

고객 맞춤형 서비스를 위한 관객 행동 기반 감정에측모형

유은정

경희대학교 경영대학
(eunjunge@khu.ac.kr)

안현철

국민대학교 경영대학 경영정보학부
(hcahn@kookmin.ac.kr)

김재경

경희대학교 경영대학
(jaek@khu.ac.kr)

정보기술의 비약적 발전에 힘입어, 오늘날 기업들은 지금까지 축적한 고객 데이터를 기반으로 맞춤형 서비스를 제공하는 것에 많은 관심을 가지고 있다. 고객에게 요구하는 맞춤형 서비스를 효과적으로 제공하기 위해서는 우선 그 고객이 처한 상태나 상황을 정확하게 인지하는 것이 중요하다. 특히, 고객에게 서비스가 전달되는 이른바 진실의 순간에 해당 고객의 감정 상태를 정확히 인지할 수 있다면, 기업은 더 양질의 맞춤형 서비스를 제공할 수 있을 것이다. 이와 관련하여 사람의 얼굴과 행동을 이용하여 사람의 감정을 판단하고 개인화 서비스를 제공하기 위한 연구가 활발하게 이루어지고 있다. 얼굴 표정을 통해 사람의 감정을 판단하는 연구는 좀 더 미세하고 확실한 변화를 통해 정확하게 감정을 판단할 수 있지만, 장비와 환경의 제약으로 실제 환경에서 다수의 관객을 대상으로 사용하기에는 다소 어려움이 있다. 이에 본 연구에서는 Plutchik의 감정 분류 체계를 기반으로 사람들의 행동을 통해 감정을 추론해내는 모형을 개발하는 것을 목표로 한다. 본 연구는 콘텐츠에 의해 유발된 사람들의 감정적인 변화를 사람들의 행동 변화를 통해 판단하고 예측하는 모형을 개발하고, 4가지 감정 별 행동 특징을 추출하여 각 감정에 따라 최적화된 예측 모형을 구축하는 것을 목표로 한다. 모형 구축을 위해 사람들에게 적절한 감정 자극영상을 제공하고 그 신체 반응을 수집하였으며, 사람들의 신체 영역을 나누었다. 특히, 모션캡처 분야에서 널리 쓰이는 차영상 기법을 적용하여 사람들의 제스처를 추출 및 보정하였다. 이후 전처리 과정을 통해 데이터의 타임프레임 셋을 20, 30, 40 프레임의 3가지로 설정하고, 데이터를 학습용, 테스트용, 검증용으로 구분하여 인공신경망 모형을 통해 학습시키고 성과를 평가하였다. 다수의 일반인들을 대상으로 수집된 데이터를 이용하여 제안 모형을 구축하고 평가한 결과, 프레임셋에 따라 예측 성과가 변화함을 알 수 있었다. 감정 별 최적 예측 성과를 보이는 프레임셋을 확인할 수 있었는데, 이는 감정에 따라 감정의 표출 시간이 다르기 때문인 것으로 판단된다. 이는 행동에 기반한 제안된 감정에측모형이 감정에 따라 효과적으로 감정을 예측할 수 있으며, 실제 서비스 환경에서 사용할 수 있는 효과적인 알고리즘이 될 수 있을 것으로 기대할 수 있다.

논문접수일 : 2013년 05월 23일 게재확정일 : 2013년 06월 19일

투고유형 : 학술대회 우수논문 교신저자 : 김재경

1. 서론

오늘날 정보사회를 맞아 정보를 이용해 가치를 창출하는 지식서비스의 중요성이 날로 높아져 가고

있다. 더불어 IT 기술의 발전에 따라 정보의 수집과 활용이 용이해지고, 기업들은 다양한 산업에서 고객 정보를 이용하여 마케팅에 활발히 적용하고 있다. 또한 오늘날 비약적인 경제성장에 힘입은 개인 소득

* 본 연구는 문화체육관광부 및 한국콘텐츠진흥원의 2011년도 콘텐츠산업기술지원사업의 연구결과로 수행되었음.

증가와 여가시간의 확대로 인해 소비자들은 전보다 폭넓은 여가소비 활동을 하게 되었다. 여러 가지 형태의 여가소비 활동 중에서도 소비자들이 가장 쉽게 접근할 수 있는 것은 문화서비스에 대한 참여이다 (You, 2010). 21세기에 들어 기업들은 문화예술 분야를 그들의 상업적 이익에 보다 밀접하게 연계한 형태로 기업이미지 제고 및 시장관리와 마케팅 차원에서 적극적으로 활용하고 있다. 또한 기업들의 보편화된 기술 차별화로 소비자를 유지하거나 유인한다는 것은 그렇게 쉽지 않기 때문에 대기업들로부터 시작된 문화예술을 도구로 한 문화서비스 활동이 중소기업에까지 점차 확대되고 있는 추세에 있다 (Korea Mecenat Association, 2013). 공연, 전시 등의 문화서비스 산업들과 소비자에게 간접적 경험을 제공하는 광고산업의 경우, 기업들이 경쟁시장에서 효과적으로 대처하기 위해 고객 경험을 새로운 마케팅 전략으로 사용하고 있다. 이에 발맞추어 개인의 특성이 담겨 있는 개인 프로파일 정보를 통해 고객에게 새로운 경험을 제공하는 고객 맞춤형 서비스의 필요성 또한 급격하게 부상하고 있다.

이처럼 고객과 콘텐츠 간의 인터랙션을 판단하고 고객의 경험과 만족을 극대화시키기 위해서는 기호, 언어, 행동, 감정 등의 다양한 고객 개인 정보를 활용하여 개인화된 맞춤형 서비스를 제공해야 한다 (Bae et al., 2012). 이러한 고객 맞춤형 서비스를 제공할 수 있는 다양한 관련 연구가 있는데 특히 최근에는 그 중에서도 고객의 감정을 이용하는 감정인식분야가 중요하게 부상하고 있다. 고객에게 콘텐츠나 서비스가 전달되는 시점, 즉 진실의 순간(MoT, moment of truth)에 그 고객이 갖고 있는 감정상태를 정확히 인지할 수 있다면, 해당 기업은 보다 효과적으로 맞춤화된 서비스를 제공할 수 있다.

지금까지 다양한 센서의 신호들을 활용해 지능적으로 고객의 감정을 인식하기 위한 다양한 연구들

이 시도되어 왔다. 특히 이들 중에는 주로 생체 신호(음성, 얼굴표정, 제스처, 뇌파, 맥박 등)들을 이용하여 실험이 이루어졌으며 특히 감정에 따른 변화가 심하고 측정이 쉬운 음성과 얼굴 표정을 이용한 연구들이 많았다(Joo et al., 2007). 그러나, 실제 고객 경험이 이루어지는 환경에서는 얼굴이나 음성을 이용하여 고객의 감정을 측정하기에 장비와 주변 환경의 제약으로 인한 몇 가지 어려움이 있다. 따라서 본 연구에서는 기존 연구의 제약을 극복하기 위해 시각 기반 인터페이스를 이용하여 관객 행동 기반의 감정 예측 모델을 개발하고자 한다. 관객 행동 기반의 감정 인식 연구도 여러 학자들에 의해 꾸준히 진행되고 있다. Ambady and Rosenthal(1992)는 인간의 의사소통적인 행동을 판단할 때 신체 행동과 얼굴의 움직임을 이용한 방법의 의존도가 다른 방법에 비해 높다고 밝혔으며, Coulson(1992)은 여섯 가지 감정(화남, 억거움, 두려움, 즐거움, 슬픔, 놀람)의 감정적 신체 자세에 귀속된다는 것을 실험을 통해 증명하였다. 또한 그의 연구에서 자세를 통한 인간의 감정 인지는 음성, 얼굴 표정과 마찬가지로 가능하다고 말하고 있으며, Coulson(1992)과 Burgoon et al.(2005)은 감정 별 행동 표현의 상관관계를 밝혀 실험에 사용하였다. 또한 Laban(1968)은 인간의 움직임 요소를 조합하여 8가지 기본 행동과 이에 해당하는 감정을 맵핑하였으며, 연속적인 동작을 통해서 자신의 감정을 표현한다고 하였다.

본 연구는 시각 기반의 차영상 기법을 이용하여 콘텐츠에 의해 유발된 사람들의 감정적인 변화와 교류를 사람들의 신체행동 변화를 통해 인지하고 예측하는 모형을 개발하고, 4가지 감정 별 행동 특징을 추출하여 각 감정에 따라 최적화된 예측 모형을 구축하는 것을 목표로 한다. 본 연구의 제안 모형은 콘텐츠에 의해 유발된 4가지 감정을 사람의 행동 변화를 통해 자동적으로 판단하는 것을 목적으로

설계되었으며, 모형 구축을 위해 한국콘텐츠진흥원 1층 로비에 부스를 설치하여 다양한 사람들에게 적절한 자극영상을 제공하고 그 신체 반응을 수집하였다. 또한 실험대상자를 9개 영역으로 신체를 구분하고 행동 및 움직임 데이터를 추출, 보정한 이후 인공지능망 모형을 통해 학습시켰다. 4가지 감정은 각각 20, 30, 40프레임의 동적 데이터를 입력변수로 사용하였으며 그 중 가장 성과가 뛰어난 모형을 채택하였다. 이를 통해 각 감정 별 감정 예측 모형의 성과 차이를 규명하고 감정 별로 최적화된 예측 모형을 제시하고자 한다.

이후 본 논문의 구성은 다음과 같다. 제 2장에서는 본 논문의 이론적 배경이 되는 관련 연구를 탐색한다. 다음 제 3장에서는 본 연구에서 제안하는 연구모형을 살펴보고, 이어 제 4장에서는 수집된 데이터를 통해 제안 모형에 기반하는 실험을 수행한 후, 감정 별 최적화 모형을 구축하고 성능을 평가한다. 마지막으로 제 5장에서는 본 연구의 결론과 함께 한계점과 향후 연구방향에 대하여 논의한다.

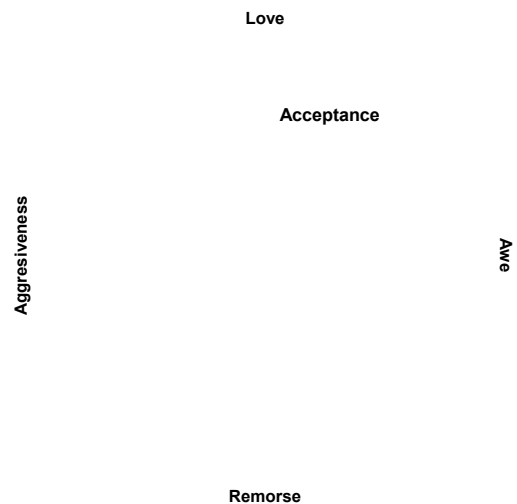
2. 이론적 배경

2.1 감정 분류 연구

감정이란 감성, 정서, 느낌, 기분, 분위기 등을 모두 포괄하는 개념으로서, 감정이라는 개념을 효과적으로 다루기 위해서는 적절한 분류 기준이 필요하다(Kim et al., 2012). 감정에 대한 분류 체계는 크게 Watson, Ekman, Plutchik의 연구가 있다(Lee, 2011). Watson은 감정은 인간에게 가장 기본적인 요소로서 두려움, 분노, 사랑의 3가지 감정으로 분류하였으며, Ekman and Friesen(1975)은 인간의 6가지 기본 감정이 놀람(surprise), 공포(fear), 혐오(disgust), 화남(anger), 행복(happiness), 슬픔(sadness)

이라고 분류하였다. 또한 이 6가지의 기본 감정들과 이들의 조합을 통해 인간의 감정을 표현할 수 있다고 주장하였다.

Plutchik(1980)의 인간의 감정을 8가지 기본 감정으로 분류한 분류 체계는 감정을 이용한 광고나 마케팅에 가장 많이 사용되는 이론이다. Plutchik(1980)은 인간과 동물이 생존에 필요한 여러 가지 적응 행동(adaptive behavior)을 하도록 만드는 8가지의 기본 감정(primary emotion)을 수용(acceptance), 두려움(fear), 놀람(surprise), 슬픔(sadness), 거부(disgust), 분노(anger), 기대(anticipation), 즐거움(joy)이라 밝혔으며, Ekman과 마찬가지로 이러한 기본 감정들이 혼합되어 여러 가지 혼합 감정을 이룬다고 주장하며 그 방법에 대한 연구를 하였다. 나아가 Plutchik(2002)은 8가지 기본 감정을 원반상으로 분류하며, 양극에 위치한 양극성 감정은 제외하고, 원반상에서 이웃한 두 감정이 혼합되어 다양한 혼합 감정을 파생한다고 주장하였다. 이 정서 원반은 수용(acceptance), 분노(anger), 슬픔(sadness)의 3개의 감정을 일차적 정서로 사용하고 이를 이용



<Figure 1> Plutchik's Wheel of Emotion

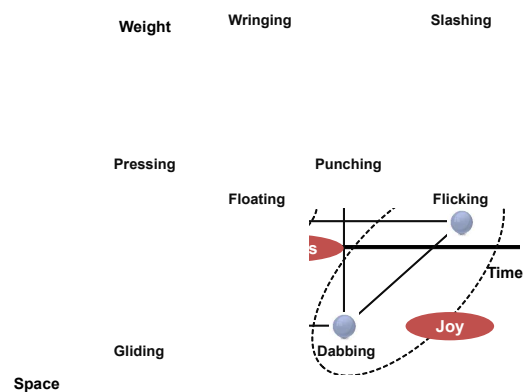
하여 다른 감정 단어들을 검색하여 146개의 정서단어로 그들간의 유사성을 판단하고, 유사성이 높은 단어들을 비슷한 위치에 배열하고, 유사성이 낮은 단어들의 단어들은 반대쪽에 배열하고 있다. 또한 그는 진화라는 관점에서 많은 동물들에게 공통적으로 나타나는 기본적 행동패턴을 8종류로 정리하여 감정의 출현이 이들의 행동 패턴과 직접 관련되어 있음을 확인하였다.

2.2 동작 인식을 통한 감정 인식 연구

기존의 시각 기반 감정 인식 기법들은 주로 영상을 기반으로 한 연구들이다. 초기 연구에서는 3차원 모델 기반의 센서와 마커를 이용한 동작 인식 기법의 연구가 주로 진행되었다. 이 방법은 인간의 신체에 센서를 부착하여 움직임 정보를 얻어낸 후 동작에 관한 데이터를 분석하여 인식에 이용하는 방법이다(Kim et al., 2009). 이러한 모션 캡처 시스템은 센서와 인식 장치의 작동 방식 또는 원리에 따라 음향식(acoustic), 자기식(magnetic), 기계식(mechanical), 광학식(optical) 등으로 나눌 수 있으나(Lee and Park, 1999), 인식 장비들이 대부분 고가이며, 제한적 환경에서만 응용이 가능하다는 한계점이 있다(Kim et al., 2009). 다른 연구들을 살펴보면, 시각 기반의 영상정보를 이용한 동작 인식 연구들 또한 활발하게 이루어지고 있다. 시각 기반의 영상 정보를 이용한 동작 인식 방법은 비 접촉식 인식 방법으로 인간의 신체에 어떤 장치도 부착하지 않고 카메라를 통해 얻어진 영상을 이용하여 외관의 특징을 기반으로 제스처를 인식한다(Dempster et al., 1977; Davis, 1996; Watanabe and Yachida, 1998; Zhu et al., 1998).

Ambady and Rosenthal(1992)은 인간의 의사소통적인 행동을 판단할 때 신체 행동과 얼굴의 움직임을 이용한 방법의 의존도가 다른 방법에 비해 높

다고 밝혔으며, Laban(1968)은 인간의 움직임 요소를 조합하여 8가지 기본 행동과 이에 해당하는 감정을 맵핑하고, 연속적인 동작을 통해서 자신의 감정을 표현한다고 하였다.



<Figure 2> Laban's 8 Basic Movements and their Mappings to Emotions

Coulson(1992)과 Burgoon et al.(2005)은 긴장, 화남, 역겨움, 무서움, 즐거움, 불안함의 6가지 감정으로 인간의 감정을 분류하고 감정을 표현하는 감정 별 행동 특징을 규정하고 데이터베이스를 구축하였고, 나아가 Coulson(1992)은 여섯 가지 감정(화남, 역겨움, 두려움, 즐거움, 슬픔, 놀람)의 감정적 신체 자세에 귀속된다는 것을 실험을 통해 증명하였다. 또한 그의 연구에서 자세를 통한 인간의 감정 인지는 음성, 얼굴 표정과 마찬가지로 가능하다고 말하고 있으며, Coulson(1992)과 Burgoon et al.(2005)은 감정 별 행동 표현의 상관관계를 밝혀 실험에 사용하였다. 국내에서는 Cho et al.(2009)이 자세(posture)와 제스처(gesture)를 이용하여 캐릭터의 기분과 현재 감정을 표현할 수 있는 시스템을 제안하였으며, 기분은 자세로, 세부적 감정은 제스처를

이용하여 복잡한 감정을 표현할 수 있다고 주장하였다.

2.3 인공신경망(Neural Network)

McCulloch and Pitts(1943)에 의해 최초로 제안된 인공신경망은 인공지능의 한 분야로써, 인간 두뇌의 생물학적 작동 형태를 모방한 개념으로 마디(node)와 고리(link)로 구성된 망 구조를 모형화하고, 과거에 수집된 데이터로부터 반복적인 학습과정을 거쳐 데이터에 내재되어 있는 패턴을 찾아내는 모델링 기법이다(Jang, 1999). 인공신경망 기법은 매우 복잡한 구조를 가진 데이터들 사이의 관계나 패턴을 찾아내는 유연한 비선형모형(flexible nonlinear model)의 하나로 패턴 및 문자 인식, 음성 분석, 기상예측, 로봇공학, 수자원 및 환경공학 등 다양한 분야에 적용된다(Jung and Kim, 2012). 경영분야의 문제해결에 있어서도 인공신경망은 고객의 신용평가, 불량거래의 색출, 의료진단 예측, 우량고객 선정, 타겟 마케팅 등 광범위한 분야에서 현재 활용되고 있다. 또한, 시계열 데이터의 패턴을 분류

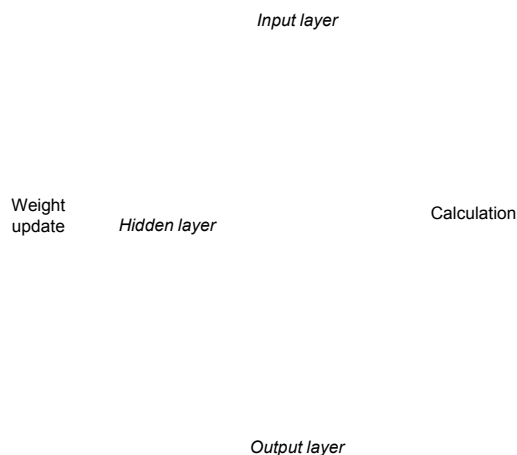
하고 매칭하는 연구에도 많이 적용되어 왔다(Ryoo et al., 2012) 일반적으로 인공신경망은 <Figure 3>과 같이 입력층(input Layer)과 출력층(output Layer) 사이에 한 개 이상의 은닉층(hidden layer)이 존재하는 형태로 구성된다.

이러한 인공신경망을 학습하기 위한 방법으로는 역전파 알고리즘(back-propagation algorithm)이 가장 널리 활용되는데, 이는 목표로 하는 결과값과 모델에 의해 계산된 결과값의 차이를 학습을 통해 계속 조정해 나감으로써 오차값을 최소화하도록 학습시키고 학습결과를 이용하여 예측 대상에 검증을 실시하는 방법이다(Kim, 1993). 인공신경망을 이용하여 감정 인식을 시도한 국내 연구로는 Hong and Woo(2005)의 연구가 있다. 이들은 사용자의 동적인 제스처를 이용하여 감정을 분석하기 위해 인공신경망을 이용하여 감정을 분석하였다.

본 연구에서는 동적인 제스처를 분석하기 위해서는 현재의 동작 정보뿐만 아니라, 이전의 움직임 정보를 고려해야 하기 때문에 이러한 시간적 정보를 효과적으로 사용하여 감정을 분류할 수 있는 인공신경망을 이용하여 모형을 구축하였다.

3. 연구 모형

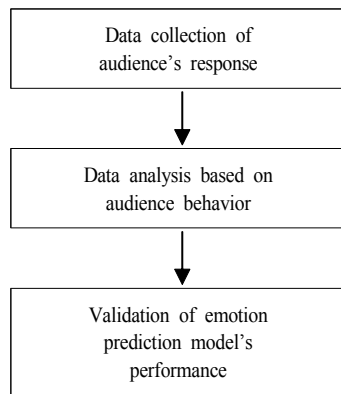
본 연구는 감정 자극에 의한 사람들, 즉 소비자의 반응을 이용하여 고객의 감정을 인식하기 위한 자동화 감정 인식 모형을 개발하는 것을 목표로 한다. 본 제안 모형의 개발을 통해, 기업과 서비스 기획자는 고객 개개인에 맞춤형된 개인화 서비스를 제공하여 고객의 경험과 만족을 극대화시키고, 맞춤형 콘텐츠를 제공함으로써 이를 통해 기업과 상품 및 서비스의 이미지 제고, 매출 증대 등의 효과를 기대할 수 있다. 감정인식모형의 개발에 관련된 연구는 지금까지 많은 연구가 진행되어 왔지만, 기존 연구에



<Figure 3> Structure of a Standard Neural Network

서는 고가의 장비와 장비 사양에 따른 물리적 제약과 고객의 동적인 움직임으로 인한 데이터 수집 범위 및 환경의 제한이 있는 문제점이 있었다. 본 연구에서는 이러한 기존 연구의 한계점을 극복하고자 (1) 실제 고객 체험 환경에 적용될 수 있는 실시간 감정 인식모형을 개발하고, (2) 다양한 감정 별 최적화 감정인식모형을 개발한다는 특징을 지니고 있다.

본 연구의 제안 모형은 <Figure 4>와 같이 세 단계에 의해 구축되었다.



<Figure 4> Steps of the Proposed Research Model

Step 1 : Data Collection

첫 번째 단계는 모형 학습에 필요한 데이터를 수집하는 단계이다. 이 단계에서는 사람들의 행동 및 움직임 데이터를 수집하기 위해, 적절한 감정 자극 영상을 제공했을 때 나타나는 사람들의 움직임을 무비카메라를 이용해 촬영하고 데이터베이스를 구축한다. 또한, 실험대상자들이 느낀 실제 감정을 설문조사를 통하여 수집하는 작업도 이 단계에서 진행된다.

Step 2 : Data analysis based on difference image method

두 번째 단계는 앞 단계에서 수집된 데이터를 이

용하여 실험대상자의 제스처와 포스처 등의 움직임을 추출하고, 그 값을 정규화하여 보정한다. 실험대상자의 움직임을 자세히 측정하기 위해 신체 영역은 얼굴(머리), 왼팔, 왼어깨, 오른어깨, 오른팔, 왼손 1, 왼손 2, 오른손 1, 오른손 2의 총 9가지 영역으로 나뉘어 진다. 각 영역별 움직임은 차영상 기법을 이용하여 그 움직임을 판별하고 그 값을 추출하게 되며 움직임 데이터는 0~1사이의 값으로 정규화 된다. 각 영역의 값은 모형의 입력변수로 사용되고, 각 자극 별 감정이 종속변수로 사용된다.



<Figure 5> Difference Image Analyzer

Step 3 : Emotion prediction model based on audience behavior

세 번째 단계에서는 이전 단계에서 수집된 관객 행동 및 움직임 데이터들을 학습용, 테스트용, 검증용 셋으로 분류하고 모형을 구축하는 단계이다. 이 단계에서는 자극 영상이 유발한 슬픔(sad), 놀람(surprised), 즐거움(joy), 혐오(disgust)의 4가지 감정 별 모형을 구축하고 시계열 구간(프레임셋)의 변화에 따라 감정 별로 성과가 어떻게 변하는가를 판단한다. 즉, 감정에 따라 신체 행동 변화를 추출하고 각 감정 별 최적 예측 모형을 구축한 후 감정 별 모형의 성과를 비교한다.

4. 실증분석

4.1 데이터 수집

본 연구에서는 일반 대중을 대상으로 데이터를 수집하였다. 이 데이터는 2012 Seoul DMC Culture Open 행사에서 이벤트 부스를 설치하여 2012년 9월 13일부터 15일까지 수집된 데이터로서, 위 행사에 참가한 일반 참가자가 자의에 의해 참여하여 준비된 감정 별 자극 영상을 보고 반응하는 영상을 촬영하고, 영상 종료 후 느꼈던 감정을 설문조사를 통해 수집하였다. 감정 자극 영상은 Adobe Premiere를 이용해 편집하였으며, 상영을 위해 고성능 프로젝터를 설치하였으며 정확한 관객 영상 촬영을 위해 외벽을 설치하고 고해상도의 카메라를 이용해 관객들의 영상을 촬영하였다. 3일간 확보된 데이터 중 모형 구축에 불량한 데이터를 제외하고 총 97개의 데이터가 수집되었으며, 슬픔, 놀람, 즐거움, 혐오의 4가지 감정 상태에 대해 촬영되었다.

각 감정 상태의 실험 대상자 영상은 신체의 9가지 영역 별로 움직임 값이 추출되었으며, 이 값은 1초당 10프레임의 간격으로 추출하였다. 또한 각 감정 별로 각각 2초(20프레임), 3초(30프레임), 4초(40프레임)의 움직임 값을 각각 입력변수로 설정하여 세 가지 모델을 구축하였다. 프레임 별 감정 예측 모형을 구축하는 것은 다양한 감정에 따라 표출하는 감정 행동의 크기와 지속시간이 다르기 때문이다. 예를 들어 놀람의 경우 짧고 강하게 감정 행동이 표출되지만, 혐오의 경우 놀람보다 비교적 느리고 천천히 감정 행동이 표출된다.

4.2 실험설계

본 연구에서는 인공지능망 모형 구축에 앞서 전체 데이터를 아래와 같이 학습용, 테스트용, 그리고

검증용 데이터셋으로 구분하였다. 학습용 데이터 집합은 인공지능망 기반 관객 감정 판단 모형을 학습하는데 사용되며, 테스트 데이터 집합은 과적합(overfitting)을 방지하기 위한 학습 중지점 탐색에 사용한다. 끝으로, 검증용 데이터 집합으로 구축된 모형의 성과를 검증한다.

<Table 1> Experimental Data Sets

Data set	Ratio	Purpose
Training	60%	To train ANN models
Test	20%	To detect optimal stopping point to avoid overfitting
Validation	20%	To validate trained ANN models using hold-out samples
Total	100%	

<Table 2> Number of Samples for Each Data Set

Data set	20 frames	30 frames	40 frames
Training	1164	1746	2328
Test	388	582	776
Validation	388	582	776
Total	1940	2910	3880

감정 예측을 위한 본 제안 모형은 인공지능망 학습 방법으로는 역전파 알고리즘을 사용하였고, 학습률(learning rate)을 10%, 모멘텀률(momentum rate)은 10%로 설정하였다. 변환함수로는 로지스틱(logistic)함수를 사용하였으며, 계층구조는 3계층 퍼셉트론, 즉, 은닉층 수는 1개로 설계하였다. 학습 중지점을 탐색하기 위해서는 테스트 데이터 집합을 기준으로 최소 오류 도달 후 50,000번 도달 시 중지하는 것으로 하였다. 인공지능망 실험을 수행하기 위한 도구로는 상용 소프트웨어인 Neuroshell2 4.0을 사용하였다.

4.3 실험결과

실험을 위하여 총 97개의 데이터를 수집 및 전처리하여 사용하였으며, 데이터는 프레임 별(20, 30, 40프레임)로 구분하여 실험하였다. 프레임 별 감정 인식모형을 각각 구축해 보았는데 다음 <Table 3>~<Table 5>가 각 프레임 별 모형 결과를 정리하여 나타내며, <Table 6>이 전체적으로 결과를 요약하여 제시하고 있다.

<Table 3> Experimental Results (%) - 20 Frames

	Sad	Surprised	Joy	Disgust	Hit Ratio
Sad	100	-	-	-	100
Surprised	17	83	-	-	83
Joy	-	-	96	4	96
Disgust	5	-	5	90	90

<Table 4> Experimental Results (%) - 30 Frames

	Sad	Surprised	Joy	Disgust	Hit Ratio
Sad	65	-	8	27	65
Surprised	12	88	-	-	88
Joy	-	-	92	8	92
Disgust	-	-	2	98	98

<Table 5> Experimental Results (%) - 40 Frames

	Sad	Surprised	Joy	Disgust	Hit Ratio
Sad	84	-	2	14	84
Surprised	13	87	-	-	87
Joy	-	-	92	8	92
Disgust	3	-	2	95	95

<Table 6> Hit Ratio of the Proposed Models(%)

	Sad	Surprised	Joy	Disgust
20F	100	83	96	90
30F	65	88	92	98
40F	84	87	92	95

<Table 6>에서 볼 수 있듯이, 슬픔(Sad)과 즐거움의 예측을 위한 인공신경망 모형은 20프레임의 입력변수를 적용했을 때 가장 우수한 성과를 보이는 것을 확인하였다. 반면, 놀람, 혐오의 예측을 위한 인공신경망 모형은 30프레임의 입력변수를 적용했을 때 가장 우수한 예측 성과를 보이고 있다. 이는 각 감정이 행동으로 표출되는 시간이 감정 별로 다양하며, 감정 행동 특징이 구별되는 구간의 길이가 다르다는 것을 뜻한다. 위 실험 결과 슬픔과 즐거움은 짧은 시간 내에 감정 행동의 특징을 나타낸다는 것을 알 수 있으며, 놀람과 혐오의 경우 감정 행동의 특징이 구별되는 시간이 앞의 두 감정보다 상대적으로 길다는 것을 보여주고 있다. 본 연구의 결과는 다양한 감정에 따라 입력변수가 최적화 되도록 조절할 수 있으며, 성과가 낮은 입력변수를 제거함으로써 예측 속도를 향상시킬 수 있다는 점을 시사한다.

또한 슬픔, 즐거움, 혐오 감정의 경우 예측성과가 우수한 것으로 나타났는데, 놀람의 경우 그 성과가 다른 감정들 보다 다소 낮은 것을 확인할 수 있다. 이는 놀람의 경우 그 감정 행동이 표출되는 시간이 극히 짧아, 놀람 이후의 행동들이 감정 예측에 과다하게 영향을 주어 잡음으로 작용하기 때문으로 생각된다. 즉, 앞서 말했듯이 감정에 따라 감정 행동의 특징과 표출 시간이 다르기 때문이라고 할 수 있다.

5. 결론

본 연구에서는 감정적 자극에 반응한 사람의 행동 및 움직임 변화를 카메라로 촬영하고, 차영상 기법을 이용하여 시간구간 별 동적 데이터 값을 추출하여 사람들의 감정을 자동적으로 예측하는 감정 별 최적 예측 모형을 제안하고 이를 일반인 대상의 실제 데이터를 이용하여 실증분석 하였다. 제안한 모형은 슬픔, 놀람, 즐거움, 혐오의 4가지 감정에 대

해 20프레임 모형에서 슬픔은 100%, 즐거움 96%의 예측정확도를 보였으며, 30프레임 모형에서 놀람 88%, 혐오 98%의 예측정확도를 보였다. 이러한 본 연구의 성과는 추후 광고/전시/공연 등의 다양한 산업에서 실제 고객들을 대상으로 고객 맞춤형 서비스를 제공하기 위한 감정판단모형을 구축할 때, 기존 연구들에 비해 비교적 간단하고 빠르게 모형을 구축할 수 있는 알고리즘이 될 수 있을 것으로 기대한다.

본 연구의 한계점 및 향후 연구방향은 다음과 같다. 첫째, 보다 다양한 감정을 예측할 수 있는 모형 구축에 대한 연구가 필요하다. 현재 본 연구에서는 슬픔, 놀람, 즐거움, 혐오의 4가지 감정만을 대상으로 실험을 진행하였는데, 다른 다양한 감정이 많은 만큼, 다양한 감정에 대한 적용에 대한 연구가 추후 이루어 질 필요가 있다. 둘째, 감정 예측 정확도 향상을 위해 신체 영역별 예측 성과를 비교하는 연구가 필요하다. 현재 본 연구에서는 신체를 9가지 영역으로 나누어 실험을 진행하였으나, 9가지 영역이 예측 성과에 미치는 영향을 규명하지 못하였다. 다양한 감정에 따라 영향이 큰 영역을 규명하여 데이터의 차원축소를 통해 더욱 신속한 계산과 성능을 향상시킬 수 있는 모형을 구축하고 그 성능을 검증해 볼 필요가 있다.

참고문헌

- Ambady, N. and R. Rosenthal, "Thin slices of expressive behavior as predictors of interpersonal consequences : A meta-analysis," *Psychological Bulletin*, Vol.111, No.2(1992), 256~274.
- Bae, J. H., S. H. So, and L. K. Choi, "Designing Mobile Framework for Intelligent Personalized Marketing Service in Interactive Exhibition Space," *Journal of Intelligence and Information Systems*, Vol.18, No.1(2012), 59~69.
- Burgoon, J. K., M. L. Jensen, T. O. Meservy, J. Kruse, and J. F. Nunamaker, "Augmenting human identification of emotional states in video," *Proceedings of the international conference on intelligent data analysis*, 2005.
- Cho, K. E., C. S. Lee, and K. H. Um, "Design of a Character' System to Express Mixed-Emotion by Gesture and Posture," *Proceeding of the conference on Korea Multimedia Society*, (2009), 191~194.
- Coulson, M., "Attributing emotion to static body postures : recognition accuracy, confusions, and viewpoint dependence," *Journal of Non-verbal Behavior*, Vol.28, No.2(1992), 117~139.
- Davis, J. W., *Appearance-Based Motion Recognition of Human Actions*, Technical Report, MIT, 1996.
- Dempster, A. P., N. M. Laird, and D. B. Rubin, "Maximum-likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society, Series B*, Vol.39, No.1 (1977), 1~38.
- Ekman, P. and W. V. Friesen, *Unmasking the face : a guide to recognizing emotions from facial clues*, Englewood Cliffs, NJ : Prentice-Hall, 1975.
- Hong, D. P. and W. T. Woo, "AR-based Tangible User Interface for I2-NEXT," *Proceedings of the conference on HCI 2005*, (2005), 1628~1633.
- Jang, N. S., *Data Mining*, Dae Chung Media, 1999.
- Joo, J. T., I. H. Jang, H. C. Yang, and K. B. Sim, "Emotion Recognition and Expression Method using Bi-Modal Sensor Fusion Algorithm," *Journal of Control, Automation and Systems Engineering*, Vol.13, No.8(2007), 754~759.

- Jung, M. K. and J. K. Kim, "The Intelligent Determination Model of Audience Emotion for Implementing Personalized Exhibition," *Journal of Intelligence and Information Systems*, Vol.18, No.1(2012), 39~57.
- Kim, D. S., *Neural Network Theory and Application*, Hitech information, 1993.
- Kim, J. H., Y. S. Yun, T. Y. Kim, and C. S. Lim, "Human primitive motion recognition based on the hidden markov models," *Journal of Korea Multimedia Society*, Vol.12, No.4(2009), 521 ~ 529.
- Kim, S. J., E. C. Ryoo, M. K. Jung, J. K. Kim, and H. C. Ahn, "Application of support vector regression for improving the performance of the emotion prediction model," *Journal of Intelligence and Information Systems*, Vol.18, No.3(2012), 185~202.
- Korea Mecenat Association, (2013), Available at <http://www.mecenat.or.kr/> (Accessed 14 June, 2013).
- Laban, R., *Modern Educational Dance*, Macdonald and Evans, 1968.
- Lee, I. H. and C. J. Park, "The status of the motion capture technology and Applications," *Journal of Korea Multimedia Society*, Vol.3, No.1(1999), 38~48.
- Lee, S. H., "The Study for Equilibrium of Emotions by Oriental Medicine Music Therapy," *Journal of Oriental Medical Classics*, Vol.24, No.6(2011), 39~47.
- McCulloch, W. S. and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *The Bulletin of Mathematical Biophysics*, Vol.5, No.4(1943), 15~33.
- Plutchik, R., *Emotion : A psycho-evolutionary synthesis*, New York : Harper and Row, 1980.
- Plutchik, R., *Emotions and life : Perspectives from psychology, biology, and evolution*, Washington, DC : American Psychological Association, 2002.
- Ryoo, E. C., S. B. Park, and J. K. Kim, "Audience reaction judgement model for interactive performance," *Proceeding of the Conference on Intelligence and Information Systems*, (2012), 121~126.
- Watanabe, T. and M. Yachida, "Real time recognition of gesture and gesture degree information using multi input image sequences," *Proceeding of the 14th International Conference on Pattern Recognition*, Vol.2(1998), 1855~1858.
- You, H. J., "The re-experience intention affects experiential traits and emotions of consumer on the performances cultural service of companies," *The e-Business Studies*, Vol.11, No.3 (2010), 519~538.
- Zhu, Y., Y. Huang, G. Xu, H. Ren, and Z. Wen, "Vision-based Interpretation of Hand Gestures by Modeling Appearance Changes in Image Sequences," *Proceedings of MVA'98 IAPR Workshop on Machine Vision Applications*, (1998), 573~576.

Abstract

The Audience Behavior-based Emotion Prediction Model for Personalized Service

Eun Chung Ryoo* · Hyunchul Ahn** · Jae Kyeong Kim***

Nowadays, in today's information society, the importance of the knowledge service using the information to creative value is getting higher day by day. In addition, depending on the development of IT technology, it is ease to collect and use information. Also, many companies actively use customer information to marketing in a variety of industries. Into the 21st century, companies have been actively using the culture arts to manage corporate image and marketing closely linked to their commercial interests. But, it is difficult that companies attract or maintain consumer's interest through their technology. For that reason, it is trend to perform cultural activities for tool of differentiation over many firms. Many firms used the customer's experience to new marketing strategy in order to effectively respond to competitive market. Accordingly, it is emerging rapidly that the necessity of personalized service to provide a new experience for people based on the personal profile information that contains the characteristics of the individual. Like this, personalized service using customer's individual profile information such as language, symbols, behavior, and emotions is very important today. Through this, we will be able to judge interaction between people and content and to maximize customer's experience and satisfaction. There are various relative works provide customer-centered service. Specially, emotion recognition research is emerging recently. Existing researches experienced emotion recognition using mostly bio-signal. Most of researches are voice and face studies that have great emotional changes. However, there are several difficulties to predict people's emotion caused by limitation of equipment and service environments. So, in this paper, we develop emotion prediction model based on vision-based interface to overcome existing limitations. Emotion recognition research based on people's gesture and posture has been processed by several researchers. This paper developed a model that recognizes people's emotional states through body gesture and posture using difference image method. And we found

* Graduate School of Business Administration, Kyung Hee University

** School of Management Information Systems, Kookmin University

*** Corresponding Author: Jae Kyeong Kim

Graduate School of Business Administration, Kyung Hee University

1 Hoegi-dong, Dongdaemun-gu, Seoul 130-701, Korea

Tel: +82-2-961-0508, Fax: +82-2-961-9355, E-mail: jaek@khu.ac.kr

optimization validation model for four kinds of emotions' prediction. A proposed model purposed to automatically determine and predict 4 human emotions (Sadness, Surprise, Joy, and Disgust). To build up the model, event booth was installed in the KOCCA's lobby and we provided some proper stimulative movie to collect their body gesture and posture as the change of emotions. And then, we extracted body movements using difference image method. And we revised people data to build proposed model through neural network. The proposed model for emotion prediction used 3 type time-frame sets (20 frames, 30 frames, and 40 frames). And then, we adopted the model which has best performance compared with other models.' Before build three kinds of models, the entire 97 data set were divided into three data sets of learning, test, and validation set. The proposed model for emotion prediction was constructed using artificial neural network. In this paper, we used the back-propagation algorithm as a learning method, and set learning rate to 10%, momentum rate to 10%. The sigmoid function was used as the transform function. And we designed a three-layer perceptron neural network with one hidden layer and four output nodes. Based on the test data set, the learning for this research model was stopped when it reaches 50000 after reaching the minimum error in order to explore the point of learning. We finally processed each model's accuracy and found best model to predict each emotions. The result showed prediction accuracy 100% from sadness, and 96% from joy prediction in 20 frames set model. And 88% from surprise, and 98% from disgust in 30 frames set model. The findings of our research are expected to be useful to provide effective algorithm for personalized service in various industries such as advertisement, exhibition, performance, etc.

Key Words : Emotion Prediction Model, Audience Experience, Gesture Recognition, Neural Network, Difference Image Method

저 자 소 개



유은정

경희대학교 경영학 학사를 취득하였으며, 현재 동 대학 경영학과 일반대학원 MIS 전공 석사 과정에 재학 중이다. 주요 관심분야는 데이터마이닝, 비즈니스 인텔리전스 분야의 인공지능, 정보시스템 기반의 행동 모형 등이다.



안현철

현재 국민대학교 경영대학 경영정보학부 조교수로 재직 중이다. KAIST에서 산업경영 학사를 취득하고, KAIST 테크노경영대학원에서 경영정보시스템을 전공하여 공학석사와 박사를 취득하였다. 주요 관심분야는 금융, 고객관계관리 및 인터넷 보안 분야의 인공지능 응용, 정보시스템 수용과 관련한 행동 모형 등이다.



김재경

서울대학교에서 산업공학 학사, 한국과학기술원에서 경영정보시스템 전공으로 석사 및 박사학위를 취득하였으며 현재 경희대학교 경영대학 교수로 재직하고 있다. 미국 미네소타 주립대학교, 그리고 텍사스 주립대학교(달라스)에서 교환교수를 역임하였다. 주요 관심분야로는 비즈니스 인텔리전스, 추천시스템, 유비쿼터스 서비스 등이다.

IEEE Transactions on Services Computing, IEEE Transactions on SMC-A, International Journal of Human-Computer Studies, International Journal of Information Management, Technological Forecasting and Social Change, Information and Management 등 다수의 학술지에 논문을 게재하였으며, 또한 한국지능정보시스템학회장, 국가과학기술위원회 서비스 R&D 전문위원, 경희대학교 경영대학 BK21 사업단장, Information Technology and Management(SSCI) AE(Associate Editor)를 역임하였다.