

## 컨덴세이션과 타원근사를 이용한 얼굴추적기

홍현석, 정명진  
한국과학기술원 전자 전산학과

### Face Tracker Using Condensation and Ellipse Fitting

Hyun Suk Hong, Myung Jin Chung  
Department of Electrical Engineering & Computer Science, KAIST

**Abstract** - 색상정보는 물체의 특성을 나타내는 고유한 특징점이 될 수 있으며, 물체를 인식하는데 중요한 정보를 제공한다. 색상정보를 이용한 얼굴영역의 추출은 얼굴의 방향이나 형태의 변화에 덜 민감하고 그 추출속도가 빠르다는 장점 때문에 많이 사용된다. 그러나 색상정보는 조명의 변화에 따라 민감하게 바뀐다는 단점을 가진다. 또한 실내환경에서 피부색과 유사한 배경이나 배경물체들이 많이 존재한다. 이러한 조명의 변화나 배경들이 존재하는 경우에 피부색을 이용한 얼굴추출은 실패하기 쉽다. 본 논문에서는 이러한 단점을 극복하기 위하여 피부색상 모델의 추적을 행하였으며, 얼굴의 움직임데이터로부터 타원근사를 이용하는 방식을 제안하였다. 또한 카메라는 팬틸트 장치에 탑재되어서 사람의 얼굴을 추적하도록 하였다.

## 1. 서 론

얼굴 추적 기술은 인간-컴퓨터상호작용에 필수적인 요소로서 다양한 응용분야를 가지며 현재까지도 계속 연구가 진행되고 있다. 또한 얼굴추적은 원격회의나 가상공간상의 아바타제어, 영상전화기, 건물출입이나 통제를 위한 개인식별 그리고 주요장소의 원격감시등에 사용되기도 한다.

색상은 물체의 특성을 나타내는 고유한 특징점이 될 수 있으며, 물체를 인식하는데 중요한 정보를 제공한다. 색상 정보를 이용하여 얼굴을 추적하는 경우 형태의 변형이나 얼굴의 방향에도 크게 영향을 받지 않는 장점이 있다[1][2]. 그러나 실내환경에서 피부색을 기반으로 얼굴을 추적하는 경우에, 조명의 변화나 얼굴과 유사한 색상의 배경물체나 배경색상에 의하여 추적하고자 하는 얼굴의 추적이 실패하는 경우가 생기게 된다.

움직임정보로부터 얻어진 물체의 외곽선 정보는 물체를 인식하는데 도움을 주기도 한다[3][4]. 그러나 카메라가 팬틸트와 같은 기구부에 의해 제어되는 시스템에서, 카메라의 움직임에 의해 정지한 배경까지도 움직이는 것으로 관찰되게 된다. 따라서 두 대의 카메라를 이용한 스테레오 매칭 등의 방법에 의해 영상의 각 점의 거리를 계산하고, 카메라를 탑재한 기구부의 동역학으로부터 카메라 움직임에 의한 배경 움직임 제거과정이 필요하기도 한다[7][14].

본 논문에서는 색상정보와 움직임정보를 이용한 얼굴추적 알고리즘과 그 실험결과를 보여준다. 클러터(clutter)의 영향에도 추적이 강인하다고 알려진 컨덴세이션 알고리즘을 피부색 색상변화의 추적에 이용하여, 조명의 효과로부터 피부색을 강인하게 찾을 수 있게 하였다. 또한 움직임의 탐색방향을 추정된 얼굴의 중심으로부터 바깥쪽으로 진행함으로써 얼굴의 움직임만을 얼굴 외곽선 데이터로 인식하게 된다. 따라서 배경의 움직임을 제거할 필요가 없게 된다. 이렇게 움직임으로부터 획득된 얼굴의 외곽선 데이터들은 타원근사를 통하여 배경물체를 얼굴과 분리시킬 수 있게 해준다. 제안된 색상

정보와 움직임에 의한 타원정보를 이용하는 얼굴추적기는 피부색과 유사한 색의 배경물체가 가까이 붙어 있는 것으로 보이는 경우에도 얼굴과 배경물체를 구분할 수 있게 한다.

얼굴 추적기는 카메라영상에서 얼굴의 위치를 화면의 중앙으로 오도록 카메라를 움직일 수 있게 해 주는 팬틸트 장치를 제어하여 화면의 중앙에 사람의 얼굴을 나타내도록 하였다.

## 2. 본 론

### 2.1 컨덴세이션 알고리즘의 이용

기본적인 추적기로는 컨덴세이션(Condensation) 알고리즘을 사용하였는데 이는 통계적인 방법에 기초하여 확률이 높은 방향으로 추적하려는 경향을 가지며, 입력 데이터에 클러터(clutter)가 존재하는 경우에도 추적이 잘 성공한다고 알려져 있다.

컨덴세이션 알고리즘은 선택(select), 예측(predict), 측정(measure)이라는 3단계의 과정을 반복하면서 진행된다. 이는 변수공간에 놓여있는 N개의 샘플들중에서 측정값이 높은 샘플들을 선택한 후, 동적모델 등을 통해서 그 다음 시간의 샘플 위치를 예측하고, 측정을 통해 각 샘플의 확률값을 부여하는 과정을 되풀이 하게 된다. 이러한 일련의 과정중에서 샘플들의 위치를 통해 물체의 위치를 추정(estimate)해내게 된다.

본 연구에서 컨덴세이션은 영상평면에서 얼굴추적과 색상공간에서 피부색추적에 사용된다.

#### 2.1.1 영상평면에서 얼굴추적

컨덴세이션 알고리즘에서는 추적하고자 하는 개체의 동적모델을 잘 알수록 추적의 성공률은 더 좋아진다고 알려져 있는데, 이를 위해서 영상평면에서의 얼굴추적에서는 2차 마르코프 동적모델을 학습을 통해 정하여 사용하였다.

시간 t에서 물체의 위치를  $\vec{x}_t = [x(t), y(t)]^T$ 로 나타낼 때, 2차 마르코프 동적모델은 식(1)처럼 표현된다.

$$\vec{x}_t = A_1 \vec{x}_{t-1} + A_2 \vec{x}_{t-2} + D + B w_t \quad (1)$$

연속적인 M개의 데이터  $\{\vec{x}_1, \dots, \vec{x}_M\}$ 로부터 MLE(Maximum Likelihood Estimation)를 이용하여 유도되는  $A_1, A_2, B, D$ 는 식(2)~(5)와 같다.

$$\hat{A}_2 = \begin{pmatrix} R_{02} - R_{01} (R_{11})^{-1} R_{12} \\ R_{22} - R_{21} (R_{11})^{-1} R_{12} \end{pmatrix} \times \quad (2)$$

$$\hat{A}_1 = (R_{01} - \hat{A}_2 R_{21}) R_{11}^{-1} \quad (3)$$

$$\hat{D} = \frac{1}{M-2} (R_0 - \hat{A}_2 R_2 - \hat{A}_1 R_1) \quad (4)$$

$$\hat{B} = \sqrt{\frac{1}{M-2} (R_{00} - \hat{A}_2 R_{20} - \hat{A}_1 R_{10} - \hat{D} R_0^T)} \quad (5)$$

$$\text{단, } R_i = \sum_{k=3}^M \vec{x}_{k-i}, \quad R_{ij} = \sum_{k=3}^M \vec{x}_{k-i} \vec{x}_{k-j}^T$$

$$R'_i = R_i - \frac{1}{M-2} R_i R_i^T$$

### 2.1.2 색상공간에서 피부색 추적

실내에서 피부색의 시간에 대한 변화는 주로 조명의 변화에 의존적이게 된다. 그런데 조명의 변화에 대한 피부색의 변화는 일정한 규칙성이 없으므로, 그 동적모형을 학습한다는 것이 의미가 없게 된다. 따라서 색상의 동적모형은 아래와 같이 두었다.

$$\vec{x}_i = \vec{x}_{i-1} + \vec{w}_i \quad (6)$$

위 식처럼 동적모형의 신빙성이 크지 않은 경우 컨덴세이션 알고리즘으로부터 좋은 추적성능을 기대하기 힘들 수 밖에 없다. 따라서 색상공간에서 피부색의 추적은 잘못된 피부색으로의 변이를 막기 위하여, 현재 적용된 색상모델에 대한 측정 가중치뿐만 아니라 일반적인 사람의 피부색 모델에 대한 측정 가중치를 샘플의 가중치로 부여하였다. 즉, 각 샘플의 측정 가중치는 다음의 식으로 사용하였다.

$$\pi_i^{(n)} = \pi_1^{(n)} \times \pi_2^{(n)} \quad (7)$$

단,  $\pi_1^{(n)} = c_1 \exp(-0.5D_M^{(n)})$ ,  $\pi_2^{(n)} = c_2 \exp(-0.5D_C^{(n)})$ 이며,  $D_M$ 은 n번째 샘플과 현재 적용된 색상모델과의 마할라노비스거리(Mahalobis distance),  $D_C$ 는 n번째 샘플과 여러 조명환경에서 구한 일반적인 피부색상모델과의 마할라노비스 거리를 나타낸다.

### 2.2 타원근사 알고리즘의 이용

배경의 물체와 얼굴을 구별하기 위해서, 얼굴은 타원형의 모양을 하고 있다고 가정하였다. 이 얼굴모양의 정보를 얻기 위하여, 연속이미지에서 두 프레임의 차이를 이용하여 움직임 정보를 획득하고, 이로부터 얼굴의외곽선을 추출해내게 된다.

얼굴의 외곽선을 추출하는데 있어서 외곽선의 탐색 방향을 바깥쪽에서 안쪽으로 두면 얼굴의 외곽선 뿐만 아니라 다른 움직임들이 탐지되게 된다. 따라서 본 연구에서는 얼굴의 중심에서부터 시작하여 바깥쪽으로 탐색하게 하였다.

타원의 방정식은 일반적인 2차 방정식의 형태로 식(8)과 같이 나타낼 수 있다.

$$ax^2 + bxy + cy^2 + dx + ey + f = 0 \quad \text{where } b^2 - 4ac < 0 \quad (8)$$

식(8)은 벡터의 형태인 식(9)로 나타낼 수 있다.

$$F(\vec{a}, \vec{x}) = \vec{a}^T \vec{x} = 0 \quad \text{where } \vec{a}^T C \vec{a} = 1 \quad (9)$$

$$\text{단, } \vec{a} = [a \ b \ c \ d \ e \ f]^T, \quad \vec{x} = [x^2 \ xy \ y^2 \ x \ y \ 1]^T$$

$$C = \begin{bmatrix} 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

타원으로 근사화 할 N개의 데이터  $\{\vec{x}_1, \dots, \vec{x}_N\}$ 가 주어져 있을 때, 타원의 방정식을 특징짓는  $\vec{a}$ 는 아래의 식(11)을 만족시키면서 식(10)을 최소로 하는 값을 선택하면 주어진 데이터를 가장 잘 근사화하는 타원을 구할 수 있게 된다.

$$\sum_{i=1}^N F(\vec{a}, \vec{x}_i)^2 = \|\vec{D}\vec{a}\|^2 \quad \text{where } \vec{a}^T C \vec{a} = 1 \quad (10)$$

$$\text{단, } D = [\vec{x}_1^T, \dots, \vec{x}_N^T]$$

제한조건을 가지는 식(10)을 만족시키는 해  $\vec{a}$ 를 구하기 위해 Lagrange multiplier를 도입하면, 식(10)

은 식(11)로 표현된다.

$$2D^T D \vec{a} - 2\lambda C \vec{a} = 0 \quad (11)$$

식(11)의 정리된 식(13)으로 표현된 eigen system을 만족하는  $\vec{a}$ 가 타원을 결정짓는 파라메타가 되는데, 식(12)에서 양수의 eigen value에 해당하는 eigen vector가  $\vec{a}$ 가 되며 그 해는 유일하다[12].

$$S \vec{a} = \lambda C \vec{a} \quad \text{where } S = D^T D \quad (12)$$

외곽선 데이터의 잡음에 의해  $\vec{a}$ 가 잘못된 값으로 근사화 되는 것을 막기 위해 식(13)와 같이 타원의 파라메타를 추정하여서 사용하였다.

$$\vec{a}_k = \sum_{i=0}^{k-1} \alpha_{k-i} \vec{a}_{k-i} \quad (13)$$

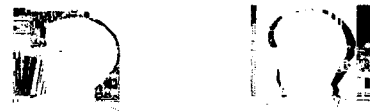


그림 1. 획득된 움직임 영역

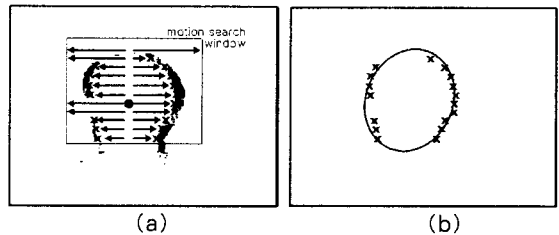


그림 2. 얼굴외곽선데이터 탐색과 타원근사. (a) 얼굴외곽선데이터 탐색, (b) 타원근사

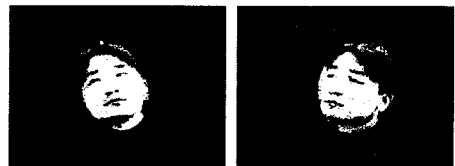


그림 3. 타원근사에 의한 얼굴 추출

그림 1은 추적 도중에 추출한 움직임 영역을 보여주며, 그림 2는 외곽선데이터 추출 방법과 추출된 외곽선 데이터로부터 얼굴을 표현하는 타원을 나타내고 있다. 얼굴외곽선 데이터를 움직임 영역으로부터 추출할 때에는 추정된 얼굴의 중심위치에서부터 바깥쪽으로 탐색하게 되며, 이는 외곽선 탐색영역 내에서만 탐색을 하게 된다. 얼굴과 관련 없는 어게션, 목 등의 외곽선 데이터를 제외하기 위하여 탐색영역은 얼굴 피부색영역을 기준으로 정해진다. 즉 탐색영역은 추정된 얼굴영역에서 여유폭을 두고 더 크게 정해지지만, 얼굴 하단부 부분만은 여유폭을 없애으로써 얼굴과 무관한 움직임을 제외하였다. 또한 이 외곽선 탐색영역은 얼굴의 크기에 비례하게 크기를 정해짐으로써 얼굴의 크기변화에도 동작하도록 하였다.

이렇게 추출된 얼굴외곽선으로부터 타원근사화를 통하여 얼굴영역을 뽑아내게 되며, 이렇게 하여 얻어진 얼굴영역과, 색상을 이용하여 추출한 얼굴영역을 비교하여 배경의 물체와 실제 얼굴영역을 추출하였다. 그림 3은 추출한 얼굴영역을 보여준다.

### 3. 실험결과

표 1은 얼굴이 상하로 출추듯이 움직이는 연속 이미지(그림 4)에 대하여 학습을 한 후 비슷한 움직임에 대한 학습 전후의 RMS(root mean square) 오차를 보여주고 있으며, 학습의 유용성을 보여준다.

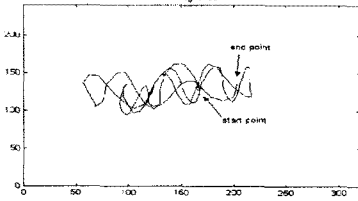


그림 4. 얼굴의 움직임 궤적

	학습전	학습후
좌우방향 RMS 오차	5.8724	2.7125
상하방향 RMS 오차	11.0760	2.8987

표 1. 학습전후의 RMS오차

그림 5는 어두운 조명상태에서 밝은 조명상태로 바뀔 때 피부색을 찾아내는 결과를 보여준다.

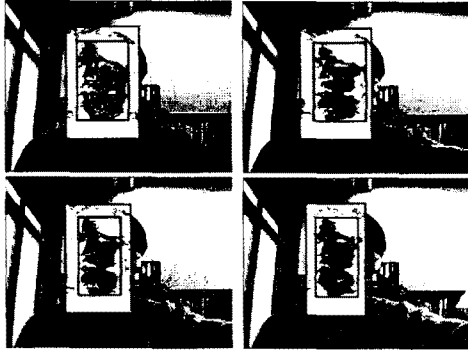


그림 5. 컨덴세이션을 이용한 색상 추적

그림 6은 배경에 추적하던 사람이 아닌 다른 사람이 추적하던 얼굴과 배경에서 겹쳐서 나타나는 경우를 보여준다. 색상정보만으로 얼굴을 찾는 경우에는 두 개의 얼굴을 한사람의 얼굴로 인식하지만, 본 연구에서는 움직임 영역 중 얼굴 외곽선부분의 타원근사화를 통하여 계속하여 기존에 추적해오던 사람만을 추적함을 보여주고 있다.



(a) (b)

그림 6. 두 사람의 얼굴이 겹치는 경우. (a) 얼굴 피부색영역, (b) 타원근사화를 이용한 얼굴탐지

마지막으로 타원근사에 의한 배경분할 전과 후의 수행시간을 비교하면 다음과 같았다.

배경분할 전	배경분할 후
약 40 ~ 80ms/frame	약 20 ~ 60ms/frame

표 2. 수행시간 비교

표 2는 배경에 피부색과 유사한 배경이 있는 경우의 수행시간을 나타내고 있다. 수행시간이 배경분할에 의해서 오히려 전체적인 추적 속도가 빨라짐을 알 수 있는데 이것은 배경으로부터 관심 있는 얼굴영역만을 추출하여 얼굴영역에 대해서만 이미지 처리를 하게 되기 때문에 시스템속도의 향상을 가져왔다.

### 4. 결 론

색상정보에 의해서 얼굴의 방향, 크기변화, 조명 변화, 부분적인 가려짐에도 얼굴을 추적할 수 있는 성능을 가지게 되었다. 색상정보만을 이용할 때에 항상 문제점으로 대두되는 배경과 얼굴의 색상이 비슷한 경우에 움직임 정보를 타원 근사화 방법에 이용함으로써, 배경에 피부색과 유사한 색상의 물체가 있더라도 기존에 추적하던 얼굴을 계속하여 추적할 수 있음을 볼 수 있었다. 움직임 영역을 이용하는데 있어서, 본 연구에서 사용한 방법은 카메라의 움직임에 대한 배경 움직임을 보상할 필요가 없으므로 계산량이 많지 않음을 알 수 있었다.

### (참 고 문 헌)

- (1) J. Yang, A. Waibel, "A Real-Time Face Tracker", Proceedings of WACV, pp.142-147, 1996
- (2) Richard J. Qian, M. I. Sezan, Kristine E. Matthews, "A Robust Real-Time Tracking Algorithm", Proceedings International Conference on Image Processing, vol. 1, pp.131-135, 1998
- (3) C. Wang and M. S. Brandstein, "A Hybrid Real-Time Face Tracking System", Proceedings of Acoustics, Speech, and Signal Processing, Volume 6, 1998
- (4) Sun, Q.B. Huang, W.M. Wu, J.K., "Face detection based on color and local symmetry information", Proceedings of Third IEEE International Conference on Automatic Face and Gesture Recognition, pp.130 135, 1998
- (5) S. J. McKenna, S. Gong, "Tracking Faces", Proceedings of FG96, pp.271-276, 1996.
- (6) T. Darrell, G. Gordon, M. Harbille, J. Woodfill, "Integrated Person Tracking Using Stereo, Color, and Pattern Detection", Proceedings of CVPR98, pp.601-609, 1998.
- (7) S. Feyrer and A. Zell, "Tracking and pursuing persons with a mobile robot", Proceedings of Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems, 1999.
- (8) 장기정, "적응 칼라 모델을 이용한 실시간 얼굴 추적", 한국과학기술원 전기 및 전자공학과, 석사학위논문, 2000.
- (9) M. Isard and A. Blake, "CONDENSATION - conditional density propagation for visual tracking", International Journal of Computer Vision, 1998.
- (10) A. Blake and M. Isard, Active Contours, Springer, 1998.
- (11) E. B. Meier and F. Ade, "Using the Condensation Algorithm to Implement Tracking for Mobile Robots", 1999 Third European Workshop on Advanced Mobile Robots, 1999.
- (12) A. Fitzgibbon, M. Pilu and R. B. Fische, "Direct Least Square Fitting of Ellipses", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 21, NO 5, May, 1999
- (13) A. R. Gourlay, G.A. Watson, Computational methods for matrix eigenproblems, John wiley & sons, 1973
- (14) Don Murray and Anup Basu, "Motion tracking with an active camera", IEEE Transactions on Pattern Analysis and Machine Intelligence, 1994.