

스테레오 영상의 변이 정보를 이용한 머리 회전 각도 추정에 관한 연구

김택훈, 장중환
배재대학교 정보통신공학과

e-mail : aroide@mail.pcu.ac.kr, jangjw@mail.pcu.ac.kr

A Study on Head Rotation Angle Estimation Using Disparity Information of Stereo Images

Taek-Hoon Kim, JongWhan Jang
Dept. of Information Communication Engineering, PaiChai University

요 약

본 논문은 스테레오 카메라를 통해 얻은 변이 정보를 이용하여, 3 차원으로 머리의 회전 각도를 추정하는 방법을 제안한다. 머리 회전에 의한 주시 방향은 사람이 관심을 가지는 방향이므로 이동을 추정하는 것에 비해 많은 중요성을 갖는다. 본 논문에서는 얼굴 영역 내의 여러 특징점들 중 3 개의 특징점들을 포함하는 여러 평면(Plane)이 머리가 이동하더라도, 그 평면들 사이의 각은 변하지 않으므로, 회전 각도 추정에 영향을 주지 않는 점을 이용하여 회전 각도를 추정하였다. 또한, 기존 논문이 카메라 보정을 통해 3 차원 위치를 측정하지만, 제안하는 방법은 변이 공간에서 회전 각도를 추정하기 때문에 카메라 보정 과정이 필요하지 않다. 변이 정보를 얻기 위한 스테레오 장비는 평행 카메라 모델로 가정하며, 얼굴 내의 특징점은 KLT(Kanade-Lucas-Tomasi) 특징 추적 알고리즘을 이용하였다. 실험결과는 기존 영상에 대하여 추정된 3 차원 각도를 나타낸다.

1. 서론

IT 산업의 발전과 더불어, 컴퓨터의 사용은 더욱 많은 필요성을 갖게 되었으며, 사람과 컴퓨터간의 좀 더 편리한 인터페이스를 위한 연구가 많이 진행 되어지고 있다.[1][2] 뿐만 아니라, 로봇 비전을 연구하는 분야에 있어서 사람이 주시하는 방향을 추정하는 것은 중요한 문제 중 하나이다.[5][6][7] 이는 핸드 프리 마우스(Hands free mouse) 를 개발하기 위한 기술로도 적용이 가능하며, 운전자 감시 시스템으로도 응용될 수 있다.[8]

사람의 얼굴이 향하고 있는 방향은 그 사람이 주시하고 있는 방향이며, 관심을 가지는 방향이라고 할 수 있다. 다시 말하면, 사람의 머리 방향을 추정함으로써 관심 객체를 추적할 수 있으며, 또한 주시 방향 등을

예측할 수 있다.

사람의 머리 방향을 추정하는 것은 곧, 머리의 회전 각도를 추정하는 것과 상통한다. 즉, 머리의 회전 각도를 예측함으로써, 사람의 관심이 되는 방향을 추정할 수 있다.

사람의 머리 자세를 추정하는 방법은 현재에 이르기까지 많이 연구 되어 지고 있다. 기존 방법에서 단일 영상을 이용하는 방법은 영상의 밝기 정보에 민감하며, 얼굴의 피부 영역과 머리카락 부분을 나누어 추정하는 영역 기반 방법은 머리털 모양이 변하면 추정이 어려워지는 문제가 있다.[1] 또한, 임의의 점에 대한 3 차원 위치 측정을 필요로 하는 방법은 실제 깊이를 계산하기 위해 카메라 보정을 통한 내부 및 외부 인자 측정이 필수적이다. 본 논문에서는 스테레오 카메라를 통해 얻은 변이 정보를 이용하여, 영상 내의 얼굴 영역과 배경 영역을 쉽게 구별하였으며, 카메라 보정 없이 변이 공간에서 머리의 회전 각도를 추정하

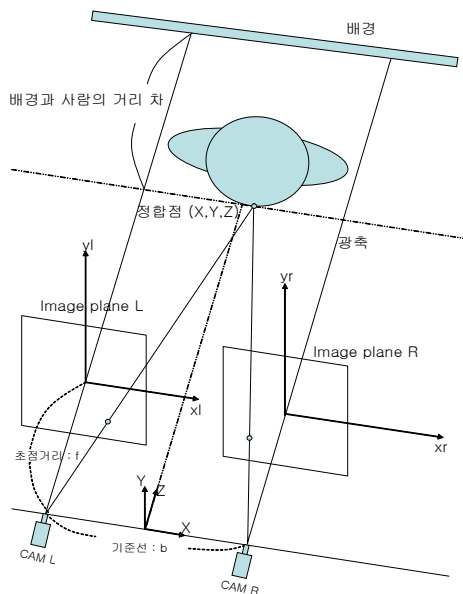
였다.

본 논문에서 회전 각도 추정의 기준이 되는 초기 영상의 얼굴 방향은 카메라를 정면으로 향하고 있다고 가정하며, 스테레오 장비는 평행 시각 스테레오 카메라 모델(A parallel looking stereo camera model)을 사용하였다. 실험결과는 기준 영상에 대하여 추정된 3 차원 각도를 나타낸다.

2. 얼굴 검출

2.1 배경 분리

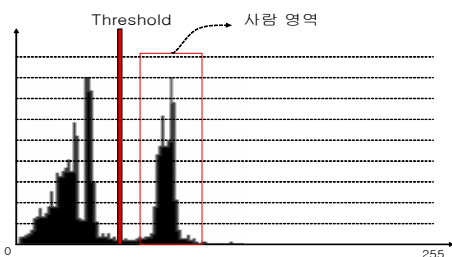
기준 영상으로부터 구해진 초기 변이 정보를 이용하면, 사람과 배경을 쉽게 구별할 수 있다. 본 논문에서 사용된 스테레오 장비는 다음과 같은 구조를 갖는다.



[그림 1] 평행 시각 스테레오 장비 구조

[그림 1]의 카메라의 두 광축이 평행하며, 렌즈 중심을 포함하는 에피폴라선(Epipolar line) 상에 있으므로, 수직 불균형(Vertical disparity)에 의한 변이 값은 항상 0 이 되며, 오직 수평 불균형(Horizontal disparity)에 의한 차이만을 고려할 수 있다.

따라서, 수평 불균형에 대한 변이 맵을 히스토그램(Histogram)으로 나타내면 [그림 2]와 같으며, 임계값(Threshold)을 기준으로 배경과의 분리가 가능하다.



[그림 2] 임계값을 이용한 사람과 배경의 분리

$$g(x, y) = \begin{cases} 1 & \text{if } d(x, y) > T \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

식 (1)의 $g(x, y)$ 는 임계값을 기준으로 이진화된 영상이며, T 는 임계값, $d(x, y)$ 는 변이 영상을 나타낸다. 즉, 변이 영상의 히스토그램을 임계값을 기준으로 나누어 이진화 함으로써, 사람 영역을 구할 수 있다.

2.2 얼굴 영역 검출

일반적으로 사람의 목 부분은 다른 부분에 비해 좁은 영역을 차지한다. 따라서, 얼굴 영역은 이진화 된 영상 $g(x, y)$ 를 가로 방향으로 한 라인씩 스캔하여 영역 합을 계산한 값 중 가장 큰 두 값의 영역 사이에서 가장 작은 값을 갖는 부분이라고 할 수 있다. 가장 큰 값을 갖는 두 영역을 p, q 라고 하면, 목 부분을 검출하기 위한 식은 다음과 같이 표현될 수 있다.

$$\begin{cases} l(y) = \sum_{x=0}^{M-1} g(x, y), & y = p+1, \dots, q-1 \\ h = \min_y \{l(y)\} \end{cases} \quad (2)$$

식(2)의 M 은 영상의 가로크기를 h 는 검출된 목의 위치를 나타낸다. 따라서, 사람 영역 내에서 검출된 목 위치로부터 그 윗부분이 얼굴 영역이 된다.

3. KLT 특징 추적 알고리즘

본 논문에서는 프레임과 프레임 사이의 특징점 (Feature Points)을 추적하기 위해 KLT 특징 추적 알고리즘을 사용한다.[3][4] 특징점은 얼굴 영역내에서 추적이 가장 잘 될 것이라고 생각되는 점들로 구성된다. 만약, 영상 시퀀스의 샘플링 속도가 충분히 빠르다고 가정하면, 영상내의 임의의 작은 영역에 대한 영상 강도의 변화는 프레임과 프레임 사이에 거의 없다고 생각할 수 있다. 따라서, 프레임과 프레임 사이의 모션은 식(3)과 같이 표현될 수 있다.

$$\varepsilon = \sum_w [I(x+d, t+\tau) - I(x, t)]^2 \quad (3)$$

식(3)의 w 는 d 를 계산하기 위한 윈도우 크기이며, $x=[u, v]^T$ 는 윈도우의 중심에 위치한다. τ 는 프레임과 프레임 사이의 시간 변화량을 나타낸다. 식(3)은 1 차 테일러 근사(First-order Taylor Approximation)에 의해 식 (4)와 같이 나타낼 수 있다.

$$Gd = e, \quad (4)$$

여기서 G, e 는

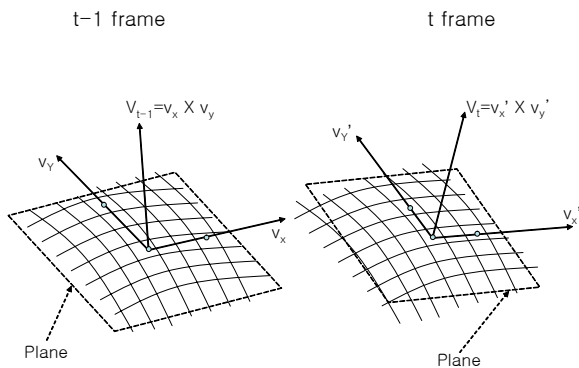
$$G = \sum_w \begin{bmatrix} I_u^2 & I_u I_v \\ I_u I_v & I_v^2 \end{bmatrix}, \quad e = -\tau \sum_w I_t [I_u \quad I_v]^T$$

이고, I_u, I_v 는 u, v 각각에 대한 편미분이며, $I_t = I_v$ 이다. 식 (4)로부터 $d = G^{-1}e$ 로 표현될 수 있으며, d 의 값을 계산함으로써 프레임과 프레임 사이의 특징점의 변화를 추적할 수 있다.

4. 제안된 회전 각도 추정 방법

4.1 법선 벡터(Normal vector)

얼굴 내에서 추출된 임의의 3 점을 포함하는 평면은 얼굴이 회전 없이 순수하게 이동만 한다고 가정하면, 그 평면과 3 차원상의 XY, YD, DX 평면의 사이 각은 변하지 않는다. 즉, 평면 사이의 각도는 그대로 유지되며, 이것은 평면의 회전 각도가 얼굴의 회전에 의해서만 영향을 받는다는 것을 의미한다.



[그림 3] 변이 공간에서의 법선 벡터

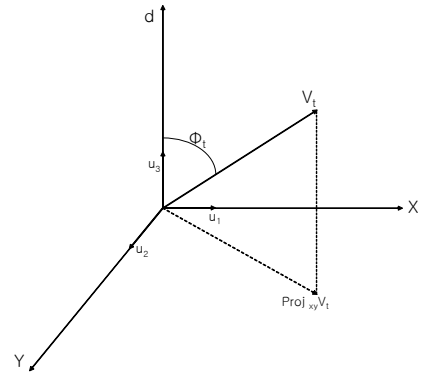
[그림 3]은 변이 공간에서 임의의 평면에 대한 법선 벡터를 나타내고 있다. V_x 와 V_y 는 한 특징점으로부터 다른 두 특징점 방향의 벡터를 의미하며, 두 벡터의 외적(Cross product)을 취함으로써 평면의 법선 벡터를 계산할 수 있다.

$$\begin{aligned} V_{t-1} &= v_x \times v_y, & N_{t-1} &= \frac{V_{t-1}}{\|V_{t-1}\|} \\ V_t &= v_x' \times v_y', & N_t &= \frac{V_t}{\|V_t\|} \end{aligned} \quad (5)$$

식(5)에서 V_{t-1} 은 $t-1$ 프레임에서의 법선 벡터를, V_t 는 t 프레임에서의 법선 벡터를 나타낸다. t 프레임의 법선 벡터 V_t 는 $t-1$ 프레임으로부터 추적된 특징점들로부터 얻어진 법선 벡터를 의미하며, 두 법선 벡터의 차이는 머리가 회전된 각도를 의미한다.

4.2 법선 벡터의 직교투영을 통한 회전 각도 계산

X, Y, d 축을 기준으로 머리의 회전된 각도를 추정하기 위해 두 벡터를 각각 XY, YD, DX 평면에 직교투영(Orthographic projection) 하여, 그 사이 각을 계산한다. 직교 투영은 [그림 4]와 같이 나타내어 진다.



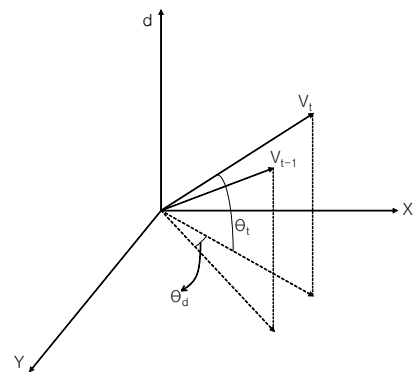
[그림 4] 변이 공간에서 XY 평면에서의 직교투영

[그림 4]는 t 프레임에서 XY 평면에 직교 투영된 벡터 $Proj_{xy}V_t$ 를 표현하고 있다. 직교투영벡터 $Proj_{xy}V_t$ 를 구하기 위해서는 V_t 와 $Proj_{xy}V_t$ 사이의 각을 알아야 하므로, d 축과 V_t 사이의 각 Φ_t 를 먼저 계산하여 각을 구한다.

$$\Phi_t = \arccos \frac{u_3 \cdot V_t}{\|u_3\| \|V_t\|} \quad (6)$$

여기서 u_3 는 d 방향의 단위벡터(Unit vector)를 의미하고, V_t 와 $Proj_{xy}V_t$ 사이의 각은 $90 - \Phi_t$ 가 되므로, $Proj_{xy}V_t$ 는 식(7)로 구해진다.

$$proj_{xy} V_t = V_t \cos(90 - \Phi_t) = V_t \sin \Phi_t \quad (7)$$



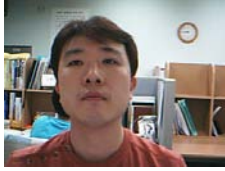
[그림 5] 직교투영 된 두 벡터의 사이 각

$t-1$ 프레임에서의 $Proj_{xy}V_{t-1}$ 또한, 식 (7)과 같은 방법으로 구해지며, 두 투영벡터 사이의 각은 d 축을 기준으로 회전된 각도를 나타낸다. 그러므로, 회전 각도는 식 (8)을 이용하여 계산할 수 있다.

$$\theta_d = \arccos \frac{(V_{t-1} \sin \Phi_{t-1}) \cdot (V_t \sin \Phi_t)}{\|V_{t-1} \sin \Phi_{t-1}\| \|V_t \sin \Phi_t\|} \quad (8)$$

5. 실험 결과

실험은 평행 카메라 모델로 구성된 스테레오 장비로부터 얻어진 변이 맵을 이용 하였으며, 3 개의 값은 기준 영상에 대해 얻어진 회전 각도를 의미한다.



(a) 기준 영상



(b) [43.2, 26.4, 10.3]



(c) [-8.1, 76.4, 2.3]



(d) [-11.5, -33.7, -4.7]



(e) [53.6, 46.3, 1.3]



(f) [-63.2, 1.3, -1.5]

[그림 6] 변이 맵을 이용한 회전 각도 실험결과

[그림 6]은 스테레오 장비를 통해 얻은 변이 맵을 이용하여 구해진 회전 각도이다. 회전 각도는 X, Y, d 축을 기준으로 회전된 각도를 의미한다. (a)는 변이 맵을 이용하여 얼굴 영역을 추출하기 위해 사용된 초기 영상이며, (b) ~ (f)는 매 프레임에 대하여 추적된 영상들로부터 구해진 회전 각도를 나타낸다. 본 논문에서 제안된 방법으로 구해진 회전 각도는 약 +70 ~ -70 도 사이에서 좋은 결과를 보였으며, 그 이상의 각도에서는 특징 추적의 어려움으로 오차가 조금씩 늘어나는 현상을 보였다.

6. 결론

본 논문은 스테레오 장비를 이용하여 얻은 변이 맵을 이용하여, 머리의 회전각도를 추정하는 알고리즘을 제안하였다. 초기 영상으로부터 얻어진 변이 맵은 얼굴 영역과 배경 영역을 구별하는데 효과적으로 이용되었으며, 얼굴 영역 내에서 추적하기에 좋은 특징을 가진 점들을 선택하여 추적하기 위해 KLT 알고리즘을 사용하였다.

얼굴 영역에서 얻어진 특징 점들은 3 개의 점을 포함하는 평면들로 구성될 수 있으며, 평면이 얼굴의 이동에 영향을 받지 않는 점을 이용하여 각 프레임 사이의 추적된 점들의 평면들에 대한 회전 각도를 계산하였다. 회전 각도는 X, Y, d 축을 기준으로 계산되었으며, 실험을 통하여 회전 각도의 추정이 가능함을 보

였다.

본 논문에서 사용된 회전 각도 추정 방법을 이용하면 머리의 이동을 고려하지 않고 머리의 포즈를 예측할 수 있으며, 이동 벡터를 구하고자 할 경우, 구해진 회전 각도를 이용하면 쉽게 계산이 가능할 것이다. 또한, 변이 맵이 아닌 깊이 정보를 얻기 위한 별도의 하드웨어 장비를 이용하면, 좀 더 나은 정확성을 기대할 수 있을 것으로 보인다.

참고문헌

- [1] Q. Chen, H. Wu etc., "3D Head Pose Estimation without Feature Tracking," Proc. 3rd Int'l Conf. on Automatic face and Gesture Recognition, Japan, 1998.
- [2] Qian Chen, Tetsuo Shimada, Haiyuan Wu, Tadayoshi Shioyama, "Head Pose Estimation Using both Color and Feature Information", ICPR 2000, 2842-2841
- [3] J.Shi and C.Tomasi. Good features to track. In Proceedings of the Conference on Computer Vision and Pattern Recognition, Seattle, Washington, USA, pages 593-600, 1994.
- [4] Carlo Tomasi and Takeo Kanade. "Detection and Tracking of Point Features", Carnegie Mellon University Technical Report CMU-CS-91-132, April 1991.
- [5] K. Nickel and R. Stiefelbogen, "Detection and Tracking of 3D-Pointing Gestures for Human-Robot-Interaction", Humanoids 2003, Karlsruhe, Germany, 2003
- [6] R.Stiefelbogen, J. Yang, A. Waibel, "Tracking Focus of Attention for Human-Robot Communication", Humanoids 2001, Tokyo, Japan, 2001
- [7] A. Agah, "Human interactions with intelligent systems: research taxonomy", Computers and Electrical Engineering, pages 71-107, 2001
- [8] Youding Zhu; Fujimura, K.; "Head pose estimation for driver monitoring", Intelligent Vehicles Symposium, pages 501-506, 2004