

모델 기반의 SVM을 이용한 시선 방향 추정

김 종 배, 김 향 준
 경북대학교 컴퓨터공학

e-mail : jbkim@ailab.knu.ac.kr, kimhj@knu.ac.kr

Model based Gaze Direction Estimation Using Support Vector Machine

Jong Bae Kim, Hang Joon Kim

Department of Computer Engineering, Kyungpook National University

Abstract - 실내 환경에서 사람의 행동을 인식하는 시스템을 만들 때 사람의 의도를 파악하는 것은 중요한 정보가 될 수 있다. 사람의 시선방향은 의도를 파악하는데 있어서 깊은 관계가 있다. 본 논문에서는 실내 환경에서 사람의 시선 방향을 모델에 기반하여 추정하는 방법을 제안하였다. 머리 모델은 얼굴 영역과 머리카락 영역을 포함하는 두 개의 겹쳐진 타원으로 표현되고, 각 타원의 파라미터는 시선 방향을 추정하는 정보로 사용된다. 시선 방향은 SVM(Support Vector Machine) 알고리즘을 사용하여 8방향중 하나로 추정된다. 이미지에서 얼굴영역과 머리영역은 색상 정보에 의해 검출된다. 사무실 환경에서 시선방향을 다양하게 변화시켜 실험을 하였고, 이를 통해 성능 평가를 수행하였다.

1. 서 론

비전 기반의 애플리케이션에서 사람의 행동을 인식하는 것은 중요한 것 중 하나이다[1][2]. 스마트 룸과 같은 환경을 만들기 위해 실내 환경에 위치한 감시 카메라를 통해 사람의 행동을 인식하는 연구가 계속 되고 있다. 실내 환경에서 사람의 정확한 행동을 인식하는데 있어서 사람의 의도와 주목하는 방향을 파악하는 것은 중요한 단서가 된다. 사람의 시선 방향은 사람의 의도와 주목하는 방향과 많은 관계가 있기 때문에, 사람의 행동을 인식하는 시스템 개발에 있어서 중요한 단서가 될 수 있다.

최근에는 사람의 지각을 이용하는 애플리케이션들이 많이 개발되고 있기 때문에 시선 방향을 추정하는 연구는 점점 더 많은 관심을 받고 있다[3][4][5]. 다양한 애플리케이션들은 각각의 주제에 맞는 이미지 해상도를 필요로 하며 별개의 기술을 필요로 한다. 지각을 이용한 인터페이스의 경우는 사람의 눈을 보고 판단하기 때문 상당한 고해상도의 이미지를 필요로 한다[6][7]. 이와는 대조적으로, 사람의 제스처나 스마트 룸 애플리케이션에서 사람은 일반적으로 이미지에서 작은 부분으로 나타나기 때문에 사람의 눈을 보고 판단하는 것은 불가능하다.

여기서는 후자의 비교적 저해상도 이미지에서 시선 방향을 추정하는 방법에 대해 언급한다. 시선 방향을 추정하는데 있어서 얼굴의 특징점이나 색상, 모양, 위치 등을 사용할 수 있다. 고해상도의 이미지를 얻을 수 있다면 얼굴의 특징점만으로도 시선 방향을 추정할 수 있지만 앞에서 말한바와 같이 항상 고해상도의 이미지를 획득하기가 쉽지 않다. 그러므로 여기선 색상이나 모양을 사용하여 시선방향을 추정한다.

본 논문에서는 사람이 있는 이미지에서 모델에 기반하여 자동으로 시선 방향을 추정하는 방법을 제시한다. 이미지에서 얼굴 영역은 색상과 모양에 기반하고, 머리카락영역은 색상에 기반하여 추출된다. 이미지에서 검출된 얼굴과 머리는 각각의 타원으로 표현되고, 겹쳐진 한 쌍의 타원으로 표현된다. 두 개 타원의 위치와 크기, 기울기는 Support Vector Machine(SVM)에서 학습되는 정보로 사용된다. SVM 알고리즘에 의해 결과는 8개(그림1)중 한 방향으로 추정된다.

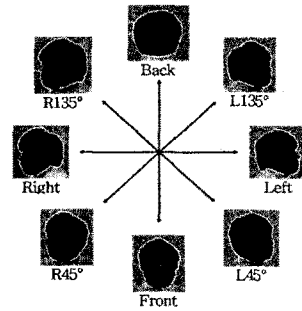
2. 머리 모델링과 검출

2.1 머리 모델링

머리는 얼굴과 머리카락 부분으로 구성된다. 그림 2와 같이 검출된 얼굴과 머리를 각각 가장 잘 표현한 타원이 여기서 사용하는 얼굴과 타원 모델이다. 따라서, 모델은 $e_f=(x_f, y_f, w_f, h_f)$, $e_h=(x_h, y_h, w_h, h_h)$ 파라미터로 구성된 각각의 타원이다. (x, y) 는 타원의 중심 좌표이고, w 와 h 는 타원이 폭과 높이의 길이이다. 또한, 타원은 회전하지 않으며, 얼굴 타원 e_f 는 항상 머리 타원 e_h 의 앞에 위치하고, 머리카락만 존재할 경우에는 얼굴타원이 머리 타원에 가린 것으로 간주하여 시선 방향은 Back으로 추정된다. 타원은 가변의 비율을 가지며, 초기 타원의 크기는 초기 프레임으로부터 계산된다. 그림 2에서 보는 것처럼 얼굴 영역은 얼굴 색상 분포를 가지며, 머리카락 영역(머리에서 얼굴영역을 뺀 부분)은 머리카락의 색상 분포를 진다.

2.2 머리 검출

이미지에서 얼굴 영역과 머리카락 영역을 합하여 머리의 외형을 추출하고, 그것을 잘 감싸는 타원이 모델이다. 머리 검출에서의 목적은 머리의 위치와 이미지의 비율을 결정하고, 그것을 잘 감싸는 타원을 나타내는 것이다.



〈그림 1〉 시선 방향에 따른 이미지 분류

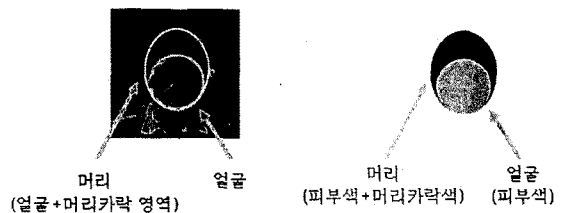
이미지에서 머리를 검출하기 위해, 먼저 얼굴 후보를 분류하고 나서 후보와 이웃하는 머리카락 영역을 검출하는 것이다. 머리카락이 존재한다면 실제 얼굴로 분류한다. 얼굴은 얼굴 영역을 모두 포함하는 안쪽 타원에 의해 표현되고, 머리는 머리 영역을 모두 포함하는 바깥쪽 타원에 의해 각각 표현된다. 하지만, 피부색만으로 얼굴을 검출하는 데는 무리가 있다. 이 결점은 모양정보를 사용하여 효과적으로 제거할 수 있다. 머리 검출에 모양 정보를 추가하기 위해, 피부색으로 검출된 피부 영역에 연결 요소 그룹화를 사용하였다. 그 영역은 피부색 영역을 모두 포함하는 가장 작은 타원으로 나타내어진다. 타원 안의 피부색 비율을 계산하고, 그 비율이 사전에 정의된 임계값보다 크면 피부색 영역은 얼굴 후보로 분류된다. 사람 얼굴의 색 분포는 채도 컬러 공간에서 작은 부분으로 분류된다[4]. 얼굴 패치로부터 CrCg 공간에서 피부색의 채도 R과 채도 G로 구분하여 나타내면 타원 형태로 나타내어진다. SmileCAM Su-320 컬러 비디오카메라로 수집된 이미지로부터 실험한 결과에서 피부색이 분류된 결과는 작은 타원 모양이다. 그러므로 멤버십 함수를 다음과 같은 타원 함수로 나타낼 수 있다.

$$S = \begin{cases} 1, & \frac{(x'-c_x)^2}{a^2} + \frac{(y'-c_y)^2}{b^2} < 1, \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

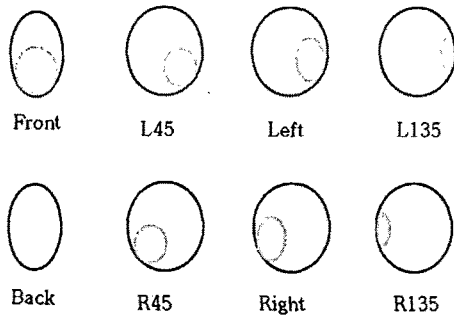
$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (2)$$

$C_x = 0.41$, $C_y = 0.32$, $a = 0.05$, $b = 0.03$, $\theta = 0.20$ (라디안)은 CrCg 공간에서 피부색 클러스터로부터 계산된다.

머리카락 영역 검출도 피부색 영역 검출에서 사용된 방법과 동일한 방법으로 검출한다. 머리카락 패치 샘플을 사용하고 식1을 사용한다. 앞에서와 같이 CrCg 공간에서 머리카락색 클러스터의 파라미터는 $C_x = 0.27$, $C_y = 0.40$, $a = 0.09$, $b = 0.03$, $\theta = 0.90$ (라디안)과 같이 계산된다. 각 시선 방향에 따른 모델은 그림 3과 같다.



〈그림 2〉 겹쳐진 타원 모델



〈그림 3〉 모델에 따른 시선 방향

3. SVM을 이용한 시선 방향 추정

최초에 SVM(Support Vector Machine)은 이진분류(binary classification)를 위하여 개발되었으며, 문자인식과 필기인식뿐만 아니라 얼굴인식 등 다양한 패턴인식 분야에 적용되어 사용되고 있다. SVM은 커널 기반 기법의 한 종류이다. 커널 함수를 사용하는 고차원의 공간에서 특징 벡터를 맵핑하고, 이 공간에서 최적의 선형 경계 분류 함수가 만들어 진다. SVM에서 정확한 커널은 명확히 정의되지는 않지만, 고차원에서 어떤 두 점 사이에 거리는 정의 되어야 한다. 하지만, 일반적으로 선형 경계가 입력 벡터를 분류하기에는 한계가 있다. 이와 같은 경우 SVM은 입력 벡터를 보다 고차원 공간 내의 벡터로 변형한 후 선형 경계 분류 함수를 찾는 문제로 변형하여 SVM을 구성하게 된다. 이와 같이 이진 분류를 위하여 설계되었지만 다른 기계학습방법에 비해 우수한 성능을 보이며, 고차원에서 데이터가 적은 경우에도 성능이 우수하다.

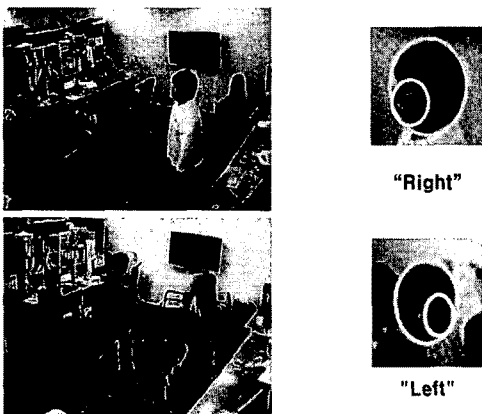
본 논문에서는 시선 방향 추정을 위해 7차원 특징 벡터 $[H_w, H_h, F_w, F_h, R, V_x, V_y]$ 를 사용하였다. H_w 와 H_h 는 각각 머리타원의 폭과 높이이며, F_w, F_h 는 얼굴 타원의 폭과 높이를 나타낸다. R 은 두 타원의 넓이 비율이며, V_x 와 V_y 는 두 타원의 중심간의 거리 벡터를 나타낸다. 얼굴 타원은 정면에서 뒤로 회전할수록 타원의 넓이가 좁아지며, 두 타원의 중심 간의 거리가 증가한다.

앞에서 정의된 7차원 특징 벡터를 3단계의 SVM을 사용하여 분류하였다. 각각의 단계는 이진 분류를 계속적으로 반복하여 사용하였다. 각 단계는 360°를 기준으로 1/2씩 줄여가며 분류된다. 첫 단계에서는 두 가지군(1~180°, 181~360°)중 하나로 분류한다. 다음 단계에서는 4군(1~90°, 90~180°, 181~270°, 271~360°)중 하나로, 마지막 단계에서는 최종적으로 8군(1~45°, 46~90°, 91~135°, 136~180°, 181~225°, 226~270°, 271~315°, 315~360°)중 하나로 시선 방향이 추정된다. 이렇게 SVM을 이용하여 이진분류를 반복하여 사용함으로써 빠르게 분류하고, 인식률을 높일 수 있다.

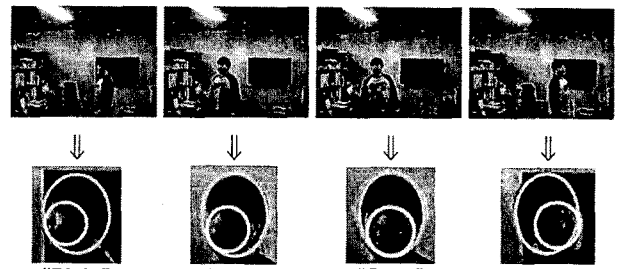
4. 실험 및 결과

실험은 연구실에서 고정된 단일 카메라로부터 획득된 이미지를 사용하였다. Visual C++를 사용하여 구현하였으며, 실험을 위해 OpcnCV에서 제공되는 함수들을 사용하였다. 320x240 크기의 24비트 이미지를 사용하였다.

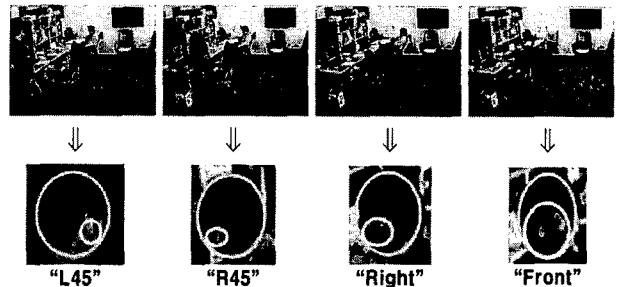
8개의 시선 방향별로 20장씩 총 160장의 이미지를 이용해 학습하였으며, 이를 성능 평가 하기 위해 각 방향별로 20장씩 총 160장의 이미지를 테스트 하였다. 입력 이미지에 대한 시선 방향은 약 75%의 인식률을 보였다. 그림 4에서 실험 결과를 보여준다. 그림 5에서는 연속 영상에서 시선 방향의 변화에 따른 결과를 보여준다. 그림 6에서와 같이 수직의 시선 방향이 차이가 클 경우에는 여러가 발생하였다.



〈그림 4〉 시선 방향 추정의 결과



〈그림 5〉 연속 영상에서의 시선방향 추정 결과



〈그림 6〉 수직의 시선방향의 변화가 클 경우

5. 결 론

스마트 롬과 같은 비전 기반의 애플리케이션을 구현하기 위해 사람의 행동 인식이 선행되어야 한다. 시선 방향 추정은 사람의 의도를 파악하는데 있어서 중요한 문제 중 하나이다. 본 논문에서는 실내 환경에서 사람의 시선 방향을 모델에 기반하여 추정하는 방법을 제안하였다. 이미지에서 사람의 머리와 얼굴을 검출하고, 각각을 잘 감싸는 타원이 모델이다. 각 타원의 파라미터는 SVM 알고리즘에서 시선 방향을 추정하는 정보로 사용된다. 타원으로부터 얻어진 7차원 특징 벡터를 SVM 알고리즘을 사용하여 8방향중 하나로 추정된다. 실험 결과, 타원 모델로 사람의 시선 방향을 추정할 수 있었다.

[참 고 문 헌]

- [1] Kim Shearer, Kirrily D Wong, Svetha Venkatesh, "Combining multiple tracking algorithms for improved general performance.", *Pattern Recognition*, Vol. 34, 1257-1269, 2001.
- [2] Michael Kass, Andrew Witkin, Demetri Terzopoulos, "Snakes: active contour models.", *Int. J. Computer Vision*, Vol. 1, 321-331, 1996.
- [3] Annie Xin Guan, Harold H. Szu, "A local face statistics recognition methodology beyond ICAand/orPCA.", *International Joint Conference on Neural Network*, 1016-1021, 1999.
- [4] Jie Yang, Alex Waibel, "A Real-Time Face Tracker.", *IEEE Workshop on Application of Computer Vision*, 142-147, 1996.
- [5] Ragini Choudhury Verma, Cordelia Schmid, and Krystian Mikołajczyk, "Face Detection and Tracking in a Video by Propagating Detection Probabilities.", *Ieee Transactions On Pattern Analysis And Machine Intelligence*, Vol. 25, No. 10, 1215-1228, 2003.
- [6] Jian-Gang Wang, Eric Sung, Ronda Venkateswarlu, "Eye Gaze Estimation from a Single Image of One Eye.", *IEEE Proceedings ICCV '03*, 136-143, 2003
- [7] Kar-Han Tan, David JKrie gman, Narendra Ahuja, "Appearance-based eye gaze estimation.", *IEEE Proceedings WACV '02*, 191-195, 2002