

컬러 SSD 알고리즘 기반 칼만 예측기를 이용한 다수의 얼굴 검출 및 추적 시스템

김병기, 한영준, 한헌수
승실대학교 전자공학과

e-mail : qsilver@ssu.ac.kr, young@ssu.ac.kr, hahn@ssu.ac.kr

Multiple Face Tracking System Using the Kalman Estimator Based on the Color SSD Algorithm

Byoung-Ki Kim, Young-Joon Han, Hern-Soo Hahn
School of Electronic Engineering
Soongsil University

Abstract

This paper proposes a new tracking algorithm using the Kalman estimator based color SSD algorithm. The Kalman estimator includes the color information as well as the position and size of the face region in its state vector, to take care of the variation of skin color while faces are moving. Based on the estimated face position, the color SSD algorithm finds the face matching with the one in the previous frame even when the color and size of the face region vary. The features of a face region extracted by the color SSD algorithm are used to update the state of the Kalman estimator.

I. 서론

감시시스템을 위해 많은 얼굴검출과 추적 기법들이 연구되어져 왔다. 특히, 얼굴추적에 관한 연구들은 크게 모델기반 추적과 움직임기반 추적 알고리즘으로 연구되어 왔다. 모델기반(model-based) 추적은 추적하고자 하는 정확한 객체의 특징을 추출하여 원형 객체모델과 정합하는 기법이다. 따라서 정확한 데이터베이스를

구축하고 보이지 않는 경우에도 복잡한 특징을 추출해야 하기 때문에 충분한 수행시간이 보장되어야 한다 [1][2]. 움직임기반(motion-based) 추적은 순차적으로 입력되는 일련의 영상에서 차영상을 이용하여 움직임 화소들을 검출하고 그들을 병합함으로써 객체 영역을 검출하는 추적기법이다. 움직임기반 추적기법은 적용하는 응용방법에 따라 얼굴이나 몸 전체를 검출하거나 칼만 필터나 mean-shift 알고리즘에 의해 추적방법이 달라진다. 그러므로 잡음에 영향을 덜 받고 움직임의 정보를 정확하게 예측하는 기법이 요구된다[3][4].

따라서, 본 논문은 컬러 SSD 알고리즘 기반 칼만 예측기를 이용하는 새로운 얼굴 추적 알고리즘을 제안한다. 제안된 방법은 얼굴 영역의 위치와 크기뿐만 아니라 색상분포까지 예측하며, 컬러 SSD 알고리즘을 이용해서 예측된 위치에서 얼굴영역을 검출한다. 컬러 SSD 알고리즘은 적응적으로 모델의 서식(template)을 갱신하기 때문에 객체의 모양이 점차적으로 변하는 경우에도 객체를 적응적으로 추적할 수 있다. 컬러 SSD 알고리즘에 의해 정합된 결과는 칼만 예측기의 상태벡터(state vector)를 갱신하는데 사용된다. 색상분포 정보를 포함하는 예측된 상태벡터는 움직이는 다수의 얼굴들 사이에서 발생하는 겹치는 문제를 해결할 수 있다. 그림 1은 제안된 얼굴 검출 및 추적 시스템의 흐름도를 보여준다.

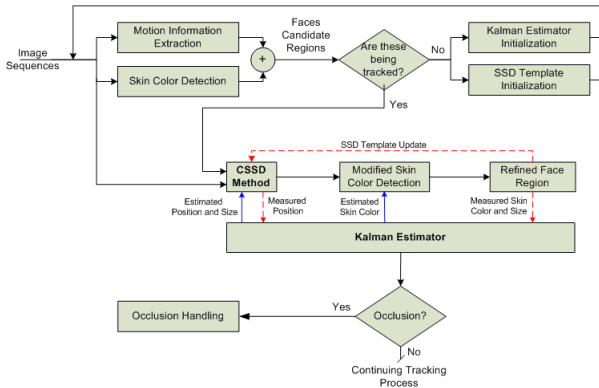


그림 1. 제안된 얼굴 추적시스템의 흐름도
II. 크기, 위치, 색상정보를 위한 칼만 예측기

본 논문에서 칼만 예측기[5]는 얼굴 검출 과정의 정확도와 속도를 향상시키기 위해서 사용된다. 칼만 예측기는 다음과 같은 수식으로 나타낼 수 있다.

$$S(t) = \Phi(\Delta t)S(t-\Delta t) + W(t)$$

$$m(t) = H(t)S(t) + R(t)$$

다른 논문들과 달리 제안된 방법에서는 칼만 예측기의 상태벡터에 색상정보가 포함되어 있다. 그러므로 위 식에 주어진 상태벡터는 얼굴 영역의 중심위치(C_x, C_y), 너비(W), 높이(H), 얼굴 영역 내의 I 와 Q 의 평균값과 각각의 일차 미분값들로 구성된다. 다시 말해서 $S(t) = [C_x, C_y, W, H, I^m, Q^m, V_x, V_y, W', H', I^{m'}, Q^{m'}]^T$ 이다.

$$\Phi(\Delta t) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & \Delta t & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & \Delta t & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & \Delta t & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & \Delta t & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & \Delta t & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & \Delta t & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} C_x \\ C_y \\ W \\ H \\ I^m \\ Q^m \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} S(t) + \begin{pmatrix} R_x(t) \\ R_y(t) \\ R_W(t) \\ R_H(t) \\ R_{I^m}(t) \\ R_{Q^m}(t) \end{pmatrix}$$

Δt 는 샘플링 간격을 뜻하고, Φ 는 상태전이행렬을

나타낸다. $m(t) = [C_x, C_y, W, H, I^m, Q^m]^T$ 은 측정벡터 (measurement vector)이고, $H(t)$ 는 측정행렬 (measurement matrix)이다. 또한 $W(t)$ 와 $R(t)$ 는 각각 정규분포를 갖는 입력과 측정의 오차를 반영한다.

다음 프레임에서의 새로운 상태벡터와 예측 오차 공분산 행렬은 아래의 수식에 의해서 예측된다. '-'는 예측된(estimated) 값을, '+'는 갱신된(predicted) 값을 의미한다.

$$S^-(t) = \Phi(\Delta t)S^+(t-\Delta t) + W(t)$$

$$P^-(t) = \Phi(\Delta t)P^+(t-\Delta t)\Phi^T(\Delta t) + R(t)$$

예측된 상태벡터는 다음 프레임의 얼굴 영역의 위치, 크기, 색상정보를 제공한다. 이 정보들에 기반하여 컬러 SSD 알고리즘을 가지고 다음 프레임에서 얼굴을 검출하게 되고 측정벡터, $m(t)$ 가 갱신된다. 측정벡터가 갱신되면 칼만 이득(gain)에 의해서 상태벡터가 보정된다.

$$K(t) = P^-(t)H^T(t)[H(t)P^-(t)H^T(t) + R(t)]^{-1}$$

$$S^+(t) = S^-(t) + K(t)[m(t) - H(t)S^-(t)]$$

$$P^+(t) = [I - K(t)H(t)]P^-(t)$$

III. 컬러 SSD 알고리즘에 기반한 측정벡터 갱신

전체 시스템의 전처리 단계로 동작하는 얼굴 검출 모드에서 얼굴 후보 영역들은 피부색상 정보와 움직임 정보를 이용해서 검출된다. 차영상을 이용해서 움직임 픽셀들이 검출되고 얼굴 특징을 갖는 것들끼리 얼굴 영역으로 병합된다. 검출된 얼굴 후보 영역의 위치, 크기와 색상정보가 수용 가능한 오차의 범위를 벗어난다면 현재 프레임에서 새로운 얼굴이 처음 검출된 것으로 볼 수 있으므로 추적을 시작하기 위해서 등록한다. 이와 달리 이전 프레임에서 존재했던 얼굴 영역이 현재 프레임에서도 검출된 것이라면 추적 모드가 활성화된다. 추적모드에서는 이전 프레임에서 검출된 얼굴을 서식(template)으로 하여 칼만 예측기에 의해서 예측된 얼굴 영역의 위치를 중심으로 컬러 SSD 알고리즘을 적용하여 얼굴 영역의 위치를 찾아낸다.

다음의 수식을 최소화하는 위치, $i^*P(k) = i^*P(k-1) + \Delta t$, 가 선택된다.

$$CSSD(i^*P(k-1), \Delta t) = \sum_{m, n \in N} [(I_{k-1}(i^*x(k-1) + m, i^*y(k-1) + n)) - I_k(i^*x(k-1) + m + u, i^*y(k-1) + n + v))$$

$$+(Q_{k-1}(i_x(k-1)+m, i_y(k-1)+n) - Q_k(i_x(k-1)+m+u, i_y(k-1)+n+v))]$$

$iF(k-1)=[i_x(k-1), i_y(k-1)]$ 은 (k-1)번째 프레임에서 템플릿의 위치이고, I_{k-1} 과 Q_{k-1} 은 (k-1)번째 프레임에서 얼굴 템플릿 안에 포함된 픽셀들의 YIQ 색상 공간에서 I와 Q의 평균값을 나타낸다. $\Delta I=[u, v]^T$ 는 예측된 위치로부터의 이동변위이다.

그림 2와 같이 검색 영역 내에서 컬러 SSD 알고리즘에 의해서 가장 정합이 잘 되는 영역의 위치가 얻어지면 칼만 예측기에 의해서 예측된 색상 정보를 이용해서 아래의 수식에 주어지는 수정된 피부색상 범위에 해당되는 피부색상 영역을 검출하여 두 영역의 크기 관계를 이용해서 얼굴 영역의 크기를 보정한다.

$$\begin{aligned} \widehat{I}^m - \Delta I < I < \widehat{I}^m + \Delta I \\ \widehat{Q}^m - \Delta Q < Q < \widehat{Q}^m + \Delta Q \end{aligned}$$

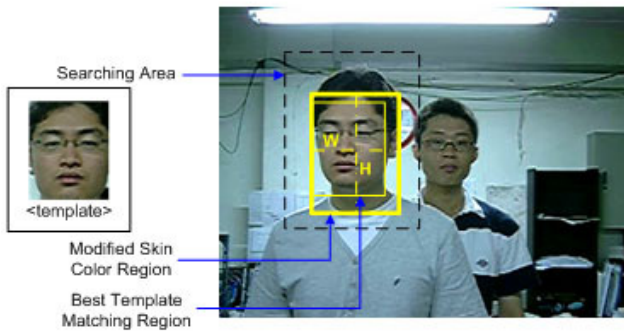


그림 2. 얼굴 영역 보정

$$W = 2 \times \min(w1, w2)$$

$$H = 2 \times \min(h1, h2)$$

$$I^m = \frac{1}{W \times H} \sum_{x, y \in \text{face region}} I(x, y)$$

$$Q^m = \frac{1}{W \times H} \sum_{x, y \in \text{face region}} Q(x, y)$$

$w1$ 과 $h1$ 은 컬러 SSD 알고리즘에 의해서 얻어진 영역의 너비와 높이이고, $w2$ 와 $h2$ 는 칼만 예측기에 의해서 예측된 색상정보를 이용해서 검출된 피부색상 영역의 너비와 높이를 나타낸다.

IV. 실험

제안된 알고리즘의 성능을 평가하기 위해서 320x240

의 CCD 카메라와 3GHz CPU를 장착한 펜티엄 4 PC를 가지고 실험을 수행하였다. 얼굴 검출을 하기 위해 차영상과 움직임 픽셀들을 구하는 과정에서 노이즈의 영향을 줄이기 위해서 표준편차가 2인 가우시안 필터를 사용해서 영상을 필터링하였고, ΔI 와 ΔQ 는 각각 12와 5로 결정하였다.

첫 번째 실험은 그림 3과 같이 한 명의 사람이 존재할 때 제안된 알고리즘이 잘 동작하는지를 검증하였다. 실험에서 카메라와 얼굴과의 거리는 2.6m에서 8.4m까지의 범위였는데 비교적 먼 거리임에도 불구하고 잘 동작하는 것을 알 수 있었다.



(a) t=0.5sec

(b) t=10sec



(c) t=19sec

(d) t=28sec

그림 3. 한 사람 추적 실험

	위치		크기	
	x	y	너비	높이
(a)	47	102	16	19
(b)	212	83	16	18
(c)	162	112	28	34
(d)	72	93	17	21

표 1. 시간에 따른 얼굴의 위치와 크기 변화

그림 4은 실제 얼굴의 위치, 크기와 칼만 예측기에 의해서 추적된 결과와의 차이를 보여준다. 에러는 0.06초부터 28초까지 측정되었다. 위치 에러는 최대 7픽셀 이하였고, 평균 2픽셀 정도였으며, 평균 크기 에러는 5픽셀 정도로 나타났다. 갑자기 방향을 바꾸는 5.9초, 15.4초, 23.1초, 27.9초와 속도 변화가 나타난 1초와 5.5초 부근에서 에러가 갑자기 커지는 것을 그래프를 통해서 알 수 있었다.

V. 결론

본 논문은 얼굴 영역의 크기와 색상의 변화, 다수의 얼굴들의 겹침 등의 다수의 얼굴 검출과 관련된 문제들의 해결책으로써 컬러 SSD 알고리즘 기반 칼만 예측기를 이용한 새로운 얼굴 검출 및 추적 알고리즘을 제안했다. 칼만 예측기는 얼굴 영역의 위치와 크기뿐만 아니라 색상 정보에 대해서도 훌륭한 예측치를 제공하며, 이전 프레임에서 검출된 얼굴 영역을 이용해 적응 템플릿을 갱신하고 예측된 얼굴 영역의 위치를 기준으로 컬러 SSD 알고리즘을 적용함으로써 현재 프레임에서 얼굴 영역의 위치를 정확히 찾아낼 수 있었다. 컬러 SSD 알고리즘이 이전 프레임에서 칼만 예측기에 의해서 예측된 얼굴 영역에 관한 정보들을 사용하기 때문에 얼굴 영역들 간의 겹침이 발생하는 것을 알 수 있으며 얼굴들의 겹침이 발생했을 때에도 겹친 정도가 심하지 않거나 가려지지 않은 가장 위의 얼굴은 검출이 가능하다.

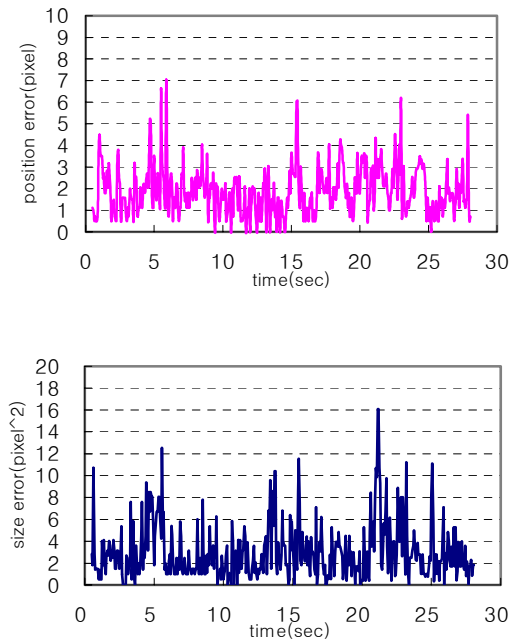


그림 4. 위치와 크기 오차

두 번째 실험에서 우리는 다수의 얼굴이 영상 내에 존재하여 겹침 혹은 가림이 발생하는 경우에도 제안된 알고리즘이 효과적으로 동작하는 것을 검증하였다. 가림 현상의 검출은 컬러 SSD 알고리즘의 CSSD 측정값의 갑작스런 변화를 찾아냄으로써 가능했다. 그림 5에서 보는 바와 같이 얼굴이 조금 가려진 정도까지 얼굴 검출 및 추적이 잘 동작하고, 얼굴이 완전히 겹친 경우엔 가려지지 않은 얼굴만 검출된다. 또한 가려 있던 얼굴이 다시 나타났을 때에도 문제없이 추적이 되고 있는 것을 볼 수 있다.

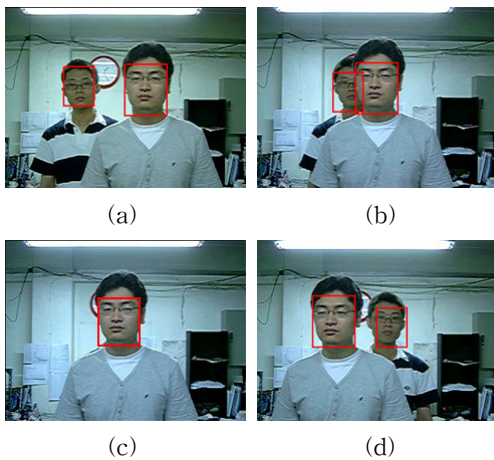


그림 5. 겹침 현상이 발생하는 경우 실험

참고문헌

- [1] Strom, J., Jebara, T., Basu, S., Pentland A.: Real time tracking and modeling of faces: an EKF-based analysis by synthesis approach. Modeling People, Proceedings. IEEE International Workshop on 20 Sept. (1999) 55-61
- [2] Dornaika, F., Ahlberg, J.: Fast and reliable active appearance model search for 3-D face tracking. Systems, Man and Cybernetics, Part B, IEEE Transactions, 34,(2004) 1838 - 1853
- [3] Jun, Wang., Achanta, R., Kankanhalli, M., Mulhem, P.: A hierarchical framework for face tracking using state vector fusion for compressed video. Acoustics, Speech, and Signal Processing, Proceedings, (ICASSP '03), IEEE International Conference on, 3,(2003) III - 209-12
- [4] Ce, Wang., Brandstein, M.S.: A hybrid real-time face tracking system. Acoustics, Speech, and Signal Processing, 1998, (ICASSP '98), Proceedings of the 1998 IEEE International Conference, 6,(1998) 3737 - 3740
- [5] Kalman R. E.: A new approach to linear filtering and prediction problems. Transactions of the ASME, Journal of Basic Engineering, (1960) 35-45