

# 개인화 전시 서비스 구현을 위한 지능형 관객 감정 판단 모형

정민규  
경희대학교 경영대학 경영학부  
(minkyuli@khu.ac.kr)

김재경  
경희대학교 경영대학 경영학부  
(jaek@khu.ac.kr)

.....

최근 기존 전시 공간 내에 유비쿼터스 환경이 구축되면서, 관객과의 상호작용을 통해 전시 효과를 배가할 수 있는 인터랙티브 전시에 많은 사람들의 관심이 집중되고 있다. 이러한 인터랙티브 전시가 보다 고도화되기 위해서는 전시물에 대한 다양한 관객 반응을 측정하고, 이를 통해 대상 관객이 어떤 감정을 느끼는지 예측할 수 있는 적절한 의사결정지원 모형이 요구된다. 이러한 배경에서 본 연구는 인터랙티브 전시 공간 내에서 수집 가능한 다양한 관객 반응 중 얼굴표정의 변화를 이용하여, 관객의 감정을 추론, 판단하는 지능형 모형을 제시한다. 본 연구에서 제시하는 모형은 무자극 상태의 관객의 표정과 자극이 주어졌을 때 관객의 표정이 어떻게 변화하는지 변화량을 측정하여, 이를 기반으로 인공지능기법을 이용해 해당 관객의 감정을 판단하는 모형이다. 이 때, 제안모형의 감정 분류체계로는 간결하면서도 실무에 적용이 용이하여 그간 기존 문헌에서 널리 활용되어 온 매력-각성(Valence-Arousal) 모형을 사용한다. 제안모형의 유용성을 검증하기 위해, 본 연구에서는 2011 서울 DMC 컬처 오픈 행사에 참여하여, 일반인을 대상으로 얼굴 표정 변화 데이터를 수집하고, 이들이 느끼는 감정 상태를 설문조사하였다. 그리고 나서, 이 자료들을 대상으로 본 연구에서 제안하는 모형을 적용해 보고, 제안모형이 비교모형으로 설정된 통계기반 예측모형에 비해 더 우수한 성과를 보이는지 확인해 보았다. 실험 결과, 본 연구에서 제시하는 모형이 비교 모형인 중회귀분석 모형보다 더 우수한 결과를 제공함을 확인할 수 있었다. 본 연구를 통하여 구축된 관객 감정 판단 모형을 실제 전시장에서 활용한다면 전시물을 관람하는 관객의 반응에 따라 시의적절하면서도 효과적인 대응이 가능하기 때문에, 관객의 몰입과 만족을 보다 증대시킬 수 있을 것으로 기대된다.

.....

논문접수일 : 2012년 03월 02일    논문수정일 : 2012년 03월 14일    게재확정일 : 2012년 03월 16일  
투고유형 : 국문급행    교신저자 : 김재경

## 1. 서론

1960년대 이후 관람객이 예술작품의 주체로서 새롭게 등장하게 됨에 따라, 현대의 전시/공연은 예술 작품의 보존과 전시 관람 장소라는 기존의 정적인 개념에서 벗어나 작품과 관람객의 상호작용을 중시하는 개념으로 변화하고 있다(김미연 등,

2008; 송현미 등, 2005; 유재엽, 2004). 이러한 경향에 따라, 최근 관객과 작품이 상호작용을 하고 관객에게 다양한 예술적 기회와 체험을 제공하는 ‘인터랙티브 전시(interactive exhibition)’가 전시 분야의 주요한 화두로 떠오르고 있다(김미연 등, 2008; 임승희 등, 2009; 최지영 등, 2010). 일명 체험 전시라고도 불리는 인터랙티브 전시는 전시물을 관람

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

자가 직접 체험해 봄으로써 시각에 의해서 얻는 정보뿐만 아니라 모든 감각기관을 통해 전시물의 무거움, 질감, 온도, 미각 등을 직접적으로 감지하여 많은 정보를 얻을 수 있는 새로운 전시기법을 말한다(김철근 등, 1996). 인터랙티브 전시 공간 내에서는 전시물과 관람자의 다양한 커뮤니케이션이 이루어지는데, 이러한 커뮤니케이션에서는 인간의 지각과 감각에서의 사고적 확장이라는 측면에서 보기(seeing), 시간(time), 공간(space) 등의 여러 구성요소의 연출표현능력이 요구된다(김미연 등, 2008; 유재엽, 2004). 특히, 인터랙티브 전시를 기획하는 기획자에게 있어, 관객들이 보다 집중하여 전시를 관람하고, 몰입을 할 수 있도록 하기 위한 개인 맞춤형 전시 기획과 환경 조성을 위한 시도는 매우 중요하다(김규정 등, 2008; 김미연 등, 2008; 임승희 등, 2009).

그런데, 전시장 내에서의 관객에게 맞춤형 개인화 서비스를 제공하기 위해서는 전시장에서 제공되는 전시물, 즉 콘텐츠에 대한 관객의 반응을 이해하는 것이 우선적으로 필요하다. 즉, 전시 콘텐츠에 대한 관람객의 반응을 이해하고, 그에 따라 적절하게 대응책을 제공하는 것은 수준 높은 인터랙티브 전시를 설계, 구현하는데 있어 매우 중요한 요소라 할 수 있다. 지금까지 관객의 얼굴 표정, 소리/음성, 몸짓 등 여러 가지 반응을 보고, 그들의 반응으로부터 대상 관객의 감정 상태를 예측하고자 시도한 연구는 이전부터 계속적으로 다양하게 진행되어 왔다. 하지만, 기존 연구에는 다음과 같은 몇 가지 공통적인 한계점이 존재한다. 첫째, 얼굴 표정이 아닌 소리나 몸짓을 이용해 파악하려는 경우, 전시의 내용이나 상황에 따라 아예 적용이 불가능한 경우가 많다. 예를 들어, 소리를 이용해 관객의 감정 상태를 파악하려는 경우에 조용한 전시 성향을 지닌 전시회에서는 적용이 불가능하다.

둘째, 사람의 감정을 예측하는 연구들은 대체로 Ekman이 제시한 6가지 감정 분류 체계(놀람, 공포, 혐오, 화남, 행복, 슬픔)를 활용하였는데, 이들 감정 중 행복 하나만 빼고는 모두 부정적인 감정들뿐이라 실제 활용에 어려움이 존재한다(Brainerd et al., 2010; Nicolaou et al., 2011; Ekman, 1972). 마지막으로 셋째, 사용자의 얼굴표정을 대상으로 감정을 추론하는 연구들 대부분이 실험실 환경에서 얻어진 소수의 데이터를 기반으로 실험이 이루어지거나, JAFFE DB같은 외국에서 만들어진 공개 데이터베이스를 기반으로 실험을 수행하였다는 한계점을 지니고 있다.

상기 지적한 기존 연구들의 한계점을 보완하고자, 본 연구에서는 수집 가능한 다양한 관객 반응 중에서 얼굴표정이 변화하는 데이터를 이용하여 관객이 전시물에 대해 어떤 감정을 느끼는지를 판단하는 새로운 모형을 제시하고자 하였다. 이 때, 감정에 대한 분류체계는 기존 연구들이 주로 사용해 온 Ekman의 6가지 분류체계가 아닌 매력-각성 (Valence-Arousal, 이하 V-A) 모형을 적용함으로써, 판단된 감정을 토대로 적절한 전시 서비스의 변화가 이루어질 수 있도록 하였다. 아울러, 관객의 감정을 판단하기 위한 모형을 학습하기 위한 기법으로는 인공지능망 기법, 그 중에서도 역전파 알고리즘을 제안한다.

이후 본 논문은 다음과 같이 구성된다. 다음 제 2장에서는 본 논문의 제안모형을 구성하는 다양한 요소들과 관련해 이론적 배경을 제시하고, 제 3장에서는 본 연구가 제안하는 연구모형과 그 절차를 설명한다. 이어, 제 4장에서는 수집된 실제 데이터를 이용하여, 제안모형의 유용성을 실증분석하기 위한 실험설계에 대해 설명하며, 제 5장에서는 실증 분석을 통한 결과를 제시한다. 끝으로, 마지막 장에서 본 연구의 결론 및 한계점에 대해서 논의한다.

## 2. 이론적 배경

### 2.1 인터랙티브 전시 관련 연구

인터랙티브란 컴퓨터 공학 분야에서는 ‘대화식’, 통신 분야에서는 ‘양방향 통신’이라는 의미로 해석되며, 인터랙티브 시스템은 사용자가 마치 컴퓨터와 대화를 하듯이 쌍방향으로 입력과 출력을 할 수 있는 시스템을 나타낸다(김규정 등, 2008). 이러한 인터랙티브를 전시 분야에 응용한 인터랙티브 전시는 전시물과 관객 간의 상호작용을 기반으로 관객이 작품을 스스로 선택하고 직접 손으로 만지고 조작(hands-on)하고 관찰하면서 작품의 내용을 창의적으로 이해하고 경험하게 되는 전시 방법이다(김철근 등, 1996; 최지영 등, 2010). 인터랙티브 전시에서는 계속적으로 전시물과 관람자의 다양한 커뮤니케이션이 이루어지기 때문에, 기존의 관람 형태의 전시와는 전시물의 기획, 주변 환경 조성 등 많은 측면에서 서로 차별화된다. 때문에, 관람 전시가 아닌 인터랙티브 전시를 대상으로 어떻게 하면 관람객의 참여와 만족도를 증가시키고 긍정적인 정서를 함양 할 수 있도록 할 수 있는지에 대한 연구가 최근 학계에서 활발하게 이루어지고 있다(임승희 등, 2009; 최지영 등, 2010).

임승희 등(2009)은 몰입 이론(flow theory)을 인터랙티브 전시에 적용하는 방안을 제시하였다. 이 연구에서는 몰입 이론에 기반하여 전시콘텐츠를 제작하고, 관람객의 몰입도를 분석하기 위해 경험 표집법(Experience Sampling Method : ESM)을 사용하여 주관적 경험을 측정하였다. 한편 김미연 등(2008)은 인터랙티브 공간의 사례 조사를 통하여, 전시 콘텐츠와 체험 방식에 대해 분석하고, 인터랙티브 전시공간의 체험 방식의 분류에 대한 접근과 분석과정을 통해, 체험을 디자인하기 위한 전시공간의 공간적 고려사항을 도출하였다. 최지영 등

(2010)은 어린이를 위한 체험 전시의 개념을 분석하고, 어린이 체험 전시에서 디지털 미디어 기술을 활용한 체험 전시를 구현함으로써, 어린이가 적극적으로 작품에 개입할 수 있고 창의적으로 체험을 학습 할 수 있는 전시에 대한 실례를 제시하였다.

하지만, 이와 같은 인터랙티브 전시 관련 기존 연구들에는 몇 가지 한계점이 존재한다. 첫째, 인터랙티브 전시 공간 내에서 관객이 전시물을 관람하는 중 해당 전시물에 대한 관객의 감정, 그 중에서도 특히 몰입도를 직접 측정하려는 연구가 거의 전무하다. 둘째, 일부 몰입도 측정에 관련된 연구가 존재하긴 하지만, 유비쿼터스 공간 구성의 어려움으로 인해 실시간으로 수집되는 데이터를 이용하지 못하는 한계점이 존재한다. 이에, 본 연구에서는 기존 논문들이 지닌 한계점을 보완하고자 인터랙티브 전시 공간 내에서 관객이 전시물을 관람하는 중 실시간으로 수집되는 데이터를 이용하여 전시물에 대한 관객의 몰입의 정도가 어떠한지 판단할 수 있는 모형을 제시하고자 한다.

### 2.2 감정 분류 관련 연구

관객의 감정을 분류하고, 체계화하기 위한 시도는 Ekman 모형으로부터 시작하여 V-A 모형에 이르기까지 다양한 연구들에서 이루어져 왔다. 사용자의 반응을 이용하여 감정을 판단하는 기존 연구들은 대체로 P. Ekman의 6가지 감정 분류 체계를 활용하였다(<표 1> 참고). <표 1>과 같이 Ekman 모형은 인간의 감정을 총 6가지로 분류하는데, 각 감정별로 어떤 표정의 변화가 야기되는지를 제시하였다는 장점이 있다. 때문에 이러한 Ekman의 감정 분류 체계는 얼굴표정을 통해 감정을 판단하려고 한 많은 기존 연구들에 의해 활발히 적용되어 왔다. 하지만, 이 감정 분류 체계는 행복감 하나를 제외하고는 모두 부정적인 감정들뿐이라, 감정

&lt;표 1&gt; P. Ekman의 감정 모형

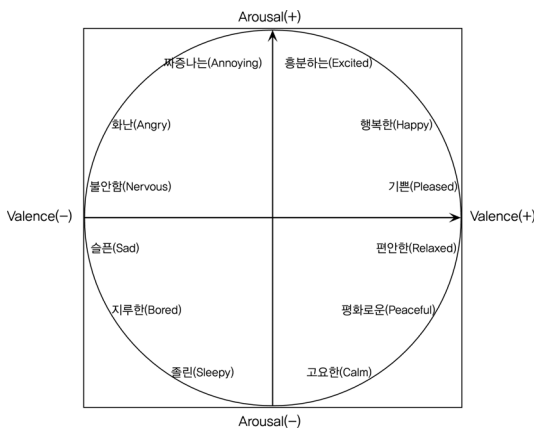
감정상태	얼굴 표정의 변화
놀람(Surprise)	<ul style="list-style-type: none"> <li>- 눈썹이 상승하여 높게 올라가고 구부러짐.</li> <li>- 눈썹 밑의 피부가 당겨짐.</li> <li>- 이마에 수평적 주름(Horizontal wrinkles)이 생겨남.</li> <li>- 눈꺼풀이 활짝 열림; 위쪽 눈꺼풀은 올라가고 아래쪽 눈꺼풀은 아래로 내려감; 눈의 공막이 홍채의 위에 보여지고, 종종 아래로 내려가기도 함.</li> <li>- 턱이 내려가고 이에 따라 입과 이가 서로 벌어짐, 하지만 입의 근육의 변화는 없음.</li> </ul>
공포(Fear)	<ul style="list-style-type: none"> <li>- 눈썹이 올라가고 내려가는 것이 함께 일어남.</li> <li>- 이마의 주름은 가운데 모이고, 이마 전체로 주름지지 않음.</li> <li>- 위쪽 눈꺼풀이 올라가고, 공막이 노출되고, 아래쪽 눈꺼풀이 아래로 당겨짐.</li> <li>- 입은 벌어지고 입술은 약간 당겨지며 내려가거나, 늘어지며 내려감.</li> </ul>
혐오(Disgust)	<ul style="list-style-type: none"> <li>- 윗입술이 상승.</li> <li>- 아랫입술도 같이 올라가고 위쪽 입술의 상승을 지지하거나, 내려가서 약간 돌출됨.</li> <li>- 코가 주름짐.</li> <li>- 뺨이 상승.</li> <li>- 아래 눈꺼풀에서 선이 생기며, 눈꺼풀이 상승하지만 긴장되지는 않음.</li> <li>- 눈썹이 내려가고, 위 눈꺼풀도 내려감.</li> </ul>
화남(Anger)	<ul style="list-style-type: none"> <li>- 눈썹이 내려감.</li> <li>- 양 눈썹 사이에 세로 주름이 생김.</li> <li>- 아래 눈꺼풀이 당겨지고 상승하기도 함.</li> <li>- 아래 눈꺼풀이 당겨지고 눈썹에 의해서 내려가기도 함.</li> <li>- 양 눈이 오래 응시 하며, 약간 돌출되기도 함.</li> <li>- 입술은 두 가지 기본 위치를 보임; 굳게 다물어 양 입가가 일자이거나 내려감 또는 소리친다면 사각형 모양으로 긴장됨.</li> <li>- 콧구멍이 벌어지지만, 화남의 감정에서 필수적인 얼굴 표정은 아니며 이는 슬픔의 감정에서도 같은 움직임을 보임.</li> </ul>
행복(Happiness)	<ul style="list-style-type: none"> <li>- 입가가 뒤로 말리며 올라감.</li> <li>- 입은 벌어질 수도 있으며, 이가 보일 수도 있음.</li> <li>- 코에서부터 입가의 가장자리까지 주름이 생성.</li> <li>- 뺨은 상승.</li> <li>- 아래 눈꺼풀 아래에서 주름이 생기며, 상승하지만 긴장되지는 않음.</li> <li>- 양 눈의 바깥 가장 자리에서 눈가의 주름(Crow's feet)이 생성.</li> </ul>
슬픔(Sadness)	<ul style="list-style-type: none"> <li>- 눈썹의 안쪽 부분이 아래로 내려감.</li> <li>- 눈썹 아래의 피부가 삼각형화 되며, 안쪽 눈썹 부분은 상승함.</li> <li>- 위 눈꺼풀의 안쪽 부분이 상승.</li> <li>- 입가가 내려가거나, 입술이 떨림.</li> </ul>

판단이 정확하게 이루어진다고 하더라도 그에 따른 적절한 대응을 취하는데 어려움이 존재한다는 문제점을 안고 있다(Nicolaou et al., 2011; Ekman, 1972).

한편, V-A 모형, 즉 매력-각성 모형은 인간이 느끼는 감정의 차원을 매력도와 각성도의 두 차원으로 구분하여 측정하는 방식을 제안한다(Mehrabian and Russell, 1974; Russell, 1980). 이 중, 매력도 차

원은 대상자의 감정이 얼마나 긍정적인지 또는 부정적인지를 나타내는 차원으로써, 행복을 기준으로 기분 나쁜 감정으로부터 기분 좋은 감정까지를 나타내고 있다. 반면, 각성도 차원은 대상자가 얼마나 무관심한지 혹은 흥미를 느끼며 대상에 몰입하는지의 여부를 반영하는 차원이다. 이 차원은 지루하고 무관심한 상태에서부터 매우 흥분된 상태까지 나타내고 있다. 심리학에서는 이러한 매력과 각

성의 두 차원이 서로 상호 연관되어있다고 말한다 (Oliveira et al., 2006; Alvarado, 1997; Lewis et al., 2007; Lane and Nadel, 2000). 더 명확히 말해, 각 차원이 나타내는 값은 상호의존적으로 계속적으로 연관되어 값이 설정된다. 다음 <그림 1>는 V-A 모형의 2개 차원과 여러 감정을 나타내는 형용사들이 이 2개의 차원으로 구성된 사분면에서 어디에 위치하는지를 나타내고 있다.



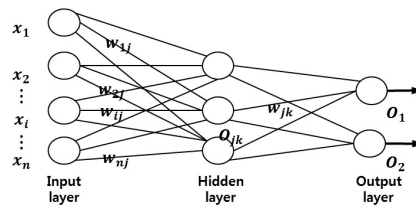
<그림 1> 매력-각성 감정 분류 모형

V-A 모형을 사용하여 사용자의 감정 및 감성 상태를 판단하기 위한 연구는 다양한 분야에서 시도되어 왔다. Nicolaou et al.(2011)은 얼굴표정, 어깨 움직임, 소리 등을 입력변수로 하여 BLSTM-NNs(the bidirectional Long Short-Term Memory neural networks), SVR(Support Vector Regression) 등의 알고리즘을 이용, 매력도와 각성도를 판단하고자 하였다. 유민준 등(2009)은 소리를 입력변수로 한 V-A 모형을 이용하여 음악을 정렬한 후, 이를 바탕으로 음악을 탐색하는 새로운 인터페이스를 제안하였다. 김용준과 조성배(2009)는 베이지안(Bayesian) 체계를 사용하여 추론된 사용자의 감정을 V-A 모형을 사용하여 표현하였고, 추론된

감정을 통해 현재 상황에 가장 적합한 장르의 벨 소리를 추천하여 사용자의 편의를 증대하였다. 이에, 본 연구 역시 이처럼 상대적으로 응용가능성이 더 높다고 평가되는 V-A 모형에 기반하여 관객의 감정을 체계화하고자 하였다.

### 2.3 인공신경망을 기반한 패턴 인식 관련 연구

1943년 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(장남식 등, 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. 경영분야의 문제해결에 있어서도 인공신경망은 고객의 신용평가, 불량거래의 색출, 의료진단 예측, 우량고객 선정, 타겟 마케팅 등 광범위한 분야에서 현재 활용되고 있다. 일반적으로 인공신경망은 <그림 2>과 같이 입력층(Input Layer)과 출력층(Output Layer) 사이에 한 개 이상의 은닉층(Hidden



<그림 2> 인공신경망 기본 모형

layer)이 존재하는 형태로 구성된다(심흥기, 김승권, 2008).

이러한 인공신경망을 학습하기 위한 방법으로는 역전파 알고리즘(Back-Propagation Algorithm)이 가장 널리 활용되는데, 이는 목표로 하는 결과값(Desire Output)과 모델에 의해 계산된 결과값(Actual output)의 차이를 학습을 통해 계속 조정해 나감으로써 오차 값을 최소화하도록 학습시키고 학습결과를 이용하여 예측대상에 검증 실시하는 방법이다(김대수, 1993; Haykin, 1999).

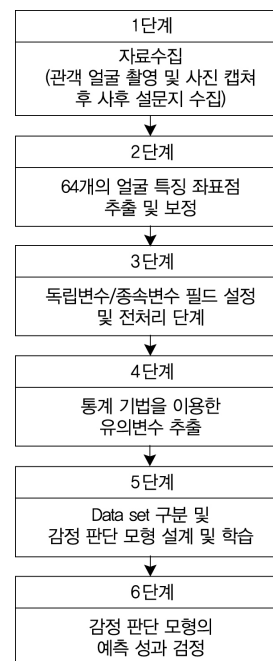
본 연구에서는 얼굴 특징으로부터 감정을 판단하기 위한 도구로 인공신경망을 적용하고자 하는데, 본 연구와 마찬가지로 얼굴 특징으로부터 특정 패턴을 인식하는 연구는 기존에도 많이 시도되어 왔다. 예를 들어, Ioanna-Ourania(2004), Mark et al. (1996) 등은 인공신경망을 이용하여 얼굴 특징값을 추출하고, 추출된 특징점의 변화 패턴을 인지하였다. 이보다 한 발 더 나아가 인지된 얼굴 특징점의 변화를 이용하여 감정을 추론하는 모형을 만들어 제안한 연구도 다수 존재한다(Bouzalmat et al., 2011; Henry et al., 1998; Huang and Huang, 1977; Mark et al., 1996; Zahid et al., 2009; Colibazzi et al., 2010). 하지만, 이들 기존 연구들은 대부분 얼굴 특징의 변화량에서 감정을 판단할 때, 앞서 소개한 Ekman의 6가지 유형만을 적용해 왔다는 한계가 존재한다. 이러한 한계점을 보완하기 위하여 본 연구에서는 인공신경망 기법으로 V-A 모형 기반의 감정을 판단할 수 있는 새로운 모형을 제안하고자 한다.

### 3. 연구모형 및 구축절차

본 연구는 인터랙티브 전시장에 참여한 관객들에게 개인화된 전시 서비스를 제공하기 위해, 관객

의 반응을 통해 그 관객의 감정을 판단할 수 있는 새로운 모형을 제안하는 것을 목적으로 한다. 인공신경망 기법에 의해 학습되는 본 연구의 제안 모형은 관객의 얼굴 특징의 변화를 통해, 그 관객이 현재 대상 전시물에 대해 얼마나 매력을 느끼고 있는지, 그리고 얼마나 몰입하고 있는지를 판단할 수 있다.

본 연구의 제안 모형은 <그림 3>과 같은 여섯 단계 절차에 의해 구축될 수 있다. 우선, 첫 번째 단계는 자료를 수집하는 단계이다. 이 단계에서는 판단 모형의 입력값과 출력값이 될 수 있는 자료 원천을 확보하는 단계인데, 본 연구의 입력값은 관객의 얼굴표정을 대상으로 하고 있고, 출력값은 매력도와 각성도 두 개로 정의되어 있다. 그러므로, 첫 번째 단계에서는 입력값에 대한 데이터 수집을 위해 주로 관객 얼굴에 대한 화상 촬영(캡처) 작업



<그림 3> 제안된 연구모형의 구축절차

이 이루어지게 되고, 출력값에 대한 데이터 수집을 위해서는 사후 설문지 작성이 이루어지게 된다. 두 번째 단계는 수집된 자료로부터 얼굴 특징점을 추출하고 보정하는 단계이다. 이 단계에서는 사람의 얼굴을 구성하는 눈, 코, 입, 입술, 눈썹, 턱선 등을 구성하는 다양한 특징점들의 위치정보를 계량화하여 산출하게 된다. 세 번째 단계는 본 연구의 학습 모형인 인공신경망 모형의 독립변수와 종속변수를 생성하고, 생성된 변수들을 전처리하는 단계이다. 즉, 이 단계에서는 앞서 2단계에서 확보된 얼굴 특징점이 구체화된 얼굴 특징으로 변환되는 작업이 이루어지게 된다. 네 번째 단계는 통계적 모형을 사용하여 종속 변수에 영향을 미치는 유의한 독립 변수를 추출하는 단계이다. 이 단계에서는 주로 상관분석 및 독립표본 t-검정과 같은 기법들이 적용된다. 다섯 번째 단계는 전처리가 완료된 자료들을 학습용 데이터 집합, 실험용 데이터 집합, 검증용 데이터 집합 세 가지 데이터 셋으로 구분한 뒤, 인공신경망 모형을 학습하는 단계이다. 사실상 이 단계까지 끝나고 나면, 관객의 감정을 판단할 수 있는 인공신경망 모형의 구축은 완료된다. 마지막 여섯 번째 단계는 앞서 다섯 단계를 통해 구축된 모형이 과연 믿을만하게 관객의 감정을 판단하는지를 검증용 데이터 집합을 이용해 검증하는 단계이다.

각 단계별 절차를 좀 더 상세히 설명하면 다음과 같다. 첫 번째 단계는 자료 수집 단계인데, 이 단계에서는 관객의 감정을 판단할 수 있도록 인공신경망 모형을 학습시킬 수 있는 원천 데이터를 수집하는 단계이다. 본 연구에서는 매력도와 각성도를 모두 예측하기 위한 모형의 입력 데이터로 사용하기 위해, 피실험자 1명 당 크게 세 가지 상태의 얼굴 표정을 캡처한 사진의 데이터를 수집한다. 본 실험에서 크게 세 가지 상태의 화상을 이용한

이유는 다음과 같다. 우선, 각성도를 예측하기 위해 자극 상태와 무자극 상태의 데이터로 나누어 수집할 필요성이 있다. 하지만, 이것만으로는 매력도까지 예측하기는 어려운 면이 있기 때문에 자극 상태를 긍정적인 자극과 부정적인 자극이 주어진 상태로 나눠서 데이터를 수집해야 한다. 따라서, 본 실험에서는 크게 세 가지 상태의 화상을 이용하여 매력도와 각성도를 예측하기로 한다. 첫 번째 상태는 중립상태로 실험장비에 입장한 직후 촬영된 얼굴의 정지화상이다. 두 번째 상태는 긍정적 각성상태로써 긍정적인 각성을 유발하는 동영상 시청할 때 촬영된 얼굴의 정지화상이다. 마지막 상태는 부정적 각성상태로 부정적 각성을 유발하는 동영상 시청할 때 촬영된 정지화상이다. 이 세 가지 상태의 입력 데이터로 이용하여 본 연구에서 제안한 관객 반응 예측 모형을 학습한다. 그리고, 본 연구에서 제안하는 모형의 출력값은 매력도와 각성도로써, 이 출력값은 사후 설문지를 통해 관객이 직접 작성함으로써 얻어진다. 출력값은 매력도와 각성도 둘 다 7점 리커트(Likert) 척도를 사용한다.

두 번째 단계는 첫 번째 단계에서 수집된 데이터로부터 얼굴 특징점을 추출하고 보정하는 단계이다. 얼굴 특징점은 코, 눈, 눈썹 등 얼굴의 주요 부위를 나타낼 수 있는 좌표들로 구성되는데, 좌표수가 많을수록 보다 정밀한 얼굴 특징 파악이 가능해진다. 특징점 추출이 완료되면, 이렇게 추출된 관객의 얼굴 특징점 좌표값은 추출된 얼굴 특징점 중 표정에 관계없이 변하지 않는 고정점(예 : 코 끝)의 위치를 기준으로 보정된다. 이는 자극영상 얼굴 전체의 이동으로 인해 얼굴 특징점 전체가 이동하는 부분을 보정하기 위해 수행되는 과정이다. 이렇게 보정작업까지 이루어지고 나면, 얼굴 특징점들을 여러 얼굴 특징으로 변환하는 작업이 이루어지게 되는데, 관객 반응 판단 모형의 독립변수

후보로 활용될 수 있는 각종 얼굴 특징 변화값 (예 : 눈썹 끝이 올라간 정도, 입술 끝이 내려간 정도 등)이 이 과정에서 산출되게 된다. 이러한 얼굴 특징 변화값 산출에는 Pantic and Rothkrantz(2000)가 제안한 사용자의 감정에 영향을 미치는 35개 지표가 적용될 수 있다.

세 번째 단계는 본 연구에서 제안하는 새로운 관객 반응 판단 모형에 대한 독립변수와 종속변수들을 설정하는 단계이다. 우선 본 연구 모형에서 사용될 종속변수는 매력도와 각성도의 2가지가 되며, 이는 피실험자에 대한 설문을 통해 리커트 척도를 이용해 측정된 값을 적용한다. 한편, 독립변수로는 앞의 두 번째 단계에서 도출된 각종 얼굴 특징 변화값들이 사용된다.

네 번째 단계는 종속변수인 매력도와 각성도에 유의한 상관관계를 갖는 특징변수들, 즉 독립변수들을 추출하기 위한 단계이다. 이 때 유의한 독립변수를 추출하기 위한 기법으로는 통계적 검정 기법, 그 중에서도 상관분석과 독립표본 t-검정을 이용해 분석한다.

다섯 번째 단계에서는 전처리 후 최종 확보된 자료들을 인공신경망 기법을 학습하기 위한 데이터 집합으로 구분하고, 실제 모형을 학습하는 과정이 이루어진다. 이 때 확보된 전체 데이터는 3개의 집합으로 구분하게 되는데, 학습을 위한 데이터 집합과 구축된 모형의 학습 중지점을 찾기 위한 테스트용 데이터 집합, 그리고 인공신경망 모형의 예측 성과를 검증하기 위한 검증용 데이터 집합으로 구분한다.

마지막 여섯 번째 단계에서는 이전 단계에서 학습된 인공신경망 모형을 검증용 데이터 집합을 사용하여 학습된 인공신경망 모형의 예측 성과를 검증한다. 이 단계를 통해, 구축된 모형이 새로운 데이터(unknown data)에 대해 얼마나 우수한 예측

력을 갖는지, 그리고 구축된 모형이 일반성(generalizability)을 갖는지 확인하게 된다.

## 4. 실증 분석

### 4.1 데이터 수집 및 전처리

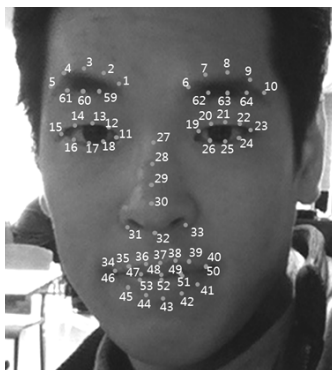
본 연구에서는 일반인을 대상으로 분석 데이터를 수집하여 제안하는 새로운 관객 반응 판단 모형의 유용성을 검증해 보고자 하였다. 이를 위해 서울 상암동에서 개최된 2011 서울 DMC 킷치 오픈 행사에 참가하여, 9월 29일부터 10월 1일까지 2박 3일간 행사 참여자들을 대상으로 데이터를 수집하였다. 실험을 위한 데이터 수집 장소는 DMC 디지털파빌리온 1층 IT KOREA관 어트랙티브 행복관이었으며, 이곳에서 대형 DID, u-테이블, 고해상도 웹캠 등 첨단 하드웨어 장비와 Flash로 프로그래밍된 소프트웨어로 구성된 자료 수집을 위한 실험 부스를 운영하였다. 본 행사에 참여한 관람객은 총 198명이었는데, 이들에 대한 인구통계학적 정보는 다음 <표 2>와 같다. 총 데이터 샘플 수는 정지 상태를 기준으로 중립 상태 데이터, 긍정적 각성 상태 데이터, 부정적 각성 상태 데이터 각각 198개로 총 594개였다. 이러한 상태 데이터들은 모두 얼굴이 찍힌 화상 이미지들인데, 본 연구에서는 이들 이미지들로부터 <그림 4>와 같은 총 64개의 얼굴 특징점을 도출해 이들을 기준으로 얼굴 특징을 산출하고자 하였다. 이 때, 얼굴 특징점을 도출하는 데에는 해당 분야에 있어 국내 최고의 기술수준을 보유한 기업인 올라웍스(<http://www.olaworks.com/>)의 솔루션을 제공받아 사용하였다.

이렇게 산출된 64개의 얼굴 특징점 좌표들은 고정점인 30번 좌표(코 끝 위치 좌표)를 기준으로 위치를 보정하였다. 그리고, 전처리 과정을 통해 최



<표 2> 피실험 대상자들의 인구통계정보

특성		참관객	
		인원 수(명)	%
성별	남	101	51.0%
	여	97	49.0%
연령대	20대 이하	46	23.5%
	30대	111	56.6%
	40대 이상	39	19.9%
학력	중학교 이하	32	16.3%
	고등학교 이하	33	16.8%
	대학교 이하	88	44.9%
	대학원 이상	43	21.9%
직업	학생	104	52.5%
	사무관리직	26	13.1%
	판매서비스직	4	2.0%
	기술기능직	4	2.0%
	단순노무직	0	0%
	전문직	27	13.6%
	자영업	6	3.0%
	공무원	9	4.5%
	주부	14	7.1%
무직	4	2.0%	
총 계		198	100.0%



<그림 4> 64개 얼굴 특징점

종적으로 실험에 활용할 데이터 집합을 구축하였는데, 특징점 도출에 실패하거나 설문에서 인지된 몰입도와 매력도에 대한 응답이 누락된 사례들을

찾아 분석 대상에서 삭제하였다. 그리고, 기존에 응답자 단위로 분류/정리되어 있던 전체 데이터 집합을 긍정적 자극과 부정적 자극 별로 구분해 각각 독립된 표본 1건으로 분리시켰다. 그 결과, 총 297건의 레코드로 구성된 전체 분석용 데이터 집합이 산출되었다.

## 4.2 실험 설계

### 4.2.1 독립변수 및 종속변수 선정

본 연구에서는 얼굴 특징점으로부터 얼굴 특징을 도출하는데 있어, Pantic and Rothkrantz(2000)의 연구결과를 참조하였다. Pantic and Rothkrantz(2000)는 인간의 얼굴 전체 영상에서 얼굴표정을 인식하고 이를 감정별로 분류하는 ISFER(Integrated System for Facial Expression Recognition) 시스템 개발에 대한 연구를 진행하였다. 이들이 진행한 연구에서는 정지된 얼굴 전체 영상에서 19개의 얼굴 특징점을 바탕으로 30개의 얼굴 특징값을 설정하고, 이를 바탕으로 하는 17개의 AU(Action Unit)를 활용하여 30개의 얼굴 특징점들의 변화를 측정하여 감정을 분류하였다. 이를 바탕으로 본 연구에서는 Pantic and Rothkrantz(2000)가 제시한 얼굴 특징점을 추출된 64개의 얼굴 특징점과 대응시켰다. 그리고 Pantic and Rothkrantz(2000)이 제시한 30개의 얼굴 특징값 중 거리와 각도에 관계있는 18개의 얼굴 특징값과 17개의 AU를 이용한 35개의 특징값을 산출하여 이들을 후보 독립변수군으로 활용하였다. 본 연구에서 사용하는 후보 독립변수들의 필드는 <표 3>과 같다.

전술했듯이, 본 연구에서는 관객 감정 판단에 사용될 종속변수로 관객의 인지된 매력도(perceived valence)와 인지된 각성도(perceived arousal)의 두 개의 변수 필드를 사용하기로 하였다. 즉, 관객

감정 판단을 위한 모형으로 매력-각성 모형을 활용한다. 매력-각성 모형은 해당 관객이 제공되고 있는 자극에 얼마나 호감을 갖고 있는지(매력도), 그리고 그 자극에 얼마나 관심을 갖고 있는지(각성도)를 동시에 측정할 수 있어, 실용성 관점에서

활용도가 아주 높은 모형이다. 때문에 현재 학계뿐만 아니라, 업계(예 : 미국 Affectiva社)에서도 매력-각성 모형에 기반한 감정 모델링 기법을 활용하고 있다. 본 연구에서 종속변수로 활용된 매력도와 각성도는 실험에 참가한 관객들에게 직접 설문

<표 3> 후보 독립변수의 필드

순번	변수명	데이터유형	변수설명	단위
1	f1	double(8)	∠ 11, 15, 59 변량	Degree
2	f2	double(8)	∠ 19, 23, 62 변량	Degree
3	f3	double(8)	15~61 거리변량	Pixel
4	f4	double(8)	23~64 거리변량	Pixel
5	f5	double(8)	A*~13 거리변량	Pixel
6	f6	double(8)	B**~21 거리변량	Pixel
7	f7	double(8)	A*~7 거리변량	Pixel
8	f8	double(8)	B**~25 거리변량	Pixel
9	f9	double(8)	13~17 거리변량	Pixel
10	f10	double(8)	21~25 거리변량	Pixel
11	f11	double(8)	32~37 거리변량	Pixel
12	f12	double(8)	11~34 거리변량	Pixel
13	f13	double(8)	19~40 거리변량	Pixel
14	f14	double(8)	32~34 거리변량	Pixel
15	f15	double(8)	32~40 거리변량	Pixel
16	f16	double(8)	34~40 거리변량	Pixel
17	f17	double(8)	37~43 거리변량	Pixel
18	f18	double(8)	32~56 거리변량	Pixel
19	au1	Boolean	f1과 f2 증가	0 : False, 1 : True
20	au2	Boolean	f1 또는 f2 증가	
21	au4	Boolean	f1과 f2 감소	
22	au5	Boolean	f5와 f6 증가	
23	au7	Boolean	f7 또는 f8 감소	
24	au10	Boolean	f11 감소	
25	au12	Boolean	f12와 f13 감소, f14와 f15 증가	
26	au13	Boolean	f12, f13, f14, f15 감소	
27	au15	Boolean	f12 또는 f13이 증가	
28	au18	Boolean	f16 감소	
29	au20	Boolean	f16 증가, f12와 f13 증가하지 않음	
30	au23	Boolean	f17 감소, f16 감소하지 않음 f12와 f13 증가하지 않음	
31	au24	Boolean	f17 감소, f16 감소	
32	au25	Boolean	f18 증가하지 않음, f17 증가	
33	au26	Boolean	f18 값이 한계점 사이에 있음(증가)	
34	au28	Boolean	f17 = 0	
35	au41	Boolean	f7 감소하지 않음, f9 감소, f5 감소 또는 f10 감소, f6 감소, f8 감소하지 않음	

을 받아 수집하였는데, 측정 단위로는 7점 리커트 척도를 사용하였다. 본 연구에서 사용하는 독립변수의 필드를 정리하면 <표 4>와 같다.

<표 4> 종속변수의 필드

순번	종속 변수명	변수설명	단위
1	Valence	관객의 인지된 매력도 수준	7점 Likert scale
2	Arousal	관객의 인지된 각성도 수준	7점 Likert scale

#### 4.2.2 유의 독립변수 선정

앞서 산출된 데이터 집합으로부터 매력도와 각성도를 예측 및 판단할 수 있는 모형을 개발하기 위해, 1차적으로 각 종속변수들과 통계적으로 유의한 상관관계를 갖는 특징변수, 즉 독립변수를 선정하였다. 본 연구에서는 실제 모형에 활용될 독립

변수를 선정하기 위해 2가지 통계기법을 이용하였다. 우선, 비율형 변수로 측정되어 있는 독립변수들에 대해서는 상관분석을 통해, 피어슨(Pearson) 상관계수를 산출하고, 이 값이 통계적으로 유의한지 검증하였다. 명목형(Boolean) 변수로 측정되어 있는 독립변수들에 대해서는 독립표본 t-검정(independent Samples t-test)을 통해, 유의변수를 추출하였다.

그 결과, 매력도의 경우, 비율형 변수로 측정되어 있는 독립변수들은 총 3개의 변수가 90% 신뢰수준 하에서 유의변수로 산출되었다. 그리고 명목형 변수로 측정되어 있는 독립변수들은 통계 검증 결과 총 7개의 변수가 90% 신뢰수준 하에서 유의변수로 산출되었다. 따라서, 총 35개의 변수 중 통계 검증 결과 매력도와 유의하다고 판단된 독립변수는 총 10개이다.

각성도의 경우, 비율형 변수로 측정되어 있는

<표 5> 최종 선택된 독립변수군

변수명	변수설명	각성도'	매력도'	단위
f1	< 11, 15, 59 변량		선정(-)	Degree
f4	23~64 거리변량	선정(-)		Pixel
f11	32~37 거리변량	선정(+)		Pixel
f14	32~34 거리변량	선정(-)		Pixel
f18	32~56 거리변량		선정(-)	Pixel
au1	f1과 f2 증가	선정(+)		0 : False, 1 : True
au2	f1 또는 f2 증가	선정(-)		
au5	f5와 f6 증가		선정(+)	
au7	f7 또는 f8 감소		선정(-)	
au10	f11 감소		선정(-)	
au13	f12, f13, f14, f15 감소	선정(-)		
au15	f12 또는 f13이 증가	선정(+)		
au18	f16 감소	선정(-)	선정(-)	
au20	f16 증가, f12와 f13 증가하지 않음	선정(+)	선정(+)	
au24	f17 감소, f16 감소	선정(-)		
au41	f7 감소하지 않음, f9 감소, f5 또는 f10 감소, f6 감소, f8 감소하지 않음		선정(-)	

독립변수들은 총 2개의 변수가 90% 신뢰수준 하에서 유의변수로 산출되었다. 그리고 명목형 변수로 측정되어 있는 독립변수들은 통계 검증 결과 총 7개의 변수가 90% 신뢰수준 하에서 유의변수로 산출되었다. 따라서, 총 35개의 변수 중 총 9개의 독립변수가 각성도와 유의한 것으로 파악되었다.

결국 본 연구에서 제안하는 관객 감정 판단 모형에 활용된 최종 독립 변수는 <표 5>와 같다. 이때, 선정 옆에 제시된 ( )안의 부호는 상관관계의 방향을 의미한다.

#### 4.2.3 인공신경망 실험설계

본 연구에서는 인공신경망 모형 구축에 앞서 전체 297건의 데이터를 아래와 같이 학습용 데이터 집합(Training Set)과 테스트 데이터 집합(Test Set), 그리고 검증용 데이터 집합(Validation Set)으로 구분하였다. 구분된 데이터 집합은 다음 <표 6>과 같다. 학습용 데이터 집합은 인공신경망 기반 관객 감정 판단 모형을 학습하는데 사용되며, 테스트 데이터 집합은 과적합(overfitting)을 방지하기 위한 학습중지점 탐색에 사용한다. 끝으로, 검증용 데이터 집합으로 구축된 모형의 성과를 검증한다.

매력도와 각성도 추정을 위한 인공신경망 모형의 경우, 학습 방법으로는 역전파 알고리즘을 사용하였고, 학습률을 10%, 모멘텀률은 10%로 설정하였다. 변환함수로는 로지스틱(Sigmoid)함수를 사용하였으며, 계층구조는 3계층 퍼셉트론, 즉, 은닉층 수는 1개로 설계하였다. 학습중지점을 탐색하

기 위해서는 테스트 데이터 집합을 기준으로 최소 오류 도달 후 50000번 도달 시 중지하는 것으로 하였다. 은닉층을 구성하는 노드의 수로는 매력도의 경우 5, 9, 14, 18의 4가지 경우에 대해, 각성도의 경우 5, 10, 15, 20의 4가지 경우에 대해 모두 실험해 본 뒤, 가장 우수한 성과를 보이는 노드수를 선택하였다. 그리고 인공신경망 실험을 수행하기 위한 도구로는 상용 소프트웨어인 Neuroshell2 4.0을 사용하였다.

#### 4.2.4 비교모형 설정

본 연구에서는 제안모형에 사용된 인공신경망 기법의 우수성을 검증하기 위해, 비교모형으로 중회귀분석 기법을 적용하였다. 이를 통해, 과연 인공신경망 기법이 관객 감정을 판단하기 위한 모형으로 적합한지 확인해 보고자 하였다.

### 5. 분석 결과

본 연구에서는 제안모형의 판단 정확도를 검증하기 위한 기준으로 MAE(Mean Absolute Error)를 사용한다. MAE는 예측된 값과 실제값의 차이들에 대해 절대값을 취해 평균한 값으로서, 그 값이 작으면 작을수록 정확도가 높은 것으로 판단할 수 있다. 본 실험에서 매력도와 각성도의 성과 기준은 검증용 데이터 집합을 기준으로 판단하였다.

우선 매력도 추정을 위한 인공신경망 모형 구축 실험에 대한 결과를 보면 다음과 같다. 본 실험은

<표 6> 데이터 집합 구분

자료 구분	표본수	비중	용도
학습용 데이터 집합	179	60.2%	인공신경망 모형 구축
테스트용 데이터 집합	59	19.9%	과적합 방지를 위한 학습중지점 탐색에 사용
검증용 데이터 집합	59	19.9%	모형의 예측 정확도 검증
합 계	297	100.0%	

은닉층 노드의 수가 각각 5개, 9개, 14개, 18개 일 경우에 대해 실험해 보고, 성과가 가장 우수한 경우의 모형을 선정하는 것으로 실험을 진행하였다. 그 결과 <표 7>과 같은 결과를 얻을 수 있었다.

<표 7> 매력도 판단을 위한 인공신경망 모형 실험 결과

자료 구분	h = 5	h = 9	h = 14	h = 18
학습용 데이터 집합	1.0674	1.0641	1.0785	1.0933
테스트 데이터 집합	0.9921	1.0047	1.0234	1.0317
검증용 데이터 집합	1.3527	1.3706	1.4244	1.4333

<표 7>에서 볼 수 있듯이, 은닉층 노드 수가 가장 적은 5개일 때 매력도 판단 모형의 정확도가 가장 높은 것을 알 수 있다. 따라서, 인공신경망 기반 매력도 판단 모형은 은닉층 노드 수가 5개인 모형을 최종적으로 선택하였다.

다음 <표 8>은 각성도를 판단하기 위한 인공신경망 모형의 실험 결과를 제시하고 있다. 각성도의 경우, 은닉층 노드의 수가 각각 5개, 10개, 15개, 20개 일 경우에 대해 실험해 보고, 성과가 가장 우수한 경우의 모형을 선정하는 것으로 실험을 진행하였다.

각성도 실험 분석 결과, 은닉층 노드 수가 두 번째로 적은 10개일 때 모형의 예측 정확도가 가장 높은 것을 알 수 있다. 따라서, 인공신경망 기반 각성도 판단 모형은 은닉층 노드 수가 10개인 모형을 최종적으로 선택하였다.

<표 8> 각성도 판단을 위한 인공신경망 모형 실험 결과

자료 구분	h = 5	h = 10	h = 15	h = 20
학습용 데이터 집합	0.9370	0.9457	0.9463	0.9460
테스트 데이터 집합	0.9697	0.9736	0.9738	0.9699
검증용 데이터 집합	0.8421	0.8298	0.8382	0.8479

한편, 비교모형으로 실험한 매력도와 각성도에 대한 중회귀모형의 예측 결과는 다음 <표 9>와 같이 산출되었다.

<표 9> 중회귀모형 실험 결과

자료 구분	매력도	각성도
	평균 MAE	평균 MAE
학습용 데이터 집합	1.0811	1.0797
검증용 데이터 집합	1.3433	1.3845

최종적으로 본 연구에서 제안한 인공신경망 기반 판단 모형과 비교모형인 중회귀분석 기반 판단 모형의 성능을 <표 10>과 같이 비교해 보았다. 비교 결과, 각성도와 매력도 모두 본 연구에서 제안하는 인공신경망 모형이 중회귀분석 모형보다 더 높은 정확도의 판단결과를 제공하고 있음을 확인할 수 있었다. 아울러, 각성도를 예측하는 모형이 매력도를 예측하는 모형에 비해 대체로 MAE가 작다는, 즉 예측오차가 더 적다는 점을 확인할 수 있다. 이는 매력도보다는 각성도를 예측하는 것이 상대적으로 더 수월하다는 점, 즉 각성도의 변화가 매력도의 변화보다 관객의 얼굴표정에 더 잘 드러

<표 10> 성과 종합 비교표

자료 구분	매력도 판단 모형		각성도 판단 모형	
	중회귀 모형	신경망 모형	중회귀 모형	신경망 모형
학습용 데이터 집합	1.2034	1.0674	0.9433	0.9457
테스트용 데이터 집합		0.9921		0.9736
검증용 데이터 집합	1.3596	1.3527	0.8682	0.8298

난다는 점을 일부 시사한다고 할 수 있다.

## 6. 결론

전시물과 관람자의 다양한 상호작용이 일어나는 인터랙티브 전시 공간 내에서 전시 관람에 보다 집중하기 위해서는 관객에게 맞춤형 전시 환경을 제공하고, 관객의 선호를 파악해 개인화된 서비스를 제공할 필요성이 있다. 이러한 관객 맞춤형 개인화 서비스를 제공하기 위해서는 우선 전시물을 관람하는 관객의 반응을 이해하고, 이를 통해 대상 관객의 감정을 판단하는 것이 필요하다. 이에 본 논문에서는 기존의 연구들을 바탕으로 인터랙티브 전시 공간 내에서 수집 가능한 관객 반응 데이터를 이용하여 관객의 감정을 판단할 수 있는 새로운 모형을 구축하기 위한 연구를 수행하였다.

구체적으로 본 논문은 인터랙티브 전시 공간 내에서 수집 가능한 다양한 관객 반응 데이터 중 얼굴 표정의 변화를 통해 대상 관객이 얼마나 대상물에 매력을 느끼고 있는지, 그리고 대상물에 몰입해 있는지를 판단할 수 있는 인공지능 기반의 관객 감정 판단 모형을 제시하였다. 이러한 지능형 모형에 의거한 인간의 감정 판단을 전시 분야에 응용한 학술적 접근은 본 논문이 사실상 최초라고 판단되며, 이것이 본 논문이 갖는 가장 큰 학술적 의의라고 사료된다. 또한 본 연구에서는 단순한 모형 제시에 머무르지 않고, 대규모로 수집된 일반인 대상의 데이터로 제안모형에 대한 검증을 시도하였는데, 이 역시 주요한 본 논문의 학술적 의의 중 하나라 할 수 있다.

실무적 관점에서 본 연구에서 제안하는 인공지능 기반 관객 감정 판단 모형은 매력-각성 모형에 기반하고 있기 때문에, 그 활용도가 높을 것으로 기대된다. 만약 본 연구의 제안모형이 실제 인

터랙티브 전시에 적용된다면, 전시물을 관람하는 관객의 매력도나 각성도가 낮다고 예측하는 경우, 전시물 관람에 조금 더 집중하기 위한 환경 조성(소리, 조명 조절 등)을 만든다거나 아니면 다른 콘텐츠를 추천하던지 등 관객에게 적합한 다양한 대응책을 마련할 수 있다. 이렇듯, 본 연구의 제안 모형은 인터랙티브 전시 분야의 원천기술로서, 그 활용도가 매우 높을 것으로 전망된다. 매력-각성 모형의 두 가지 차원 중 특히, 각성도 차원은 그 실무적 활용도가 더 높을 수 있을 것으로 전망되는데, 추후 교육, 광고 등의 분야로 확대 적용 가능하다. 또한, 감정을 예측하는 기존 연구들은 대부분 외국 환경, 특히, 미국 환경에 최적화 되어 있어 일반화하기에는 한계점을 지니고 있다 하지만, 본 연구는 한국인을 대상으로 모형을 구축했다는 점에서 또 하나의 실무적 의의가 있다.

반면 연구의 한계점을 살펴보면, 본 연구에서 구축한 인공지능 기반 관객 감정 판단 모형은 비교모형인 중회귀모형보다 예측 정확도가 높기는 하나, 아직까지 전반적으로 예측 정확도가 많이 떨어진다. 그 원인은 다음과 같이 세 가지 부분에 있다고 사료된다. 첫째, 자료 수집을 일반인 대상으로 하다 보니 통계가 어렵다는 한계점을 지닌다. 즉, 실험 환경에 최적화하지 못하여 수집한 데이터에 노이즈가 많아 전반적으로 예측력이 떨어지는 현상이 나타나고 있는 것으로 해석된다. 둘째, 관객반응 분석을 위한 정보를 얼굴 표정 변화에 국한한 점도 본 연구의 한계점 중 하나이다. 추후 연구에서 소리/음성, 동작 등이 반영되면, 더 나은 결과를 얻을 수 있을 것으로 기대된다. 마지막으로 현재 제시한 모형이 저해상도 정지화상에 기반하고 있다는 점도 한계점으로 지적된다. 이로 인해 현재 얼굴 특징점 좌표값에 대한 위치 변화량이 크지 않은데, 이것이 예측 정확도를 높이는

데 주요한 방해요인이 되고 있는 것으로 보인다. 때문에 추후 수행될 실증분석에서는 보다 고해상도의 얼굴 화상을 수집, 저장할 필요가 있다. 아울러, 정지 화상이 아닌 동영상 분석을 시도할 경우, 더 예측성고가 좋아진다는 기존 연구(예 : Nicolaou et al., 2011)가 존재하므로, 시간의 흐름에 따른 변화량 분석에 대한 연구가 추후 이루어져야 할 것이다.

## 참고문헌

- 김규정, 김봉화, 이명학, 한혜정, “음성인식 기반 인터랙티브 미디어 아트의 연구 1-소리-시각 인터랙티브 설치미술 불꽃 뮤직을 중심으로-”, *기초조형학연구*, 9권 5호(2008), 27~35.
- 김대수, 신경망 이론과 응용, 하이테크정보, 1993.
- 김미연, 김정현, 최진원, “인터랙티브 체험형 전시 공간 디자인을 위한 사례분석 연구”, *대한건축학회*, 24권 1호(2008), 11~18.
- 김용준, 조성배, “감정 모델을 통해 휴대폰의 벨소리를 추천하는 상황 인식 추천 시스템”, *한국정보과학회 2009 한국컴퓨터종합학술대회 논문집*, 36권 1호(2009), 162~165.
- 김철근 외, *전시이론과 기법 연구집*, 국립중앙과학관, 학술총서 12권(1996).
- 송현미, 최준혁, 임채진, “첨단과학기술 체험전시를 위한 연출기법 및 전시구성체계에 관한 연구-첨단과학분야(NT, MEMS, BT)의 체험전시를 중심으로-”, *한국실내디자인학회 학술대회논문집*, 7권 1호(2005), 182~186.
- 심흥기, 김승권, “인공신경망을 이용한 대대전투간 작전지속능력 예측에 관한 연구”, *대한산업공학회/한국경영과학회 춘계공동학술대회*, 2008.
- 유민준, 김현주, 이인권, “감성모델을 이용한 음악 탐색 인터페이스”, *한국 HCI 2009 학술대회*, (2009), 707~710.
- 유재엽, “영상미디어 연출 특성에 따른 공간 표현에 관한 연구”, *한국실내디자인학회논문집*, 13권 6호(2004), 175~183.
- 임승희, 배경수,곽수정, 박인석, 박지수, “체험전시 콘텐츠의 몰입도 분석을 위한 주관적 경험 측정”, *디자인학 연구*, 22권 4호(2009), 19~30.
- 장남식, 홍성완, 장재호, *데이터마이닝*, 대청, 1999.
- 최지영, 노승석, 박진완, “디지털 미디어를 활용한 어린이 체험 전시 작품 구현 방안 연구”, *디자인학 연구*, 23권 5호(2010), 233~242.
- Alvarado, N., “Arousal and Valence in the Direct Scaling of Emotional Response to Film Clips”, *Motivation and Emotion*, Vol.21(1997), 323~348.
- Bouzalmat, A., N. Belghini, A. Zarghili, J. Kharroubi, and A. Majda, “Face Recognition Using Neural Network Based Fourier Gabor Filters and Random Projection”, *International Journal of Computer Science and Security (IJCSS)*, Vol.5, No.3(2011), 376~386.
- Brainerd, C. J., R. E. Holliday, V. F. Reyna, Y. Yang, and M. P. Toglia, “Developmental reversals in false memory : Effects of emotional valence and arousal”, *Journal of Experimental Child Psychology*, Vol.107(2010), 137~154.
- Colibazzi, T., J. Posner, Z. Wang, D. Gorman, A. Gerber, and S. Yu, “Neural systems subserving valence and arousal during the experience of induced emotions”, *Emotion*, Vol.10, No.3(2010), 377~389.
- Ekman, P., “Universals and Cultural Differences in Facial Expressions of Emotion. In J. Cole (Ed.)”, *Nebraska Symposium on Motivation*, Vol.19(1972), 207~282.
- Haykin, S., “Neural Networks : A Comprehensive Foundation”, Prentice Hall : Upper Saddle

- River, NJ, USA, 1999.
- Henry, A. R., B. Shumeet, and K. Takeo, "Neural Network-Based Face Detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.20, No.1(1998), 23~38.
- Huang, C. L. and Y. M. Huang, "Facial Expression Recognition Using Model-Based Feature Extraction and Action Parameters Classification", *Journal of Visual Communication and Image Representation*, Vol.8, No.3(1977), 278~290.
- Lane, R. and L. Nadel, *Cognitive Neuroscience of Emotion*. Oxford Univ. Press, 2000.
- Lewis, P. A., H. D. Critchley, P. Rotshtein, and R. J. Dolan, "Neural Correlates of Processing Valence and Arousal in Affective Words", *Cerebral Cortex*, Vol.17, No.3(2007), 742~748.
- Luigina, C., "Situating 'Place' in Interaction Design : Enhancing the User Experience in Interactive Environments", *Dept. of Computer Science and Information Systems*, University of Limerick, 2004.
- McCulloch, W. S. and W. H. Pitts, "A logical calculus of ideas immanent in nervous activity", *Bulletin of Mathematical Biophysics*, Vol.9 (1943), 115~133.
- Mark, R., Y. Yaser, and S. D. Larry, "Human Expression Recognition from Motion Using : a Radial Basis Function Network Architecture", *IEEE Transactions on Neural Networks*, Vol.7, No.5(1996), 1121~1138.
- Mehrabian, A. and J. Russell, *An Approach to Environmental Psychology*. MIT Press, 1974.
- Nicolaou, M., H. Gunes, and M. Pantic, "Continuous Prediction of Spontaneous Affect from Multiple Cues and Modalities in Valence-Arousal Space", *IEEE Transactions on Affective Computing*, Vol.2, No.2(2011), 92~105.
- Oliveira, A. M., M. P. Teixeira, I. B. Fonseca, and M. Oliveira, "Joint Model-Parameter Validation of Self-Estimates of Valence and Arousal : Probing a Differential-Weighting Model of Affective Intensity", *Proc. 22nd Ann. Meeting Int'l Soc. for Psychophysics* (2006), 245~250.
- Pantic, M. and L. J. M. Rothkrantz, "Expert system for automatic analysis of facial expressions", *Image and Vision Computing*, Vol.18(2000), 881~905.
- Russell, J. A., "A Circumplex Model of Affect", *J. Personality and Social Psychology*, Vol.39(1980), 1161~1178.
- Stathopoulou, I. O. and G. A. Tsihrintzis, "An Improved Neural Network-Based Face Detection and Facial Expression Classification System", *IEEE International Conference on Systems, Man, and Cybernetics 2004, The Hague, The Netherlands*(2004).
- Xu, M., J. S. Jin, S. Luo, and L. Duan, "Hierarchical movie affective content analysis based on arousal and valence features", *MM '08 : Proc. of the 16th ACM Mult*, (2008), 677~680.
- Zahid, R., G. Suat, B. Micheal, and R. Bernd, "A Unified Features Approach to Human Face Image Analysis and Interpretation", *Affective Computing and Intelligent Interaction(ACII)*, (2009), 41~48.



Abstract

## The Intelligent Determination Model of Audience Emotion for Implementing Personalized Exhibition<sup>\*</sup>

Min Kyu Jung<sup>\*\*</sup> · Jae Kyeong Kim<sup>\*\*</sup>

Recently, due to the introduction of high-tech equipment in interactive exhibits, many people's attention has been concentrated on Interactive exhibits that can double the exhibition effect through the interaction with the audience. In addition, it is also possible to measure a variety of audience reaction in the interactive exhibition. Among various audience reactions, this research uses the change of the facial features that can be collected in an interactive exhibition space. This research develops an artificial neural network-based prediction model to predict the response of the audience by measuring the change of the facial features when the audience is given stimulation from the non-excited state. To present the emotion state of the audience, this research uses a Valence-Arousal model. So, this research suggests an overall framework composed of the following six steps. The first step is a step of collecting data for modeling. The data was collected from people participated in the 2012 Seoul DMC Culture Open, and the collected data was used for the experiments. The second step extracts 64 facial features from the collected data and compensates the facial feature values. The third step generates independent and dependent variables of an artificial neural network model. The fourth step extracts the independent variable that affects the dependent variable using the statistical technique. The fifth step builds an artificial neural network model and performs a learning process using train set and test set. Finally the last sixth step is to validate the prediction performance of artificial neural network model using the validation data set. The proposed model is compared with statistical predictive model to see whether it had better performance or not. As a result, although the data set in this experiment had much noise, the proposed model showed better results when the model was compared with multiple regression analysis model. If the prediction model of audience reaction was used in the real exhibition, it will be able to provide countermeasures and services appropriate to the audience's reaction viewing the exhibits. Specifically, if the arousal of audience about Exhibits is low, Action to increase arousal of the audience will be taken. For instance, we recommend the audience another preferred contents or using a light or sound to focus on these exhibits. In other

---

\* This research is supported by Ministry of Culture, Sports and Tourism(MCST) and Korea Creative Content Agency(KOCCA) in the Culture Technology(CT) Research and Development Program 2011.

\*\* The School of Management, Kyunghee University

words, when planning future exhibitions, planning the exhibition to satisfy various audience preferences would be possible. And it is expected to foster a personalized environment to concentrate on the exhibits. But, the proposed model in this research still shows the low prediction accuracy. The cause is in some parts as follows : First, the data covers diverse visitors of real exhibitions, so it was difficult to control the optimized experimental environment. So, the collected data has much noise, and it would results a lower accuracy. In further research, the data collection will be conducted in a more optimized experimental environment. The further research to increase the accuracy of the predictions of the model will be conducted. Second, using changes of facial expression only is thought to be not enough to extract audience emotions. If facial expression is combined with other responses, such as the sound, audience behavior, it would result a better result.

**Key Words** : Emotion Determination Model, Interactive Exhibition, Valence-Arousal Model, Facial Features, Artificial Neural Network

## 저 자 소개



**정민규**

경희대학교 전자공학 학사를 취득하였으며, 현재 동 대학 경영학과 일반대학원 MIS 전공 석사 과정에 재학 중이다. 주요 관심분야는 추천시스템, 데이터마이닝, 사회 네트워크 분석 등이다. 지능정보연구, IJSSH에 논문을 게재하였으며, 2011년 2월 27일에는 싱가포르에서 열린 2011 International Conference on Social Science and Humanity(ICSSH 2011) 국제학술대회에서 논문을 발표하였다.



**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, 사회 네트워크 분석 등이다.

저탄소녹색성장국민포럼 그린IT분과 위원, 경희대학교 경영대학 BK21 사업단장, Information Technology and Management(SSCI)저널의 AE(Associate Editor)를 역임중이다.