

단일 비전 기반 인체의 3차원 운동 해석

한 영 모[†]

요 약

본 논문은인간-컴퓨터 인터페이스, 가상현실의 의학 응용, 환자의 원격 모니터링과 같은 실시간 응용 분야에 적합한 인체 운동의 시각적 해석 알고리즘 (visual analyzer algorithm)을 제안한다. 본 알고리즘을 사용할 때의 비용을 줄이기 위해서, 단수의 카메라를 사용하도록 설계한다. 그리고 제안한 알고리즘을 좀 더 편리하게 사용할 수 있도록 하기 위해서 광학적 표시자의 사용을 피한다. 제안하는 알고리즘이 실시간 사용에 편리하도록 하기 위해서, 폐쇄적 형태가 되도록 설계한다. 폐쇄적 형태의 알고리즘을 설계하기 위해서, 인체의 각 관절을 기존의 3차원 관절 모델 대신 어떤 형태의 근사화도 사용하지 않고도 2차원 관절 모델로 공식화하는 아이디어를 제안한다. 그리고 이 폐쇄적 형태의 알고리즘이 높은 정확도를 갖게 하기 위해서, 계산 알고리즘을 최적화 문제로 공식화한다. 이렇게 해서 설계된 알고리즘을 인체의 각 관절에 차례대로 적용한다.

키워드 : 인체 운동 해석, 폐쇄적 형태 알고리즘, 국부적 최적해

Understanding of 3D Human Body Motion based on Mono-Vision

Youngmo Han[†]

ABSTRACT

This paper proposes a low-cost visual analyzer algorithm of human body motion for real-time applications such as human-computer interfacing, virtual reality applications in medicine and telemonitoring of patients. To reduce cost of its use, we design the algorithm to use a single camera. To make the proposed system to be used more conveniently, we avoid from using optical markers. To make the proposed algorithm be convenient for real-time applications, we design it to have a closed-form with high accuracy. To design a closed-form algorithm, we propose an idea that formulates motion of a human body joint as a 2D universal joint model instead of a common 3D spherical joint model, without any kinds of approximation. To make the closed-form algorithm has high accuracy, we formulates the estimation process to be an optimization problem. Thus-designed algorithm is applied to each joint of the human body one after another. Through experiments we show that human body motion capturing can be performed in an efficient and robust manner by using our algorithm.

Keywords : Human Motion Understanding, Closed-Form Algorithm, Local Optimal Solution

1. 서 론

카메라로 관측한 인체 운동을 분석하는 시각적 해석 알고리즘(visual analyzer algorithm)들은 일반적으로 두 개의 부분으로 나누어진다. 한 부분은 각각의 카메라 영상에서 인체 부위의 2차원 위치를 추적하는 기능을 하고, 나머지 한 부분은 주어진 2차원 위치들로부터 인체의 3차원 위치를 복원하는 기능을 한다. 각각의 카메라 영상에서 2차원 위치 추적을 쉽게 하기 위해서, 많은 기존의 알고리즘들은 인체에 광학적 표시자를 부착한다. 그리고 3차원 복원 과정

을 단순화하기 위해서, 많은 기존의 알고리즘들은 다수의 카메라를 사용한다. 그 결과로서, 거추장스러운 장비들로 인해 인체 운동 해석 과정은 비용도 많이 들고, 그 장비들을 수용할 특수 공간이 필요하므로 공간에 제한을 받게 되고, 많은 수의 광학적 표시자를 부착해야 하므로 사용하기 번거롭다.

인체 운동 해석 과정의 비용을 낮추기 위해서, 본 논문은 다수의 카메라 대신 단수의 카메라를 사용하여 인체 운동을 관측하고, 관측된 영상을 PC에서 분석할 수 있는 시각적 해석 알고리즘을 설계한다. 다수의 카메라를 사용하는 경우에 비해 단수의 카메라를 사용하는 경우의 시각적 해석 알고리즘은 설계가 어렵다. 단수의 카메라를 사용하는 기존의 시각적 해석 알고리즘 중 많은 수는 반복 계산 형태 (iterative

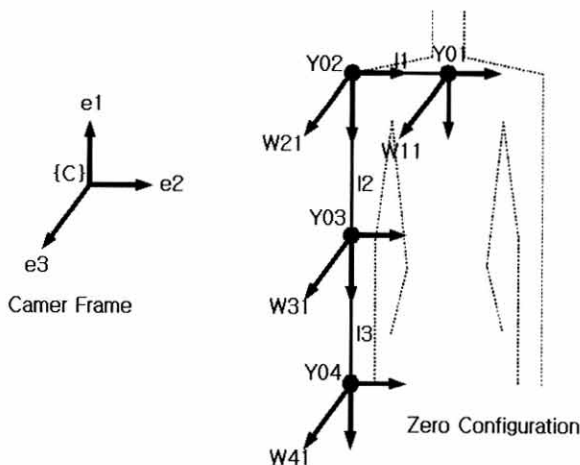
[†] 정 회 원 : 한양사이버대학교 컴퓨터공학과 조교수
논문접수 : 2011년 1월 27일
수 정 일 : 1차 2011년 5월 2일
심사완료 : 2011년 5월 3일

computing form)이다[1], [2]. 이러한 알고리즘은 대개 최적화된 해를 제공하지만, 실시간 사용에는 다소 불편함이 있다. 단수 카메라를 사용하는 기존의 시각적 해석 알고리즘 중 일부는 폐쇄적 형태도 있지만, 이들은 특정 종류의 근사화 방법을 사용하거나 [3], [4], 최적화 해가 아니라서 [5] 그 결과의 정확도가 높지 않다.

이러한 단점들을 보완하기 위해서, 본 논문은 어떤 형태의 근사화도 사용하지 않으면서 관절별로 최적화된 국부적 최적 (local optimal) 수준의 정확도를 갖는 폐쇄적 형태 (closed-form)의 알고리즘을 제안한다 (본 논문에서 폐쇄적 형태(closed-form)는 유한 개의 함수를 사용하여 해석적으로 표현될 수 있는 형태의 의미로, 반복 계산 형태(iterative computing form)는 폐쇄적 형태로 표현될 수 없는 수치해석적인 반복 계산 형태의 의미로 사용한다.)

2. 인체의 각 관절에 적용할 시각적 해석 알고리즘

본 절에서는 각 관절에 적용할 시각적 해석 알고리즘을 제안하고자 한다. 시각적 해석 알고리즘 중에서 2차원 관절 위치는 기존의 영상 추적 알고리즘을 사용하여 주어진 것으로 가정한다. 따라서, 본 절의 주제는 주어진 관절 2차원 영상 위치로부터 관절의 3차원 운동을 복원하는 것이다. 시각적 해석 알고리즘을 설계할 때, (그림 1)에 주어진 모델을 사용하여 설명한다. (그림 1)에서, 인체는 1, 2, 3차원 회전관절의 사슬로 표현된다. 이 그림에서, {C}는 카메라 좌표를 나타내고, e_1, e_2, e_3 는 카메라 좌표의 축을 나타낸다. 특히, e_3 는 카메라 광축과 평행한 방향의 축을 나타낸다.



(그림 1) 영점 자세에서의 인체의 회전 관절 사슬 모델

본 논문에서 제안하는 인체 운동 추정엔 영점 자세 (zero configuration)에 대해 표현된다. 여기서, 영점 자세는 모든 관절각이 0인 자세를 말한다. 링크 길이 l_i 가 주어진 경우에는 영점 자세를 쉽게 선택할 수 있다. 이 영점 자세 선택의 한 예가 (그림 1)에 주어져 있다. 이 그림에서, 관절 위치

Y_{0i} 는 링크 벡터 $Y_{0(i+1)} - Y_{0i}$ 가 l_i 의 크기를 갖고, e_3 에 수직하도록 선택되었다. 이 그림에 W_{i1}, W_{i2}, W_{i3} 는 관절 위치 Y_{0i} 에 위치한 몸통 좌표계 (body frame)의 세 축을 나타낸다.

이렇게 주어진 영점 자세에 대해, 위치가 확인된 현재관절 (Y_i)과 연결된 다음관절 (Y_{i+1})의 위치는 다음 식과 같이 표현된다.

$$Y_{i+1} - Y_i = \exp([\hat{w}_{i1}] \theta_{i1}) \exp([\hat{w}_{i2}] \theta_{i2}) \exp([\hat{w}_{i3}] \theta_{i3}) (Y_{0(i+1)} - Y_{0i}) \quad (1)$$

여기서, \hat{w}_{ij} 는 회전축 W_{ij} 방향의 단위 벡터이고, θ_{ij} 는 \hat{w}_{ij} 축을 중심으로 하는 회전의 회전각이다.

W_{i3} 를 링크 벡터 $Y_{0(i+1)} - Y_{0i}$ 에 평행하게 선택하면, 다음 관계가 성립한다.

$$\exp([\hat{w}_{i3}] \theta_{i3}) (Y_{0(i+1)} - Y_{0i}) = (Y_{0(i+1)} - Y_{0i}) \quad (2)$$

식 (2)를 사용하면, 식 (1)이 다음과 같이 간단해진다.

$$Y_{i+1} - Y_i = \exp([\hat{w}_{i1}] \theta_{i1}) \exp([\hat{w}_{i2}] \theta_{i2}) (Y_{0(i+1)} - Y_{0i}) \quad (3)$$

결과적으로, 식 (1)로부터 한 개의 회전각 θ_{i3} 이 제거된다. 이 회전각 변수의 감소는 폐쇄적 형태의 알고리즘 (analytical solution type algorithm) 설계를 가능하게 한다. 더구나, 이러한 최적 회전각 변수의 선택은 제안하는 알고리즘의 중복성을 없애고, 알고리즘의 효율도 높여준다.

인체의 전체 운동을 한번에 해석하는 것은 매우 복잡하므로, 계산 시간도 많이 걸리고 폐쇄적 형태의 해를 구하기도 어렵다. 실시간 사용의 편의를 생각해서, 본 논문에서는 인체의 운동 복원 문제를 운동을 일련의 관절 운동 복원 문제로 나누어 생각한다. 먼저 한 개 관절의 회전각을 계산한 후에, 이를 기반으로 다음 관절의 회전각을 차례대로 계산한다. 이러한 접근법에서는 Y_i 는 이전 관절의 계산 과정 중 이미 알려져 있는 상태가 된다.

지금부터는 식 (3)에서 θ_{j1}, θ_{j2} 를 계산하는 알고리즘을 설계한다. 식 (3)의 양변에 \hat{w}_{i1}^T 를 곱하면 다음 식을 얻는다.

$$\hat{w}_{i1}^T (Y_{i+1} - Y_i) = \hat{w}_{i1}^T \exp([\hat{w}_{i1}] \theta_{i1}) \exp([\hat{w}_{i2}] \theta_{i2}) (Y_{0(i+1)} - Y_{0i}) \quad (4)$$

Rodrigues의 식 [6]: $\exp([\hat{w}_i] \theta_i) = I + \sin(\theta_i)[\hat{w}_i] + (1 - \cos(\theta_i))[\hat{w}_i]^2$ 와 관계식: $\hat{w}_{i1}^T [\hat{w}_{i1}] = 0$ 을 사용하면, 식 (4)를 다음과 같이 고쳐 쓸 수 있다.

$$\hat{w}_{i1}^T (Y_{i+1} - Y_i) = \hat{w}_{i1}^T \exp([\hat{w}_{i2}] \theta_{i2}) (Y_{0(i+1)} - Y_{0i}) \quad (5)$$

$\hat{w}_{i1} = e_3$ 으로 선택하고, 관계식: $e_3^T Y_{i+1} = z_{i+1}$ 을 사용하면, 식 (5)는 다음과 같이 표현될 수 있다.

$$z_{i+1} = \hat{w}_{i1}^T \exp([\hat{w}_{i2}] \theta_{i2}) (Y_{0(i+1)} - Y_{0i}) + \hat{w}_{i1}^T Y_i \quad (6)$$

이 식이 Y_{i+1} 의 깊이 (영상 평면에 수직한 방향 성분, i.e., $e_3^T Y_{i+1} = z_{i+1}$)에 대한 운동 방정식이다.

영상 위치에 대한 운동 방정식을 얻기 위해서, 영상위치와 깊이 사이의 관계식인 카메라 투영 모델을 도입한다. 참고문헌[7]에서 표현한 바와 같이, 경사 투영 모델 (perspective projection model)은 가장 많이 사용되는 카메라 투영 모델 중에 하나이다. 경사 투영 모델은 다음 식과 같이 표현된다.

$$q_{i+1} = \frac{f(Y_{i+1} - Y_i)}{z_{i+1} - e_3^T Y_i}, \quad z_{i+1} = e_3^T Y_{i+1} \quad (7)$$

여기서, Y_i 는 카메라의 광학적 중심 (optical center)이고, f 는 카메라의 초점 거리이다.

식 (3)과 식 (6)을 식 (7)에 대입하면, Y_{i+1} 의 영상 좌표를 다음과 같이 표현할 수 있다.

$$q_{i+1} = f(\exp([\hat{w}_{i1}] \theta_{i1}) \left(\frac{\exp([\hat{w}_{i2}] \theta_{i2}) (Y_{0(i+1)} - Y_{0i})}{z_{i+1} - e_3^T Y_i} \right) + \frac{f(Y_i - Y_i)}{z_{i+1} - e_3^T Y_i} \quad (8)$$

식 (8)이 Y_{i+1} 의 영상 위치에 대한 운동 방정식이다.

식 (6)과 식 (8)을 풀어서, θ_{j1} , θ_{j2} 를 계산할 수 있다. 하지만, 관절별로 최적화된 국부적 최적 (local optimal) 정도의 좀 더 정확한 결과를 얻기 위해서, 계산 알고리즘을 다음과 같은 최적화 문제로 정식화한다.

제안하는 알고리즘의 첫 번째 단계는 다음에 주어진 깊이 목적 함수를 최소화하여 θ_{j2} 를 구하는 것이다.

$$J_D = \|\tilde{z}_{i+1} - z_{i+1}(\theta_{i2})\|^2 \quad (9)$$

여기서, $z_{i+1}(\theta_{i2})$ 는 식 (6)에 주어지고, \tilde{z}_{i+1} 는 링크 길이의 제약식인 식 (10)을 풀어서 z_{i+1} 의 측정값이다.

$$\|\tilde{Y}_{i+1}(\tilde{z}_{i+1}) - Y_{i+1}\|^2 = \|Y_{0(i+1)} - Y_{0i}\|^2 \quad (10)$$

여기서, $\tilde{Y}_{i+1}(\tilde{z}_{i+1})$ 는 $Y_{i+1} = \tilde{Y}_{i+1}(\tilde{z}_{i+1})$, $z_{i+1} = \tilde{z}_{i+1}$, $q_{i+1} = \tilde{q}_{i+1}$ (측정된 영상 위치)를 식 (7)에 대입한 값이다. 제안하는 알고리즘의 2번째 단계는 식 (11)에 주어진 영상 목적 함수를 최소화하는 θ_{j1} 를 찾는 것이다.

$$J_I = \|\tilde{q}_{i+1} - q_{i+1}(\theta_{i1})\|^2 \quad (11)$$

여기서, $q_{i+1}(\theta_{i1})$ 는 식 (8)에 주어지고, \tilde{q}_{i+1} 는 q_{i+1} 의 측정값이다.

지금부터는 식 (9)와 식 (11)에 주어진 목적함수를 푸는 해법 알고리즘에 대해 살펴본다. 먼저, 식 (9)에 주어진 깊이 목적함수를 살펴본다. Rodrigues의 식:

$$\exp([\hat{w}_{i2}] \theta_{i2}) = I + \sin(\theta_{i2})[\hat{w}_{i2}] + (1 - \cos(\theta_{i2}))[\hat{w}_{i2}]^2 \quad \text{과 관계식:}$$

$$\cos(\theta_{i2}) = \frac{1 - t_{i2}^2}{1 + t_{i2}^2}, \quad \sin(\theta_{i2}) = \frac{2t_{i2}}{1 + t_{i2}^2},$$

$t_{i2} = \tan(\theta_{i2}/2)$ 을 사용하여, 식 (6)을 다음과 같이 고쳐 쓸 수 있다.

$$z_{i+1} = \frac{m_{11}t_{i2}^2 + m_{12}t_{i2} + m_{13}}{n_{11}t_{i2}^2 + n_{12}t_{i2} + n_{13}} \quad (12)$$

여기서, $m_{12} = 2k_{13} \in R$, $m_{13} = C_2 + k_{11} + k_{12} \in R$, $n_{11} = 1$, $n_{12} = 0$ and $n_{13} = 1$, $C_2 = \hat{w}_{i1}^T Y_{i+1} \in R$, $k_{11} = \hat{w}_{i1}^T (I + [\hat{w}_{i2}]^2) (Y_{0(i+1)} - Y_{0i}) \in R$, $k_{12} = -\hat{w}_{i1}^T [\hat{w}_{i2}]^2 (Y_{0(i+1)} - Y_{0i}) \in R$,

$$k_{13} = \hat{w}_{i1}^T [\hat{w}_{i2}] (Y_{0(i+1)} - Y_{0i}) \in R$$

식 (9)의 최적화 1차 필요 조건, i.e., $\frac{\partial J_D}{\partial t_{i2}} = 0$ 을 풀어서 정리하면 다음과 같은 4차 다항식을 얻는다.

$$t^4 + c_1 t^3 + c_2 t^2 + c_3 t + c_4 = 0 \quad (13)$$

여기서, $c_1 = (a_1^T b_2 + a_2^T b_1) / (a_1^T b_1)$, $c_2 = (a_1^T b_3 + a_2^T b_2 + a_3^T b_1) / (a_1^T b_1)$, $c_3 = (a_2^T b_3 + a_3^T b_2) / (a_1^T b_1)$, $c_4 = (a_3^T b_3) / (a_1^T b_1)$, $a_1 = (m_{11} - n_{11} \tilde{z}_{i+1})$, $a_2 = (m_{12} - n_{12} \tilde{z}_{i+1})$, $a_3 = (m_{13} - n_{13} \tilde{z}_{i+1})$, $b_1 = (m_{11} n_{12} - m_{12} n_{11})$, $b_2 = 2(m_{11} n_{13} - m_{13} n_{11})$, $b_3 = (m_{12} n_{13} - m_{13} n_{12})$.

위 다항식은 네 개의 폐쇄적 형태의 해를 준다 (구체적인 해의 형태는 참조 문헌 [8]에 주어지고 있다).

다음은 식 (11)에 주어진 영상 목적 함수를 살펴본다.

Rodrigues 식: $\exp([\hat{w}_{i1}] \theta_{i1}) = I + \sin(\theta_{i1})[\hat{w}_{i1}] + (1 - \cos(\theta_{i1}))[\hat{w}_{i1}]^2$

$$\text{과 관계식: } \cos(\theta_{i1}) = \frac{1 - t_{i1}^2}{1 + t_{i1}^2}, \quad \sin(\theta_{i1}) = \frac{2t_{i1}}{1 + t_{i1}^2},$$

$t_{i1} = \tan(\theta_{i1}/2)$ 을 사용하여 식 (8)을 다음과 같이 고쳐 쓸 수 있다.

$$q_{i+1} = \frac{m_{21}t_{i1}^2 + m_{22}t_{i1} + m_{23}}{n_{21}t_{i1}^2 + n_{22}t_{i1} + n_{23}} \quad (14)$$

여기서 $m_{21} = C_q + k_{21} - k_{22} \in R^3$, $m_{22} = 2k_{23} \in R^3$, $m_{23} = C_q + k_{21} + k_{22} \in R^3$, $n_{21} = 1$, $n_{22} = 0$ and $n_{23} = 1$,

$$C_q = \frac{f(Y_i - Y_i)}{z_{i+1} - e_3^T Y_i} \in R^3$$

$$k_{21} = f(I + [\hat{w}_{i1}]^2) \left(\frac{\exp([\hat{w}_{i2}] \theta_{i2})(Y_{0(i+1)} - Y_{0i})}{z_{i+1} - e_3^T Y_i} \right) \in R^3$$

$$k_{22} = -f[\hat{w}_{i1}]^2 \left(\frac{\exp([\hat{w}_{i2}] \theta_{i2})(Y_{0(i+1)} - Y_{0i})}{z_{i+1} - e_3^T Y_i} \right) \in R^3$$

$$k_{23} = f[\hat{w}_{i1}] \left(\frac{\exp([\hat{w}_{i2}] \theta_{i2})(Y_{0(i+1)} - Y_{0i})}{z_{i+1} - e_3^T Y_i} \right) \in R^3$$

식 (14)의 최적화 1차 필요 조건, i.e., $\frac{\partial J}{\partial t_i} = 0$ 을 풀어서 정리하면 다음과 같은 4차 다항식을 얻는다.

$$t^4 + c_1 t^3 + c_2 t^2 + c_3 t + c_4 = 0 \quad (15)$$

여기서, $c_1 = (a_1^T b_2 + a_2^T b_1) / (a_1^T b_1)$, $c_2 = (a_1^T b_3 + a_2^T b_2 + a_3^T b_1) / (a_1^T b_1)$, $c_3 = (a_2^T b_3 + a_3^T b_2) / (a_1^T b_1)$, $c_4 = (a_3^T b_3) / (a_1^T b_1)$

$a_{21} = (m_{21} - n_{21} \tilde{q}_{i+1})$, $a_{22} = (m_{22} - n_{22} \tilde{q}_{i+1})$, $a_{23} = (m_{23} - n_{23} \tilde{q}_{i+1})$

$b_{21} = (m_{21} n_{22} - m_{22} n_{21})$, $b_{22} = 2(m_{21} n_{23} - m_{23} n_{21})$,

$b_{23} = (m_{22} n_{23} - m_{23} n_{22})$

위 다항식은 네 개의 폐쇄적 형태의 해를 준다 (구체적인 해의 형태는 참조 문헌 [8] 에 주어져 있다).

제안 1: 인체 관절의 3차원 운동 복원

주어진 $Y_{0(i+1)}$, Y_{0i} , Y_i 에 대해, 다음 알고리즘은 θ_{i1} , θ_{i2} , z_{i+1} , q_{i+1} 를 구한다.

단계 1: 식 (13)을 만족하는 t_{i2} 의 해 (폐쇄적 형태의 해) 를 계산한다.

단계 2: 관계식: $t_{i2} = \tan(\theta_{i2}/2)$ 를 사용하여 θ_{i2} 를 계산하고, 식 (12)를 사용하여 z_{i+1} 를 계산한다.

단계 3: 조건식: $z_{i+1} > 0$ 과 관절각 θ_{i2} 에 대한 제한을 고려하여, 적절한 결과값을 선택한다.

단계 4: 식 (15)를 만족하는 t_{i1} 의 해 (폐쇄적 형태의 해) 를 계산한다.

단계 5: 관계식: $t_{i2} = \tan(\theta_{i2}/2)$ 을 사용하여 θ_{i1} 를 계산하고, 식 (14)를 사용하여 q_{i+1} 를 계산한다.

단계 6: 관절각 θ_{i1} 에 대한 제한을 고려하여, 적절한 결과값을 선택한다.

3. 인체의 관절 사슬에 적용할 시각적 해석 알고리즘

본 절에서는 앞 절에서 제안한 알고리즘을 인체 운동을 해석하는 데 어떻게 적용할 수 있는지 살펴본다. 그 핵심

아이디어는 앞 절에서 제안한 알고리즘을 인체 관절 사슬의 각 관절에 round-robin 방식으로 적용하는 것이다.

이 알고리즘을 설명하기 위해서, (그림 1)에 주어진 오른 팔 사슬의 예를 살펴본다. 이 그림에서 Y_{01} , Y_{02} , Y_{03} , $Y_{04} \in R^3$ 는 정의된 영점 자세에서의 관절 위치를 나타낸다 (앞에서도 설명하였듯이, 영점 자세는 모든 관절각이 0인 자세이다). 이 관절 위치들이 각각 $(\theta_{11}, \theta_{12}, \theta_{13})$, $(\theta_{21}, \theta_{22}, \theta_{23})$, $(\theta_{31}, \theta_{32}, \theta_{33})$ 만큼의 회전에 의해 $Y_1, Y_2, Y_3, Y_4 \in R^3$ 로 움직인다고 생각하자. 영점 자세에서 관절들의 축이 $(\hat{w}_{11}, \hat{w}_{12}, \hat{w}_{13})$, $(\hat{w}_{21}, \hat{w}_{22}, \hat{w}_{23})$, $(\hat{w}_{31}, \hat{w}_{32}, \hat{w}_{33}) \in R^3$ 로 표시된다고 하자. 그러면 Y_1, Y_2, Y_3, Y_4 는 아래에 주어진 지수들의 곱 (product-of-exponentials) 식으로 표현된다.

$$Y_1 = Y_{1zero}, Y_2 = R_1 Y_{2zero}, Y_3 = R_2 R_1 Y_{3zero},$$

$$Y_4 = R_3 R_2 R_1 Y_{4zero}$$

여기서,

$$R_1 = \exp([\hat{w}_{11}] \theta_{11}) \exp([\hat{w}_{12}] \theta_{12}) \exp([\hat{w}_{13}] \theta_{13})$$

$$R_2 = \exp([\hat{w}_{21}] \theta_{21}) \exp([\hat{w}_{22}] \theta_{22}) \exp([\hat{w}_{23}] \theta_{23}) \quad (16)$$

$$R_3 = \exp([\hat{w}_{31}] \theta_{31}) \exp([\hat{w}_{32}] \theta_{32}) \exp([\hat{w}_{33}] \theta_{33})$$

그리고, $(\hat{w}_{11}, \hat{w}_{12}, \hat{w}_{13})$, $(\hat{w}_{21}, \hat{w}_{22}, \hat{w}_{23})$, $(\hat{w}_{31}, \hat{w}_{32}, \hat{w}_{33})$, $Y_{02}, Y_{03}, Y_{04} \in R^3$ 는 (그림 1)의 배치와 관절들 사이의 거리를 고려하면 쉽게 구할 수 있다. 그러면, 인체의 자세 복원 과정은 다음에 주어진 제안 2와 같이 정리될 수 있다.

제안 2: (그림 1)에 주어진 인체 관절 사슬의 3차원 운동 복원:

단계 1: z_1 값을 선택하고, 식 (7)을 사용하여 Y_1 을 계산한다.

단계 2: (그림 1)의 배치와 일치하도록, \hat{w}_{i1} 을 e_3 방향의 단위 벡터로 선택하고, \hat{w}_{i3} 를 Y_{0i} 와 $Y_{0(i+1)}$ 사이의 링크 방향의 단위 벡터로 선택한다. (그림 1)의 배치와 각 관절 위치 사이의 거리를 고려하여, $Y_{01}, Y_{02}, Y_{03}, Y_{04} \in R^3$ 을 선택한다. 이때 주의할 점은 $Y_{01} = Y_1$ 가 되도록 선택한다.

단계 3: $Y_1 = Y_{01}$ 로 놓고, $i = 1$ 부터 $i = 3$ 까지 단계 4와 단계5를 반복한다.

단계 4: 주어진 Y_i 에 대해서, 제안 1을 사용하여 $(\theta_{i1}, \theta_{i2})$, z_{i+1} , q_{i+1} 를 구한다.

단계 5: 식 (7)을 사용하여 Y_{i+1} 를 구한다.

단계 6: 인체의 각 관절 사슬에 대해 단계 1부터 단계 5를 반복한다.

제안 2의 단계 1에서, 정확한 z_1 값을 알면 인체의 정확한 위치를 구할 수 있다. z_1 값을 구하는 간단한 방법 중에 하나는 해당 부위의 투영된 영상 크기와 실제 크기를 비교하는 것이다.

본 절을 마치기 전에, 관절의 영상 위치 \tilde{q}_{i+1} 를 구하는 방법을 설명한다. 비전 문헌에서 많은 종류의 2차원 영상 위치 추적 알고리즘들이 제안되었다 [9], [10]. 이러한 기존의 영상 위치 추적 알고리즘을 영상 시퀀스 상의 각 관절에 적용하면, 각 영상프레임에서의 \tilde{q}_{i+1} 를 구할 수 있다. 이 중에서 3차원 자세를 알고 싶은 영상 프레임에 대해서 제안하는 알고리즘 (제안 2)를 적용한다.

본 논문에서 제안한 알고리즘이 제대로 동작하기 위한 영점 자세의 최소 조건은 다음과 같은 관절 좌표축을 선택할 수 있는 것이다: 1) 수직 (6)에서 사용된 바와 같이 $\hat{w}_{i1} = e_3$ 로 선택

2) 수식 (2)에서 사용된 바와 같이 W_{i3} 를 링크 벡터 $Y_{0(i+1)} - Y_{0i}$ 에 평행하게 선택

이 두 가지 조건을 만족하는 관절 좌표축을 선택할 수 있으면, 카메라의 영상 평면에 평행한 한 평면에 모든 관절이 포함되는 자세를 영점 자세로 선택하는 것이다. 이러한 영점 자세의 한 예는 (그림 2)처럼 카메라를 정면으로 바라보며 차려 자세로 서는 방법이 있다. 다음 절의 실험에서는 이 자세를 영점 자세로 선택하였다.

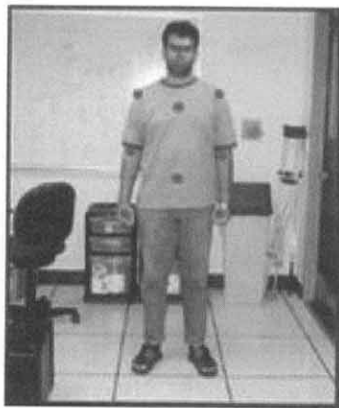
4. 결 과

본 절에서는 제안한 알고리즘의 성능을 검증한다. 이를 위해서 몇 가지 2차원 영상 추적의 예를 실험 시나리오로 선택하여, 3차원 배치를 알고자 하는 영상 프레임을 골라서 본 논문에서 제안한 알고리즘을 적용한다. 이 실험 시나리오의 클럭 속도가 1.20 GHz이고 RAM이 540 메가 바이트인 펜티엄 PC에서 MATLAB을 사용하여 수행된다. 이 실험 시나리오는 3차원 운동 검출과 3차원 운동 추적의 2종류로 구성되어 있다.

4.1 실험 시나리오와 질적 결과

첫 번째 시나리오는 운동 검출이다. (그림 2-(a), (c))는 참고 문헌 [9]에서 제시된 Gritai 알고리즘의 영상 추적 결과이다. 제안된 알고리즘을 (그림 2-(a), (c))의 관절에 적용한 결과 (그림 2-(b), (d))와 같이 3차원 배치가 복원되었다.

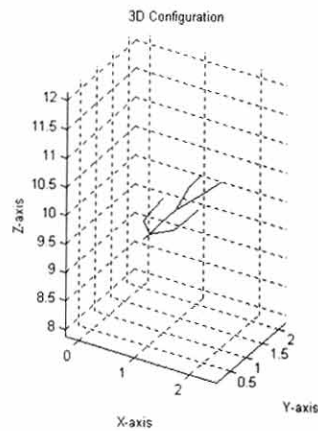
두 번째 시나리오는 운동 추적이다. (그림 3-(a), (c), (e), (g))는 참고 문헌 [9]에서 제시된 Gritai 알고리즘의 2차원



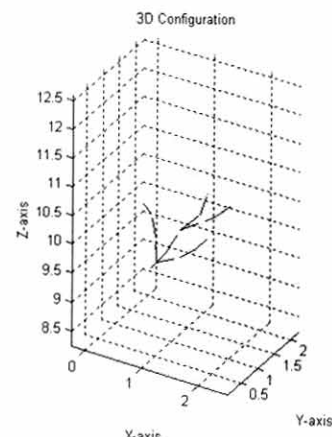
(a)



(c)



(b)

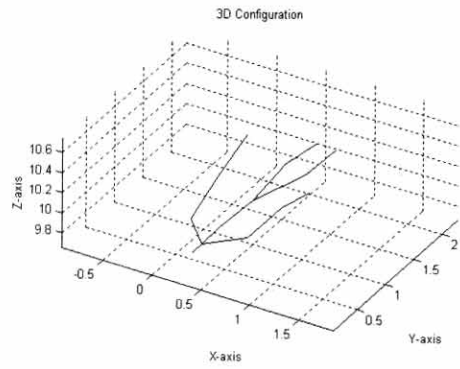


(d)

(그림 2) 첫 번째 실험 시나리오: 3차원 운동 검출. (a), (c)는 참고 문헌 [9]에서 제시된 Gritai의 알고리즘의 2차원 관절 위치의 추적 결과이다. (b), (d)는 제안된 알고리즘을 통해 복원한 3차원 관절 위치들이다.



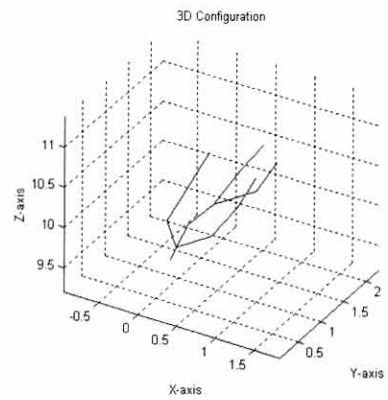
(a)



(b)



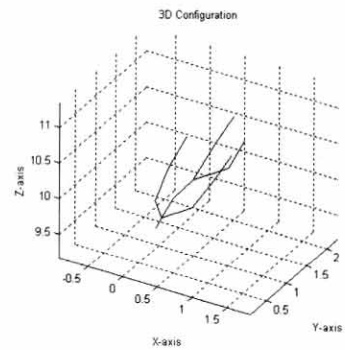
(c)



(d)



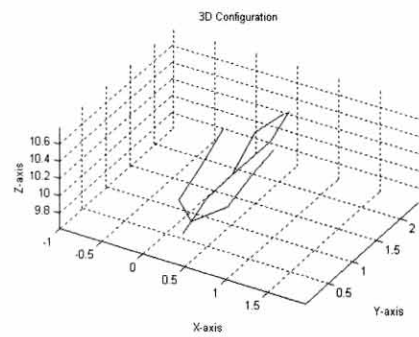
(e)



(f)



(g)



(h)

(그림 3) 두 번째 실험 시나리오: 3차원 운동 추적. (a), (c), (e), (g)는 참고 문헌 [9] 에서 제시된 Gritai의 알고리즘의 2차원 관절 위치의 추적 결과이다. (b), (d), (f), (h)는 제안된 알고리즘을 통해 복원한 3차원 관절 위치들이다.

운동 추적 결과이다. 제안된 알고리즘을 (그림 3-(a), (c), (e), (g))의 관절에 적용한 결과 (그림 3-(b), (d), (f), (h))와 같이 3차원 배치가 복원되었다.

4.2 양적 분석

다음으로, 복원된 3차원 배치의 에러를 분석한다. 정확한 에러 분석을 위해서는, 모의 실험이 효과적이다. 그 이유는 에러 계산에 필요한 정확한 참값을 알 수 있고, 측정 잡음을 조절할 수 있기 때문이다. 이런 이유로, (그림 2)의 실험 조건을 반영하도록 설계된 모의 실험을 고려한다. 이 모의 실험에서, (그림 2-(b), (d))의 복원된 3차원 관절 위치는 영상 평면으로 투영된다. 이렇게 투영된 관절의 영상 위치에 잡음을 조절해서 더하면서 에러 분석을 수행한다. 이 모의 실험에서, 가우시안(Gaussian) 잡음을 사용한다. 이 잡음 모델에서, k-화소 잡음은 평균이 0이고, 표준편차가 k-화소인 잡음을 말한다. 같은 모의 실험 과정이 (그림 3)에도 적용된다.

테이블 I, II 는 각각 (그림 2, 3)의 복원 오차 비율을 퍼센트로 표현 (계산 시간을 초로 표현)한 값을 보여주고 있다. 여기서, 한 개 관절에 대한 복원 오차 비율은 다음과 같이 정의된다.

$$\frac{\sum_{i=1}^n \| Y_i - Y_r \|}{(n \| Y_r \|)}$$
 여기서 Y_i 와 Y_r 는 각각 계산된 3차원 위치와 그 참값을 나타내고, n 은 추적점의 개수를 나타낸다. 이 표들에서, 제안한 알고리즘의 결과를 기존의 알고리즘인 Lee의 알고리즘 [11] 과 Difrancio의 알고리즘 [12] 의 결과와 비교한다. Lee의 알고리즘은 폐쇄적 형태의 알고리즘이고, Difrancio의 알고리즘은 반복 계산 형태의 알고리즘이다.

<표 1> 첫 번째 실험 시나리오에서 복원된 3차원 관절 위치의 평균 오차 비율 (초 단위 계산 시간)

영상 잡음	제안하는 방법	Lee	Difrancio
0.3	0.0097 (0.064)	0.0180 (0.019)	0.0118 (0.258)
0.5	0.0108 (0.066)	0.0228 (0.020)	0.0136 (0.303)
0.7	0.0125 (0.064)	0.0250 (0.018)	0.0184 (0.309)
0.9	0.0152 (0.062)	0.0305 (0.019)	0.0220 (0.315)
평균	0.0121 (0.064)	0.0241 (0.019)	0.0165 (0.297)

<표 2> 두 번째 실험 시나리오에서 복원된 3차원 관절 위치의 평균 오차 비율 (초 단위 계산 시간)

영상 잡음	제안하는 방법	Lee	Difrancio
0.3	0.0091 (0.062)	0.0135 (0.018)	0.0068 (0.273)
0.5	0.0126 (0.064)	0.0176 (0.021)	0.0111 (0.324)
0.7	0.0129 (0.062)	0.0206 (0.018)	0.0151 (0.264)
0.9	0.0138 (0.061)	0.0236 (0.021)	0.0187 (0.222)
평균	0.0121 (0.0623)	0.0188 (0.020)	0.0129 (0.271)

이 테이블로부터 제안된 알고리즘은 기존의 폐쇄적 형태의 알고리즘에 비해 훨씬 정확한 복원 결과값을 준다는 사실을 확인할 수 있다. 제안된 알고리즘의 복원 결과의 정확도는 기존의 반복 계산 형태의 알고리즘에 가깝다. 더구나, 제안된 알고리즘의 정확도는 기존의 두 알고리즘에 비해 영상 잡음에 덜 민감하다. 영상 위치를 정확하게 찾아 내기가 쉽지 않다는 사실을 감안한다면, 제안된 알고리즘의 이러한 특성은 매우 유용하다고 할 수 있다.

계산 시간의 견지에서 보면, 제안하는 알고리즘이 기존의 반복 계산 형태의 알고리즘에 비해 계산 시간이 훨씬 짧다. 제안하는 알고리즘의 계산시간은 기존의 폐쇄적 형태의 알고리즘과 큰 차이가 나지 않는다. 기존의 반복 계산 형태의 알고리즘은 1초에 3~4 프레임 밖에 계산하지 못하는데, 이는 실시간으로 사용하기에는 느린 속도이다. 반면, 제안하는 알고리즘은 이에 비해 약 4배 정도 빠르다.

5. 결론

본 논문에서는 단수의 카메라를 사용하는 인체 운동의 시각적 해석 알고리즘을 제안하였다. 실시간 응용에 편리하도록 3차원 복원 알고리즘을 폐쇄적 형태로 설계하였다. 성능 검증을 통해 제안한 알고리즘은 기존의 반복 계산 형태의 알고리즘에 비해 훨씬 적은 계산 시간을 가졌고, 기존의 폐쇄적 형태의 알고리즘에 비해 훨씬 높은 정확도를 가졌음을 확인하였다. 또한 제안하는 알고리즘은 기존의 반복 계산 형태의 알고리즘과 폐쇄적 형태의 알고리즘에 비해 영상 잡음에 덜 민감하였다. 이러한 특성은 관절의 영상 위치를 정확하게 찾아 내기가 쉽지 않다는 점을 고려할 때 매우 유용한 특성이라 할 수 있다.

참고 문헌

[1] K. Rohr, "Incremental recognition of pedestrians from image sequences," Proc. IEEE Comput. Vision and Pattern Recognition, pp.8-13, New York City, June, 1993.

[2] J. Zhao, "Moving posture recognition from perspective projections of jointed figure motion," PhD thesis, University of Pennsylvania, 1993.

[3] 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.

[4] C. Bregler, J. Malik and K. Pullen, "Twist based acquisition and tracking of animal and human kinematics", International Journal of Computer Vision, vol.56, No.3, pp.179-194, 2004.

[5] C. J. Taylor, "Reconstruction of Articulated Objects from

Point Correspondences in a Single Uncalibrated Image," Computer Vision and Image Understanding,, Vol.80, No.10, pp.349-363, October, 2000.

[6] R. M. Murray, Z. Li, and S. S. Sastry, "A mathematical introduction to robotic manipulation," CRC Press, pp. 28, 1994.

[7] A. Zomet, D. Feldman, S. Peleg, and D. Weinshall, "Mosaicing new views: the crossed-slits projection," IEEE T. Pattern Analysis and Machine Intelligence, Vol.25, No.6, pp.741-754, June, 2003.

[8] I. Z. Emiris, "Solving polynomial equations: foundations, algorithms, and applications," Springer Verlag, pp.33, July, 2005.

[9] A. Gritai and M. Shah, "Tracking of human body joints using anthropometry," International Conference of Multimedia and Expo, pp.1037-1040, 2006.

[10] Neil J. Gordon, Simon Maskell, Thiagalingam Kirubarajan, "Efficient particle filters for joint tracking and classification", Signal and Data Processing of Small Targets 2002, pp.439-449, 2002.

[11] S. U. Lee and I. Cohen, "3D Hand and Fingers Reconstruction from Monocular View," Proc. 17th Int. Conf. Pattern Recognition (ICPR04), pp.231-236, 2004.

[12] D. E. Difranco, Tat-Jen CHAM, and James M. REHG, "Reconstruction of 3-D Finger Motion from 2-D Correspondences," Proc. Conf. Computer Vision and Pattern Recognition, pp.307.314, 2001.



한 영 모

e-mail : ymhan123@hanmail.net

1992년 서울대학교 물리교육학과(이학사, 준우등 졸업)

1995년 서울대학교 제어계측공학과(공학사, 차석 졸업)

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, 정보 기술을 위한 통합 과학적 접근법