연속적인 얼굴 표정 변화 인식 방법에 관한 연구

박호식*, 배철수*, 나상동**
*관동대학교 전자통신공학과, **조선대학교 컴퓨터공학부
e-mail:mediana@netsgo.com

A Study of Recognizing Degree of Continuous Facial Expression Change

Ho-Sik Park*, Cheol-Soo Bae*, Sang-Dong Na**
*Dept of Electronic Communication Eng., Kwandong University
**Dept of Computer Eng., Chosun University

요약

본 논문에서는 영상에서 연속적인 얼굴 표정 변화 인식 방법에 대하여 제안하였다. 제안된 방법은 가중 결합으로 정합 된 분류 그래프를 이용한 얼굴 특징점 추적과 특징 궤도와 표정 변화 모델의 비교에 의한 얼굴표정 인식의 두 가지 부분으로 구성된다. 표정 변화 모델은 얼굴 특징 동작과 표정의 변화사이의 연관 관계를 표현한 B-spline 곡선을 이용하여 구성된다. 유형뿐만 아니라 표정의 변화 단계까지 인식 할 수 있다. 더욱이 획득된 표정 정보는 다음 프레임의 추적의 지침으로 피드백 됨에 따라 추적의 검색 시간을 감소시킴으로서 교점 확산 구간을 최소화 할 수 있다.

1. 서 론

얼굴 표정의 인식 기술은 모델 기반의 비디오 코딩, 인간과 기계의 인터페이스, 영상 통신 등 활용분야가 넓은 중요한 학문이다.

이전의 얼굴 표정 인식 에 관한 연구들은 정적인 단일 영상을 기반으로 한 것이었다, 그러나, 얼굴 표정은 고정되어진 것이 아니라 시시각각으로 변화되어지므로 연속적인 얼굴 표정의 변화를 인식하는 것이 바람직하다. 최근 들어 동적인 영상에서 얼굴 표정을 인식하는 분야가 연구되어지고 있으며 대다수의 연구는 얼굴 근육의 움직임을 추적한 것이다.[2,4,6]. 그러나, 이러한 접근 방식으로는 정밀한 얼굴의 변화에서 세밀하지 않은 얼굴 특징의 이동을 분리하여 찾는 것은 쉽지 않다.

본 논문에서는 얼굴 특징점 이동을 기반으로 한 얼굴 표정 변화 인식 방법을 제안하고자 한다. 표정 변화에 영향을 받지 않는 일련의 특징점을 사용하여, 얼굴에서 변화하는 얼굴 특징의 이동 동작도 쉽게 분리 해 낼 수 있다. 첫 번째로, 순차적으로 입력된 영상에서 가중 결합과 정합 된 분류 그래프(labeled graph)를 이용하여 얼굴 특징을 찾아내면 표정 변화 모델과 특징 궤도가 정합 함으로서 순차 영상의 모든 프레임의 표정 변화의 유형과 단계를 정할 수 있다. B-spline 곡선은 얼굴 특징점의 이동과 표정변화의 상관관계를 명백하게 묘사한 표정 변화 모델을 구성하는데 사용된다. 현재 프레임에서 얻어진 표정 정보는 다음 프레임의 특징점 위치를 예상하기 위하여 피드백 된다.

2. 얼굴 특징점의 추적

2.1 얼굴 특징점의 정의

그림 1과 같이, 눈썹, 눈, 코 그리고 입 근처에 있는 19개 지점을 얼굴 특징점(FFPs : Facial Feature Points)으로 선택하였다. 그중 12개의 특징

점(그림 1의 |)은 실제로 얼굴 표정을 인식하기 위하여 사용되어진다. 그것을 얼굴 표정 특징점(FEFPs : facial expression feature points)으로 구분하였다. 남은 7개의 특징점(그림1의 ○)은 표정인식에 사용되지 않고 표정 변화에 민감하지 않은 지점이다. 그것을 얼굴 참조 지점(FRPs : facial reference points)으로 구분하였다. 얼굴참조지점(FRP)은 얼굴 특징점을 추적함에 있어 적절한 지점을 선정하고 향후 얼굴의 정밀한 동작을 계산하는데 사용되어진다.

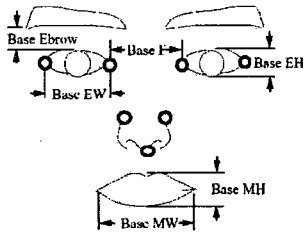


그림1. 19개의 얼굴 특징점과 정규화 된 기점
Figure 1. 19 facial features points and the normalization bases

2.2 얼굴 특징점 추적

얼굴 표정 특징의 명암과 모양을 추적함에 있어 일반적인 형태 비교는 그리 효율적이지 못했다. 추적의 신빙성을 개선하기 위해서 J. Buhmann[1] 에 의하여 제안된 정합 된 분류그래프 방법을 사용한 얼굴 특징점 추적시스템을 사용하였다.

2.2.1 분류그래프 (Labeled graph)

그림 2에서와 같이 얼굴 특징점 추적시스템은 메모리 층인 M 과 입력 층인 I 로 2개의 층으로 구성되어있다. 각 층은 교점과 결합으로 구성된 분류그래프이다. 각 교점은 분류된 얼굴 특징점에 대응하게 되므로 교점 부근에서 17×17 회색 등급으로 이루어진 형태로 구성되어진다. 위상적 그래프의 기능을 개선하기 위하여 얼굴 특성과 관련하여 가중 변수를 결합 부위에 추가하였다.

순차 영상 $n-1$ 프레임에서 n 프레임으로의 얼굴 특징점 추적은 분류그래프의 정합 문제를 발생시킨다. $n-1$ 프레임은 알려진 교점의 메모리 층으로 n 프레임은 찾고자 하는 교점과 정합 된 입력 층으로 다룬다.

2.2.2 그래프 정합

손실 효용을 최소화하는 그래프 정합을 실현하여 다음과 같이 정의하였다:

$$C_{total} = \sum_{i} (C_s(i^M, i^I) + \lambda C_T(i^M, i^I)) \qquad (1)$$

여기서 $C_s(i^M, i^I)$ 는 얼굴 특징점 i^M 과 i^I 의 정합 된 쌍 사이의 유사 형태를 측정한 "유사 손실"이고, $C_T(i^M, i^I)$ 는 두 영상에서 얼굴 특징점의 공간적 관계를 보전한 "위상 손실"이다. 영상 그래프의 정밀성은 계수 λ 를 제어한다.

손실 효용의 최적화는 결합 진행을 기반으로 한 교점 손실의 동적 처리방법으로 이루어진다.

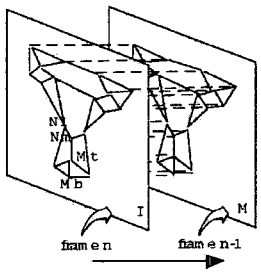


그림 2. 얼굴 특징점 추적 시스템
Figure 2. The FFPs tracking system

3. 얼굴 표정의 인식

표정 변화 모델과 특징 궤도에 대한 정합에 의해서 순차 영상의 모든 프레임의 표정 변화의 유형과 단계를 정의 할 수 있다.

3.1. 표정변화 모델의 구성

얼굴 표정 변화의 단계는 중립적인 얼굴의 12개의 얼굴 표정 특징점의 변위에 달려 있다. 얼굴 표정 특징점의 이동은 표정의 변화에 불규칙적인 비선형 이동이므로, 본 논문에서 제안된 모델은 B-spline 곡선에 의한 12개의 얼굴 표정 특징점의 변위와 얼굴 표정 변화의 단계 사이의 관계를 나타내고 있다.

B-spline 곡선은 세 가지(곡선 조정 지점 P_i, 곡선의 순서 k, 교점 벡터 t_i)매개변수에 의한 완벽한 자유형태의 곡선으로 아래와 같이 정의된다.

$$C(t) \equiv [x(t), y(t)] = \sum_{i=0}^{n} B_{i,k}(t) P_i \ , \qquad (2)$$

여기서 $B_{i,k}$ 는 B-spline 기본 작용이다.

본 논문에서는 B-spline 곡선을 표정 변화의 단계를 표시하는 변수 t 와 함께 얼굴 표정 특징점의 변

위를 묘사하는데 사용하였다. 제어 지점 P_i는 얼굴 표정 특징점의 변위에 의하여, 교점 t_i 각 제어 지점과 결합되어진 얼굴 표정의 변화 단계의 목록에 의하여 정의하였고, 곡선의 순서 k는 상수 3으로 고정하였다.

얼굴 표정의 각각의 유형을 위하여, 12개의 얼굴 표정 특징점의 B-spline 곡선을 사용하여 표정 모델을 구성하였다. 중요한 표정인 기쁨, 놀람, 노염의 세 가지 표정 변화 모델을 구성하였다.

이젠 그림 3과 같이 세 가지 표정 변화 모델을 구성한다. 각각은 12개의 얼굴 표정 특징점 의 B_spline 곡선 $C^*(t)$의 평균으로 구성된다. 각각의 $C^*(t)$ 는 얼굴 표정 특징점의 변위와 표정의 변위의 단계 사이의 연관성을 나타낸다.

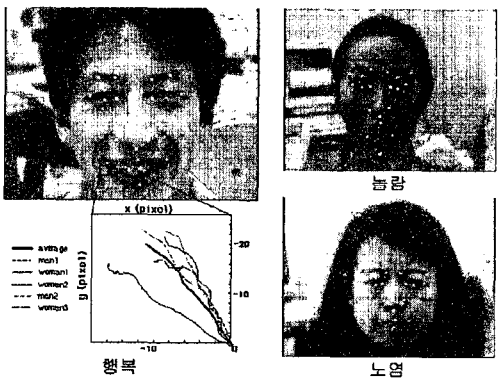


그림 3. 세 가지 표정 변화 모델 : 기쁨, 놀람, 노염

Figure 3. The expression change models of happiness, surprise, and anger

3.2. 표정의 인식

얼굴 표정 특징점이 이동할 때 3개의 연속적인 프레임이 탐지되고, 표정의 변화를 추정한다. 순차 입력 영상에서 얼굴 표정을 인식하기 시작한다. 세 프레임들의 첫 번째 프레임에 앞선 프레임을 표정 변화의 시작 프레임으로 간주한다.

3.2.1 얼굴 표정 유형의 인식

얼굴 표정 모델 i 에 대한 얼굴 표정 특징점의 B-spline 곡선 j 에 대한 평균을 C_{ij}^* 라 하면 표정 변화의 프레임을 $k=0$, 현재 프레임을 $k=m$으로 할 수 있다. 그리고, 12개의 얼굴 표정 특징점의 실제 궤도와 표정 모델 i 에 상응하는 추정지 사이의 실제거리를 다음과 정의 할 수 있다.

$$Y_m^i(d) = \sum_{k=0}^{m}\sum_{j=1}^{12}(C_{i,j}^*(k\cdot d) - \overline{P}_{jk})^2 \qquad (3)$$

여기서 $\overline{P}_{jk}$ 는 프레임 k 에서의 중립적 얼굴로부터 얼굴 표정 특징점 j 를 정규화 한 변위이다.

Brent 방법에 의하여 세 가지 표정 모델을 위한 d 를 고려하여 거리 효율을 최소화한다. 그리고 만약 $Y_{m,\min}^i = \min(Y_{m,\min}^1, Y_{m,\min}^i < \varepsilon_T, Y_{m,\min}^2, Y_{m,\min}^3)$ 과 $d_{\min} > 0$ 이라면 순차 입력 영상 i 에서 표정의 유형 후보를 결정한다.

3.2.2 얼굴 표정 단계의 인식과 피드백

프레임 m 에서 순차 입력된 표정의 유형을 인식한 후, 뒤따르는 프레임 n $(n>m)$ 의 d_n 의 표정의 유형을 결정하여 변화의 단계를 정할 수 있다. d_n 으로 결정된 유형의 모델과 순차 입력 영상 사이의 시간적 차이를 조정할 수 있다. 프레임 m 이전의 프레임 k 에서의 표정 변화의 단계는 $k\cdot d_{\min}(k\le m)$로 추정 할 수 있다.

프레임 $n-3$ 에서 프레임 n 의 각 프레임에서 표정 변화의 단계가 d_n 에 의하여 증가한다고 하면 식(3)은 다음과 같이 수정 할 수 있다.

$$Y_n^i = \sum_{k=n-2}^{n}\sum_{j=1}^{12}(C_{i,j}^*(d_{n-3} + (k-n-3)\cdot d_n) - \overline{P}_{jk})^2, \qquad (4)$$

d_n 을 고려하여 오차 효율을 $Y_n^i d_n$ 을 최소화한다. 현재 프레임 n 에서의 표정 변화 단계는 $d_n = d_{n-3} + 3\cdot d_n$ 이다. d_n 의 유사성으로서 d_{n-1} 과 d_{n-2} 또한 갱신한다.

이젠 확정된 표정 변화 단계 d_n 으로 다음 프레임 $n+1$ 에서의 얼굴 표정 특징점의 지점을 예측할 수 있다.

프레임 $n+1$ 의 표정 변화의 단계를 $d_{n+1} = d_n + d_n$ 이라 하면, 얼굴 표정 특징점 j 의 위치 $\overline{P}_{j,(n+1)}$ 는 다음에 의하여 계산 할 수 있다.

$$\overline{P}_{j,(n+1)} = \overline{P}_{j,n} + (C_{i,j}^*(d_{n+1}) - C_{i,j}^*(d_n)) \quad (j=1,\cdots,12) \qquad (5)$$

획득된 얼굴 표정 특징점의 위치는 추적 시스템의 입력 층을 초기화하는데 사용되어진다.

4. 실험 결과

실험은 3명의 얼굴을 고정시키고 5분 동안 각각 4번씩 순차적으로 취득한 후 그 중 한가지를 얼굴 표정 모델을 구성하는데 사용하였다. 이러한 얼굴 영상들은 일반적인 조명에서 초당 30 프레임으로 기록하였고 640×480 화소로 저장하였으며 세 가지 유형의 표정을 순차적으로 캡쳐하였다.

결합 온도를 달리하여 5000회 정도 반복하여 초기치 보다 추적시스템의 효율이 떨어지지 않을 때, 얼굴 특징점은 이전의 프레임에서 현재의 프레임으로 움직이지 않는다고 결론지었으며, 이러한 정지된 지점에서, 얼굴 표정의 시작과 마지막 프레임을 자동으로 정할 수 있었다. 순차 영상에서 결정된 프레임 길이는 19 프레임(행복), 12 프레임(놀람)과 9 프레임(노염)이였으며, 확정된 표정의 처음과 마지막 프레임을 그림 4에서 나타내었다. 19개의 얼굴 특징점의 추적 결과를 작은 흰 점으로 표기하였으며 인식 결과를 표 1에 나타내었다.

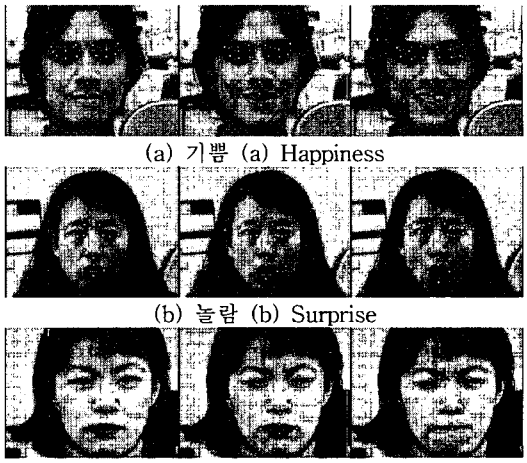
표정	성공	실패	백분율	오차
기쁨	11	0	100%	3.07
놀람	11	0	100%	2.58
노염	6	1	85.7%	3.78

표 1. 8 명의 피실험자에 대한 29개의 순차 입력 영상의 인식 결과

Table 1. Recognition results for the 29 image sequences from 8 subjects

5. 결론

본 논문에서는 얼굴 표정들의 유형과 순차영상에서 연속적인 얼굴 표정 변화를 인식하는 새로운 방법을 제안하였다. 이 제안된 연구방법은 개개인의 일반적인 얼굴 표정 모델과 정밀하게 얼굴 특징점을 찾아내는 기술을 사용하였으며, 일련의 피검자로부터 평균 B-spline 곡선으로 표정 모델을 구성함으로서 어떠한 피검자의 표정 변화도 인식할 수 있었다. 또한 얼굴 특징점의 지역 위상과 얼굴의 특징을 유지함으로서 가중 결합과 정합 된 분류그래프를 사용한 추적 시스템은 얼굴 특징점을 정밀하게 추적 할 수 있었다. 또 12개의 얼굴 표정 특징점 전체의 이동 정보를 이용하여 동시에 얼굴 표정을 인식함에 따라 얼굴 표정의 유형을 인식할 수 있었고 미묘한 표정 변화의 단계까지 추정할 수 있었다. 또한, 획득된 표정 정보가 피드백 됨에 따라 추적의 검색 시간을 감소시킴으로서 교점 확산 구간을 최소화 할 수 있었다.



(a) 기쁨 (a) Happiness

(b) 놀람 (b) Surprise

(c) 노염 (c) Anger

그림 4. 19개의 얼굴 특징점들의 추적결과에 대한 세 가지 순차 표정 영상

Figure 4. Three expressional image sequences with the tracking result of 19 FFPs

참고문헌

[1] J. Buhmann, J. Lange, and C. von der Malsburg. Distortion invariant object recognition by matching hierarchically labeled graphs. *IJCNN*, pages 155-159, 1989.

[2] I. A. Essa and A. Pentland. Facial expression recognition using a dynamic model and motion energy. In fifth international Conference on Computer Vision, pages 360-367, 1995.

[3] H. kobayashi and F. Hara. Recognition of six basic facial expressions and their strength by neural network. In IEEE International Workshop on Robot and Human Communication, pages 381-386, 1992.

[4] K. Mase. Recognition of facial expressions for optical flow. IEICE, 74(10), 1991.

[5] K. Matsuno, C. W. Lee, S. Kimura, and S. Tsuji. Automatic recognition of human facial expression. In fifth International Conference on Computer Vision, pages 352-359, 1995.

[6] Y. Yacob and L. Davis. Computing spatio-temporal representations of human faces. In International Conference on Computer Vision and Pattern Recognition, pages 70-75, 1994.