논문 2005-42SC-4-2

인간-컴퓨터 상호 작용을 위한 인간 팔의 3차원 자세 추정 - 기계요소 모델링 기법을 컴퓨터 비전에 적용

(3D Pose Estimation of a Human Arm for Human-Computer Interaction - Application of Mechanical Modeling Techniques to Computer Vision)

한 영 모*

(Youngmo Han)

요 약

인간은 의사 표현을 위해 음성언어 뿐 아니라 몸짓 언어(body languages)를 많이 사용한다. 이 몸짓 언어 중 대표적인 것은, 물론 손과 팔의 사용이다. 따라서 인간 팔의 운동 해석은 인간과 기계의 상호 작용(human-computer interaction)에 있어 매우 중요하다고 할 수 있다. 이러한 견지에서 본 논문에서는 다음과 같은 방법으로 컴퓨터비전을 이용한 인간팔의 3차원 자세 추정 방법을 제안하다. 먼저 팔의 운동이 대부분 회전 관절(revolute-joint)에 의해 이루어진다는 점에 착안하여, 컴퓨터 비전 시스템을 활용한 회전 관절의 3차원 운동 해석 기법을 제안한다. 이를 위해 회전 관절의 기구학적 모델링 기법(kinematic modeling techniques)과 컴퓨터 비전의 경사 투영 모델(perspective projection model)을 결합한다. 다음으로, 회전 관절의 3차원 운동해석 기법을 컴퓨터 비전을 이용한 인간 팔의 3차원 자세 추정 문제에 응용한다. 그 기본 발상은 회전 관절의 3차원 운동 복원 알고리즘을 인간 팔의 각 관절에 순서 대로 적용하는 것이다. 본 알고리즘은 특히 유비쿼터스 컴퓨팅(ubiquitous computing)과 가상현실(virtual reality)를 위한 인간-컴퓨터 상호작용(human-computer interaction)이라는 응용을 목표로, 고수준의 정확도를 갖는 폐쇄구조 형태(closed-form)의 해를 구하는데 주력한다.

Abstract

For expressing intention the human often use body languages as well as vocal languages. Of course the gestures using arms and hands are the representative ones among the body languages. Therefore it is very important to understand the human arm motion in human-computer interaction. In this respect we present here how to estimate 3D pose of human arms by using computer vision systems. For this we first focus on the idea that the human arm motion consists of mostly revolute joint motions, and then we present an algorithm for understanding 3D motion of a revolute joint using vision systems. Next we apply it to estimating 3D pose of human arms using vision systems. The fundamental idea for this algorithm extension is that we may apply the algorithm for a revolute joint to each of the revolute joints of human arms one after another. In designing the algorithms we focus on seeking closed-form solutions with high accuracy because we aim at applying them to human computer interaction for ubiquitous computing and virtual reality.

Keywords : human-computer interaction, 3D pose estimation, human arm , revolute joint, computer vision, closed-form solution.

I. 서 론

최근 들어 유비쿼터스 컴퓨팅 (ubiquitous computing)과 가상현실 (virtual reality)의 연구가 활발히 진행되고 있는데 그 내용을 살펴보면 다음과 같다.

유비쿼터스 컴퓨팅은 생활 환경안의 도처에 통신 기능을 가진 컴퓨터를 눈에 띄지 않게 묻은 후, 이들로 하여금 인간의 움직임을 검출시키거나 컴퓨터를 서로 제휴시키는 방법을 통해서, 인간의 생활을 지원하려고 하는 것이다. 결국 유비쿼터스 컴퓨팅에서 제

* 정회원, 이화여자대학교 정보통신공학과
(Dept. of Information Electronics Ewha Womans University)

※ 이 논문은 2005년도 두뇌한국21사업에 의하여 지원되었음.
접수일자: 2005년1월25일, 수정완료일: 2005년6월15일

안하는 환경은 컴퓨터가 여러 가지 가전 기기와 결합하고 또 그 컴퓨터 간에 network을 형성하여, 인간의 편의를 위한 종합적인 서비스를 제공하는 것이라 할 수 있다. 이러한 유비쿼터스 컴퓨팅을 실현시키기 위해서는 컴퓨터의 소형화 - 마이크로컴퓨터를 가전제품 안에 내장시키기 위해, 즉 내장형 시스템(embedded system) 형태의 구현을 위해, 통신 및 네트웍 기술의 발달 - 예를 들면 홈 네트워킹 (home networking)과 같은 - 과 함께 또한 컴퓨터와 사람 간의 의사소통 즉, 인간-컴퓨터 상호작용(human-computer interaction)의 기술 발달이 필요하게 된다. 그리고 인간-컴퓨터 상호작용을 구현하기 위해서는 컴퓨터 비전을 활용한 인체의 공간 추적 (human body spatial tracking) 및 인간의 의지 읽기 (human intention reading), 가정용 홈 로봇 (home robot), 촉각 제어 (haptic control) 등의 핵심기술의 발달이 필요하게 된다.

가상현실 기술은 3차원 컴퓨터 그래픽스 기술을 활용한 것으로서, 단지 컴퓨터 기술 뿐 아니라 사람의 모든 감각-시각, 청각, 촉각-과 연결되어 실제와 같은 3차원 환경을 구성하게 하는 데 그 목적이 있다. 이를 위해서는 사용자의 움직임 정보를 추적할 수 있는 공각 추적 장치 (spatial trackers), 3차원 환경을 보여주기 위한 Head-Mount Display(HMD)를 비롯한 display 장치, 3차원 환경에서의 소리의 움직임을 재현할 수 있는 3차원 음향장치, 감촉이나 온도, 힘이나 토오크(torque)를 느낄 수 있게 해주는 촉각 디스플레이(haptic display) 장치 등이 필요하다. 이때 사용자의 움직임 정보를 추적할 수 있는 공간 추적 장치의 구현은 컴퓨터 비전을 이용한 인간-컴퓨터 상호작용 기술을 통해 효과적으로 구현될 수 있다.

이처럼 근래에 활기를 띄고 있는 유비쿼터스 컴퓨팅과 가상현실의 구현을 위한 공통적인 핵심 기술이 바로 컴퓨터 비전을 활용한 인간-컴퓨터 상호작용이라 할 수 있다. 이때 주목할 부분은 실제 인간의 몸짓 언어와 활동의 상당 부분은 주로 손, 팔, 다리 등에 그 중점이 맞추어져 있다는 점이다. 따라서 인간-컴퓨터 상호 작용의 주 대상체가 바로 손, 팔, 다리 등에 맞추어 지는 것이 효과적이라 할 수 있다. 이러한 견지에서 본 논문에서는 컴퓨터 비전 기술을 활용한 인간의 팔 운동 해석 기법을 다루고자 한다.

컴퓨터비전을 이용한 인체의 운동해석 문제는 그 동안에도 많은 학자들에 의해 연구된 바 있는데, 이들은 크게 운동 검출(motion detection)과 운동 추적(motion tracking)의 두 부류로 나누어 질 수 있다. 전자는 이산적(discrete)으로 주어진 영상 사이의 상대적인 운동 값을 직접적으로 추출해 내는 것이고, 후자는 연속적으로 주어지는 영상정보로부터 운동의 변화량을 추출한 후 이를 더해서 현재의 운동 값을 간접적으로 추정해 내는 기법이다.

일반적으로 운동검출(motion detection)을 구현할 때에는 모델 정합 방법(model-fitting approach)을 많이 사용 한다[1], [2], [3], [4], [5], [6], [7], [8]. 이는 인체의 운동 모델을 설정한 후 주어진 영상정보와 비교하여 모델 파라미터를 찾아내는 일종의 3차원 등록(3D registration) 기법이라 할 수 있다. 그러나 인체의 운동 모델은 흔히 심각한 비선형(non-linear)형태로 나타나고 그 매개 변수의 개수도 많아서 일반적으로 모델 정합에 많은 계산 량이 필요하게 된다. 그리고 그 해도 여러 개가 존재하게 될 뿐만 아니라, 그 해조차도 수치해석(numerical analysis)을 통해서만 얻을 수 있게 되는 불편함이 있다. 이 문제를 해결하고자 Taylor는 주어진 인체의 링크(link) 길이로부터 각 링크의 운동 파라미터를 순서 데로 하나씩 복원하는 방법을 주창하였다[9]. 그러나 Taylor가 사용한 방법은 지나치게 단순화된 모델을 사용함으로써 그 결과의 정확성이 떨어질 수 있다는 단점이 있다.

반면 운동추적(motion tracking)을 구현할 때에는, 일반적으로 인체의 운동 모델의 미분형태 모델을 사용하여 운동의 증분(increment)를 찾아내게 된다.[10],[11] 흔히 이 미분 형태 모델은 선형 형태로 나타나기 때문에 계산 량이 적다는 장점이 있다. 그러나 이 운동의 증분을 적분해야만 우리가 찾고자 하는 인체 운동 값을 구할 수 있다는 단점이 있다. 왜냐하면 적분을 계산하는 불편함과 측정에러가 적분을 통해 누적될 우려가 있기 때문이다. 그리고 이 누적된 오차는 결국 인체 운동 해석의 실패로 이어질 위험이 있다. Yamamoto 등은 다양한 운동의 제약조건을 운동 추적 문제에 포함시킴으로써 측정오차를 줄여 보고자 하였다[12]. Bregler 와 Malik는 트위스트(twist)와 지수형 맵(exponential map)의 개념을 도입하여 인체 모델의 파라미터의 개수를 최적화하려 하였다[10].

본 논문에서는 인간 팔의 운동검출을 위한 효과적인 컴퓨터 비전 알고리즘을 제안하고자 한다. 본 알고리즘은 특히 유비쿼터스 컴퓨팅과 가상현실을 위한

인간-컴퓨터 상호 작용이라는 응용을 목표로 하기 때문에, 실시간 사용이 가능해야 하고 작은 몸짓도 읽을 수 있는 높은 신뢰도가 요구된다. 따라서 고 수준의 정확도를 갖는 폐쇄구조 형태(closed-form)의 알고리즘 구현이 필요하게 되는 데, 이를 위해 다음과 같은 접근법을 사용한다. 먼저 팔의 운동이 대부분 회전 관절(revolute-joint)에 의해 이루어진다는 점에 착안하여, 비전을 이용한 회전 관절의 운동추출 알고리즘을 제안한다. 이를 위해 회전관절의 기구학적 모델링 기법(kinematic modeling techniques)과 컴퓨터 비전의 경사투영 모델 (perspective projection model)을 결합한다. 다음으로, 회전 관절의 운동추출 알고리즘을 인간 팔의 각 관절에 순서 데로 적용하는 방법을 통해, 인간 팔의 운동추출 알고리즘으로 확장한다.

Ⅱ. 컴퓨터 비전을 이용한 회전 관절 (Revolute Joint)의 운동추출 (Motion Detection) 알고리즘

이 절에서는 컴퓨터 비전을 이용한 회전관절의 3차원 운동 복원 방법을 제안하고자 한다. 이를 위해 회전관절의 기구학적 모델(kinematic model)과 카메라의 경사투영 모델(perspective projection model)을 결합한다.

1. 회전 관절의 기구학적 모델

그림 1의 회전 관절을 살펴보자. 이 그림에서 회전관절 상의 모델 기준점 $Y_r \in R^3$ 이 회전축 $w \in R^3$ 주위로 $\theta \in R$ 만큼의 관절 회전에 의해 $Y \in R^3$ 로 이동하였다고 하자. 그리고 C를 회전축 상의 임의의 점이라 하자. 이 경우의 기하학적 구조를 살펴보면, 벡터 $(Y-C)$의 w에 평행한 방향 성분이 다음과 같이 주어짐을 알 수 있다.

$$\hat{w}\hat{w}^T(Y_r-C)$$

여기서, $\hat{w} = w/\|w\|$로서 w 방향의 단위 벡터를 나타낸다.

또한 벡터 $(Y-C)$의 w에 수직한 방향 성분은 두 개의 벡터 $(I-\hat{w}\hat{w}^T)(Y_r-C)$ 와 $\hat{w}\times\{(I-\hat{w}\hat{w}^T)(Y_r-C)\}$의 선형결합으로 표현됨을 알 수 있다. 따라서 두 개의 스칼라 매개변수 α_1, α_2에 대해 다음과 같은 관계식을 얻을 수 있다.

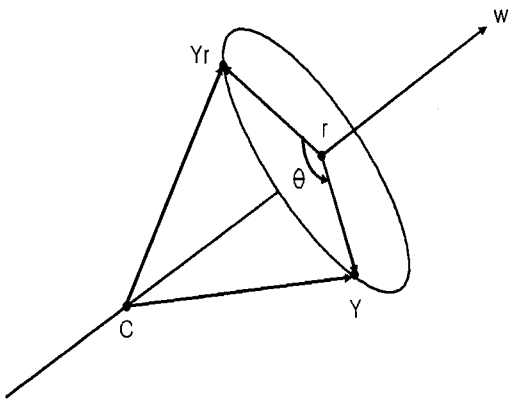

그림 1. 회전 관절의 운동. C는 회전축 w상의 임의의 점이다.

Fig. 1. Revolute Joint Motion. C is an Arbitrary Point on the Rotation Axis w.

$$(Y-C) = k_1 + \alpha_1 k_2 + \alpha_2 k_2 \tag{1}$$

여기서,

$k_1 = \hat{w}\hat{w}^T(Y_r-C)$, $k_2 = (I-\hat{w}\hat{w}^T)(Y_r-C)$,

$k_3 = \hat{w}\times(I-\hat{w}\hat{w}^T)(Y_r-C)$ 이다.

이 때, $\{k_1, k_2, k_3\}$는 R^3 상에서 직교집합을 이룸을 쉽게 확인할 수 있다.

또한 회전관절은 강체 운동을 하므로 다음의 관계식이 성립해야 한다.[13]

$$\|Y - C\|^2 = \|Y_r - C\|^2 \tag{2}$$

식 (1)을 식 (2)에 대입하면 다음 식을 얻게 된다.

$$\alpha_1^2 + \alpha_2^2 = 1 \tag{3}$$

2. 비전(Vision) 시스템의 카메라 투영 모델

그림 2에서 묘사된 컴퓨터 비전 시스템을 생각해 보자. 좌표계 {CF}는 카메라 좌표계를 나타내고 카메라에 부착되어 있다. 그리고 카메라 좌표계의 3축은 각각 I, J, K로 표현되어 있는데, 이 중 K축은 카메라 렌즈의 중심을 통과하고, 렌즈 중심점에서 렌즈에 수직하게 놓여 있다. {IF}는 이미지 평면에 부착되어 있는 이미지 좌표계이다. 이미지 좌표계의 3축은 각각 i, j, k로 표현되어 있는데, 이 중 k축은 {CF}의 K축과 일치하게 놓여 있다. 그리고 {CF}의 원점과 {IF}의 원점 사이 거리를 초점 거리 f라 한다.

주어진 강체(rigid body) 상의 한 점 $Y \in R^3$이 경사

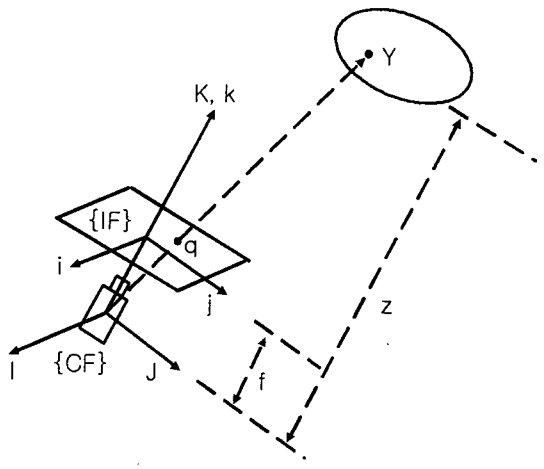

그림 2. 비전(vision) 시스템
Fig. 2. Vision System.

투영(perspective projection)에 의해 이미지 상에 q 점으로 투영되었다고 하자. 이 경우 다음의 관계식이 성립한다.[14]

$$Y=(x,y,z)^T$$
$$q=f(Y/z)=f(x/z,y/z,1)$$

여기서 x,y,z는 각각 Y의 세 축 I,J,K 방향(카메라 좌표계 {CF}에 대해) 성분을 나타낸다.

본 논문에서는 편의상 scaled depth z/f를 depth z로 표현하고자 한다. 그러면 카메라의 경사 투영 모델은 다음과 같이 표현된다.

$$q=(Y/z)=(x/z,y/z,1) \tag{4}$$

3. 알고리즘

식 (4)을 식 (1)에 대입한 후 간단한 수학적 조작을 거치면 다음 관계식을 얻게 된다.

$$\alpha_1=m_{11}z+m_{12},\ \alpha_2=m_{21}z+m_{22} \tag{5}$$

여기서,

$m_{11}=(k_2^Tq)/\|k_3\|^2$, $m_{12}=-(k_2^TC)/\|k_3\|^2$,

$m_{21}=(k_3^Tq)/\|k_3\|^2$, $m_{22}=-(k_3^TC)/\|k_3\|^2$

이다.

식 (4) 와 식 (5)을 식 (1)에 대입하여 다음 관계식을 얻는다.

$$q=n_1+n_2a \tag{6}$$

여기서, $a=1/z$, $n_1=m_{11}k_2+m_{21}k_3$,

$n_2=C+k_1+m_{12}k_2+m_{22}k_3$ 이다.

또한 식 (4) 와 식 (5)을 식 (1)에 대입하여 다음 관계식을 얻는다.

$$s_1+2s_2a+s_3a^2=0 \tag{7}$$

여기서,

$$s_1=m_{11}^2+m_{21}^2\ ,\ s_2=m_{11}m_{12}+m_{21}m_{22}\ ,$$
$$s_3=m_{12}^2+m_{22}^2-1$$

결과적으로, 비전 시스템을 사용한 회전 관절의 3차원 운동 복원 문제는 다음에 주어진 제약조건을 갖는 목적 함수의 최소화 문제로 귀결될 수 있다.

- 목적함수 : $J(a)=\|\tilde{q}-(n_1+n_2a)\|^2$ (8)
- 제약조건 : $s_1+2s_2a+s_3a^2=0$ (9)

여기서, $\tilde{q}$는 q의 측정값을 나타낸다.

일반적으로 제약조건을 갖는 최소화 문제는 Lagrangian multiplier를 사용하여, 제약조건을 최적화 문제 안에 함께 묶을 수 있다.[15] 이를 위해 다음의 Hamiltonian을 정의하자.

$$H(a,\lambda)=\|\tilde{q}-(n_1+n_2a)\|^2 + \lambda(s_1+2s_2a+s_3a^2)$$

여기서, $\lambda\in R$은 Lagrangian multiplier이다.

최적화 일차 필요조건 $\partial H/\partial a=0$, $\partial H/\partial\lambda=0$을 풀면 다음 식을 얻는다.

$$a=(n_2^T(\tilde{q}-n_1)-\lambda s_2)/(\|n_2\|^2+\lambda s_3) \tag{9}$$

여기서,

$\lambda=-\gamma_2+\sqrt{\gamma_2^2-\gamma_1\gamma_3}$ 또는 $\lambda=-\gamma_2-\sqrt{\gamma_2^2-\gamma_1\gamma_3}$

의 값을 갖는다.

그리고

$$\gamma_1=\|n_2\|^4s_1+2n_2^T(\tilde{q}-n_1)\|n_2\|^2s_2 + (n_2^T(\tilde{q}-n_1))^2s_3\ ,$$
$$\gamma_2=\|n_2\|^2s_1s_3+s_2(n_2^T(\tilde{q}-n_1)s_3-\|n_2\|^2s_2) - n_2^T(\tilde{q}-n_1)s_2s_3\ ,$$
$$\gamma_3=s_1s_3^2-s_2^2s_3$$

일단 a가 구해지면 $Y=q/a$을 계산할 수 있다.

편의상 $u'=(I-\hat{w}\hat{w}^T)(Y_r-C)$,

$v'=(I-\hat{w}\hat{w}^T)(Y-C)$ 라 놓자. 그러면 그림 1의 기하학적 구조로부터, 관절의 회전각 θ는, $u'=0$이 아닌 경우에 다음 관계식을 만족함을 알 수 있다 ($u'=0$인 경우는 Y_r이 w축 상에 놓여 있는 경우로서, 우리의 관심 대상이 아니다).

$$w^T(u'\times v') = w\sin(\theta)\|u'\|\|v'\|$$

$$u'^T v' = \cos(\theta)\|u'\|\|v'\|$$

위의 두 식을 연립하여 풀면 관절의 회전각을 구할 수 있다.

$$\theta = atan2(\hat{w}^T(u'\times v'), u'^T v') \quad (10)$$

Ⅲ. 컴퓨터 비전을 이용한 인간 팔(Human Arm)의 운동추출(Motion Detection) 알고리즘 및 결과

기본적으로 인간의 팔은 여러 개의 링크가 회전 관절로 연결된 구조로 간주될 수 있다. 예를 들어 팔꿈치는 회전 관절 1개, 손목과 어깨는 회전 관절 3개가 서로 수직하게 놓여있는 것으로 모델링 될 수 있다. 따라서 이 회전 관절의 3차원 운동을 순서 데로 복원하면 결국 팔의 자세를 구할 수 있게 되는 것이다. 그리고 각 회전 관절의 3차원 운동은 앞 절에서 살펴 본 컴퓨터 비전을 이용한 회전관절의 3차원 운동 복원 알고리즘을 사용하여 구할 수 있다.

본 알고리즘의 성능 검증을 위해 그림 3에 주어진 실험 예제를 고려해 보자. 본 논문에서 제안한 알고리즘을 적용하기 위해서는 먼저 팔의 골격(skeleton)을 구해야 하는데, 이는 몇 가지 영상처리를 거쳐 이루어 질 수 있다. 일반적으로 골격은 윤곽선 검출(edge detection)이 행해진 영상에서 구하기가 더 용이하므로, 그림 3의 영상에 윤곽선 검출(edge detection) 과정을 적용한다. 그림 4에 그 결과가 주어져 있다. 그림 4에서 볼 수 있듯이 일반적으로 윤곽선 검출을 행한 영상에는 작은 영상의 잔해가 남게 된다. 이 잔해는 다시 thresholding을 거침으로써 대부분 제거 될 수 있다(그림 5 참조). 이제 팔의 윤곽선이 찾아졌으므로 팔의 윤곽선의 중심점들을 찾아 연결한 후, 직선 정합(line fitting)을 통해 골격을 구할 수 있다 (물론 Zhang-Suen Skeletonizing 알고리즘[16]을 직접 적용하는 것도 가능하다).

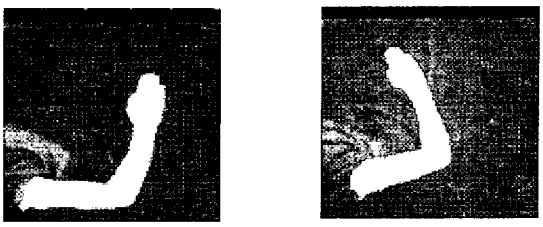

그림 3. 인간 팔의 운동 영상
Fig. 3. Images of Human Arm Motion.

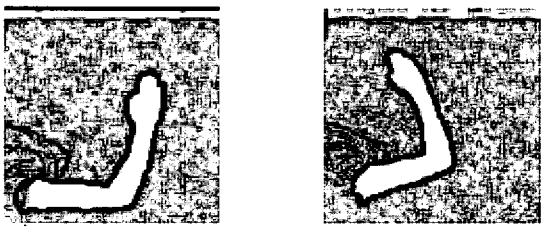

그림 4. 윤곽선 추출 후의 영상
Fig. 4. Images after Edge Detection.

그림 5. Thresholding 후의 영상
Fig. 5. Images after Thresholding.

그림 6. 팔의 골격이 점선으로 표시되어 있다.
Fig. 6. The skeletons of the human arms are denoted by the dotted lines.

찾아진 골격은 모두 3개의 링크 (link)로 표현 되고, 이들은 다시 어께, 팔꿈치, 손목의 회전관절로 연결되어 있다. 물론 손목과 어깨는 3방향의 회전이 가능한 구형 관절 (spherical joint)이라는 특징이 있는데, 이는 3개의 회전 관절이 서로 수직하게 놓여있는 것으로 간주될 수 있다. 이와 같은 문제 해석을 통해 2절에서 제안한 알고리즘을 차례로 어깨 -> 팔꿈치 -> 손목 순으로 세 개의 관절에 적용할 수 있다. 그 결과가 표 1에 주어져 있다.

이 표에서 성능평가와 비교를 위해서 본 논문에서 제

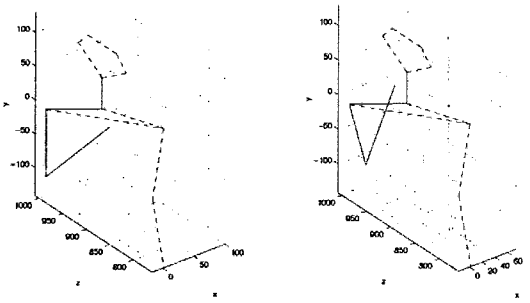

그림 7. 아령 들기 시뮬레이션의 2 프레임.
Fig. 7. Two frames in simulations of lifting dumbbells.

표 1. 측정한 값과 본 논문의 알고리즘을 사용해 계산된 값을 퍼센트 에러로 표현하였다. 비교를 위해 본 논문에서 제안한 것과 같은 종류의 해 (closed-form solution)를 제안한 Taylor의 알고리즘[9]도 함께 고려하였다.

Table 1. Percentage Errors between the measured values and the computed ones by using our algorithm are presented. For performance comparison the results in the case of using Taylor's algorithm[9] are also presented.

	어깨	팔꿈치	손목
본 논문의 방법	6.25	9.86	9.38
Taylor의 방법	24.53	23.49	21.83

표 2. 그림 7의 2 프레임에 연속하여 본 논문의 알고리즘을 적용해 계산된 값과 참값 사이의 퍼센트 에러 값 (잡음 수준이 $\sigma = 0.3$인 경우). 비교를 위해 본 논문에서 제안한 것과 같은 종류의 해 (closed-form solution)를 제안한 Taylor의 알고리즘[9]도 함께 고려하였다.

Table 1. Percentage Errors between the true values and the computed ones by applying our algorithm successively to the two frames in Fig. 7 are presented. For performance comparison the results in the case of using Taylor's algorithm[9] are also presented.

	어깨	팔꿈치	손목
본 논문의 방법	4.28	5.47	4.64
Taylor의 방법	9.91	9.64	9.29

안한 알고리즘과 같은 종류의 해 (closed-form solution)를 제안한 Taylor의 알고리즘[9]도 함께 고려하였다. 표 1의 결과를 살펴보면, 본 논문에서 제안한 방법이 Taylor의 방법에 비해 보다 정확한 결과를 주는 것을 알 수 있다.

좀 더 정교한 성능 평가를 위해서 그림 7에 주어진 시뮬레이션을 고려해 보자. 일반적으로 시뮬레이션은 실험

표 3. 그림 7의 2 프레임에 연속하여 본 논문의 알고리즘을 적용해 계산된 값과 참값 사이의 퍼센트 에러 값 (잡음 수준이 $\sigma = 1.5$인 경우). 비교를 위해 본 논문에서 제안한 것과 같은 종류의 해 (closed-form solution)를 제안한 Taylor의 알고리즘[9]도 함께 고려하였다.

Table 3. Percentage Errors between the true values and the computed ones by applying our algorithm successively to the two frames in Fig. 7 are presented. For performance comparison the results in the case of using Taylor's algorithm[9] are also presented.

	어깨	팔꿈치	손목
본 논문의 방법	5.52	10.23	10.26
Taylor의 방법	46.23	43.28	41.39

에 비해 파라미터 값을 정확히 통제할 수 있고, 그 결과 참값을 정확히 산출해 낼 수 있다는 장점이 있다. 따라서 각 파라미터에 대한 알고리즘의 민감도를 확인할 때 유용하다. 그림 7은 왼팔로 아령을 드는 사람의 운동을 2개의 프레임으로 표현한 것이다. 이 2개의 프레임에 본 알고리즘을 적용하여 오차 값을 확인하려 한다. 이때 영상의 측정 오차 값을 조절하여, 측정 오차 값에 따른 알고리즘의 정확도를 비교해 보고자 한다. 이 시뮬레이션에서 측정 오차는 표준편차가 σ인 가우시안 잡음 (Gaussian noise)으로 선택하였다.

$\sigma = 0.3$과 $\sigma = 1.5$인 경우의 결과가 각각 표2와 표3에 주어져 있다. 이 표에서 볼 수 있듯이 본 알고리즘은 Taylor의 방법에 비해 잡음 수준 σ에 덜 민감하고, 잡음 수준이 높을 때 상대적으로 정확한 결과 값을 주는 것을 알 수 있다.

Ⅳ. 결론 및 추후 연구

본 논문에서는 인체가 기본적으로 회전관절로 구성되어 있다는 것에 착안하였다. 그래서 먼저 컴퓨터 비전을 사용한 회전 관절의 3차원 운동 복원 알고리즘을 닫힌 형태(closed-form)로 제안하였다. 그리고 다시 이를 확장해서 컴퓨터 비전을 사용한 인간 팔의 3차원 자세 추정 문제에 적용하였다. 이 과정에서 컴퓨터 비전의 경사투영 모델 (perspective projection model)과 회전 관절의 기구학적 모델 (kinematic model)을 분석해 결합하였다.

그리고 동종 형태의 해 (닫힌 형태의 해)를 제안한 Taylor의 방법과 비교 검토한 결과 본 논문에서 제안한 방법이 보다 정확한 3차원 자세 추정 결과를 주는 것을

알게 되었다.

이러한 결과의 차이는 Taylor의 방법에서 사용된 인체 운동 모델이 지나치게 단순해서 인체 운동의 현실성과 상당한 차이가 있는데 반해, 본 논문에서 제안한 방법에서는 인체 운동의 기구학을 세밀히 검토한 정교한 모델을 사용한 것이 그 원인이라 생각된다.

이는 결국 컴퓨터 비전을 활용한 인체의 자세 복원 문제에서 기계요소 모델링 기법이 매우 중요하게 사용될 수 있음을 시사하는 결과이기도 하다. 즉, 컴퓨터 비전의 인체 운동 복원 문제에서는 컴퓨터, 전기, 기계, 물리 등의 많은 인접 분야의 상호 보완적 협력이 절실히 요구됨을 시사하는 것이라 할 수 있다.

앞으로의 추후 진행되어야 할 연구를 살펴보면 다음과 같다. 먼저, 본 논문에서는 인체의 모든 관절을 회전관절로 표현하였다. 그래서 손목과 어깨의 구형 관절(spherical joint)도 회전관절 3개가 수직으로 배치된 형태로 간주하였다. 그러나 이 모델은 손목의 구형관절에는 비교적 잘 맞으나, 어깨의 경우는 조금 차이가 있을 수 있다. 왜냐하면 어깨는 3방향의 회전 뿐 아니라 움츠리고 펼 수 있는 운동도 약간이나마 가능하기 때문이다. 따라서 추후 연구에서는 어깨 부위의 모델링을 위해 회전 뿐 아니라 전이 변위의 가능성 까지 고려해서 모델링 - 예를 들면 회전관절과 전이관절(prismatic joint)의 결합 형태, 또는 병렬 메카니즘(parallel mechanism)을 사용한 형태 - 하는 연구가 진행되어야 할 것이다.

다음으로, 본 연구에서는 인간팔의 인식 (recognition)이 사전에 이루어져 있다고 가정 하였다. 그러나 실제 응용에서 좀더 유용하게 본 알고리즘이 사용되기 위해서는 초기 상태에서의 팔의 형상 인식 (pattern recognition) 알고리즘과 결합되는 것이 바람직하다. 보통 인체의 형상 인식 문제에서는 크게 색채에 기반 한 방법(color-based approach)과 형태에 기반 한 방법(shape-based approach)을 생각할 수 있다. 이 중 색채에 기반 한 방법은 매우 빠른 인식속도가 장점이지만, 옷을 입고 있는 부위는 인식이 어렵다는 한계가 있다. 반면 형태에 기반 한 방법은 옷을 입어도 인식이 가능하다는 장점이 있는 반면, 인식 속도가 매우 느려서 실시간 적용이 어렵다는 단점이 있다. 그 이유를 살펴보면, 이 접근법은 흔히 템플릿 비교(template matching)을 통해 이루어지는데 이 템플릿 비교의 속도가 매우 느리기 때문이다. 이를 보완할 수 있는 한 가지 방법으로서 형상 인식 모듈을 VLSI로 구현하는 것을 생각할 수 있다. 즉 하드웨어적인 접근법으로 알고리즘의 실행 속도를 향상시키고자 하는 것이다. 이러한 견지에서 형상인식 모듈을 VLSI로 구현하여, 본 논문에서 제안한 알고리즘의 초기 인식 용도로 활용하는 연구가 필요하다 할 수 있다.

마지막으로, 본 논문에서는 카메라가 공간에 고정된 수동 비전(passive vision system)을 가정하였다. 이 경우, 사람이 걸으면서 몸짓을 취하게 되면 영상이 카메라 렌즈의 외곽 부분으로 밀려가서 렌즈의 방사형 왜곡(radial distortion) 문제에 영향을 받게 된다. 심지어 시야 밖으로 영상이 빠져 나가 팔의 자세 추정 자체를 실패할 위험마저 생기게 된다. 이를 위한 해결 방안은 크게 전방향성 렌즈(omni-directional lens)를 사용한 비전 시스템과 카메라의 시선을 조종하는 능동비전(active vision system)을 사용하는 방법이 있다. 하지만 전자의 경우 전방향성의 렌즈를 통과한 빛의 경로가 심각한 비선형으로 나타남으로써 변형해석(deformation mapping) 해석이 매우 어렵다는 단점이 있다. 그에 비해 후자는 제어의 기술을 응용함으로써 비교적 쉽게 구현될 수 있는 장점이 있다(참고 문헌 [16]에 그 예가 주어져 있다.) 따라서 참고문헌 [16]에서 제안한 알고리을 적용하여 본 논문에서 제안한 인간 팔의 3차원 자세 추정용 비전 시스템을 능동적인 비전 시스템(active vision system)으로 확장하는 연구가 필요하다 하겠다.

참 고 문 헌

[1] D. M. Gavrila and L. S. Davis, "3-D model-based tracking of humans in action: a multiple-view approach," Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition, pp. 73-80, 1996.

[2] L. Goncalves, E. D. Bernardo, E. Ursella, and P. Perona, "Monocular tracking of the human arm in 3D," Proc. IEEE Int. Conf. Computer Vision, pp. 764-770, 1995.

[3] I. Haritaoglu, D. Harwood and L. S. Davis, "W^4: Who? When? Where? What? A real time system for detecting and tracking people," Proc. IEEE Int. Conf. Automatic Face Gesture Recognition, pp. 222-227, 1998.

[4] I. A. Kakadiaris and D. Metaxas, "Model-based estimation of 3D human motion with occlusion based on active multi-viewpoint selection," Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition, pp. 81-87, 1996.

[5] M. K. Leung and Y. H. Yang, "First sight: A human body outline labeling system," IEEE

Trans. Pattern Analysis and Machine Intelligence, vol. 17, no. 4, pp. 359-377, April 1995.
[6] J. M. Rehg and T. Kanade, "DigitEyes: vision-based human tracking for human-computer interaction," Proc. IEEE Int. Conf. Motion of Non-Rigid and Articulated Objects, pp. 16-22, 1994.
[7] K. Rohr, "Towards model-based recognition of human movements in image sequences," CVGIP: Image Undersanding, vol. 59, no. 1, pp. 94-115, 1994.
[8] C. Wren, A. Azarbayejani, T. Darrel, and A. Pentrand, "Pfinder: real-time tracking of the human body," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp.780-785, July 1997.
[9] C. J. Taylor, "Recognition of articulated objects from point correspondences in a single uncalibrated image," IEEE CVPR 2000, pp. 677-684, 2000.
[10] C. Bregler and J. Malik, "Tracking people with twists and exponential maps," Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition, pp. 8-15, 1998.
[11] M. Yamamoto and K. Koshikawa, "Human motion analysis based on a robot arm model," Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition, pp. 664-665, 1991.
[12] M. Yamamoto, T. Ohta, T. Yamagiwa, K. Yagishita, H. Yamanaka, and N. Ohkubo, "Human action tracking guided by key-frames," Proc. IEEE Int. Conf. Automatic Face and Gesture Recognition, pp. 354-361, 2000.
[13] R. M. Murray, Z. Li, and S. S. Sastry, "A Mathematical introduction to robotic manipulation," CRC Press, 1993.
[14] D. H. Ballard, C. M. Brown, "Computer Vision," Prentice-Hall, Inc. 1982.
[15] R. Fletcher, "Practical methods of optimization," A Wiley-Interscience Publication, 1987.
[16] G. X. Ritter and J. N. Wilson, "Handbook of computer vision algorithms in image algebra," CRC Press, 1996.
[16] 한영모, "능동적인 비전 시스템에서 카메라의 시선 조정 : 컴퓨터 비전과 제어의 융합테마", 전자공학회 논문지, 41권 SC편 4호, 2004년 7월.

저 자 소 개

한 영 모(정회원)
1992년 서울대학교 물리교육학과 학사 졸업.
1995년 서울대학교 제어계측공학과 학사 졸업.
1998년 서울대학교 전기공학부 석사 졸업.
2002년 서울대학교 기계항공공학부 박사 졸업
2002년~2003년 단국대학교 기계공학과 시간강사,
세종-록히드마틴 우주항공연구소 연구원, 연구교수
2004년~현재 이화여자대학교 정보통신공학과 연구전임강사
<주관심분야: 컴퓨터비전, 컴퓨터비전 관련 영상처리 및 신호처리, Human-computer Interaction, robotic system control, 전기, 기계, 물리 분야의 융합테마(fusion theme)>