세를 추정하는 3차원 모델 정합으로 구분된다. 시스템의 Flow Chart는 그림 2와 같다.

영상 시퀀스에서 발 객체를 검출하고 발 객체의 움직임을 찾아 기하학적 수식에 의해 3차원 좌표값을 계산하는 것 까지가 2차원 영상분석이고, 2차원 영상분석에 의해 계산되어진 3차원 좌표값과 신발 모델 좌표 값과의 자세를 맞추고 기하학적 수식에 의해 다시 3차원 좌표 값을 2차원 좌표 값으로 재투영 하는 것이 3차원 모델의 정합이다.

매 프레임 마다 움직이는 발을 강건하게 찾기 위해서는 영상의 전처리 과정이 필요하다. 2차원 영상분석에서 발을 검출하기 위한 첫 단계로 배경영상을 제거(Background Subtraction)한다. 배경이 제거 된 영상은 조명이나 그림자, 반사 등에 영향을 받아 적게는 배경의 잡음으로, 크게는 객체에 구멍이 생기는데, 이러한 문제를 보완하기 위해 모폴로지(Morphology)연산을 수행한다. 이렇게 찾아낸 움직임 객체에서 발 이라는 것을 구체화 하기위해 미리 지정한 발의 색상인 흰색만을 찾아 발 색상 영역을 검출하고 실루엣을 생성한다. 생성된 실루엣은 발의 형태학적 특징과 움직임에 따라 발의 골격과 관절 포인트를 찾고 그것들의 2차원 이동 정보를 계산한다. 2차원 이동 정보를 기하학적 수식에 적용하여 3차원 좌표를 계산한다.

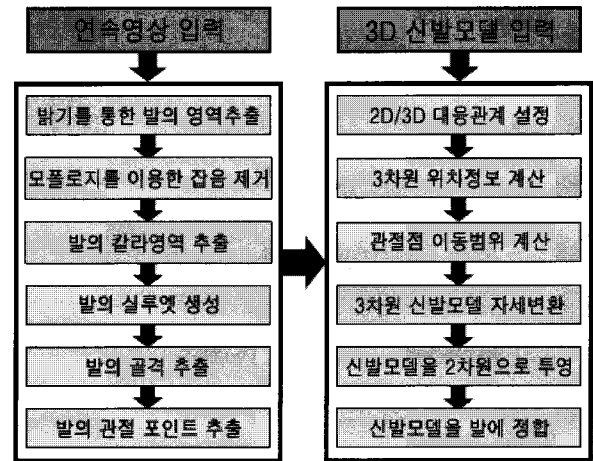


그림 2. 시스템 흐름도  
Fig. 2. Flow Chart.

기하학적 수식에 의해 계산 되어진 3차원 좌표값을 가지고 발의 자세를 추정(Pose Estimation)한다. 자세 추정에 앞서 3차원 신발모델은 3D Tool(3D Max, Maya 등)에 의해 미리 만들어진 데이터이다. 계산 된 3차원 좌표값과 3차원 신발 모델을 가지고 상호 유사한 형태로 자세를 보정하면서 3차원 신발모델의 데이터와 좌표체계를 갱신한다. 최종적으로 자세 추정이 완료된 3차원 신발 모델을 기하학적 수식에 적용하여 2차원 좌표를 다시 계산해 내고 이를 영상에 재투영 하면 신발 모델과 발이 정합된다.

III. 발의 움직임 추적

1. 전처리 과정

대부분의 움직임 추적(Motion Tracking) 알고리즘은 강건한 객체 추적을 위해 어떤 상황에서도 변하지 않는 특징점을 찾고 추적하는 것이 중요하다. 인체 움직임 분석(Human Motion Analysis)에서도 인체의 움직임을 추적하기 위해 먼저 특징점을 추출하고 매 프레임마다 그것에 대응되는 특징점을 추적한다. 인체 움직임을 추적하기 위한 방법에는 크게 4가지가 있는데, 인체 모델 기반(Human Model-based) 추적 방법이나, 광류(Optical Flow), 실루엣(Silhouette)기반 추적 방법, 인체의 역운동학(Inverse Kinematics) 추정 방법 등이 있다. 이 4가지 방법들 중 인체 모델 기반 추적이나 광류, 역운동학 추정 방법은 매 프레임마다 인체의 형태학적 고정점이 존재했을 때 유용하게 사용될 수 있다. 인체 부위 중 고정점이 존재하는 손(손가락 비율, 손가락 관절 마디)이나 얼굴(눈, 눈썹, 코, 입)과는 달리 발의 모

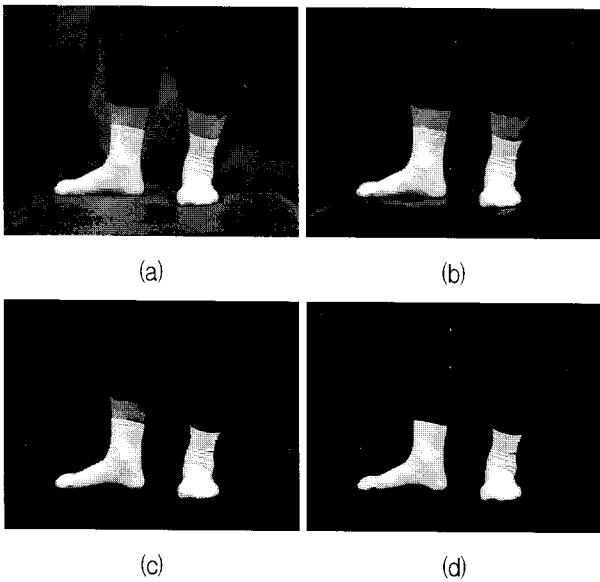


그림 3. 전처리 과정 (a)배경영상의 제거, (b)Gray 임계 값 조정, (c)잡음제거, (d)발의 색상에 의한 최종 검출  
 Fig. 3. Image Processing (a)Background Subtraction, (b)Gray Scale Threshold, (c)Remove Noise, (d)Detection of foot object

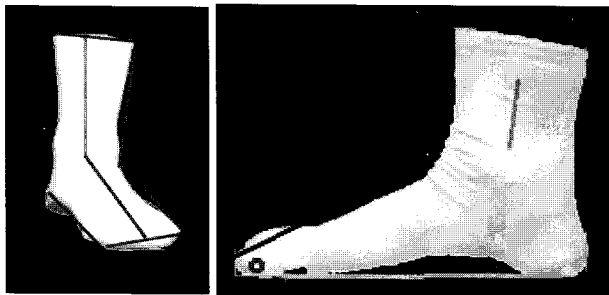


그림 4. 골격선분과 관절점  
 Fig. 4. Skeleton Line and Joints.

양은 카메라에서 투영되는 위치와 자세에 따라 그 형태가 무한하게 변화하기 때문에 인체 비율에 의한 추적이나 광류, 역운동학 보다는 형태학적 분석이 가능한 실루엣이 적합하다.

발의 실루엣을 생성하기 위한 초기 과정으로 영상의 전처리(Pre-processing)는 필수 과정이다. 먼저, 배경영상 제거기법(Background Subtraction)으로 영상에서 블루 스크린 배경을 제거한다. 배경영상 제거 기법은 배경 영상과 입력 영상의 Gray Scale 차를 구하는 것으로 실시간으로 움직이는 객체를 추출하기 위해 움직임 분석 기술에서 보편적으로 쓰이는 방법이다<sup>[11]</sup>. 그러나 두 영상의 Gray Scale 차이 값으로 분리하기 때문에 조명, 그림자, 반사 등의 외부 영향에 따라 발에 구멍이 생기거나 적잖은 잡음이 발생할 수 있다. 이러한 문제점을 보완하기 위해, 모폴로지 (Morphology)기법을 사

용한다. 모폴로지 기법은 침식(Erode), 팽창(Dilate)의 정도에 따라 객체의 구멍을 메우거나 배경영상의 잡음을 제거 할 수 있다. 이러한 전처리 과정의 결과는 그림 3에서 볼 수 있다.

### 2. 실루엣 기반 골격과 관절점 추출

발은 무릎에서 발목에 이르는 관절과 발목에서 발끝까지의 관절이 서로 연결 된 굵은 형태로, 정면과 측면에서 보는 시점에 따라 영상에 투영되는 영역이 매 프레임마다 다르다. 그러므로 발의 특징을 실루엣에서 형태학적 분석으로만 추출하려고 한다면 각 시점에 따라 기하학 제한요소를 많이 주어야 하므로 추적 알고리즘이 복잡해진다.

인체는 관절의 움직임에 제한을 가지고 있다. 인체 구조학적으로 관절은 제한된 각도 이내에서만 회전을 할 수 있으며, 회전축에 따라 골격이 이동한다. 관절체인 발 또한 움직임에 제한을 받으며, 이것은 관절의 회전과 골격 이동의 두 움직임으로 구분될 수 있다. 그러므로 발의 움직임 정보를 취득하려면, 영상에서 발의 관절과 골격의 위치를 취득해야 한다.

실루엣의 중심선을 찾기 위해 골격화를 우선 수행하여 골격을 생성한다. 골격화는 실루엣으로부터 영역 내부에 서로 교차되는 지점을 계산해 내는 방식이다. 골격에 대해 라인 피팅을 수행하여 골격의 기울기를 계산한다. 골격의 기울기는 실루엣 라인을 그림 4에서 보는 바와 같이 외곽 라인을 생성하는 기준이 된다.

찾아낸 골격선들 간의 서로 만나는 지점에서 꺾어지는 정도와 발의 형태학적 비율에 따라 관절점이라 지정하고 발목의 골격선과 발등의 골격선이 만나는 지점을 발의 중심점, 발등의 중심점 반대를 발끝점이라 하여 두 점이 위치하는 2차원 좌표값을 찾아 매 프레임마다 추적한다.

### 3. 2차원에서 3차원 좌표로의 계산

일반적인 자세 추정 시스템은 다수의 카메라를 사용하여 3차원 데이터를 취득하기 때문에 2차원 좌표의 점을 3차원 좌표로 계산하기 위해서 영상 교정 과정을 거친다. 본 논문에서는 한 대의 카메라로 3차원 정보를 취득하므로 영상 교정을 대체할 카메라 정보가 필요하다. 한 대의 카메라를 통해 3차원 정보를 취득하기 위해서는 실제 사용할 환경 조건에 맞춰 카메라의 기하학적 관계를 구성한 다음, 카메라 정보에 대한 수식을 유추해 낼 수 있다.

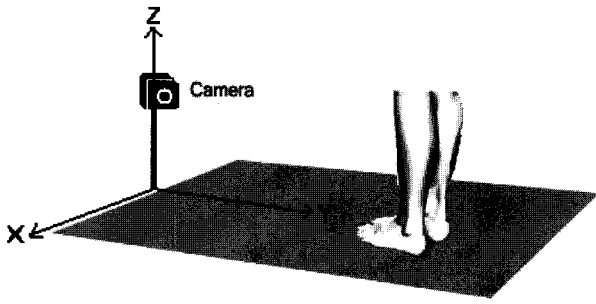


그림 5. 카메라와 발 객체의 실제 좌표계  
Fig. 5. Real Coordination between camera and foot.

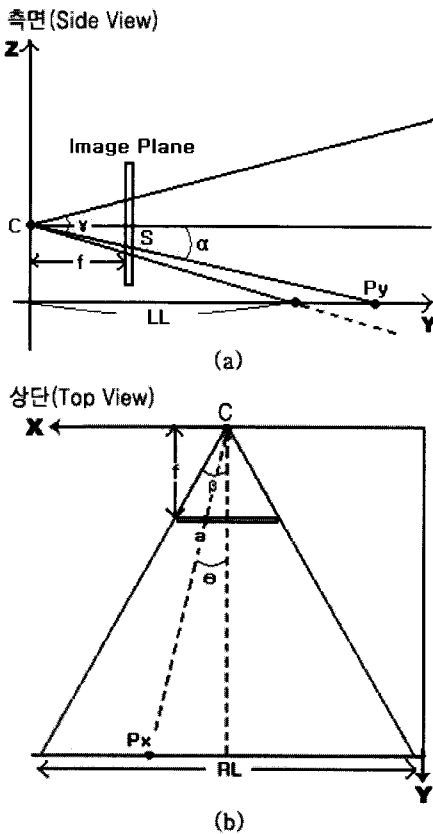


그림 6. 3차원 위치의 계산 (a)Y축에서의 3차원 위치,  
(b)X축에서의 3차원 위치  
Fig. 6. Calculate 3D Positions (a)Y-Coordinate Position,  
(b)X-Coordinate Position.

환경에 따라 카메라 정보를 기하학으로 계산하는 알고리즘은 주로 한 대의 카메라에서 인체의 높이를 계산하는 연구에 적용되고 있다<sup>[12]</sup>. Ronald Poppe의 기하학적 수식은 카메라 한 대를 머리보다 상단에 설치하여 지면을 기준으로 인체의 높이를 쉽게 계산한다. 그러나 지면을 기준으로 3차원 위치값을 취득하기 때문에, 카메라에 투영될 수 있는 유효한 실제 영역의 제한이 크다<sup>[13]</sup>. 본 논문에서 제안한 수식은 발의 정면에 카메라를 설치하여 위치를 추정하는 것으로 계산 방법이 쉬운

것과 더불어 지면과 같은 기준평면이 필요 없다.

2차원 좌표와 3차원 좌표를 대응시키는 기하학적 관계를 가지도록 수식을 구성하기 위해서 3차원 요소를 대신할 최소 개수의 실제측정 파라미터가 필요하다. 이러한 파라미터가 어떤 것인지 알아내고 관계식을 구하기 위해서 구현 환경 조건에 맞춰 기하학적 수식을 구성해야 한다. 그림 5와 6에서 보는바와 같이 카메라 위치를 원점으로 지면에 평행하게 향하고 카메라 양옆을 나누는 좌표성분을 X좌표로, 지면에 평행하면서 카메라가 투영하는 방향을 Y좌표로, 그리고 지면에 수직으로 향하는 정규벡터(Normal Vector) 성분을 Z좌표로 지정했다.

카메라에서 발이 위치하는 Y축의 점  $P_y$ 를 계산하기 위해 이미 알고 있는 영상의 세로 크기(IH: Image Height)와 카메라 높이(h), 카메라에서 카메라 하단에 보이는 위치까지의 거리(LL: Lower Length) 그리고 영상에서의 사이즈(S)를 사용한다. Y축 발의 위치를 계산하기 위해 필요한 카메라의 내부 정보(intrinsic camera information)를 구하는 식은 다음과 같다.

$$\gamma = 2 \times \tan^{-1} \left( \frac{h}{LL} \right) \tag{1}$$

$$f = \frac{IH/2}{\tan(\frac{\gamma}{2})} \tag{2}$$

$$\alpha = \tan^{-1} \left( \frac{S}{f} \right) \tag{3}$$

따라서  $P_y$ 는 다음과 같다.

$$P_y = \frac{h}{\tan \alpha} \tag{4}$$

계산된 점  $P_y$ 는 실제 영상에 투영되는 가로길이(RL: Real Length)와 영상의 가로 크기(IW: Image Width)를 사용하여 점  $P_x$ 를 계산할 수 있다.

$$\theta = \tan^{-1} \left( \frac{|(IW/2) - a|}{f} \right)$$

$$P_x = RL/2 + P_y \frac{a - (IW/2)}{f} \tag{5}$$

위의 식에 따라 시스템에서 필요로 하는 최소 실제 측정 데이터는 카메라 하단거리(LL: Lower Length)와 발을 원점으로 하는 카메라 높이(h) 두 가지 임을 알 수 있다. 이들 식에 대한 자세한 설명은 부록 I에 첨부한다.

#### IV. 재투영

##### 1. 자세 추정

기하학적 수식에 의해 관절점 들의 3차원 위치가 결정되면, 프레임 간의 관절의 회전각과 골격의 실제 이동 값을 계산하여 유사한 자세 정보를 가지도록 3차원 신발모델 데이터를 갱신해야 한다.

3차원 신발의 자세 정보를 갱신하기 위해 이전 프레임의 관절 위치( $F_{prev}$ )를 원점으로, 현재 프레임의 관절 위치( $F_{next}$ )에 이르는 회전각( $\theta$ )을 구한다. 수식은 아래와 같다.

$$\theta = \cos^{-1}\left(\frac{F_{prev} \cdot F_{next}}{|F_{prev}| |F_{next}|}\right) \quad (6)$$

또한 골격의 이동정보를 알아내기 위해 이전 프레임과 현재 프레임 간의 3차원 차이 값을 계산하고, 카메라에 들어오는 발의 골격선과 신발 모델의 골격선에 대한 비율을 구한다.

이렇게 계산해 낸 자세 정보에 맞춰 원점이동, 크기(Scale), 회전(Rotation), 이동(Transformation)의 순서로 변경하면서 신발 모델을 실제 발에 맞게 갱신한다. 신발 모델의 평균 중심점을 찾아 중심점을 원점으로 이동하고, 실제 발의 골격선과 신발 모델의 골격선 사이의 비율에 의해 크기를 조정한다. 앞서 계산한 관절 회전각으로 자세를 변경하고 골격의 이동 정보에 의해 실제 위치로 이동 시킨다.

신발 모델의 3차원 좌표체계와 본 논문에서 제안하는 기하학적 수식의 3차원 좌표체계는 서로 다르므로 투영하기 전에 신발 모델의 좌표계를 기하학적 수식의 좌표계로 변경해야한다.

##### 2. 투영

3차원 신발 모델의 좌표 갱신과 실제 좌표체계로의 변경이 완료되면, 실제로 신발을 신은 것처럼 영상 내의 사용자 발에 신발 모델을 투영해야 한다. 다시 말하면, 3차원 신발 모델을 기하학적 수식에 의해 다시 2차원 좌표값으로 변환하여 2차원 영상에 투영해야 한다.

2차원 영상좌표로 투영하기 위한 3차원 좌표를 점  $R(R_x, R_y, R_z)$ 라하고 기하학적 수식을 다시 적용하여 영상의 가로 크기( $IW$ :Image Width), 점 R에서의 영상에 보이는 실제 길이( $RL'$ : Real Length)와 함께 Y좌표 값  $I_y$ 를 구하는 식은 다음과 같다.

$$I_y = \frac{IH}{2} + f \frac{h}{R_y} - \frac{R_z}{R_y} \max(\text{SIGN}(R_z - h), 0) \quad (7)$$

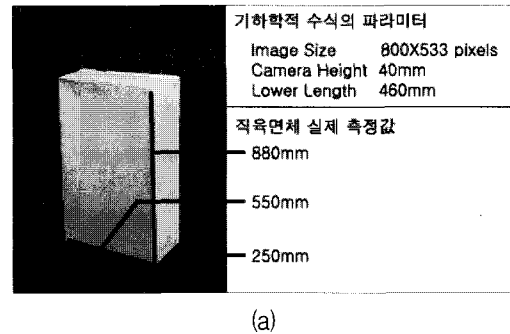
또한, X좌표값  $I_x$ 을 계산하는 식은 다음과 같다.

$$I_x = \frac{IW}{2} + f \frac{R_x - (RL'/2)}{R_y} \quad (8)$$

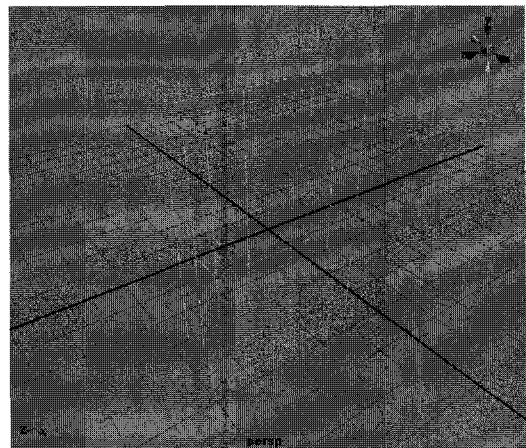
위 식에 관한 전개 과정은 부록II에 첨부한다.

#### V. 실험 결과

본 논문에서 제안하는 3차원 신발 모델 정합 시스템은 영상에 들어오는 발과 3차원 신발 모델의 데이터가 유사하게 정합 되는가에 그 목적을 두고 있다. 3차원 모델 정합(Registration)의 정확도를 결정하는 요소는



(a)



(b)

그림 7. 실험 환경 (a) 직육면체 실제측정값 (b) 3차원 직육면체 모델 데이터

Fig. 7. Experiment environment (a) measured sizes of rectangular parallelepiped (b) 3D Model Data of rectangular parallelepiped.

기하학적 수식과 3차원 자세 추정에 의한 좌표 변환이므로 기하학적 수식에 의한 3차원 변환과 자세 추정이 완료된 3차원 좌표의 재투영으로 나누어 실험했다.

정확도를 위한 실험방법은 추적 대상 물체의 실제 사이즈와 카메라를 원점으로 한 위치정보를 측정된 뒤 논문에서 제안한 방법으로 계산한 값과 비교하여 오차를 측정했다. 그러나 실제 표면 측정이 어려운 발로 정확도를 측정하는데 문제점이 많다는 것을 감안, 실제 사이즈가 측정 가능한 직육면체 모델을 사용하였다.

직육면체를 사용하여 실험한 환경은 다음과 같다. 직육면체 실험 영상은 해상도 800×533으로 촬영된 정지 영상(Still Image)으로 발과 유사한 동작에서 정확도를 측정하기 위해 좌우 회전, 좌우 이동 연속 영상으로 나누어 실험하였다.

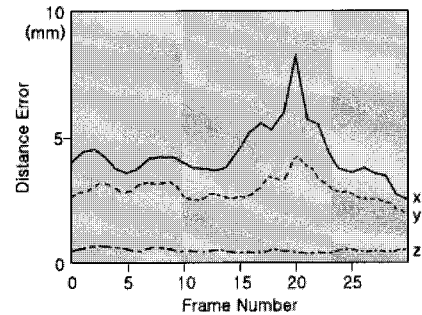
또한 직육면체 3차원 모델은 바닥을 제외한 5개의 평면과 모서리에 일정한 간격으로 24개의 3차원 좌표값을 가지고 있으며, 직육면체에서 실제 측정된 값의 비율로 구성되어 있다.

그림 7은 연속적인 영상에서 검출한 직육면체 2차원 데이터를 기하학적 수식을 사용하여 3차원 좌표값을 계산한 뒤 실제 측정된 사이즈와 비교하여 모든 모델 좌표에 대한 거리차를 x, y, z로 각각 평균을 나타낸 그래프이다. 첫 번째 그래프 (a)는 좌우 회전에 대한 평균 거리차 그래프이며, 두 번째 그래프 (b)는 연속 영상 간 이동에 대한 평균 거리차 그래프이다.

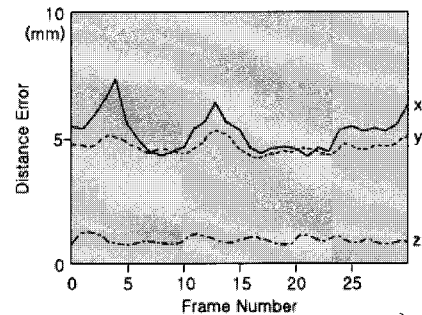
x, y좌표의 오차 범위에 영향을 주는 Lower Length와는 달리, 지면에 고정된 카메라와 발이 지면에 있어야 한다는 움직임 제한은 바닥에 수직인 z좌표의 오차 범위를 줄이는 가장 큰 요소이다. 또한 x와 y값이 거의 유사한 그래프로 오차가 나타나고 있는데, x와 y값 오차값 사이의 비율을 측정하면 실험 영상의 종횡비와 유사한 비율로 나타나는데, 이는 두 좌표가 종횡비에 일정하게 오차를 나타내고 있음을 확인할 수 있다.

대체적으로 두 그래프는 실제 측정값과의 차이가 1cm이내로 거의 실제값과 유사하게 계산되고 있으며, 이는 기하학적 수식이 3차원 좌표값 계산에 효과적으로 사용되고 있음을 확인할 수 있다.

그림 8은 기하학적 수식으로 계산된 3차원 좌표값으로 3차원 직육면체 모델을 변경하고, 그 변경된 값으로 기하학적 수식을 적용해 투영하는 값과 실제 영상에서의 위치 값과의 평균 화소(Pixel)차를 각각 x, y좌표에 대하여 나타낸 그래프이다. 그래프에서 볼 수 있듯이, 거의 유사한 위치에 투영됨을 확인할 수 있다. 기하학

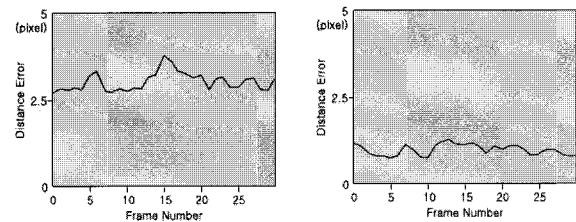


(a)



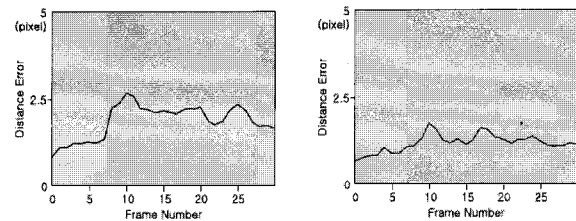
(b)

그림 8. 직육면체의 계산값과 실제길이 차이값 평균 그래프 (a)회전 시 평균 오차 (b)이동 시 평균 오차  
Fig. 8. Average of error distance graph for a rectangular parallelepiped (a) Average of error distance at a rotation (b) Average of error distance at a movement.



(a)

(b)



(c)

(d)

그림 9. 직육면체의 특징점 화소차 평균 (a) X축: 회전시 평균 오차 (b) Y축: 회전시 평균 오차 (c) X축: 이동시 평균 오차 (d) Y축: 이동시 평균 오차  
Fig. 9. Average of error distance for rectangular parallel data (a) X coordinate at a rotation (b) Y coordinate at a rotation (c) X Coordinate at a movement (d) Y Coordinate at a movement.

적 수식을 이용한 3차원 데이터 계산과 투영의 직육면체 실험 결과는 그림 9와 같다. (a)와 (b) 영상은 직육면체가 각각 반시계, 시계방향으로 회전 했을 때, 미리 제작된 3차원 포인트를 제안한 방법에 따라 투영한 결과이다.

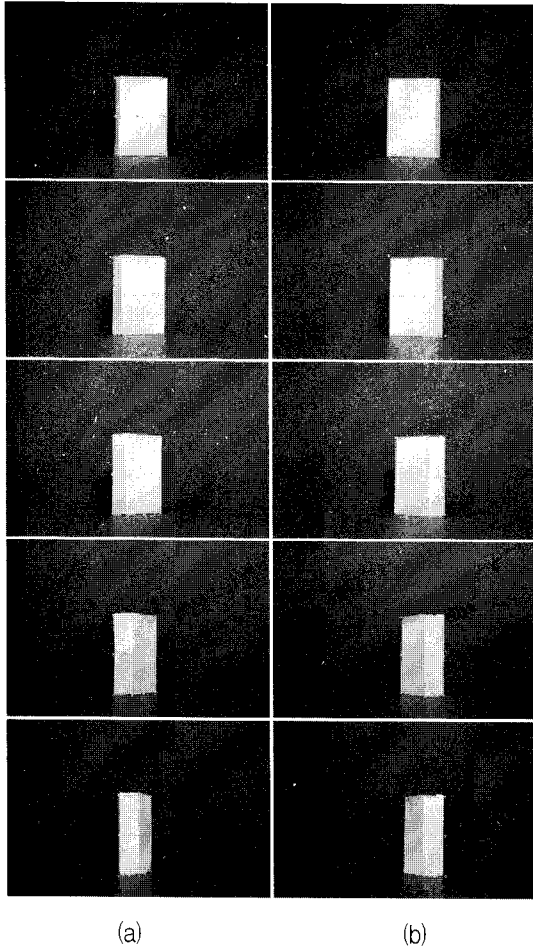


그림 10. 3차원 직육면체 모델 실험결과  
 (a) 반시계 방향 회전 (b) 시계 방향 회전  
 Fig. 10. Experiment Result of Rectangular parallelepiped.  
 (a) counterclockwise rotation  
 (b) clockwise rotation



그림 11. 3차원 신발모델과 실험환경  
 Fig. 11. 3D Shoe Model and Experiment Environment.

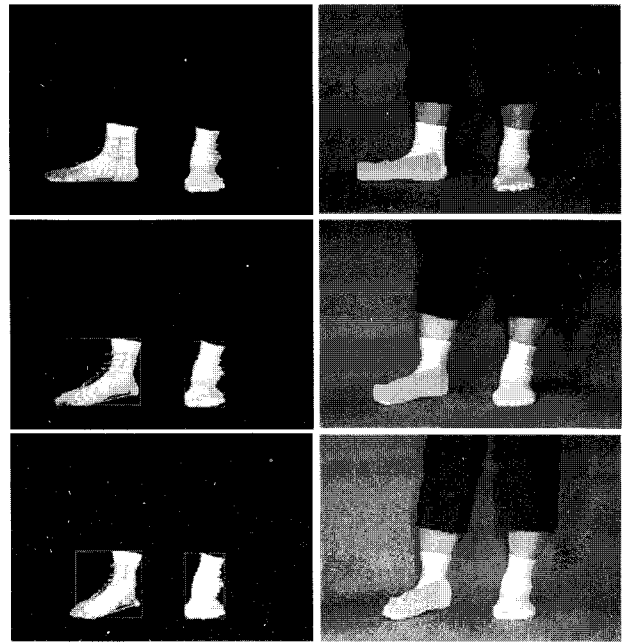


그림 12. 3차원 데이터 투영 결과 (a) 발 모델 투영 (b) 신발 모델 투영  
 Fig. 12. Projection Result for a foot and shoe data (a) Projection for Foot model (b) Projection for shoe model.

직육면체 모델을 사용하여 정확도 측정이 완료된 시스템에 실제 발을 이용하여 정합의 정도를 시뮬레이션 하고 그 결과를 가시적으로 확인했다. 실제 실험 환경은 그림 10과 같다.

실험 영상은 해상도 720×480의 동영상이며, 신발모델 데이터는 실제 신발을 바탕으로 Maya나 3D Max 등의 3D 툴로 생성이 가능한 파일 확장자 obj 파일을 사용하여 제작 하였다.

실험에 앞서 신발모델 데이터가 정합되는 정도를 눈으로 확인하기 쉽도록 몇 가지 동작을 미리 지정했다. 영상에서 발 객체를 쉽게 검출하기 위해 발의 안쪽 면이 보이는 방향을 첫 위치로 지정하였다. 실험 영상에서 사용자의 발은 항상 바닥에 붙어있으며, 매 프레임마다 발이 바닥에 붙도록 하여 큰 움직임 없이 이동한다는 조건하에 실험했다.

실제 발과 신발 모델을 사용하여 나타난 결과를 그림 11에서 확인할 수 있다.

## VI. 결 론

본 논문에서는 정보 통신의 발달 양상에 대응할 수 있는 개별화 된 인터페이스의 개발을 목적으로 사용자가 직접 신발을 신지 않고도 시착(Visual Fitting)할 수



있는 3차원 신발모델 정합 시스템을 제안하고 구현하였다.

시스템 구현에 앞서 실제 상점에서 사용할 수 있는 최소 환경을 설정하고, 환경에 맞게 인체 움직임 분석 기술을 적용하여 전체적인 시스템을 구성하였다. 추적 대상인 발이 특징점이 없는 점을 보완하기 위해 실루엣을 기반으로 골격선과 관절점을 검출, 추적하였고 본문에서 제안하는 기하학적 수식으로 자세 추정과 함께 3차원 신발 모델이 재투영 되는 모습을 실험 결과로 확인할 수 있었다. 기존 인체 움직임 분석 기술들과 달리 발을 중심으로 한 연구라는 특성과 함께 좀 더 정확한 카메라 정보와 발의 형태학 정보를 분석하여 실제 상용화 할 수 있도록 실시간으로 개발이 되어야 할 것이며 이에 대한 연구가 진행 중이다.

## 부 록 I

카메라에서 점 P에서의 실제 가로길이(RL:Real Length)를 둘로 나눈 길이에 대한 각을  $\beta$  라 하면,

$$\beta = \tan^{-1} \left( \frac{IW/2}{f} \right)$$

실제 길이는

$$RL = 2 \times P_y \times \tan \beta$$

구하고자 하는 3차원 x값이 영상의 중심에서 가로축으로 a만큼 위치한다고 할 때,

if( $a \geq 0$ ) 이면,

$$\theta = \tan^{-1} \left( \frac{(IW/2) - a}{f} \right)$$

$$P_x = RL/2 - P_y \times \tan \theta$$

if( $a < 0$ ) 이면,

$$\theta = \tan^{-1} \left( \frac{a - (IW/2)}{f} \right)$$

$$P_x = RL/2 + P_y \times \tan \theta$$

두 가지 경우를 종합하면, 식(5)로 정리 할 수 있다.

## 부 록 II

기하학적 수식의 3차원 y축 길이를 영상의 2차원 y

축 길이(Y)에 맞춰 카메라 높이(h)를 기준으로 3차원 좌표들을 나누어 계산한다고 할 때,

if( $R_z \leq h$ ) 이면,

$$\theta = \tan^{-1} \left( \frac{h}{R_y} \right)$$

$$Y = f \times \tan \theta$$

$$I_y = \frac{IH}{2} + Y$$

if( $R_z > h$ ) 이면,

$$\theta = \tan^{-1} \left( \frac{R_z - h}{R_y} \right)$$

$$Y = f \times \tan \theta$$

$$I_y = \frac{IH}{2} - Y$$

위의 식으로  $I_y$ 를 구할 수 있다.

$I_x$ 를 구하기 위해서 점R이 위치하는 곳의 영상에 들어오는 실제 길이(RL':Real Length)와 점R의 영상에서의 길이값(X)이 필요하며, 이를 이용해 계산되는  $I_x$ 는 다음과 같다.

$$RL' = \frac{IW \times R_y}{f}$$

if( $R_x \leq RL'/2$ ) 이면,

$$\theta' = \tan^{-1} \left( \frac{(RL'/2) - R_x}{R_y} \right)$$

$$X = f \times \tan \theta'$$

$$I_x = \frac{IW}{2} - X$$

if( $R_x > RL'/2$ ) 이면,

$$\theta = \tan^{-1} \left( \frac{R_x - (RL'/2)}{R_y} \right)$$

$$X = f \times \tan \theta$$

$$I_x = \frac{IW}{2} + X$$

## 참 고 문 헌

- [1] 정다운, 마정연, 서용호, 최중수, "실루엣 정보를 이용한 실시간 발의 움직임 추적 시스템," 신호처

- 리합동학술대회 논문집, 제19권, 제 1호, 249쪽, 2006년 9월
- [2] 이동욱, 조주희, 안현상, 이만재, “지능형 거울 시스템을 이용한 넥타이 스타일리스트,” HCI2005, pp. 360-364, 2005년 1월
- [3] H. Saito, J. Hoshino, “A match moving technique for merging CG and human video sequences,” Proc. of IEEE ICASSP '01, vol. 3, pp.1589-1592, 2001
- [4] L. Wang, W. Hu, T. Tan, “Recent Developments in human motion analysis,” Pattern Recognition, vol 36, pp. 585-601, 2003
- [5] S.L. Phung, A. Bouzerdoun, D. Chai, “Skin Segmentation Using Color Pixel Classification: Analysis and Comparison,” IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 148-154, 2005 January
- [6] J.K. Aggarwal, Q. Cai, “Human motion analysis: a review,” Proc. of IEEE Workshop on Motion of Non-Rigid and Articulated Objects, pp. 90-102, 1997 June
- [7] D. Tolani, A. Goswami, N.I. Balder, “Real-Time Inverse Kinematics Techniques for Anthropomorphic Limbs,” Graphics Models, vol. 62, no. 5, pp. 353-388, 2000
- [8] C. Wren, A. Azarbayejani, T. Darrell, and A. Pentland, “Pfinder: Real-time Tracking of the human body,” IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 780-785, 1997
- [9] S.E. Kim, R.H. Lee, C.J Park, I.H. Lee, “MIMIC: Real-time marker-free motion capture system to create an agent in the virtual space,” Proceedings of the ICCE '02, vol. 1, pp. 48-49, 2002 December
- [10] D. Tolani, A. Goswami, N.I. Badler, “Real-time Inverse kinematics techniques for anthropomorphic limbs,” Graphical Models and Image Processing, vol. 62, no. 5, pp. 353-388, 2000 September
- [11] M. Piccardi, “Background Subtraction Techniques: a review,” International Conference on Systems, Man, and Cybernetics, vol. 4, pp. 3099-3104, 2004
- [12] P. Eu, L. Ling, “Human skeleton proportions from monocular data,” LNCS Journal of Zhejiang University SCIENCE A, vol. 7, no. 7, pp. 1266-1274, 2006 July
- [13] Ronald Poppe, “Real-time Pose Estimation from monocular image sequences using Silhouettes,” Human Media Interaction Group (HMI), University of Twente the Netherlands, Master's thesis, 2004 April

저 자 소 개



정 다 운(학생회원)  
 2001년 덕성여자대학교 전산학과  
 학사 졸업.  
 2007년 중앙대학교 첨단영상대학  
 원 영상공학과 석사 졸업.  
 2007년~현재 중앙대학교  
 첨단영상대학원  
 영상공학과 박사 과정.

<주관심분야 : 컴퓨터비전, 영상합성, 컴퓨터 그  
 래픽스>



최 종 수(평생회원)  
 1975년 인하대학교 전기공학과  
 학사 졸업.  
 1977년 서울대학교 대학원  
 전자공학과 석사 졸업.  
 1981년 Keio University(일본)  
 전기공학과 박사 졸업.

1981년 2월~1981년 9월 Aloka Co. Ltd 연구소  
 연구원

1981년 9월~1999년 8월 중앙대학교 전자공학과  
 교수

1999년 9월~현재 중앙대학교 첨단영상대학원  
 영상공학과 교수

<주관심분야 : 영상신호처리, 컴퓨터 비전, 증강  
 현실, 영상합성, 영상통신>



서 응 호(학생회원)  
 2002년 중앙대학교 전자전기  
 공학부 학사 졸업.  
 2005년 중앙대학교 첨단영상대학  
 원 영상공학과 석사 졸업.  
 2005년~현재 중앙대학교 첨단영  
 상대학원 영상공학과 박  
 사과정 수료.

<주관심분야 : 영상처리, 컴퓨터 비전>