

# 손 동작 인식을 통한 인간-컴퓨터 인터페이스용 저가형 비주얼 모션 데이터 글러브

한 영 모<sup>†</sup>

요 약

모션 데이터 글러브는 손의 움직임을 측정하여 컴퓨터에 입력하는 대표적인 인간과 컴퓨터간의 인터페이스 도구로서, 홈 오토메이션, 가상 현실, biometrics, 모션 캡처 등의 컴퓨터 신기술에 사용되는 필수 장비이다. 본 논문에서는 대중화를 위하여, 별도의 특수 장비 없이 사용 가능한 저가형 비주얼 모션 데이터 글러브를 개발하고자 한다. 본 방식의 특징은 기존의 모션 데이터 글러브에 사용되었던, 고가의 모션 센싱 섬유를 사용하지 않음으로써, 저가형으로 개발이 가능하다는 것이다. 따라서 제작이 용이하고 대중화에 크게 기여할 수 있다는 장점을 가진다. 본 방식에서는 모션 센싱 섬유를 사용하는 기계적인 방식대신 광학적 모션 캡처 기술을 개량한 비주얼 방식을 채택한다. 기존의 비주얼 방식에 비해 본 방식은 다음과 같은 장점과 독창성을 가진다. 첫째, 기존의 비주얼 방식은 가려짐 현상을 제거하고 3차원 자세 복원을 위해 많은 수의 카메라와 장비를 사용하는 데 비해, 본 방식은 모노비전 방식을 채택하여 장비가 간소하고 저가형 개발이 가능하다. 둘째, 기존의 모노비전 방식은 가려짐 현상에 취약하여 영상에서 가려진 부분은 3차원 자세 복원이 어려웠다. 하지만 본 논문은 독창적으로 설계된 막대 모양의 지지자를 사용하여, 영상에서 가려진 부분도 3차원 자세 복원이 가능하다. 셋째, 기존의 모노 비전 방식은 비선형 수치해석 형태의 영상 해석 알고리즘을 사용하는 경우가 많아서 초기화나 계산시간 면에서 불편하였다. 하지만, 본 논문에서는 독창적인 공식화 방법을 사용하여 닫힌 형태의 영상해석 알고리즘을 도출함으로써 이와 같은 불편을 개선하였다. 넷째, 기존의 닫힌 형태의 알고리즘은 공식화 과정에서 근사화 방법을 도입하는 경우가 많아서 정확도가 떨어지고 특이점에 의한 응용분야에 제한이 있었다. 하지만 본 방식은 오일러 각과 같은 국부적인 매개변수나 근사화 등을 사용하는 대신 지수형태의 트위스트좌표계를 사용하는 독창적인 공식화 방법을 사용하여, 공식화 단계에서의 근사화 방법 없이 닫힌 형태의 알고리즘을 도출함으로써 이 문제들을 개선하였다.

키워드 : 모션 데이터 글러브, 손 동작 인식, 인간과 컴퓨터의 상호 작용, 광학적 지지자, 닫힌 형태의 해석적 알고리즘

## Inexpensive Visual Motion Data Glove for Human-Computer Interface Via Hand Gesture Recognition

Youngmo Han<sup>†</sup>

ABSTRACT

The motion data glove is a representative human-computer interaction tool that inputs human hand gestures to computers by measuring their motions. The motion data glove is essential equipment used for new computer technologies including home automation, virtual reality, biometrics, motion capture. For its popular usage, this paper attempts to develop an inexpensive visual-type motion data glove that can be used without any special equipment. The proposed approach has the special feature; it can be developed as a low-cost one because of not using high-cost motion-sensing fibers that were used in the conventional approaches. That makes its easy production and popular use possible. This approach adopts a visual method that is obtained by improving conventional optic motion capture technology, instead of mechanical method using motion-sensing fibers. Compared to conventional visual methods, the proposed method has the following advantages and originalities. Firstly, conventional visual methods use many cameras and equipments to reconstruct 3D pose with eliminating occlusions. But the proposed method adopts a mono vision approach that makes simple and low cost equipments possible. Secondly, conventional mono vision methods have difficulty in reconstructing 3D pose of occluded parts in images because they have weak points about occlusions. But the proposed approach can reconstruct occluded parts in images by using originally designed thin-bar-shaped optic indicators. Thirdly, many cases of conventional methods use nonlinear numerical computation image analysis algorithm, so they have inconvenience about their initialization and computation times. But the proposed method improves these inconveniences by using a closed-form image analysis algorithm that is obtained from original formulation. Fourthly, many cases of conventional closed-form algorithms use approximations in their formulations processes, so they have disadvantages of low accuracy and confined applications due to singularities. But the proposed method improves these disadvantages by original formulation techniques where a closed-form algorithm is derived by using exponential-form twist coordinates, instead of using approximations or local parameterizations such as Euler angles.

Keywords : Motion Data Glove, Hand Gesture Recognition, Human-Computer Interface, Optical Indicator, Closed-Form Analytical Algorithm

<sup>†</sup> 정 회 원 : 한양사이버대학교 컴퓨터공학과 조교수  
논문접수 : 2008년 10월 28일  
수 정 일 : 1차 2009년 5월 18일, 2차 2009년 6월 10일  
심사완료 : 2009년 6월 18일

### 1. 서 론

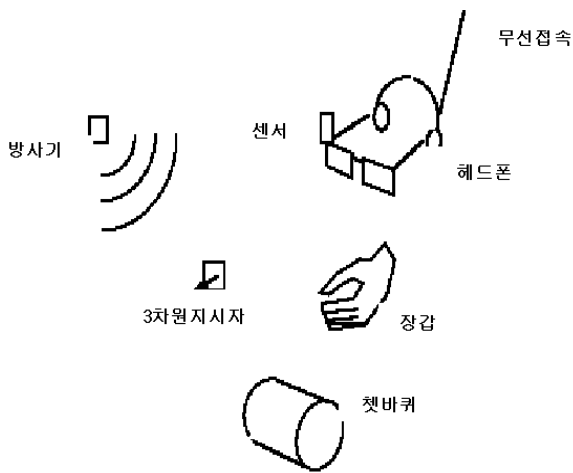
최근 들어 홈 오토메이션, 가상 현실, biometrics, 모션 캡

쳐 등의 미래지향적 컴퓨터 신기술들이 대중적인 관심을 끌고 있다. 이 두 기술은 인간과 컴퓨터 간의 인터페이스 기술을 필요로 하는데, 이 목적으로 사용되는 대표적인 장비들이 (그림 1)에 주어져 있다.

특히 손동작은 자연스럽게 장비에 구애 받지 않는 인간과 컴퓨터간의 인터페이스 수단을 제공함으로써[1], 표현 언어 인식, 가상현실, biometrics, 모션 캡처 등의 다양한 분야에 유용하다[2]. 하지만 인간과 컴퓨터간의 인터페이스 장비가 대부분 너무 고가라서 개인적으로 구입하여 사용하기가 쉽지 않다. 이와 같은 인간과 컴퓨터간의 인터페이스 장비의 고가 문제는 결과적으로 이 기술을 사용하는 홈오토메이션, 가상현실, biometrics, 모션 캡처 등의 차세대 신기술의 발전 및 대중화에 걸림돌이 되고 있다.

이러한 장비의 고가 문제를 해결하기 위해서 별도의 장비 없이 순수하게 비전 시스템만을 사용한 시각적인 접근법이 연구되기도 하였다[3-4]. 하지만 순수하게 비전만을 사용한 경우는 고가의 장비를 사용한 방법에 비해 성능이 좋지 않아서 상용화에는 어려움이 있다. 이러한 취지에서 본 연구에서는 비전 시스템과 저가의 간편한 보조 장비만을 보완적으로 사용하여 상용화가 가능한 저가형 고성능 시스템을 개발하고자 한다.

상용화된 모션 데이터 글러브는 고가의 모션 센싱 섬유를 사용하므로[3, 6], 그 가격이 매우 비싸서 개인이 구입하여 사용하기는 어렵다. 이와 같은 이유로 인해서 모션 데이터 글러브가 개인용으로 널리 사용되지는 못하고 있다. 따라서 모션 데이터 글러브의 대중화를 위해서 그 가격을 낮추는 연구가 필요하다.



(그림 1) 3D 인터페이스 장비[5]

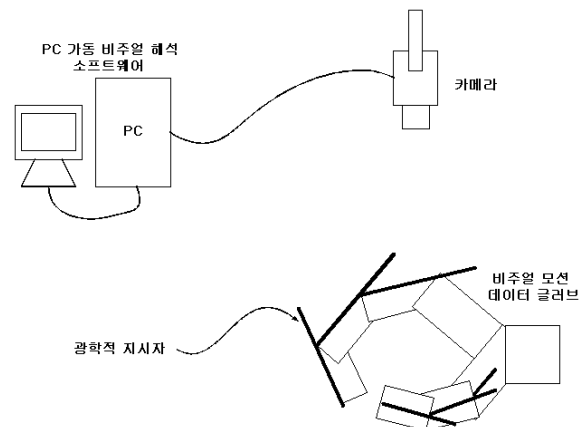
## 2. 제안하는 비주얼 모션 데이터 글러브 시스템의 구성 및 특징

본 연구에서는 고가의 모션 센싱 섬유의 사용을 피하기 위해서 비전 기술을 사용한 비주얼 타입의 접근법을 채택한다. 그리고 비전 시스템의 비용을 낮추기 위해서 흔히 사용

되는 멀티 비전 대신 모노 비전 방식을 채택한다. 일반적으로 모노 비전 방식은 멀티 비전 방식에 비해 비용이 싸지만, 가려짐 현상에 취약하고, 영상 해석 알고리즘이 난해한 약점이 있다. 특히 사람의 손에는 여러 개의 손가락이 있어서 가려짐 현상이 매우 심각하게 나타나는 특징이 있다.

이와 같은 난점을 극복하기 위해서 본 연구에서는 가는 막대 모양의 광학적 지시자를 설계하고, 손가락 각 링크에 부착함으로써 가려짐 문제를 해결하고자 한다. 이는 기존에 많이 사용되어 오던 점이나 선 형태의 광학적 지시자는 대상체의 작은 일부분이 가려지는 경우에는 효과적이지만, 손가락과 같이 많은 부분이 동시에 가려지는 심각한 가려짐의 경우에는 그 기능을 제대로 발휘하기 어렵다는 점을 보완하기 위한 선택이다. 본 논문에서 제안하는 비주얼 모션 데이터 글러브 시스템의 구성이 (그림 2)에 주어져 있다.

멀티 비전의 영상해석 알고리즘이 간단한 선형 형태로 표현되는데 비해, 모노 비전의 3차원 정보 복원 알고리즘은 난해한 비선형 형태로 표현되는 경우가 많다[7-9]. 그리고 난해한 비선형 형태의 알고리즘은 계산 시간이 길어서 실시간 사용에 불리하고, 수렴까지의 반복 계산 횟수가 일정하지 않아서 초당 처리 가능한 영상 프레임의 개수를 일정하게 하기 어렵다. 이는 결과적으로 안정된 성능을 얻기 어려움을 의미한다. 뿐만 아니라 입력 매개변수의 값과 결과 해간의 관계를 특정 짓기 어려워서 성능의 분석 및 예측이 쉽지 않다. 이와 같은 비선형 형태의 알고리즘의 단점을 극복하기 위해서 본 연구에서는 적절한 모델의 근사화를 통해서 단순 형태의 해석적 알고리즘을 구하고자 한다. 이와 같은 취지의 접근법으로는 Taylor의 연구[10]와 Lee의 연구[11]가 있는데, 그 기본 아이디어는 모델을 링크의 연결 형태로 단순화하고, 링크를 연결하는 관절점 중 하나의 거리가 사전에 알려져 있다고 가정한 후, 링크 길이 정보를 활용하여 차례대로 나머지 관절점의 3차원 위치를 계산해 낸다. 이러한 알고리즘은 계산량이 매우 적다는 장점이 있지만, 복원된 결과의 정확도가 높지 않고, 자세 인식에 중요한 관절각을 직접 찾지 못하고 관절점의 위치만을 찾게 된다는 단점이 있다.



(그림 2) 제안하는 비주얼모션 데이터 글러브 시스템

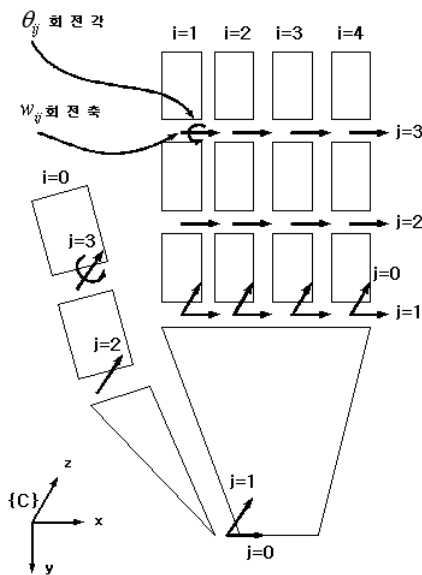
이러한 기존 방식의 단점을 극복하기 위해서 본 연구에서는 관절각을 직접 구하고 복원된 정보의 정확도를 높이기 위한 연구를 수행한다. 그 기본 아이디어는 손의 3차원 움직임을 1차원 회전 관절의 조합으로 분해하여 모델링하고, 그 회전 관절의 회전각을 차례대로 계산하는 방식을 채택한다. 그리고 회전각 복원 결과의 정확도를 높이기 위해서 최적화 기법을 도입하고, 회전 관절이 한 방향으로만 회전한다는 제약 조건을 활용한다.

### 3. 광학적 지시자의 영상해석 알고리즘

본 절에서는 (그림 2)에 주어진 광학적 지시자의 영상 위치로부터 손의 3관절각을 찾아내는 영상 해석 알고리즘을 설계한다. 이를 위해서는 먼저 손의 운동 모델을 설정하는 것이 필요하다. 그림 3은 손의 회전 관절 운동 모델을 보여주고 있다. 여기서,  $i$ 와  $j$ 는 각각 손가락 번호와 관절 순번을 나타낸다. 화살표는 회전축 방향을 표시하고,  $w_{ij}$ 는 회전축 벡터이고,  $\theta_{ij}$ 는 회전축  $w_{ij}$  방향의 회전각을 나타낸다. 이 그림을 보면 손의 3차원 운동은 1차원 회전 관절의 조합으로 분해될 수 있음을 알 수 있다. 따라서 1차원 회전 관절의 움직임을 해석하면 손의 운동을 해석할 수 있게 된다.

그럼 지금부터 (그림 3)을 구성하는 1차원 회전 관절의 운동을 해석해 보도록 한다. (그림 4)는 회전 관절의 운동을 보여주고 있다. 여기서,  $r$ 은 회전의 중심이고,  $C$ 는 회전축  $w$  상의 임의의 한 점이다. 이 그림에서 회전 축 회전  $w$ 을 중심으로 회전각  $\theta$  만큼의 회전을 보여주고 있다. 그리고 이 회전의 결과로서 회전관절 상의 한 점  $Y_r$ 이 새로운 위치  $Y$ 로 이동하였다.

만약  $Y$ 의 영상 위치와 거리(카메라 광축 방향의 좌표)가 각각  $q$ 와  $z$ 이고, 카메라의 투영모델이  $Y = zq$ 로 주어지는



(그림 3) 손의 회전관절 모델

경우에는, 다음의 관계식이 성립한다.

$$zq = \exp([\hat{w}]\theta)Y_r \tag{1}$$

여기서,  $\hat{w}$ 는 회전축  $w$  방향의 단위벡터이다. 회전축 상의 점은 회전에 대해 불변이므로, (1)식은 다음과 같이 고쳐 쓸 수 있다.

$$(zq - C) = \exp([\hat{w}]\theta)(Y_r - C) \tag{2}$$

(2)식은 지수형 이라서 사용하기가 곤란하므로 좀 더 편리한 형태로 전개할 필요가 있다. 기존의 방식에서는 흔히 Euler 각을 사용하게 되는데, 이와 같은 국부적인 매개변수화는 특이점을 만들게 되고, 공식의 비선형성을 가중시키게 되어 닫힌 형태의 답을 도출하기 어려워진다. 이 문제를 보완하기 위해서 본 논문에서는 (2) 식의 지수형 트윈스트 좌표계( $w, \theta$ )를 근사화나 국부적인 매개변수화 없이 좀 더 편리한 형태로 전개할 수 있는 Rodrigues's 포뮬러를 사용한다. Rodrigues's formula[12]에 주어지고, 다음과 같이 표현될 수 있다.

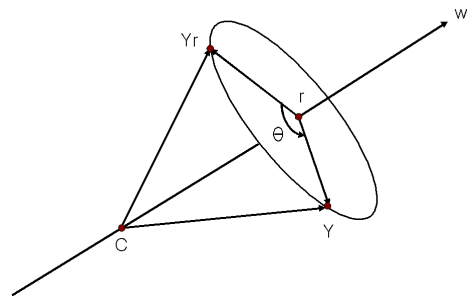
$$\exp([\hat{w}]\theta) = I + \sin(\theta)[\hat{w}] + (1 - \cos(\theta))[\hat{w}]^2$$

이 식을 살펴보면, Euler 각을 사용하여 전개했을 때에 비해, 특이점을 갖지 않고,  $\theta$ 에 대해서 선형으로 표현된 것을 볼 수 있다. 따라서, 변수 대입:  $t = \tan(\theta/2)$ 를 사용하면, 식(2)를 다음과 같이 분자와 분모가 각각 선형으로 표현되는 편리한 형태로 고쳐 쓸 수 있게 되어 닫힌 형태의 답을 도출할 수 있게 된다.

$$q = \frac{m_1 t^2 + m_2 t + m_3}{n_1 t^2 + n_2 t + n_3} \tag{3}$$

여기서,

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad m_1 = C + k_1 - k_2 \in R^3, \\ m_2 = 2k_3 \in R^3, \quad m_3 = C + k_1 + k_2 \in R^3, \\ n_1 = e_3^T(C + k_1 - k_2) \in R, \quad n_2 = e_3^T(2k_3) \in R,$$



(그림 4) 회전 관절의 운동

$$n_3 = e_3^T(C + k_1 + k_2) \in R, \quad k_1 = (I + [\hat{w}]^2)(Y_r - C) \in R^3, \\ k_2 = -[\hat{w}]^2(Y_r - C) \in R^3, \quad k_3 = [\hat{w}](Y_r - C) \in R^3,$$

$e_3 \in R^3$ 은 광축 방향의 단위벡터이다(이 식에서 표기법  $e$ 는 단위벡터임을 나타내고, 편의상 본 논문에서는 카메라 좌표계의 Z-축을 광축 방향으로 설정한다.  $k_1, k_2, k_3$ 는  $m_1, m_2, m_3, n_1, n_2, n_3$ 의 표현이 너무 길고 공통된 부분이 많아서, 이 식들을 일목요연하게 정리하기 위해 편의상 사용한 표기법이다).

식(3)은 회전 관절의 회전과 회전 관절점의 영상 위치 사이의 관계를 보여주고 있다. 바꾸어 말하면, 회전 관절점의 영상 위치가 주어지면 식(3)을 이용하여 회전 관절의 회전 각을 계산해 낼 수 있다. 이 문제는 영상 잡음을 고려하면 최적화 문제로 공식화할 필요가 있다. 물론 다양한 형태의 최적화 함수가 도입될 수 있지만, 본 논문에서는 가장 많이 사용되는 형태 중에 하나인 최소제곱 최적화 목적 함수를 예로 들어 보겠다.

$$J = \|\tilde{q} - q\|^2 \tag{4}$$

여기서,  $q$ 는 식(3)에 의해 주어지고,  $\tilde{q}$ 는  $q$ 의 측정값이다. 식(4)에 의해 주어진 최적화 문제를 풀기 위해, 최적화를 위한 1차 필요조건  $\frac{\alpha J}{\alpha t} = 0$ 을 풀어 보면 다음과 같은 식을 얻을 수 있다.

$$t^4 + c_1 t^3 + c_2 t^2 + c_3 t + c_4 = 0 \tag{5}$$

$$c_1 = (a_1^T b_2 + a_2^T b_1) / (a_1^T b_1), \\ c_2 = (a_1^T b_3 + a_2^T b_2 + a_3^T b_1) / (a_1^T b_1), \\ c_3 = (a_2^T b_3 + a_3^T b_2) / (a_1^T b_1), \quad c_4 = (a_3^T b_3) / (a_1^T b_1), \\ a_1 = (m_1 - n_1 \tilde{q}), \quad a_2 = (m_2 - n_2 \tilde{q}), \quad a_3 = (m_3 - n_3 \tilde{q}) \\ b_1 = (m_1 n_2 - m_2 n_1), \quad b_2 = 2(m_1 n_3 - m_3 n_1), \\ b_3 = (m_2 n_3 - m_3 n_2)$$

식(5)에 주어진 다항식 방정식은 4개의 닫힌 형태의 해를 가진다[13].

식(5)로 표시되는 1차원 회전 관절의 영상해석 알고리즘을 각 관절에 차례대로 적용함으로써 전체 손의 움직임을 해석할 수 있다. 식(5)에서 제안한 1차원 회전 관절의 영상해석 알고리즘이 적은 계산량을 가진 닫힌 형태의 알고리즘이기 때문에, 결과적으로 전체 손의 영상 해석 알고리즘도 적은 계산량의 닫힌 형태의 알고리즘이 된다.

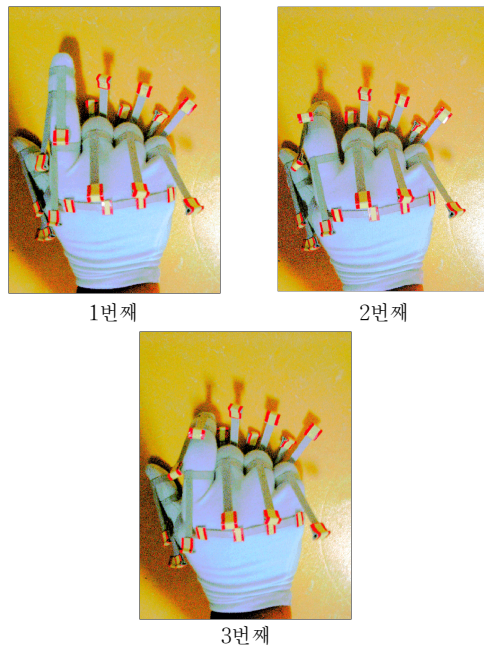
물론, 손의 전체 운동을 동시에 해석하는 방법을 생각할 수도 있다. 하지만 이 방법은 매우 복잡하기 때문에, 계산 시간이 많이 걸리고, 닫힌 형태의 해석적 알고리즘을 구하기 어렵다는 단점이 있다. 반면 각각의 1차원 회전 관절의 운동을 순차적으로 해석하는 방법은 전체 회전 관절을 동시에 최적화하는 것보다는 조금 정확도가 떨어질 수는 있다.

특히 연결된 관절의 수가 많을 경우 각 관절마다 생긴 오차가 누적되어 마지막 관절로 갈수록 오차가 커질 수 있다. 하지만 다행히도 손의 경우는 각 손가락에 연결된 회전 관절의 수가 적기 때문에, 이러한 오차로 인해 복원 에러가 커질 가능성은 많지 않다. 이 부분에 대해서는 다음 절의 실험을 통해 오차 검증을 실행하도록 한다.

### 4. 결 과

본 절에서는 2절과 3절에서 제안한 비주얼 모션 데이터 클러브의 하드웨어와 소프트웨어를 구현하여 성능을 검증한다. (그림 5)는 2절에서 제안한 비주얼 모션 데이터 클러브 본체를 구현한 예를 보여주고 있다. 저가형 흰 장갑에 색으로 표시된 막대 모형의 광학적 표시자가 관절마다 부착된 형태를 가지고 있다. 이렇게 제작된 비주얼 모션 데이터 클러브 본체는 한 개의 카메라를 사용하여 위쪽 방향에서 촬영된다. 그리고 컬러 검출을 통해 광학적 지지자의 영상 위치를 구하고, 3절에서 제안한 영상해석 알고리즘을 각 광학적 표시자의 영상 위치에 적용하여, 해당하는 관절의 회전 각을 계산한다. 본 실험에서는 카메라는 0.34 m 떨어진 위치에 고정되어 있고, 영상 크기는 512 x 384 화소이다. 실험에 사용된 시나리오는 마우스의 클릭 문제이다. 이 시나리오는 1.20GHz의 클럭 사이클을 가진 펜티엄 PC에서 MATLAB을 사용해 수행하였다.

(그림 5)의 1번째 그림은 마우스를 클릭하기 전의 준비 상태이고, 2번째 그림은 왼쪽 마우스를 클릭한 상태이고, 3번째 그림은 오른쪽 마우스를 클릭한 상태이다. 이 세 그림



(그림 5) 실험 시나리오. 1번째, 2번째, 3번째 그림은 순서대로 준비상태, 왼쪽 마우스 클릭, 오른쪽마우스 클릭 상태를 보여준다.

을 살펴보면, 손가락이 굽어져서 시야에서 사라진 관절 부분도 막대 모양의 광학적 지시자에 의해 그 위치가 추정 가능함을 확인할 수 있다.

다음으로 이 세 가지 마우스 클릭 상태에 대해 3절에서 제안한 영상 해석 알고리즘을 적용하여, 손 형태를 복원한 결과 그림이 (그림 6)에 순서대로 주어졌다. (그림 6)에서 실선은 제안하는 영상해석 알고리즘을 사용해서 복원한 손의 형태이고, 점선은 측정을 통해 구한 손의 형태이다. 이 둘 사이의 차이값을 좀 더 정확히 분석하기 위해서 표 1에 평균 오차율을 제시하였다. 여기서, 오차율은 복원된 값과 측정된 값의 차이를 측정된 값으로 나누어 계산하였다. 기존의 방식과의 비교를 위해서, 비선형의 반복형 알고리즘인 Difrancio의 방식[14]과 단형 형태의 해석적 알고리즘인 Lee의 방식[15]을 적용했을 경우의 결과값도 함께 제시하였다. 이 표의 결과를 보면, 제안하는 방식은 동종인 Lee의 방식에 비해 오차율이 매우 작고, 거의 비선형의 반복형 알고리즘의 정확도에 가까운 것을 볼 수 있다.

비주얼 모션 데이터 클러브를 사용한 손동작 인식 문제에서 정확도 못지않게 중요한 부분이 계산 시간이라 할 수 있다. 그 이유는 실시간 적용을 목표로 하기 때문이다. 이러한 취지에서, <표 1>에서 제시한 결과값을 얻기 위해 소요된 계산 시간을 <표 2>에 제시하였다.

<표 2>를 보면, 제안한 방식의 계산 시간은 동종의 방식인 Lee의 방식과 계산시간이 거의 같고, 비선형의 반복형

방식인 Difrancio의 알고리즘에 비해 매우 적은 것을 볼 수 있다. <표 1>과 <표 2>의 결과를 종합해 보면, 본 논문에서 제안한 방식이 단형 형태의 해석적 알고리즘 방식만큼 적은 계산시간으로도 비선형의 반복형 알고리즘 방식만큼 높은 정확도를 가짐을 확인할 수 있다.

<표 1> 관절각 복원 값의 오차율

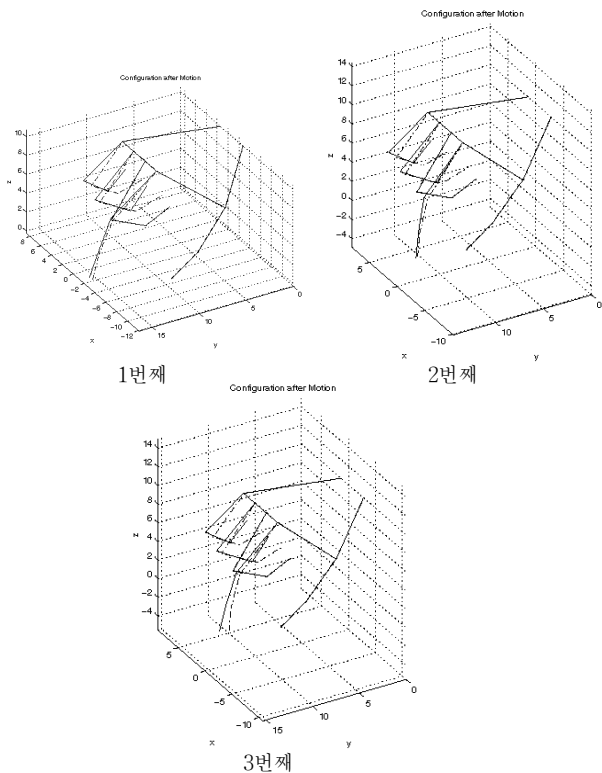
	제안 방식	Lee	Difrancio
Ready	0.1509	0.3325	0.1253
Left-push	0.1791	0.4137	0.1419
Right-push	0.1510	0.4144	0.1120

<표 2> 관절각 복원의 계산 시간

	제안 방식	Lee	Difrancio
Ready	0.1226	0.1350	0.4917
Left-push	0.1263	0.1264	0.4832
Right-push	0.1214	0.1387	0.4794

### 5. 결 론

본 논문에서는 저가형 비주얼 모션 데이터 클러브를 제안하였다. 본 방식의 특징은 다음과 같다. 첫째, 손의 3차원 움직임을 1차원 회전 관절의 움직임만으로 분해 해석하고 모델링하였다. 복잡한 손의 3차원 움직임을 기계 요소 모델링 기법을 적극 사용하여 단순한 1차원 회전 관절의 움직임만으로 분해 해석하고 모델링함으로써, 대량 생산의 용이성을 높이고, 자세 인식 시 사용되는 회전 관절을 직접적인 복원 값으로 설정할 수 있었다. 그리고 결과값의 정확도를 높이기 위해 사용될 회전 관절의 제약 조건을 얻을 수 있었다. 둘째, 결과값의 정확도를 높이기 위해서 제약 조건이 있는 최적화를 도입하였다. 기존의 단형 형태의 해석적인 접근법은 링크 길이로부터 최적화 없이 직접적으로 결과값을 얻으로써, 결과값의 정확도가 낮은 단점을 가지고 있었다. 하지만 본 연구에서는 결과값의 정확도를 높이기 위해서, 부분적인 최적화 형태로 영상해석 알고리즘을 공식화하고 회전 관절의 제약 조건을 최적화 문제에 적극 사용하였다. 셋째, 손가락의 가려짐 현상을 극복하기 위해서 긴 막대 형태의 광학적 지시자를 도입하였다. 기존의 비주얼 타입 접근법에서 흔히 사용되던 점이나 선 형태의 광학적 지시자로는 손가락과 같은 심각한 가려짐 현상 극복이 어렵다는 점에 주목하여, 본 연구에서는 독창적인 형태의 긴 막대 형태의 광학적 지시자를 도입하고, 손가락의 심각한 가려짐도 효과적으로 극복할 수 있도록 배치하였다. 본 논문에서 제안한 방식을 마우스 클릭에 관련된 3가지 손 동작에 적용한 결과, 손가락이 완전히 굽어서 시야에서 관절이 사라진 심각한 가려짐 상황에서도 막대 모양의 지시자에 의해 가려진 관절의 위치를 추정할 수 있음을 확인하였다. 그리고, 기존의 단형 형태의 해석적 알고리즘 방식만큼 적은 계산시간으로도, 비



(그림 6) 제안하는 영상해석알고리즘을 적용해 복원한 결과. 1 번째, 2 번째, 3번째 그림은 순서대로 준비상태, 왼쪽 마우스 클릭, 오른쪽 마우스 클릭 상태를 보여준다.

선형의 반복형 알고리즘 방식만큼 높은 정확도를 갖는 결과 값을 얻을 수 있음을 확인하였다.

### 참 고 문 헌

- [1] Jakub Segen, Senthil Kumar, "Shadow gestures: 3D hand pose estimation using a single camera," IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol.1, pp.485, 1999.
- [2] Vassilis Athitsos, Stan Sclaroff, "An appearance-based framework for 3D hand shape classification and camera viewpoint estimation," Proceedings, Face and Gesture Recognition, 2002.
- [3] Chung-Lin Huang, Sheng-Hung Jeng, "A model-based hand gesture recognition system," Machine Vision and Applications, Vol.12, pp.243-258, 2001.
- [4] Stefano Pellegrini, Luca Iocchi, "Human posture tracking and classification through stereo vision," Proceedings, International Conference, Computer Vision Theory and Applications (VISAPP), pp.1-8, 2006.
- [5] "VIRTUAL REALITY," www.cg.its.tudelft.nl/~fwj/aux/4-Virtual reality.pdf.
- [6] Takahashi T, Kishino F, "A hand gesture recognition method and its application," Systems and Computers in Japan, Vol.23, No.3, pp.38-48, 1992.
- [7] K. Rohr, "Incremental recognition of pedestrians from image sequences," Proceedings, IEEE Computer Vision and Pattern Recognition, pp.8-13, New York City, 1993.
- [8] J. Zhao, "Moving posture recognition from perspective projections of jointed figure motion," PhD thesis, University of Pennsylvania, 1993.
- [9] David E. DIFRANCO, Tat-Jen CHAM, James M. REHG, "Reconstruction of 3-D Figure Motion from 2-D Correspondences," Proceedings, Conference, Computer Vision and Pattern Recognition, pp.307-314, 2001.
- [10] C. J. Taylor, "Reconstruction of Articulated Objects from Point Correspondences in a Single Uncalibrated Image," IEEE CVPR 2000, pp.677-6684, 2000.
- [11] Sung Uk Lee, Isaac Cohen, "3D Hand and Fingers Reconstruction from Monocular View," Proceedings, 17th International Conference, Pattern Recognition (ICPR04), pp.231-236, 2004.
- [12] R. M. Murray, Z. Li, S. S. Sastry, "A mathematical introduction to robotic manipulation," CRC Press, 1994.
- [13] I. Z. Emiris, "Solving polynomial equations: foundations, algorithms, and applications," Springer Verlag, 2005.
- [14] D. E. DIFRANCO, Tat-Jen CHAM, James M. REHG, "Reconstruction of 3-D Finger Motion from 2-D Correspondences," Proceedings, Conference, Computer Vision and Pattern Recognition, pp.307-314, 2001.
- [15] S. U. Lee, I. Cohen, "3D Hand and Fingers Reconstruction from Monocular View," Proceedings of 17th International Conference on Pattern Recognition (ICPR04), pp.231-236, 2004.
- [16] "5DT Data Glove Ultra Series," www.5dt.com/downloads/dataglove/ultra/5DT Data Glove Ultra-Manual.pdf.



### 한 영 모

e-mail : ymhan123@hanmail.net

1992년 서울대학교 물리교육학과 졸업(이학사)

1995년 서울대학교 제어계측공학과 졸업(공학사)

1998년 서울대학교 전기공학부(공학석사, 신호처리, VLSI 설계 전공)

2002년 서울대학교 기계항공공학부(공학박사, 컴퓨터비전, 인체 영상의 자세 인식, 로봇 제어 전공)

2002년~2003년 세종-록히드마틴 우주항공연구소 전임연구원, 전임연구교수

2004년~2005년 이화여자대학교 정보통신공학과 연구전임강사, 연구교수

2006년~현 재 한양사이버대학교 컴퓨터공학과 전임강사, 조교수

2007년~현 재 International Biographical Centre (England),

Deputy Director General 및 Honorary Director General

2007년~현 재 World Congress of Arts, Science and Communications (England), Vice President

2009년~현 재 American Biographical Institute Research Association (USA), Deputy Governor

관심분야: 컴퓨터비전 응용 멀티미디어 및 생체 영상 인식, 모바일 및 로봇 임베디드 시스템 소프트웨어, 인간과 컴퓨터의 시각적 인터페이스, 정보기술을 위한 융합테마 등