

표정인식을 활용한 고객피드백에 관한 연구

송은지*, 강민식**

*남서울대학교 컴퓨터학과

**남서울대학교 산업경영공학과

A Study on Customer Feedback using Facial Expression

Eun-Jee Song*, Min-Shik Kang**

*Dept. of Computer Science , Namseoul University

**Dept. of Industrial Management Engineering, Namseoul University

sej@nsu.ac.kr

요 약

최근 감성ICT산업은 성숙기 IT산업을 새롭게 도약시킬 핵심 산업으로 인식되면서 관련 산업계의 주목을 받고 있으며 다양한 분야에 접목되고 있다. 특히, 인간표정을 인식하는 IT기술은 사회적 측면에서 인간중심의 미래생활 패러다임의 변화가 감성이해를 통한 사용자 친화적 솔루션으로 발전하고 있다. 효율적인 경영을 위해 기업은 고객의 요구사항을 정확히 파악하는 것이 중요한데 본 연구에서는 이러한 감성ICT 기술을 이용한 새로운 커뮤니케이션의 사례로서 고객의 감정 중에 특히 얼굴표정을 파악하여 고객중심의 맞춤형 서비스 기능을 제공할 수 있도록 얼굴표정에 의해 호감도를 특정할 수 있는 알고리즘을 제안한다. 이것은 기존의 7개 표정인식 알고리즘을 이용하여 고객만족도를 특정할 수 있도록 한 것이다.

키워드

Emotional Information and Communication Technology , Facial Expression, Customer feedback

I. 서 론

감성ICT산업은 일상생활에서 제공되는 제품과 서비스에 인간의 감성을 자동 인지하고 하용상황에 맞게 감성정보를 처리하여 사용자 감성 맞춤형 서비스를 제공하는 감성ICT기술이 융합된 산업으로 정의하고 있다. 최근 감성ICT산업은 성숙기 IT산업을 새롭게 도약시킬 핵심 산업으로 인식되면서 관련 산업계의 주목을 받고 있으며 다양한 분야에 접목되고 있다. 감성ICT산업은 사회적 측면에서는 인간중심의 미래생활 패러다임의 변화가 감성이해를 통한 사용자 친화적 솔루션을 요구하는데서 등장하였고 그 필요성이 증가하고 있다. 또한 시장적 측면에서도 감성 자극 기반인 스마트 디바이스의 보급과 확산이 되고 있다. 기술별로 살펴보면 사용자의 음성으로부터 감정을 인식하는 휴대폰을 비롯하여 얼굴표정이나 문자로부터 감정을 인식하는 휴대폰과 체온, 맥박 등의 생체정보로부터 감정을 인식하는 휴대폰 등 다양한 방식의 감정인식 휴대폰 기술들이 출원되고 있다.

서비스 산업의 경쟁력을 높이기 위해서는 고객 중심의 피드백 분석 및 평가가 필요하며 고객의 의견을 실시간으로 모니터링하고 문제가 발생하였을 때 이를 알리고 적절한 대안을 지원하는 시스템이 필요하다. 본 논문에서는 사용자의 감정 기호를 파악하여 고객 중심의 맞춤형 서비스 기능을 제공할 수 있도록 인간의 표정을 인식하는 알고리즘을 이용하여 만족도를 측정하는 알고리즘을 개발 한다.

II. 제안하는 방법

얼굴의 표정 변화는 일상적인 사회적 상호작용에서 근본적인 요소이다. 표정 변화는 주의, 정서, 의도, 반응 등을 나타내는 주요한 단서이다. 본 연구는 자동적인 방법으로, 자연스러운 사용환경에서 사용자 얼굴영상에서 나타나는 표정, 정서 혹은 호감도를 인식하는 것을 목적으로 한다[1]. 기존의 표정 인식 연구는 대부분 컨트롤된 환경에서 획득된, 정면 얼굴만을 대상으로 하기 때문에, 얼굴의 자세 변화, 조명 변화 등을 수반하는

경우에 성능이 급격하게 저하되는 문제가 있다. 본 연구에서는 얼굴의 자세 변화, 국소적인 변형, 조명에 강인한 영상의 특징을 고려하여 자연스러운 환경에서, 높은 정확도로 표정을 인식하는 것을 목적으로 한다[2,3].

본 연구에서는 다음과 같은 방법을 제안하고, 개발하였다. 첫째, 입의의 자세로 주어진 얼굴영상에 대하여, 정면영상생성 기법을 적용하여, 정면영상으로 렌더링 한 후에, 표정해석 기법을 적용함으로써, 자세변화에 불변하는 특징요소를 추출하고 비교하는 것이 가능하다. 주어진 영상에 대하여, 얼굴을 검출하고, 얼굴 특징점을 추출한 후에, 3차원 모델을 이용하여, 3차원 얼굴포즈를 추정하고, 모델링한 후에, 컴퓨터 그래픽스 렌더링 기법을 적용하여 정면영상을 생성한다. 둘째, 정규화된 정면영상에 대하여, 대표적인 딥러닝신경망 학습방법의 일종인, 콘볼루션신경망 (이하 CNN) [51] 을 적용하여 표정분류에 적합한 특징을 추출하고 표정을 분류한다. 영상을 기반으로 특징을 추출하기 때문에, 표정 변화시 발견되는 미묘한 영상텍스처의 변화를 잡아낼 수 있으며, CNN의 적용으로, 조명 등의 변화에 강인한 특징요소의 추출이 가능하다. 이와 같은 접근 방법으로, 제약이 없는 자연스러운 사용환경에서 (in the wild) 적용할 수 있도록, 자세변화, 조명변화에 강인한 표정인식 시스템을 구현하는 것이 가능하다. 정면영상생성 기법과, CNN을 이용한 표정인식 모듈이 각각 구현되었고, 개별적으로 테스트 되었다. CNN을 이용하여 개발된 영상기반의 표정인식 알고리즘을 대표적인 표정 데이터베이스를 활용하여 테스트 한 결과, 7가지 표정에 대하여, 92.3%의 평균 인식성능을, 호감도 분류 (like, dislike, neutral) 에서는 86.6%, 각성수준의 분류 (aroused)에서는 89.6%의 인식성능을 보였다.

III. 실험결과

제안하는 방법의 성능을 2가지의 베이스라인 방법과 비교하였다. 표1은 입력영상에 대하여 가우시안 필터와 히스토그램 평활화를 적용한 후에, 벡터로 나타낸 후에 K-NN (k=20) 방법을 사용한 결과를 나타낸다. 표2는 입력영상 대신에 추출된 2차원의 특징요소를 정규화 한 후에, 벡터로 표현한 후에 동일한 K-NN을 적용하여 실험한 결과를 나타낸다. 영상픽셀을 처리하여 적용한 방법과, 특징요소를 직접 사용하여 인식한 종래의 방법에 비하여, 제안하는 CNN 방법이 더 높은 인식성능을 나타내었음을 확인하였다. 표 3에서는 표준적인 7개의 표정으로 분류한 결과를, 선호도 (좋음 - 보통 - 나쁨) 분류기와 각성수준 (각성 - 보통) 분류기를 사용한 결과들과 비교하였다. 선호도 분류기는 동일한 데이터베이스내의 영상자료들을 새로운 기준을 사용하여 분류한 후에, 학습 및 검사에 사용하였다. 즉, joy 표정은 “좋음” (like) 로, neutral은 보통으로, 나머지 모든 표정들은

“나쁨” (dislike)으로 재 그룹하였다. 이와 유사한 방식으로, 각성수준 분류기는 anger, disgust, fear, joy, surprise를 “각성” (aroused)으로 neutral과 sadness를 “보통” (neutral)로 분류하였다 선호도 분류기와 각성수준 분류기는 각각 일관된기준으로 분류된 검사자료에 대하여 86.6%와 89.6%의 인식성능 (정확도)을 보였다. 이렇게 새롭게 분류된 표정 분류기들은, 설문조사와 연동하거나, 학습환경에서 사용자의 각성 정도를 파악하고 분류하는데 활용될 수 있을 것이다.

표 1. 표정인식기반 호감도 및 각성도 인식률

	accuracy	classification categories	positive data	negative data	applications
Standard 7 expressions	92.3%	anger, disgust, fear, joy, neutral, sadness, surprise	one expression	all others	baseline research
"like" classifier	86.6%	like, neutral, dislike (or positive/negative)	joy	anger, disgust, fear, sadness, surprise	survey feedback, user monitoring
"aroused" classifier	89.6%	aroused, neutral (or active/inactive)	anger, disgust, fear, joy, surprise	neutral, sadness	survey feedback, user monitoring

IV. 결론 및 향후과제

본 연구에서는 감성ICT기술 중 표정인식기술을 이용하여 고객의 만족도를 평가하는 알고리즘을 제안하고 실험결과에서 높은 인식률을 보였다. 즉, 영상픽셀을 처리하여 적용한 방법과, 특징요소를 직접 사용하여 인식한 종래의 방법에 비하여, 제안하는 CNN 방법이 더 높은 인식성능을 나타내었음을 확인하였다. 표준적인 7개의 표정으로 분류한 결과를 이용하여 선호도분류기와 각성수준 분류기를 사용한 결과들과 비교하여 좋은 결과를 얻었다. 이것은 설문이나 SNS를 통해 얻은 고객만족도에 관한 정보의 정확도를 높이는데 응용될 수 있다. 설문조사와의 결합 혹은 수업환경에서의 표정해석 등에 실제 활용될 수 있도록 분류방법을 개선하는 것이 향후 과제이다.

참고문헌

- [1] Liao, W.-K., 3D Face Tracking and Expression Inference from a 2D Sequence using Manifold Learning. IEEE Conference on Computer Vision and Pattern Recognition, 2008.
- [2] M. Hamouz, "Ane-invariant face detection and localization using gmm-based feature detector and enhanced appearance model. In Proceedings of the Sixth IEEE international conference on Automatic face and gesture recognition, Washington, DC, USA, IEEE Computer Society, 2004..
- [3] T. F. Cootes, "Active shape models : their training and application." Comput. Vis. Image Underst., 61(1):38{59, 1995.