

# 얼굴 특징점 추적을 통한 사용자 감성 인식

이용환\*·김흥준†

\*원광대학교 디지털콘텐츠공학과, †경남과학기술대학교 컴퓨터공학과

## Emotion Recognition based on Tracking Facial Keypoints

Yong-Hwan Lee\* and Heung-Jun Kim†

\*Dept. of Digital Contents, Wonkwang University,

†Dept. of Computer Science and Engineering, Gyeongnam National University of Science and Technology

### ABSTRACT

Understanding and classification of the human's emotion play an important tasks in interacting with human and machine communication systems. This paper proposes a novel emotion recognition method by extracting facial keypoints, which is able to understand and classify the human emotion, using active Appearance Model and the proposed classification model of the facial features. The existing appearance model scheme takes an expression of variations, which is calculated by the proposed classification model according to the change of human facial expression. The proposed method classifies four basic emotions (normal, happy, sad and angry). To evaluate the performance of the proposed method, we assess the ratio of success with common datasets, and we achieve the best 93% accuracy, average 82.2% in facial emotion recognition. The results show that the proposed method effectively performed well over the emotion recognition, compared to the existing schemes.

**Key Words** : Facial Emotion Recognition, Active Appearance Model, Facial Keypoints, Facial Feature Tracking, Emotion Classification Model

### 1. 서 론

얼굴은 일상 생활에서 매우 중요한 대상이 된다. 우리는 얼굴을 통해서 사람의 신원을 파악하고, 그 사람에 대한 여러 정보를 얻을 수 있기 때문이다[1]. 얼굴 인식 및 신원 인증 기술이 다양한 분야에서 활용되고 있는 이유이다[2]. 얼굴 인식 분야 이외에, 사람의 감정을 인식하는 기술이 다방면에서 활용되고 있으며, 최근 다양한 기술과 기계의 발달로 인하여, 소셜 로봇이 화제가 되기도 한다 [3]. 과거의 기계적인 움직임을 통해 사람의 육체적 노동을 대신하는 기존 로봇과는 다르게, 소셜 로봇은 사람과의 커뮤니케이션 능력을 갖추고, 사람의 감정 상태를 분

석하여 정서적인 상호작용을 수행한다. 또한 사용자-컴퓨터 인터페이스에서도 감성 인식이 활용된다. 기존에 사용자 피드백을 취하는 가장 보편적인 형태가 마우스나 키보드를 활용하는 방법이다. 최근에는 카메라를 통해 사용자의 활동, 제스처를 자동으로 인식하여 지능적으로 상호 동작을 수행하기도 하지만, 스마트폰에 장착된 카메라를 활용하는 경우, 한 손만의 자유도를 얻는 단점이 있고, 이를 극복하는 방법으로 스마트 안경을 포함한 웨어러블 장비에서의 상대 얼굴 인식 기술이 활용되고 있다.

사람의 감정을 분석하는 감성 인식 기술은 크게 영상 인식 기반 기술과 음성 인식 기반 기술이 있다[1, 4]. 사람의 얼굴 영상을 인식하고 분석하여 감정을 인지하는 영상 기반 기술이 주변 환경 영향에 보다 강인하여 많은 연구가 수행되고 있다. 그러나, 다양한 응용분야에도 불구하고 사람의 감정을 분석하는 데에 많은 어려움으로 시

†E-mail: thinkthe@gntech.ac.kr

장 확산이 늦어지는 경향을 보인다.

본 논문에서는 입력된 영상을 분석하여 사용자의 얼굴 감성을 인식하는 방법을 제안하고 이를 구현한다. 제안 알고리즘은 입력 영상에서 얼굴 영역을 검출하고 AAM기반의 얼굴 특징 랜드마크를 추출한다. 얼굴 변화에 따른 감정 인식 모델을 수립하고, 특징점을 기반으로 사용자의 기본 감정을 인식한다.

본 논문의 구성은 다음과 같다. 2장에서 기존의 AAM 알고리즘과 얼굴 영상을 기반한 감정 인식에 대한 연구 동향을 살펴본다. 3장에서 얼굴 감정 인식 모델 수립 및 제안하는 영상 기반의 감정 인식 알고리즘을 설명한다. 4장에서는 실험을 통해 제안 알고리즘의 성능을 측정하고, 5장에서 결론으로 마무리한다.

## 2. 관련 연구

본 절에서는 얼굴 특징을 찾는 AAM에 대해 설명하고, 얼굴 기반의 감정 인식 기술에 대한 연구 동향을 살펴본다.

### 2.1 Active Appearance Model

Active Appearance Model(AAM)은 Shape모델과 Appearance 모델로 구성되며, 두 모델을 결합하고 각 파라미터의 조합으로 여러 얼굴 생김새를 표현한다[5].

Shape모델은 사람의 얼굴을 포함하는 이미지에  $n$ 개의 랜드마크 위치를 이용하여 구성되며[6], 이를 수식으로 표현하면 다음과 같다.

$$x = (x_1, y_1, x_2, y_2, \dots, x_n, y_n)^T \quad (1)$$

Shape는 표정과 생김새에 따라 다양한 벡터가 생성되며, 좌표상의 크기, 위치, 기울기가 일정하지 않으므로, Procrustes Alignment를 수행하여 벡터들을 정렬한 후, PCA(Principal Component Analysis)를 이용하여 Shape모델을 구성한다. 구성된 Shape모델은 수식(2)로 표현되며,  $\bar{x}$ 는 평균 Shape벡터,  $\Phi_i$ 는 Shape파라미터이다. 사람의 얼굴 생김새와 표정에 따라 절대적인 얼굴 각 부위의 좌표는 다르지만, 위상은 동일하게 나타난다.[6].

$$x_i = \bar{x} + \sum_{i=1}^k \Phi_i b_i \quad (2)$$

Appearance 모델은 피부, 눈, 입술 등의 색 정보로 구성된다. Appearance 벡터가 평균 Shape 벡터에 대응되기 위해 Delaunay Triangulation을 통해 매쉬를 구성하고, 생성된 매쉬를 Piecewise Affine Warping를 수행하여 정규화한다. Appearance

벡터  $A_i(x)$ 는 PCA를 통해 최종적인 외형 모델을 생성하며,  $A_0(x)$ 는 평균 Appearance 벡터를,  $\lambda_i$ 는 Appearance 파라미터를 나타낸다.

$$A(x) = A_0(x) + \sum_{i=1}^m \lambda_i A_i(x) \quad (3)$$

AAM Fitting에서는 입력 영상과 모델 사이의 오차를 최소화하는 파라미터를 계산한다. 입력 영상과  $p \leftarrow p + \Delta p$ 로 파라미터를 갱신하는 매개변수  $\Delta p$  사이의 오차를 최소화하는  $\Delta p$ 를 찾으며, AAM Fitting은  $\Delta p$ 의 반복적인 수정을 통해 오차가 더 이상 변하지 않거나 일정한 값 이하가 되었을 때, 완료된다[6].

### 2.2 감성 인식 연구 동향

사람의 감정을 시스템적으로 분석하고 인식하는 다양한 연구가 진행되고 있으며[8], Ekman은 얼굴 표정을 통해 심리학적 관점에서 얼굴 표정 감성을 인식하는 방법을 처음 제안하면서 감성 인식 분야의 새로운 개척자로 자리매김하고 있다[9]. Ira은 비디오 시퀀스에서 사람 얼굴 표정을 자동으로 구분하고 인식하기 위한 Hidden Markov Model을 제안하였다[10]. Yashnari는 열 화상 처리와 음성 인식 시스템을 결합하여 사람의 표정 인식 방법을 연구하였으며[11], 해당 논문에서는 감정을 5가지 범주로 분류하고 말하는 동안의 얼굴 표정 변화를 열처리 알고리즘으로 구분하여 음성 인식을 통한 감성 인식의 성공률을 향상시켰다. Sprios는 적절한 얼굴 특징을 추출하고 사용자의 감정 상태를 인식하는 여러 표정 변화를 연구하였다[12]. 추정 정확도의 신뢰도를 측정하기 위해 ISO MPEG-4 표준에 따라 정의된 정면 얼굴의 비디오에서 변화량을 추출하였다. 얼굴 정면 사진 또는 프로파일 얼굴 영상을 기반으로 얼굴 표정 단위를 인식하는 몇몇 프로토타입 시스템도 발표되었다[13-14]. 이들 시스템은 대부분 전문가 규정과 신경망 네트워크와 같은 기계 학습 방법을 포함하며, 특징-기반 이미지 표현 또는 모양-기반 이미지 분석 기법을 활용하고 있다. Valster은 확률과 통계를 기반으로 학습 기법을 적용하는 감성 인식 시스템을 제안하였다[15].

## 3. 제안하는 얼굴 감정 인식 모델

본 연구에서는 AAM을 통해 입력 영상의 랜드마크를 검출하고 표정 변화에 따른 랜드마크의 움직임과 변화로 감정을 분류하는 방법을 제안한다. 제안 알고리즘에 대한 블록 다이어그램은 Fig. 1과 같으며, 구현하는 감정 인식

체계는 기존에 알려진 AAM알고리즘을 통해 얼굴 특징 랜드마크를 추출하고[17], 제안하는 감정 인식 모델에 따라 랜드마크의 변화량을 계산하여 감정을 판단한다.

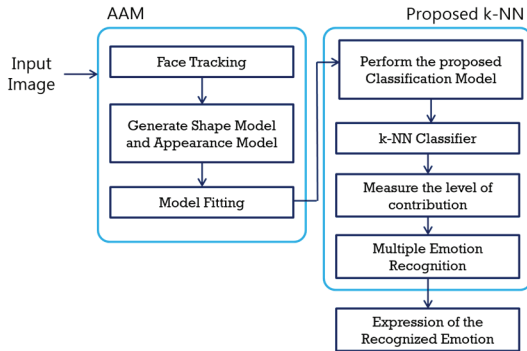


Fig. 1. Diagram of the proposed recognition system.

사람의 얼굴 감정을 분류하기 위해서는 먼저 얼굴 제어점의 움직임에서 감정 분류 표현식을 결정해야 한다.

Ekman 은 모든 얼굴 표정 단위의 집합을 정의하는 모델을 정의하면서, 얼굴 표정 단위에 기반하여 시각적으로 식별 가능한 얼굴 움직임을 표현하는 FACS(Facial Action Coding System) 시스템을 연구 개발하였다[9]. 해당 연구에서는 얼굴 표정 변화를 46개 단위(AU, Action Unit)로 구분하고 AU 조합 규칙을 통해 감정 지정 표현을 정의하는 표현식을 제안하였다. 본 연구에서의 얼굴 감정 판별 모델은 FACS의 기본 AU를 기반하여, 감정 인식을 위한 표현 규칙을 정의하고, 특징 랜드마크의 상대적 위치와 랜드마크의 변화량을 추적하여 사용자의 감성을 분류하여 판별한다. Fig. 2는 AAM을 통해 추출한 랜드마크 위치와 특징 벡터상의 랜드마크 순번이며, Table 1은 사용자 감정 별 랜드마크 변화에 대해 제안하는 분류체계이다.

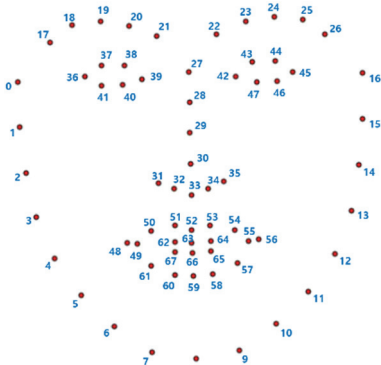


Fig. 2. Number of Facial Landmark.

Table 1. Landmark Variation for Facial Emotion

감정	표정특징	변하는 특징점
기쁨 (Happy)	눈이 작아짐	d(37,41), d(38,40)
	입이 약간 벌어짐	d(48,56), d(62,67)
	입꼬리가 올라감	△(31,4,18), △(48,50,57), △(56,54,61)
슬픔 (Sad)	입꼬리가 내려감	△(31,4,18), △(48,56,59)
	눈썹 끝이 내려감	△(0,21,16), d(17,21)
	입술이 찌그러짐	d(59,63), △(49,52,55), ◇(62,64,65,67)
	눈이 살짝 감김	d(38,40), d(36,39)
화남 (Angry)	눈썹이 내려감	△(0,21,16), d(21,22)
	눈꼬리가 올라감	d(17,21)
	입이 약간 벌어짐	d(48,56), ◇(62,64,65,67)

입력 데이터의 분류를 결정하기 위해 각각의 이웃마다 기여도를 할당하여 기여에 따른 정보를 추가하여 분류하는 Fuzzy k-NN(Neighbor Nearest) 알고리즘을 적용한다. 개별 클래스에 대한 기여도를 결정하는 Mahalanobis Distance를 계산하고 가까운 거리에 있는 클래스에 대한 기여도에 가중치를 적용한다. 클래스 c에 대한 이웃들의 기여도는 수식(4)로 계산된다.

$$\mu_{nc} = \frac{(x - m_c)^T \sum_c^{-1} (x - m_c)}{\sum_{c=1}^n (x - m_c)^T \sum_c^{-1} (x - m_c)} \quad (4)$$

n은 클래스 개수를,  $\sum_c$  는 c번째 클래스의 공분산 행렬을,  $m_c$ 는 c번째 클래스의 평균 벡터이다.

AAM 학습에 사용된 표정 Shape 벡터는 무표정을 기초하여 감정 모델에 따라 기쁨, 슬픔과 화남 표정에 대한 Procrustes Alignment를 수행하여 변화량을 측정하고 표정별 평균 Shape 벡터를 획득하고 분류한다. 입력 Shape 벡터를 Fuzzy k-NN 계산 분류에 따라 입력 영상에 대한 변화량을 구하여 가장 인접한 k개를 구한다. 이에 대한 기여도에 따라 표정의 정도를 계이자로 표시하고 근접한 감정을 인식한다.

제한한 감정 인식 모델의 성능적 평가를 위해 감정 인식 분야에서 많이 활용되는 CK+(Extended Cohn-Kanade Dataset)[16]을 대상으로 실험하였다.

실험 결과는 Table 2와 같다. 제안된 감정 인식 모델에 따라 프로토타입 시스템에서 82.2% 수준의 인식 성공률을 달성하였다. 실험 결과에 따르면 기쁨을 위한 제어점과 분산은 정규 표현식에서 보다 현저한 변화 차이를 보이며, 인식 성공률에서 가장 우수하게 나타났다. 얼굴 특징점의 변화에 민감한 감정에서 보다 우수한 인식률을 보였으며, 슬픔 감정의 경우 다른 감정에 비해 얼굴 특성을 묘사하는데 어려움이 있어, 얼굴 특징 랜드마크의 변화 상태를 정의하고 이를 추정하는 비율에서 상대적으로 다른 감정보다 낮은 인식 수준을 보였다.

**Table 2.** Emotion Recognition Rate among Different Facial Emotion

Emotion	Happy	Sad	Angry	Normal
Happy	<b>82%</b>	3%	57%	72%
Sad	3%	<b>75%</b>	68%	15%
Angry	57%	68%	<b>81%</b>	26%
Normal	72%	15%	26%	<b>80%</b>

#### 4. 결 론

본 논문에서는 사람 얼굴을 기반으로 감정을 분류하는 감성인식 모델을 제안하였다. 제안 시스템에서는 기존의 얼굴 특징 랜드마크를 추출하는 AAM 알고리즘을 활용하고 사람 감정에 따라 랜드마크의 변화량을 추정하여 감정을 인식한다. 추정 과정에서 랜드마크의 이웃간 기여도를 할당하고 가중치가 적용된 기여도 정보를 분류하는 Fuzzy k-NN 알고리즘을 적용하였다. 얼굴 감성 인식의 응용 데이터로 CK+를 이용하여 성능을 측정하였으며, 기쁨, 슬픔, 화남과 평범한 얼굴 감정을 대상으로 실험 결과 82% 수준의 감정 인식률을 보였다.

향후 제안 알고리즘을 보다 정교하게 구분하는 감정 인식 분류 모델로 확장하고, 4개의 감정 이외에 보다 많은 감정을 분류하도록 연구 확장하며, 응용 프로그램에서의 적용을 고려할 예정이다. 사람과 컴퓨터의 상호작용에 적용하여 사용자의 감정에 따라 컴퓨터의 서로 다른 서비스 지원, 사용자 상황 및 상태 감정에 따른 인지적 커뮤니케이션을 수행하고 분별력 있는 정보 디스플레이가 하나의 서비스 사례가 될 수 있다.

#### 감사의 글

이 논문은 2018-2020년도 경남과학기술대학교 대학회계 연구비 지원에 의하여 연구되었음.

#### 참고문헌

1. Aleix Martinez and Shichuan Du, "A Model of the Perception of Facial Expressions of Emotion by Humans: Research Overview and Perspectives," *Journal of Machine Learning Research*, Vol. 13, pp. 1589-1608, 2012.
2. Lacopo Masi, Yue Wu, Tal Hassner and Prem Natarajan, "Deep Face Recognition: A Survey," *International Conference on Graphics, Patterns and Images*, pp. 471-478, 2018.
3. Jamy Li, "Social Robots as Interactive Technology Agents: Supporting Design with Exploratory Assessment," *International Conference on Human-Robot Interaction*, pp. 629-630, 2016.
4. Changrong Yu, Jiehan Zhou and Kukka Rieki, "Expression and Analysis of Emotions: Survey and Experiment," *Workshops on Ubiquitous, Autonomic and Trusted Computing*, pp. 428-433, 2009.
5. G.J. Edwards, C.J. Taylor and T.F. Cootes, "Interpreting Face Images using Active Appearance Models," *International Conference on Automatic Face and Gesture Recognition*, 1998.
6. Eui-Young Cha, Jung-Hwa Lee and Hyun-Jun Park, "Facial Expression Recognition based on AAM using Backpropagation," *Journal of Korea Multimedia Society*, Vol. 13, No. 1, pp. 227-230, 2010.
7. T.F. Cootes, G.J. Edwards, C.J. Taylor, "Active Appearance Models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 6, pp. 681-685, 2001.
8. G. H. and P. M., "Automatic Temporal Segment Detection and Affect Recognition from Face and Body Display," *IEEE Transactions on Systems, Man, and Cybernetics - Part B*, Vol. 39, No. 1, pp. 64-84, 2009.
9. P. Ekman and W. Friesen, *Facial Action Coding System (FACS): Investigator's Guide*, Consulting Psychologists Press, 1978.
10. I. Cohen, A. Garg, and T. S. Huang, "Emotion Recognition from Facial Expressions using Multilevel HMM," *Workshop on Affective Computing*, 2000.
11. Y. Koda, Y. Yoshitomi, M. Nakano, and M. Tabuse, "Facial Expression Recognition for Speaker using Thermal Image Processing and Speech Recognition System," *International Symposium on Robot and Human Interactive Communication*, 2009.
12. S. V. Ioannou, A. T. Raouzaoui, V. A. Tzouvaras, T. P.

- Mailis, K. C. Karpouzis, and S. D. Kollias. Emotion, "Recognition through Facial Expression Analysis based on a Neurofuzzy Network", *Neural Networks*, Vol. 18, No. 4, pp. 423-435, 2005.
13. P. M. and I. Patras, "Dynamics of Facial Expression: Recognition of Facial Actions and their Temporal Segments from Face Profile Image Sequences," *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 36, No. 2, pp.433-449, 2006.
14. P. M. and L. Rothkrantz, "Facial Action Recognition for Facial Expression Analysis from Static Face Images," *IEEE Transactions on Systems, Man and Cybernetics - Part B*, Vol. 34, No. 3, pp. 1449-1461, 2004.
15. M. Valster, P. M., Z. Ambadar, and J. Cohn, "Spontaneous vs. Posed Facial Behavior: Automatic Analysis of Brow Actions," *ACM*, pp. 162-170, 2006.
16. P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, "The Extended Cohn-Kanade Dataset (ck+): A Complete Dataset for Action Unit and Emotion-Specified Expression," *Conference on Computer Vision and Pattern Recognition*, pp. 94-101, 2010.
17. Y.H. Lee and H.J. Kim, "Face Detection using AdaBoost and ASM," *Journal of the Semiconductor and Display Technology*, Vol. 17, No. 4, 2018.

---

접수일: 2019년 3월 19일, 심사일: 2019년 3월 25일,  
 게재확정일: 2019년 3월 25일