

퍼지 신경망과 강인한 영상 처리를 이용한 개인화 얼굴 표정 인식 시스템

김 대 진, *김 중 성, 변 증 남
한국과학기술원(KAIST) 전자전산학과,
*한국전자통신연구원(ETRI)
전화 : 042-869-8019 / 핸드폰 : 016-9595-6689

Personalized Facial Expression Recognition System using Fuzzy Neural Networks and Robust Image Processing

Dae-Jin Kim, *Jong-Sung Kim, Zeungnam Bien
Dept. of Electrical Engineering and Computer Science, KAIST,
*VR Center, ETRI
E-mail : djkim@mail.kaist.ac.kr

Abstract

This paper introduce a personalized facial expression recognition system. Many previous works on facial expression recognition system focus on the formal six universal facial expressions. However, it is very difficult to make such expressions for normal person without much effort and training. And in these days, the personalized service is also mainly focused by many researchers in various fields. Thus, we propose a novel facial expression recognition system with fuzzy neural networks and robust image processing.

I. 서론

얼굴 표정 인식 시스템은 최근 각광받고 있는 생체 인식 기술의 한 가지 부류로써, 비접촉 방식의 비전 센서를 이용하여 사용자 식별 및 감정 모니터링 등 다양한 응용 분야를 가지고 있다. 많은 생체인식 기술이 그러하듯, 얼굴 표정 역시 각 개인별로 다양한 특징이 존재하며, 다수 사용자에게 공통으로 사용되는 특징을 찾기가 힘들다. 기존의 연구에서는 이러한 개인화 측면보다는 인지 심리학에서 알려진 인종, 문화, 성별에

무관한 기본적 얼굴 표정을 인식하는 시스템에 초점을 맞추어 왔다. 하지만, 실세계에서의 다양한 적용 분야를 고려해 볼 때 개인화 측면을 고려하지 않은 얼굴 표정 인식 시스템은 큰 의미를 갖지 못한다.

본 논문은 개인의 특성을 가장 잘 반영할 수 있으며 다양한 조건의 얼굴 영상에 대한 처리가 가능한 방법과의 조화를 통해 개인화 얼굴 표정 인식 시스템을 제안하고자 한다. 특별히, 개인화된 얼굴 표정 인식 시스템을 구현하기 위하여 퍼지 규칙의 자동화된 학습이 가능한 퍼지 신경망을 이용한다. 다양한 조건의 얼굴 영상 처리를 위해, 컬러 히스토그램에 기반한 영상 내의 얼굴 위치 추정, 그룹화된 투영 정보 및 Deformable Template Model (DTM) 기반 얼굴 요소 추출, 가버 필터 및 인간 시각 시스템에 기반한 특징 추출 등을 제안한다.

본 논문은 다음과 같이 구성된다. 2장에서는 개인화 서비스의 구조와 본 논문에서 지향하는 전체 시스템의 구조를 도시화하고, 3장에서는 전체 시스템 중에서 영상 처리 부분에 해당하는 강인한 영상 처리 부분에 대해 설명한다. 4장에서는 개인화 얼굴 표정 인식 시스템의 핵심을 담당하는 퍼지 신경망 및 구분 지수에 대하여 기술하고, 영상 처리 부분과의 결합을 통해 얻어지 결과를 살펴본다. 마지막으로, 5장으로 결론을 맺는다.

II. 개인화 얼굴 표정 인식 시스템

2.1 개인화 서비스 구조

개인화 서비스는 서비스 대행자(Service Agent)의 입장에서 각 개인에 대한 프로파일(Profile)을 작성하고, 이 프로파일에 근거하여 개인의 취향이나 선호도에 적합한 서비스를 제공하는 형태를 의미한다 [1].

그림 1은 본 논문에서 제안하는 개인화 서비스의 구조를 개략적으로 도시한 것이다. 즉, 사용자는 특정한 행위를 기반으로 서비스 대행자로 하여금 개인 식별이 가능하도록 정보를 제공하고, 서비스 대행자는 행위 정보를 기반으로 개인 식별을 행한다. 개인에 대한 ID 정보가 얻어지면 이를 이용하여 기존의 서비스풀(Service Pool) 내에 구축된 다양한 서비스 중에서 특정한 서비스를 선정하고, 현재 식별된 사용자에게 대하여 개인화된 서비스를 제공하게 된다.

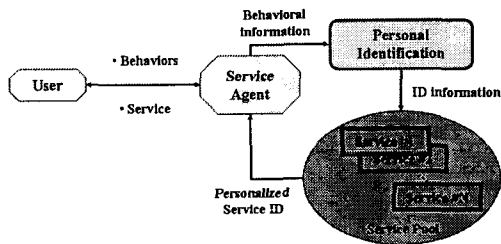


그림 3 개인화 서비스 구조

본 논문에서 대상으로 하는 개인화 서비스는 영상 기반의 얼굴 표정 인식 시스템이며, 각 개인에 적합한 서비스를 제공하기 위해서는 영상 처리에서의 강인성 이외에도 인식기 자체의 개인화 구조 역시 필수적이다.

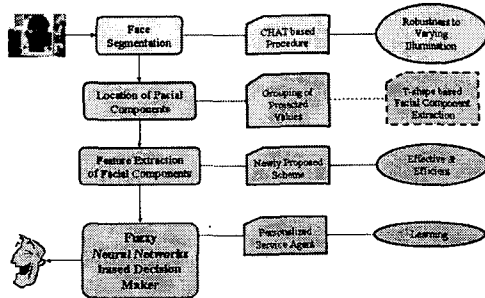


그림 4 개인화 얼굴 표정 인식의 흐름도

2.2 개인화 얼굴 표정 인식 시스템의 구조

개인화 얼굴 표정 인식 시스템은 크게 영상 처리를 담당하는 부분, 표정 인식을 담당하는 부분으로 구분

된다(그림 2). 영상 처리를 하는 부분은 얼굴 위치의 추정, 얼굴 영상에서의 각 요소(눈, 눈썹, 입 등) 추출, 추출된 각 요소에 대한 특정 추출 과정으로 세분화된다. 표정 인식을 담당하는 부분에서는 개인화 서비스에 초점을 맞추어 각 개인에 대한 학습과 적용이 가능한 구조를 취하게 된다.

III. 강인한 얼굴 영상 처리 기법

3.1 얼굴 영상 분할

본 논문에서는 얼굴 영상 분할을 위하여 I1I2I3 컬러 공간을 적용, 컬러 정보와 움직임에 기반한 얼굴 영상 분할 기법을 사용하였다. '컬러 히스토그램 기반 적응형 임계치(Color histogram based Adaptive Threshold; CHAT)' 방법으로 불리우는 이 방법은 각 프레임마다 얻어진 얼굴 영상 부위를 기반으로 하여 컬러 히스토그램 분석을 통해 가장 적합한 임계치를 얻는 과정을 의미한다 [3]. 그림 3은 CHAT 기법을 적용하여 얼굴 영상 분할을 행한 결과를 시간적으로 보여준 것이다. 초기에 선정된 얼굴 및 주변 영상을 중심으로 임계치가 변경되어, 최종적인 얼굴 영상만을 효과적으로 추출하는 것을 볼 수 있다.

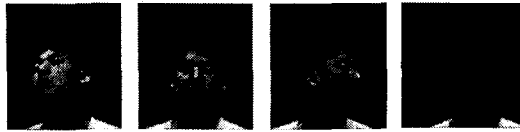


그림 3 얼굴 영상 분할 (CHAT 기법 적용)

3.2 얼굴 요소 추출

얼굴 요소의 추출을 위해서는 여러 가지 방법이 있을 수 있겠으나, 본 논문에서는 얼굴 요소 추출을 그 단계에 따라 다음과 같이 나누어 접근하였다.

먼저 전체적인 얼굴 요소의 배치 형태를 추정하기 위하여 대부분의 얼굴 영상에서의 얼굴 요소의 배치 형태가 T자형 모양을 보인다는 점에 착안하여, T자형 기반의 가변형 템플릿 매칭(Deformable Template Matching)을 행한다(그림 4). 그 후, 수평 에지 분석을 통해 각 세부 요소를 추출한다. 수평 에지 분석에 사용된 방법은 투영 그룹화(Grouping by Projection) 방법이며, 이것은 수평 에지를 수직면으로 투영한 후 임계치에 의한 그룹화를 행하고, 여러 가지 기하학적 정보(각 그룹의 수직면 상의 길이, 집중도 등)에 기반하여 세부 요소를 추출해 내는 과정을 의미한다(그림 5).



(a) 원영상



(b) 수평에지영상



(c) T자형 기반 가변 템플리트 매칭
그림 4 얼굴 요소의 대략적 추정



(a) 수평 에지의 수직 투영 (b) 투영 정보의 그룹화



(c) 눈, 눈썹의 수평 위치 추정

(d) 눈, 눈썹의 수직 위치 추정

그림 5 투영 그룹화에 의한 얼굴 요소의 세부 추정

3.3 얼굴 특징 추출

3.2 절에서 제안된 방법을 이용해 추출된 각 세부 요소에 대한 특징 추출을 위해 본 논문에서는 두 가지 점에 착안하였다. 우선, 기본적인 특징 추출 방법을 인간의 시각 시스템에 기반한 최근의 연구 경향을 반영하였으며, 특징 추출에 있어서는 전역적(Global) 특징과 국부적(Local) 특징을 결합하고자 하였다.

(가) 입에 대한 특징 추출

입에 대한 특징으로써, 전역적인 특징은 3.2절의 요소 추출에서 얻어지는 얼굴 영상에 대한 입 영상의 높이비와 면적비를 사용하였다.

입의 상태는 크게 입을 여느냐 닫느냐로 나뉘어 질 수 있다. 입의 여닫음은 앞서 언급한 전역적 특징으로 구분할 수도 있으나, 입술이 두꺼운 사람이나 높이비, 면적비 등이 사람에 따라 다른 경우가 많으므로 국부적 특징이 필수적이다.

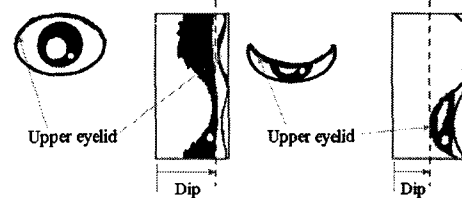
국부적 특징의 추출을 위해 본 논문에서는 가버-가우시안(Gabor-Gaussian: GG) 특징을 사용하였다. 가

버-가우시안 특징은 가버 필터(Gabor filter) 응답에 대한 수직 투영 및 가우시안 가중치 적용에 의한 특징으로써 식 (1)과 같이 표현 가능하다.

$$f_G = \frac{\sum_{i=1}^{H-1} w_G(i) dy_{proj}(i)}{\sum_{i=1}^{H-1} w_G(i)} \quad (1)$$

(나) 눈에 대한 특징 추출

눈의 상태는 눈을 뜨느냐 감느냐로 대비될 수 있다. 이러한 정보를 효과적으로 얻기 위해 여러 가지의 영상을 관찰한 결과, 눈의 위쪽 눈꺼풀의 모양에 따라 눈을 뜨는지 감는지 알 수 있다는 사실에 착안하여, 공간 변화 영상을 구현하기 위한 대표적 기법의 하나인 로그-폴라 영상과 로그-폴라 영상 특징 중의 하나인 딥(Dip) 특징 [6]을 사용하여 눈을 뜨고 감는지 추정이 가능하다.



(a) 눈을 뜬 경우

(b) 눈을 감은 경우

그림 6 로그-폴라 영상과 딥 특징을 이용한 눈에 대한 특징 추출

전역적 특징으로는 그림 6에 나타난 특징을 사용하였으며, 국부적 특징으로는 가버 필터 응답을 이용하여 효과적인 눈에 대한 특징 추출이 가능하였다.

(다) 일시적(Transient) 얼굴 요소에 대한 특징 추출

일시적 얼굴 요소란, 눈이나 눈썹 등 영구적(Permanent)으로 존재하는 얼굴 요소와 달리, 경우에 따라 나타나기도 하고 금방 사라지기도 하는 얼굴 요소를 지칭한다 [5]. 대표적인 일시적 얼굴 요소로는 까마귀 발 주름(Crow's feet wrinkle), 입 주변 주름(Nasolabial Furrow) 등이 있다 (그림 7). 본 논문에서는 특징으로는 가버 필터 계수를 사용하였다.

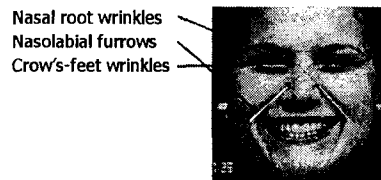


그림 7 일시적 얼굴 요소 [5]

IV. 퍼지 신경망에 의한 얼굴 표정 인식

4.1 퍼지 신경망

퍼지 신경망은 퍼지 의사 결정 시스템의 신경망 기반 구현 형태로써, 기존의 퍼지룰에 의한 전문가의 지식 표현 기능에 학습 기능을 추가한 형태로 볼 수 있다 [7]. 본 논문에서는 [7]에서 제안된 퍼지 신경망의 구조를 기본적으로 채용하고, 퍼지 신경망의 출력에 대한 평가를 위해 구분 지수(Separation Index)를 별도로 제안하여 표정 인식을 위한 척도로 사용하였다.

4.2 구분 지수

구분 지수는 기본적으로 패턴 인식 문제에서 서로 다른 클래스(Class) 간의 구분 정도를 나타내기 위한 척도로 간주될 수 있다. 구분 지수가 갖추어야 할 요건은 다음의 세 가지로 볼 수 있다.

- ① 목적요소의 값이 v_H 에 가까울수록 크다.
- ② 목적요소와 타 요소간의 값 차이가 클수록 크다.
- ③ 타 요소들의 값이 v_H 에 가까울수록 작다.

여기서, 퍼지 신경망의 출력은 N차원 벡터이며, 목적은 N개의 클래스를 구분하는 것이라 할 때 학습 출력 벡터는 $[v_L \ v_L \dots v_H \dots v_L]$ 과 같이 단 1개의 요소만 v_H 값을 갖는다. 또한, 목적요소(Target element)라 함은 학습이 완료된 퍼지 신경망의 출력 벡터 중에서 n 번째 요소의 값을 의미한다. 여기서, $1 \leq n \leq N$ 이다. 목적 요소는 임의의 입력에 대한 출력을 평가하여 현재 가장 적합한 클래스를 판별하기 위한 것이다.

4.3 실험 결과

본 논문에서는 제안된 방법의 유용성을 검증하기 위하여 우선적으로 Ekman 등이 제안한 6가지 대표적인 표정(기쁨, 슬픔, 분노, 역겨움, 공포, 놀람) 및 무표정을 대상으로 하여 얼굴 표정 인식 시스템을 학습시키고 그 결과를 얻어 보았다 (그림 8).

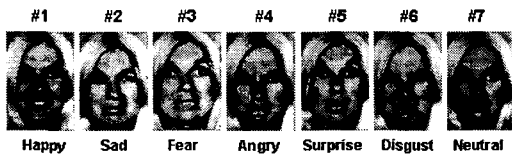


그림 8 실험에 사용된 7가지 얼굴 표정 영상 [2]

표 1의 가로축은 퍼지 신경망의 출력 노드 값을 의미한다. 즉, 그림 8의 각 영상에 대한 퍼지 신경망의 출력 노드 값으로부터 현재 입력된 영상이 어떠한 표정인지를 인식할 수 있는 것이다. 표 1의 마지막 열(Column)은 4.2절에서 제안된 구분 지수 값을 의미한다. 결과에 따르면, #2(Sad)의 경우 목적요소(#2)의 값이 다른 요소들에 비해 크긴 하지만 구분 지수가 낮아서 신뢰도가 낮으며, #7(Neutral)의 경우 #2(Sad)보다

도 낮은 출력값을 가지고 구분 지수도 낮으므로 표정 인식 시스템의 인식 대상으로 적합하지 않음을 알 수 있다. 즉, #2(Sad)와 #7(Neutral)에 대한 구분 정도가 낮으므로 나머지 5개의 표정에 대한 서비스를 제공하는 것이 유리하다.

표 5 개인화 얼굴 표정 인식 결과

출력값 입력영상	#1'	#2'	#3'	#4'	#5'	#6'	#7'	S.I.
#1	0.78	0.22	0.47	0.2	0.17	0.21	0.2	1.2054
#2	0.18	0.22	0.17	0.21	0.12	0.17	0.2	0.0571
#3	0.2	0.22	0.47	0.2	0.32	0.2	0.2	0.6488
#4	0.2	0.23	0.37	0.8	0.17	0.21	0.2	1.2249
#5	0.19	0.23	0.31	0.18	0.63	0.18	0.2	0.9083
#6	0.22	0.22	0.17	0.22	0.1	0.8	0.23	1.1975
#7	0.18	0.22	0.18	0.18	0.11	0.16	0.2	-0.0042

V. 결론

이 논문에서는 개인화 얼굴 표정 인식을 위하여, 퍼지 신경망을 이용한 얼굴 표정 인식기를 제안하고 이와 함께 강인한 영상 처리 방법을 이용한 얼굴 영상 처리 방법을 제안하였다. 제안된 방법은 인지 심리학 연구의 표준으로 사용되는 얼굴 영상 DB를 기준으로 하여 그 유용성이 검증되었다.

본 논문은 개인 식별이 필요한 공간에서의 개인화된 서비스를 제공하기 위해 필수적인 요소로 사용될 것으로 확신된다. 예를 들자면, 미래의 지능형 주택의 경우 각종 생체인식 기술을 통해 개인 식별이 가능하게 될 것이며 이에 부합되는 개인화된 서비스의 한 가지로 특정 사용자에 대한 표정 인식 시스템을 구현하는데 효과적일 것이다.

참고문헌

- [1] S. Abramowski et al., "Service management for personalized services," *Intelligent Network '94 Workshop*, vol. 1, pp. 239-251, 1994.
- [2] P. Ekman et al., "Facial Action Coding System", 1978.
- [3] 변중남 외, "지능형 비주얼 서보잉과 근전도 신호 인식 기술을 이용한 재활 로봇의 제어 기술," *2001 Service Robot Workshop: Dependable Service Robot Technology*, pp. 72-92, 2001.
- [4] 김대진 외, "인간친화형 인터페이스를 위한 사용자 얼굴에서의 효과적인 의도 파악," *대한전자공학회 추계종합학술대회*, vol. 3, pp. 25-28, 2000.
- [5] Y.-L. Tian et al., "Recognizing action units for facial expression analysis," *IEEE Transactions on PAMI*, vol. 23, no. 2, pp. 97-115, 2001.
- [6] C. F. Weiman, "Tracking algorithm using log-polar mapped image coordinates," *SPIE: Intelligent Robots and Computer Vision VIII: Algorithms and Techniques*, vol. 1192, pp. 843-853, 1989.
- [7] C.T. Lin, *Neural Fuzzy Control Systems with Structure and Parameter Learning*, World Scientific, 1994.