

# 표정변화에 따른 얼굴 표정요소의 특징점 추적 \*

최명근<sup>o</sup>, 정현숙, 신영숙, 이일병

연세대학교 컴퓨터과학과

{pobie912, hsch, hellogen, yblee}@csai.yonsei.ac.kr

## Tracking of Facial Feature Points related to Facial Expressions

Myounggeun Choi<sup>o</sup>, Hyunsook Chung, Youngsuk Shin, Yillbyung Lee

Dept. of Computer Science, Yonsei Univ.

### 요 약

얼굴 표정은 사람의 감정을 표현함과 동시에 그것을 이해할 수 있는 중요한 수단이다. 최근 이러한 얼굴 표정의 자동인식과 추적을 위한 연구가 많이 진행되고 있다. 본 연구에서는 대략적인 얼굴영역을 설정하여 얼굴의 표정을 나타내는 표정요소들을 찾아낸 후, 각 요소의 특징점을 추출하고 추적하는 방법을 제시한다. 제안하는 시스템의 개요는 입력영상의 첫 프레임에서 얼굴영역 및 특징점을 찾고, 연속되는 프레임에서 반복적으로 이를 추적한다. 특징점 추출과 추적에는 템플릿 매칭과 Canny 경계선 검출기, Gabor 웨이블릿 변환을 사용하였다.

### 1. 서론

컴퓨터 시각 분야에서 인체와 관련된 연구 중에는 얼굴, 제스처, 표정 등을 인식하거나 얼굴의 움직임, 표정의 변화, 몸 동작에 대한 추적과 애니메이션하는 방법에 대한 많은 연구가 수행되고 있다. 이 연구들은 각각 세부적인 부분으로도 나누어 지는데 얼굴의 움직임은 회전(rotation), 변환(translation)에 대해서, 표정은 눈, 눈썹, 입술에 대해서, 몸 동작은 손, 다리, 걸음등에 대해서 각각 독립적으로 연구가 이루어지고 있다.

표정 인식은 사람의 감정을 컴퓨터가 이해할 수 있다는 점에서 매우 중요하면서도 흥미로운 분야라 할 수 있다. 반면에 표정의 종류는 수없이 많으며 같은 표정에 대해서도 사람마다 각기 다르게 표현하기 때문에 표정을 인식한다는 것은 매우 어려운 문제이기도 하다.

얼굴 표정의 변화는 의사 소통에 있어서 중요한 영향을 미치는 요소 중 하나이다. 그렇지만 그 중요성 못지않게 어려움도 많이 따르고 있다. 다양한 표정의 변화는 의사소통에 있어서 감정을 풍부하게 전달해 줄 수 있는 반면에, 분류할 수 있는 기준을 더욱 찾기 어렵게 만든다. 현재 가장 널리 사용되고 있는 FACS (Facial Action Coding System)는 Ekman과 Friesen [1]에 의해 개발되었으며, 표정의 변화를 일으키는 기본 단위인 "action unit"로 구성된다. 그러나 FACS 모델도 시간적, 공간적 변화에 대한 정보가 부족하다는 것이 단점으로 지적되고 있다. 또한 각기 다른 개개인의 표정에 대한 얼굴 변화도 일정하지 않고, 얼굴의 회전이나 이동이 발생할 때에는 더욱 어려움이 따른다.

\* 본 연구는 과기부 감성공학기술개발 과제의 지원으로 수행되었음.

얼굴 표정요소의 움직임들을 추적하는데 있어서 고려되어야 할 사항은 정확성과 속도이다. 일반적으로 이 두 가지는 서로 상충되는데, 즉 정확성이 높아지면 속도는 느려지고 속도가 빨라지면 정확성이 낮아진다는 것이다. 그러나 화상회의와 같은 경우에는 속도와 정확성 모두 필수적인 요소이기 때문에, 결국 이 두 가지를 어떻게 하면 함께 높일 수 있는가 하는 것이 해결해야 할 문제이다. 본 연구에서는 이러한 점을 고려해서 표정요소의 움직임 추적을 위한 특징점으로 FAP(Facial Animation Parameter)에 해당하는 점을 사용한다. 이것은 표정을 합성하여 자연스러운 애니메이션을 하는데도 사용되어질 수 있을 것이다.

본 연구에서는 동영상에서 얼굴의 다양한 표정변화를 추적할 수 있는 방법을 제안한다. 제안하는 시스템은 (그림 1)과 같이 전처리 단계, 특징점 설정 단계, 특징점 추적 단계의 세 단계로 구성된다. 전처리 단계에서는 눈, 눈썹, 입술에 대한 템플릿과 각각의 특징점을 구성한다. 특징점 설정 단계에서는 대략적인 얼굴영역을 찾아낸 후 눈, 눈썹, 입술의 위치를 템플릿 매칭 방법을 이용해서 구하고 Canny 경계선 검출기와 Gabor 웨이블릿 변환을 적용해서 특징점을 설정한다. 특징점 추적 단계에서는 이전 프레임에서 찾아낸 특징점의 위치를 다음 프레임에서 특징점 후보 영역으로 사용함으로써 반복적으로 위치를 추적한다.

### 2. 관련연구

얼굴의 특징을 추출하기 위한 방법으로는 템플릿 매칭이나 스네이크, 신경망, 정규화된 얼굴의 대칭, 컬러정보, 광류, Gabor 웨이블릿 등의 방법이 있다. 템플릿 매칭은 얼굴

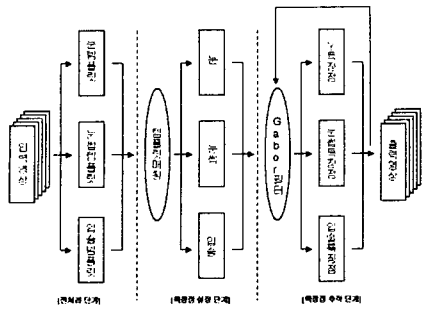


그림 1: 시스템 개요도

특징에 대한 템플릿을 작성하여 입력영상에 대해 템플릿과 매칭시킨 후 가장 유사한 모양의 특징영역을 찾아가는 방법이고, 정규화된 얼굴의 대칭은 얼굴 경계선의 각 화소중 그 주변 화소에 대하여 대칭성이 많은 지점을 특징영역으로 사용하는 방법이다. 신경망은 얼굴 특징영역을 신경망을 통해서 학습하는 방법이고, 컬러정보를 이용한 방법은 얼굴 피부 색상과 특징영역의 색상차를 이용해서 특징영역을 찾아내는 방법이다. Gabor 웨이블릿은 Gabor 웨이블릿 필터를 적용하여 얼굴내의 경계선 중 곡률이 높은 곳의 점을 특징점으로 추출하는 방법이다.

시간에 따른 표정의 변화를 추적하기 위한 연구는 최근 많이 진행되고 있다. 그 중에서 적외선 표시장치나 탐침을 사용해서 표정의 변화를 추적하는 방법, 얼굴에 표시를 함으로써 얼굴 표면의 3D 변위를 추적하는 방법, 3차원 얼굴 모델의 파라미터를 추정함으로써 특징점을 선형적으로 추적하는 방법등이 있는데, 이러한 시스템들의 단점은 카메라로 얼굴의 표정을 추적하기 위해서 사람의 얼굴에 특별한 표시가 필요하다는 것이다.

Azarbeyajani [2]는 눈이나 입의 모서리부분과 같은 얼굴의 작은 특징점을 추적하면서 재귀적으로 추정하는 방법을 소개했다. 그러나 이것은 전체 이미지 중에서 같은 점이 존재하는 부분에서만 사용할 수 있다는 단점이 있다. Saulnier [3]는 템플릿기반 방법을 사용했다. 여기서 제한점은 기본적인 학습이 필요하고, 추적할 수 있는 얼굴 움직임의 범위가 좁다는 것이다. Basu [4]는 정규화된 광류(regularized optical-flow)에 3D 모델을 사용함으로써 보다 정확한 결과를 얻었다. 그러나 일정한 움직임에 대해서만 좋은 결과를 보이고, 그렇지 않은 움직임에 대해서는 좋지 않은 결과를 보이는 단점이 있다. Jebara [5]는 3D 상에서 얼굴을 찾아내고, 모델링하고, 추적하는 시스템을 제안했다. 이것은 추적을 위해서 같은 점을 찾는다라는 단점이 있는데, [6]은 최적의 점을 뽑는 방법으로 이를 보완했다.

### 3. 특징 추출과 추적

얼굴의 표정변화를 추적하기 위해서는 표정과 관련된 요소들을 추출해서 움직임을 추적해야 하는데, 얼굴에서 표정

의 변화를 일으킬 수 있는 요소 중 눈, 눈썹, 입술이 가장 중요하므로 이 부분들을 정확하고 빠르게 추적할 수 있는 방법을 제안한다.

#### 3.1. 특징영역 설정

얼굴표정요소(눈, 눈썹, 입술)를 특징영역으로 설정하기 위한 단계로는 전처리 단계, 특징점 설정단계가 이에 해당한다. 전처리 단계에서는 각 모델영상에 대해 얼굴표정요소를 추출해내고 이에 대한 템플릿을 구성한다. 그리고 각 템플릿마다 FAP에 해당하는 점에서 Gabor 웨이블릿 변환 [7]을 수행한다. FAP는 MPEG-4 SNHC에서 정의한 얼굴 애니메이션을 위한 파라미터로 여기에 해당하는 점들 중 (그림 2(a))와 같이 눈썹(6), 눈(8), 입술(4)의 점들을 특징점으로 사용한다. 이러한 점들은 표정인식, 얼굴 애니메이션에도 사용되어질 수 있는 중요한 점들로써 여기서는 얼굴표정요소의 추적에 적용해 보았다.

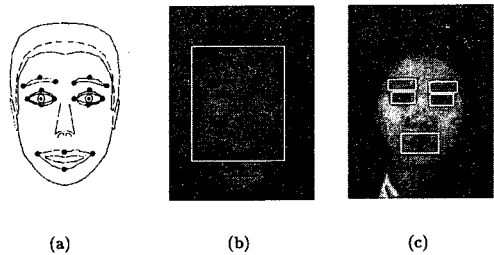


그림 2: (a) FAP에 해당하는 눈썹, 눈, 입술의 특징점. (b) 대략적인 얼굴영역 설정 결과. (c) 특징영역 설정 결과.

특징점 설정 단계는 영상이 연속적인 프레임으로 입력될 때 첫번째 프레임에서 특징영역과 특징점을 설정하는 단계로서 다시 특징영역 설정단계와 특징점 추출단계로 세분화된다. 특징영역 설정 단계에서는 대략적인 얼굴영역을 설정한 후에 특징영역을 설정하는데, 대략적인 얼굴영역(그림 2(b))은 입력영상의 히스토그램 평활화를 취하고 이를 다시 임계값으로 이진화시킨다. 이렇게 해서 얻어진 영상에서 높이의 가운데 지점 주위 좌/우의 픽셀을 흰 영역이 나올 때 까지 증가/감소 시키면서 얼굴의 좌/우영역을 찾아낸다. 얼굴의 상/하영역은 앞에서 찾아낸 얼굴의 좌에서 우까지 높이의 가운데 지점에서 픽셀을 검은 영역이 나올 때 까지 감소/증가 시키면서 찾아낸다. 특징영역(그림 2(c))은 입력영상의 첫번째 프레임에서 템플릿 매칭방법을 사용하여 설정한다. 템플릿의 매칭은 상관계수를 사용하였고, 각 특징영역의 주변에서 계수 값이 가장 큰 영역을 특징영역으로 설정하였다.

#### 3.2. 특징점 추출

얼굴의 특징점을 추출하는데 널리 이용되어온 경계선 검출 방법은 얼굴 각 부분의 위치나 움직임에 대한 많은 정보를 제공한다. 그렇지만 이 방법을 입 주위에 적용했을 때 나

타나는 경계선이 너무 많기 때문에 어떤 경계선이 입술의 어디에 해당하는지를 식별하기 어렵다. 그리고 조명이 달라짐에 따라서 경계선간의 차이가 발생하기 때문에 파라미터를 변경해야 하는 단점이 있다.

이러한 단점을 보완하기 위해서 계곡 윤곽선을 찾아내는 방법이나 Gabor 웨이블릿 변환 방법을 함께 이용할 수 있다. 입 주위에 너무 많이 나타나는 에지 방법과는 달리, 계곡 윤곽선을 찾아내는 방법은 두 입술 사이에 세 개 정도 존재하는 장점이 있는 반면에 입술의 위치에 대한 정보가 적다는 것이 단점이 있다. Gabor 웨이블릿 변환 방법은 입술의 위치나 모양을 잘 찾아내는 반면에 각도가 큰 변화에 대해서는 정확한 정보를 제공하지 못하는 단점이 있다.

특징점 추출 단계에서는 특징영역 설정 단계에서 얻어낸 특징영역에 Canny 경계선 검출기를 적용해서 경계선을 찾아내고, 상하좌우의 끝점을 특징점 후보점으로 정한다. 이렇게 해서 얻어진 특징점 후보점의 주위영역을 특징점 후보영역으로 설정하고 이 영역에 Gabor 웨이블릿 변환을 수행한다. 전처리 단계에서 획득한 Gabor 웨이블릿 변환값과의 상관계수를 이용해서 가장 유사한 점을 (그림 3(a))와 같이 특징점으로 설정한다.

### 3.3. 특징점 추적

특징점 추적 단계에서는 첫 프레임에서 추출된 특징점을 기준으로해서 연속되는 프레임에 대해 반복적으로 Gabor 웨이블릿 변환값과의 상관계수를 비교함으로써 변하는 위치를 (그림 3(b))와 같이 추적한다. 이 단계에서는 이전 프레임에서의 위치정보를 이용함으로써 특징영역을 설정할 필요 없이 특징점 후보영역을 설정할 수 있다. 즉, 이전 프레임에서의 특징점 주변 위치가 다음 프레임에서의 특징점 후보영역에 해당하는 것이다. 이 영역에 대해서 Gabor 웨이블릿 변환을 수행하고 상관계수를 비교함으로써 특징점의 위치를 추적한다.



(a)

(b)

그림 3: (a) 첫 프레임 특징점 설정 결과. (b) 변화된 표정의 특징점 추적 결과.

### 4. 결론 및 향후연구

본 연구는 동영상에서 얼굴표정 인식을 위한 전 단계로 얼굴표정요소의 움직임을 추적하는 작업을 수행하였다. 표정의 변화를 추적하기 위해서는 첫 프레임에서 원하는 특징점을 정확하게 설정해야 하고, 다음 프레임부터는 반복적으로 그 위치를 추적해야 한다. 여기서는 템플릿 매칭으로 얼굴표정요소를 추출한 후 Canny 에지 디텍터, Gabor 웨이블릿 변환을 사용해서 특징점을 설정하고, 그 위치를 다음 프레임에서의 특징점 후보영역으로 사용함으로써 특징점의 위치를 추적하였다.

표정관련요소의 움직임을 추적하는데 그 영역의 모든 점을 특징점으로 설정하지 않고 FAP에 해당하는 점만을 특징점으로 설정하였다. 이것은 표정의 변화를 충분히 반영하므로 표정인식에도 사용될 수 있고, 입력영상의 표정변화를 흉내내는 얼굴 애니메이션에도 사용될 수 있을 것이다.

얼굴표정요소의 추출과 특징점 추출시 사용된 상관계수를 이용하는 방법은 좋은 결과를 보여주는 반면 수행영역이 넓어지면 계산량이 많다는 단점이 있다. 즉, 움직임이 큰 변화에 대해서는 영역설정의 어려움이 따른다. 이러한 단점을 극복할 수 있는 방법을 추적단계에 적용하는 연구가 진행되고 있다.

### 참고 문헌

- [1] P.Ekman and W.V.Friesen. Facial Action Coding System. Consulting Psychologists Press Inc., 577 College Avenue, Palo Alto, California 94306, 1978.
- [2] A.Azarbayejani and A.Pentland. Recursive estimation of Motion, Structure, and Focal Length. *IEEE Pattern Analysis and Machine Intelligence*, 1995.
- [3] A.Saulnier, M.L.Viaud, and D.Geldreich. Real-time facial analysis and synthesis chain. In *International Workshop on Automatic Face and Gesture Recognition*, 86-91, Zurich, Switzerland, 1995.
- [4] Sumit Basu, Irfan Essa, Alex Pentland. Motion Recognition for Model-based Head Tracking. In *International Conference on Pattern Recognition*, 1996.
- [5] T.Jebara and A.Pentland. Parameterized Structure from Motion for 3D Adaptive Feedback Tracking of Faces. *IEEE Conference on Computer Vision and Pattern Recognition*, 1997.
- [6] J.Ström, T.Jebara, S.Basu and A.Pentland. Real Time Tracking and Modeling of Faces : An EKF-based Analysis by Synthesis Approach. In *Proceedings of International Conference on Computer Vision*, 1999.
- [7] Daugman J. High-Confidence Visual Recognition of Persons by a Test of Statistical Independence. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 15(11):1148-1161, 1993.