

## 신경회로망 기반의 관객 반응 추출 시스템

백영태\*, 유은순\*\*, 박승보\*\*\*

# A System for Extraction of Audience Reaction Based on Neural Network

Yeong-Tae Baek\*, Eun-Soon You\*\*, Seung-Bo Park\*\*\*

### 요 약

관객의 감정 반응은 개별 관객들의 반응을 관측하여 판단할 수 있다. 영상 카메라를 이용한 시각적 데이터를 분석하여 활용하는 방법을 이용할 경우, 관객의 반응을 적은 비용으로 데이터를 추출하여 판단할 수 있는 것이 가능하다. 하지만, 영상 카메라를 이용할 경우 다수의 관객으로부터 발생하는 대규모의 영상 신호로부터 관객의 특징값들을 분석하여 반응을 판단하기 위한 방법과 시스템이 필요하다. 이를 위해 본 논문에서는 관객의 영상 데이터를 신경회로망 기법을 이용하여 관객의 반응을 관측하여 판단하는 방법과 시스템을 제안하고, 기존의 관객석을 고정적인 영역으로 분할하여 데이터를 취합하여 분석하기 때문에 발생하는 문제를 개선하기 위해 얼굴 검출을 통해 자동적으로 관객 영역을 분할하는 방법을 제안한다. 또한 실험을 통해 기존의 고정적으로 관객 영역을 지정하는 방법에 비해 10.5%(7.75의 Hit Ratio)의 성능향상이 있음을 확인하였다. 이를 통해 제안하는 방법이 실제 관객 반응을 판단하는데 더 적합하다는 것을 알 수 있다.

▶ Keywords : 관객 반응, 신경회로망, 관객 영역, 공연, 얼굴검출

### Abstract

Emotional reaction of audience can be decided by observing reaction of audience for content. We can use a method to analyze visual data from video camera to detect reaction of audience fast and economically. This paper proposes the method and system to observe audience reaction from visual data of audience and define via neural network. Also we propose a new method to detect automatically an area for audience reaction with face detection to improve a fixed area assignment method which has a limitation

•제1저자 : 백영태 •교신저자 : 유은순

•투고일 : 2015. 2. 11, 심사일 : 2015. 2. 16, 게재확정일 : 2015. 2. 26

\* 김포대학교 스마트콘텐츠과(Dept. of Smart Contents, Kimpo University)

\*\* 단국대학교 미디어콘텐츠연구원(Institute of Media Content, Dankook University)

\*\*\* 인하대학교 IUT 사업단(IUT, Inha University)

※ 이 논문은 2014학년도 김포대학교의 연구비 지원에 의하여 연구되었음.

not to adapt depending on audiences. Additionally, the evaluation is implemented to show that the proposed method and system is effective. The proposed method showed the performance elevation of 10.5 % (7.75 hit ration) compared to a fixed area assignment method.

▶ Keywords : Audience reaction, Neural network, Audience area, Performance, Face detection

## I. 서 론

관객의 반응을 수집하여 공연의 무대장치나 조명에 반영할 경우 관객의 몰입을 강화시켜서 공연의 재미를 증가시킬 수 있다. 더욱이 공연을 관람하는 관객의 감정반응을 파악하게 되면 콘텐츠에 대한 몰입도 뿐만 아니라 만족도를 높여줄 수 있는 인터랙션을 공연 중간이나 이후에 관객에게 제공할 수 있을 것이다. 또한 기존에는 없는 다양한 형태의 새로운 공연을 개발할 수도 있어 새로운 형태의 공연을 기획하고 생활할 수 있을 것이다.

관객의 감정 반응은 개별 관객들의 반응을 관측하여 판단할 수 있을 것이다[1]. 공연에서 관객의 반응을 측정하기 위해서는 빠르게 관객의 반응을 판단하여야 한다. 또한 어두운 관객석으로부터 반응을 판단하여야 하기 때문에 제한된 데이터만을 이용할 수 있다. 객석이나 관객에게 센서를 설치하여 추가적인 반응 데이터를 추출할 수도 있겠지만, 관객의 몰입 감 방해나 설치 비용문제로 인해 활용되기 어려운 것이 현실이다. 따라서 관객의 반응을 적은 비용으로 데이터를 추출하여 빠르게 판단할 수 있는 방법은 영상 카메라를 이용한 시각적 데이터를 분석하여 활용하는 방법뿐이다[2].

하지만 기존의 관객반응 방법 또는 시스템이 가지는 문제점은 관객으로부터 발생하는 신호를 자동으로 추출하여 관객 반응을 통합적으로 컴퓨터를 이용하여 분석하여 활용하는 방법이나 시스템이 제안되어 개발된 경우가 드물다는 것이다. 이것을 위해서는 관객들의 영역을 자동으로 감지하고 검출하는 자동화 방법이 필요하다.

영상 신호로부터 차영상을 계산하여 관객의 움직임과 파악할 수 있으며, 이를 통해 본 논문에서는 관객의 영상 데이터를 신경회로망 기법을 이용하여 관객의 반응을 관측하여 판단하는 방법과 시스템을 제안한다. 추가적으로 관객의 영역을

자동적으로 할당하기 위하여 관객의 기준 영역을 얼굴 검출 기술을 이용하는 방법을 제안한다.

관객의 반응은 관객의 행동을 통해 표출되어진다. Ambady와 Rosenthal은 인간의 의사소통을 위한 행동은 신체 행동과 얼굴 표정에 의해 판단되어질 가능성이 높다고 하였다[3]. Coulson은 화남, 억겨움, 두려움, 즐거움, 슬픔, 놀람과 같은 감정이 신체 자세에 귀속된다는 것을 실험을 통해 증명하였다[4]. Laban은 자신의 감정을 표현하기 위해 연속적인 동작을 사용한다고 하였다[5]. 이렇듯 관객의 반응을 추출하기 위해 본 논문에서는 관객의 행동을 분석하여 판정하는 방식을 사용한다.

본 연구는 시각 기반의 차영상 기법을 이용하여 관객의 움직임 관측한다. 영상 콘텐츠에 의해 유발된 사람들의 행동 변화를 통해 4가지 감정 별 행동 특징을 추출하여 각 감정에 따라 최적화된 감정 반응 모형을 구축한다. 모형 구축을 위해 한국콘텐츠진흥원 1층 로비에 실험 부스를 설치하고 다양한 사람들에게 적절한 자극영상을 제공하였다. 이때 사람들의 행동 반응을 영상 카메라를 통해 수집하였다. 또한 실험대상자의 영상을 9개 영역으로 구분하고 행동 데이터를 추출, 보정한 이후 신경회로망 모형을 통해 학습시키는 기존의 연구 방법을 활용하였다[6].

기존의 관객 반응을 분석하는 연구는 관객의 영역을 추출하기 위해 수작업으로 영역을 지정하는 방법을 이용하였다[6]. 관객반응을 추출하기 위해서 수작업으로 관객의 영역을 지정하는 방법은 실제 공연환경에 적용하기에는 많은 문제점이 있다. 그래서 고정된 관객석을 기반으로 하여 사전에 미리 균일한 관객영역을 지정하는 방법이 사용될 수 있다. 관객석이 고정되어 있고, 균일한 간격으로 관객석이 배치되어 있기 때문에 균일한 관객 영역을 지정하는 것이 분석을 위해 특정 환경에서 적합할 수 있다. 하지만 관객별로 체격이나 앉는 자세의 개인차가 발생하여 획일적인 영역 설정 방식은 반응 판정에 오류를 발생시키는 원인이 될 수 있다. 또한 관객이 공

연 중간에 자세를 바꿀 경우 획일적인 영역 지정방식은 관객 영역을 재배치 못하는 한계를 갖고 있다. 그래서 기존의 관객의 반응을 관측하여 판단하는 방법이 관객석을 고정적인 영역으로 분할하여 데이터를 취합하여 분석하기 때문에 발생하는 문제를 개선하기 위해 얼굴 검출을 통해 자동적으로 관객 영역을 분할하는 방법을 제안하고 실험을 통해 성능을 검증한다.

제안하는 연구의 설명과 검증을 위해 구성된 본 논문은 다음과 같이 이루어져 있다. II 장에서는 관객 반응 행동 추출과 관련된 연구를 소개하고 III 장에서는 논문에서 제안하는 관객 반응 행동 추출 방법에 대해 설명한다. 그리고 IV 장에서는 제안하는 방법에 대한 실험 방법에 대해 설명하고 성능 평가에 대해 고찰한다. 그리고 결론 및 향후 연구에 대해 기술한다.

## II. 관련 연구

### 1. 차영상 (Differential Image)

차영상은 카메라를 통하여 입력되는 연속적인 영상 프레임(Frame)의 값 변화량을 의미한다. 앞뒤 프레임 간의 차이를 추출하는 것으로 프레임 내의 특정 대상물의 움직임이 있을 경우 차영상이 발생하게 된다. 차영상은 검출 방법이 간단하며, 처리 속도가 빨라서 이동객체 추적 시스템 분야에서 사용된 이후에 [7]과 보안 분야[8]과 로봇 분야, 영상 인식 분야에서 활용된 후에 다양한 분야에서 이용되고 있다. 본 논문은 그림 1과 같이 차영상 기법에 의해 우측의 하얗게 표시된 검출된 움직임을 총합하여 시간적인 변화량으로 나타내 주는 방식을 사용하여 관객의 움직임량을 검출한다[2].

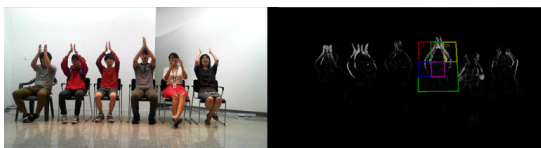


그림 1. 관객 영상에 적용된 차영상  
Fig. 1. Differential Image Applied to Audience Video

### 2. 동작 인식을 통한 감정 인식 연구

영상정보를 이용한 동작 인식 연구는 영상 카메라만을 이용해 정보를 빠르게 수집할 수 있기 때문에 활발한 연구가 이루어지고 있다[9,10,11].

Ambady와 Rosenthal는 인간이 의사소통 시에 행동과

얼굴의 움직임을 통해 의사 전달한다고 연구를 통해 밝혔으며 [3]. Laban은 인간의 움직임을 조합한 8가지 행동에 대한 감정과의 관련도를 계산하는 연구를 진행하였다[4]. Coulson은 감정(화남, 억거움, 두려움, 즐거움, 슬픔, 놀람)이 감정적 자세와 관련되었다는 연구를 진행하였으며[12], 자세를 통한 인간의 감정인지가 가능하다고 결론지었다. 조경은은 애니메이션의 캐릭터 표현을 위해 기분은 자세로 감정은 행동을 이용하여 표현할 수 있다는 연구를 진행하였다[13]. 유은정 등은 콘텐츠에 대한 관객의 감정 반응을 행동을 이용하여 분석하는 연구를 진행하였다[6].

### 3. 관객 반응 행동 연구

류준모 등은 공연에서 관객의 반응 정도를 판단하기 위해 행동 동기화 기법을 이용하여 관객 그룹의 몰입도를 판단하는 연구를 진행하였다[2]. 공연으로부터의 자극에 대한 관객의 반응이 시간적으로 일치하여 나타날 경우 관객 전체의 몰입도가 증가하는 것으로 판단하여 관객 전체의 반응 정도를 판단하였다. 하지만 이 방법은 관객의 몰입도를 통한 반응의 정도만을 판단하는 방법으로 반응의 종류를 판단하기에는 적합하지 않은 방법이다.

유은정 등은 관객의 반응의 종류를 파악하기 위해 신경회로망을 이용하는 방법을 적용하였다. 자극 영상에 의해 발생된 관객 반응을 슬픔(sad), 놀람(surprised), 즐거움(joy), 혐오(disgust)의 4가지로 구분하는 감정 모형을 구축하고 동적인 움직임 값의 입력 변수가 시계열 구간의 변화에 따라 감정별로 어떻게 관련되는지에 대한 연구를 진행하였다[6]. 이 연구는 관객의 반응의 종류를 판별하기 때문에 관객의 반응에 따른 적절한 인터랙션을 제공할 수 있는 장점이 존재한다. 하지만 이 연구는 어떻게 관객의 반응을 분류할 것인가에만 집중된 나머지 관객 반응 영역을 지정하기 위한 방법을 수작업으로 지정해야 하는 문제가 있다.

### 4. 신경 회로망(Neural Network)

McCulloch와 Pitts에 의해 최초로 제안된 신경 회로망은 인간의 뇌의 뉴런의 작동 방식을 모방하여 생성한 개념이다[14]. 신경 회로망은 노드(Node)와 가중치(Weight)를 가진 라인(Line)으로 구성된 망(Network) 구조로 그림 2와 같이 모형화된다. 또한 과거에 수집된 데이터로부터 가중치들에 대한 반복적인 교사 학습과정을 거쳐 데이터에 숨어 있는 패턴을 학습을 통해 찾아낸다. 신경 회로망 기법은 영상 분석 및 인식, 음성 분석, 기상 예측, 로봇 제어, 경영 분석 등 다양한 분야에 적용된다. 신경 회로망은 일반적으로 한 개씩의

입력층(Input layer)과 출력층(Output layer) 그리고 한 개 이상의 은닉층(Hidden layer)이 존재하는 그림 2와 같은 형태로 구성된다(6).

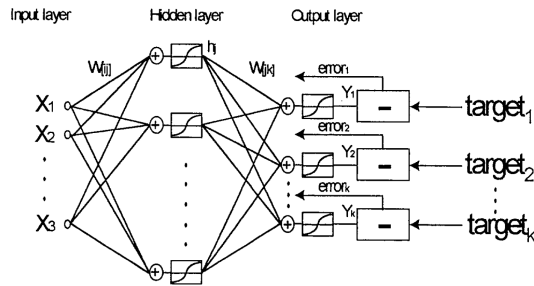


그림 2. 신경 회로망의 구조  
Fig. 2. Structure of a standard neural network

그림 2와 같은 신경 회로망 학습에는 역전파 알고리즘이 많이 활용된다. 역전파 알고리즘은 목표치(target)와 신경회로망의 출력(y)의 오차분을 출력층을 통하여 앞으로 역전이 시켜주며 가중치들을 변경시켜주는 방식이다(15,16).

신경 회로망을 이용하여 감정을 인식한 연구는 홍동표의 연구(17)과 유은정 등의 연구(6)이 있다. 이 연구들은 인간의 행동을 이용하여 감정을 분석하기 위해 신경 회로망 기법을 이용하였다.

### III. 관객 반응 추출 시스템

본 연구에서 제안하는 관객 반응 분석 시스템은 그림 3과 같이 비디오 카메라 인터페이스(Interface)와 관객 영역 할당 모듈(Assignment Module of Audience Area), 데이터 분석기(Data Analyzer), 관객 감정 검출기(Emotion Detector of Audience)의 네 부분으로 이루어져 있다.

비디오 카메라 인터페이스는 관객을 촬영하는 비디오 카메라로부터 영상 신호를 수집하는 모듈이다. 관객영역은 일반적으로 어둡기 때문에 적외선 카메라를 활용하여 촬영을 진행하게 되지만 그 출력신호는 일반적인 카메라의 신호와 동일하게 취급되기 때문에 인터페이스는 범용의 코덱(CODEC)을 사용하여 프레임별로 변환하여 관객 영역 할당 모듈로 전송하게 된다.

관객 영역 할당 모듈은 전송되는 프레임별로 관객의 영역을 추출하고 이전 프레임과의 차영상을 계산하여 관객의 시계열적 움직임 값(Moving Value Graph)을 수집하여 데이터 분석기로 전송하는 역할을 한다. 관객의 영역은 얼굴 검출을

이용하여 자동으로 지정하게 된다.

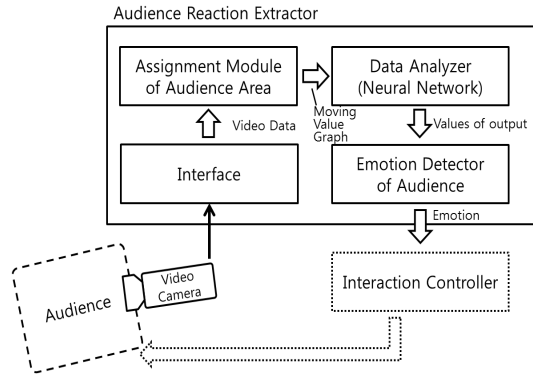


그림 3. 관객 반응 추출 시스템의 구조도  
Fig. 3. System Architecture for Detecting Audience Reaction

데이터 분석기는 신경 회로망 기반으로 이루어져 있으며, 역전파 학습으로 학습을 진행하게 된다. 관객의 움직임 값은 시계열적으로 변화하는 값으로 신경 회로망의 입력노드로 입력된다. 데이터 분석기의 신경회로망은 네 가지 감정상태가 출력 노드로 지정되어 있다. 따라서 데이터 분석기의 출력은 네 가지 노드 중에 하나가 선택되는 분류기가 된다.

관객 감정 검출기는 데이터 분석기로부터 선택된 출력에 대해 해당하는 감정 상태를 결정하는 역할을 한다. 예외 상황이 발생할 경우에 대한 처리도 규칙 기반으로 지정되어 있다. 관객 감정 검출기의 출력은 공연 시스템의 인터랙션 제어기(Interaction Controller)로 전송되어 공연을 제어하는 데 이용된다.

본 논문에서 제안하는 관객 반응 추출 시스템에서 네 가지 모듈 중에서 범용적인 인터페이스 모듈을 제외한 관객 영역 할당 모듈과 데이터 분석기와 관객 감정 검출기에 대해 설명한다.

#### 1. 관객 영역 할당 모듈

관객 영역 할당 모듈은 그림 4와 같이 비디오 데이터로부터 관객 개개인에 대한 영역을 자동으로 할당하고, 영역을 다수의 셀로 분할하는 역할을 한다. 각 셀들에 대한 차영상을 히스토그램으로 변환하여 특정 밝기 이상의 값들만을 취합하여 차영상의 대푯값으로 변환한다. 이것을 시계열적으로 연결하여 셀의 움직임 값의 집합으로 표현한 후에 데이터 분석기로 전송한다. 관객의 영역을 할당하기 위해 본 연구에서는 얼굴 검출 방법을 활용하였다.

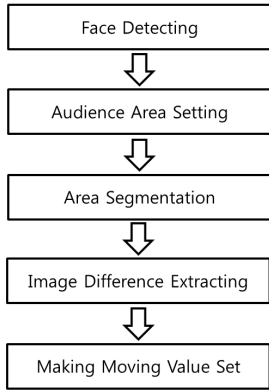


그림 4. 관객 영역 할당 흐름도  
Fig. 4. Flow Chart to Assign Audience Area

OpenCV의 Haar 기반의 얼굴 검출 API[18,19]을 활용하여 얼굴을 검출한 후에 그림 5와 같이 검출된 얼굴영역을 기준으로 관객영역을 검출하였다.

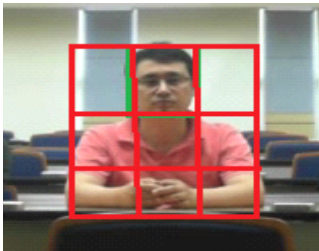


그림 5. 얼굴 검출 기반의 관객 영역 할당 결과  
Fig. 5. A Result Screen to Assign Audience Area based on Face Detection

얼굴 영역에 대해 좌우로 1배씩 확장하고, 아래로 2배를 확장하여 그림 5와 같이 기존 연구와 같은 9개의 셀로 구성된 관객 영역을 할당하였다[6].

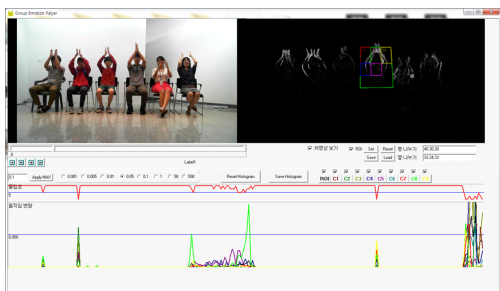


그림 6. 관객 영역의 각 셀별 시계열 움직임 그래프  
Fig. 6. Moving Value Graph for Each Cell of Audience

이 셀들에 대해 해당하는 움직임 값을 추출한 결과는 그림 6의 하단의 그래프와 같이 시계열 적으로 표현되어진다. 시스템에서는 각 셀들에서 나타나는 시계열적인 그래프들을 동시에 추출하여 보여주도록 하였으며, 추출된 값들을 데이터 분석기에서 활용할 수 있도록 텍스트 파일로 저장하여 호출 할 수 있도록 하였다.

## 2. 데이터 분석기

관객의 반응을 검출하기 위해 본 연구에서는 신경 회로망을 데이터 분석기로 활용하였다. 신경 회로망의 설계는 그림 7과 같이 입력층과 출력층 그리고 은닉층의 3층 구조로 구성하였다.

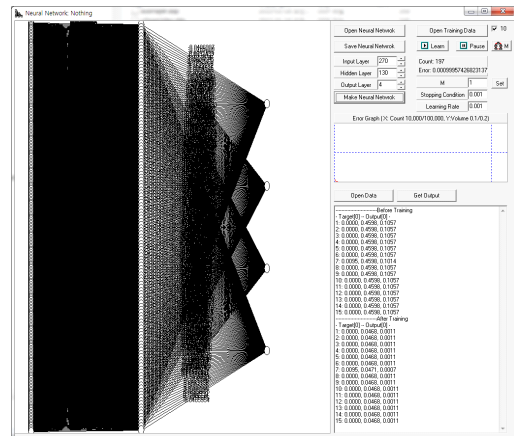


그림 7. 신경 회로망 기반의 데이터 분석기  
Fig. 7. Data Analyzer based on Neural Network

입력층의 노드 개수는 9개 셀에서 발생하는 30개 프레임에 대한 움직임 값을 한번에 처리하기 위해 270개의 입력 노드로 구성하였다. 저장된 영상이 1초당 10프레임이기 때문에 30프레임은 3초의 영상을 의미한다. 즉 3초 이내에 관객 반응이 완료되기 때문에 3초의 영상으로 입력을 결정하였다 [6]. 출력층은 네 개의 감정 상태(슬픔, 놀람, 기쁨, 억겨움)를 판별하기 위해 4개의 출력 노드로 구성하였다. 4개의 노드 중 하나는 1의 값을 갖도록 학습을 진행하였으며, 모두 0의 값을 가질 경우 아무런 반응이 없는 것으로 설정하였다. 4개의 노드 중에 하나에서 1이 검출될 경우 해당하는 감정 상태에 있는 것으로 판단하였다.

## 3. 관객 감정 검출기

관객의 반응은 4가지 감정 상태로 검출되어진다. 데이터

분석기의 신경 회로망의 출력 노드 중 1개의 노드에서만 1이 검출될 경우 해당하는 감정 반응이 있는 것으로 판단한다. 하지만 출력 노드 중 2개 이상에서 1의 값이 검출될 경우 데이터 분석기가 올바르게 작동하지 않은 것으로 판단하여 분석 실패로 지정되게 된다. 출력 노드 4개 모두에서 0이 출력될 경우 감정이 표출되지 않은 중립 상태로 판단한다. 정리하면 다음 표 1과 같다.

표 1. 관객 감정 검출기 판단 규칙  
Table 1. Decision Rule for Emotion Extraction of Audience

Output Node				Emotional State
1	2	3	4	
1	0	0	0	Sad
0	1	0	0	Surprised
0	0	1	0	Joy
0	0	0	1	Disgust
0	0	0	0	Neutral
Etc.				Error

#### IV. 실험 및 평가

##### 1. 데이터 수집 및 실험 설계

본 연구를 위해 일반인을 대상으로 하는 실험을 진행하였다. 데이터는 2012년 9월 13일부터 15일까지 콘텐츠진흥원 1층 로비에 부스를 설치하여 수집하였다. 대상자에게 감정 별 영상을 노출시킨 후에 반응하는 대상자의 영상을 촬영하였다. 이후에 영상 시청 시에 느꼈던 감정을 장면별로 설문조사를 통해 수집하였다. 오류 데이터를 제외하고 총 97개의 데이터를 슬픔, 놀람, 즐거움, 혐오의 4가지 감정 상태에 대해 촬영되었다. 실험 데이터는 기존 방법과의 비교를 위해서 유은정 등이 논문에서 사용한 실험데이터를 사용하였다[20].

실험 대상자의 영상은 1초당 10프레임의 간격으로 촬영하였으며, 각 감정별로 3초(30프레임)의 움직임 값을 데이터 분석기의 입력변수로 설정 하였다. 이것은 유은정 등의 연구 결과[6]에서 나타난 최상의 프레임수를 활용하였기 때문이다. 따라서 9개의 신체 영역에 대해 30개의 값이 만들어지므로, 총 270개의 입력 노드를 갖도록 신경회로망을 구성하였다. 감정 판단을 위한 신경 회로망 학습 방법으로는 역전파 알고리즘을 사용하였다. 그리고 신경 회로망의 학습을 위한 파라미터들인 학습률(Learning rate)과 모멘텀률(Momentum rate)은 각각 10%로 설정하였다. 각 노드에

존재하는 변환함수로는 로지스틱(Sigmoid)함수를 사용하였으며, 계층구조는 한 개씩의 입력층, 출력층, 은닉층으로 구성된 3계층 구조로 설계하였다. 학습 중지점을 탐색하기 위해 테스트 데이터 집합을 기준으로 학습시에 최소 오류 도달 후 50,000번 도달 시 중지하도록 하였다[6]. 신경 회로망의 설정에 대한 방법은 유은정 등의 연구방법을 따랐다. 시스템 개발을 위해 신경 회로망은 인프라이즈 사의 델파이 7을 이용하여 개발하였다.

##### 2. 실험결과 및 고찰

제안하는 시스템의 성능을 검증하기 위해 관객의 좌석에 대해 고정된 관객 영역 할당 방법과 얼굴 검출에 의한 자동 관객 영역 할당 방법을 비교하였다. 데이터의 분석기의 학습은 수작업으로 영역을 할당하여 발생시킨 데이터를 이용하여 진행하였다. 데이터 분석기의 학습 후 성능을 최선으로 유지하기 위해 수작업으로 영역을 할당한 데이터에 의해 학습을 진행하였다. 실험 결과는 표 2와 같이 나타났다.

표 2. 실험 결과  
Table 2. Experimental Result

	Sad	Surprised	Joy	Disgust	Hit Ratio
Sad	51	1	10	35	51
Surprised	21	72	3	1	72
Joy	-	1	88	8	88
Disgust	-	3	10	84	84

(a) Fixed Audience Area  
(a) 고정 관객 영역 방식

	Sad	Surprised	Joy	Disgust	Hit Ratio
Sad	62	-	6	29	62
Surprised	13	81	3	-	81
Joy	-	-	90	7	90
Disgust	-	-	4	93	93

(b) Automatic Audience Area  
(b) 자동 관객 영역 방식

	Method of Assigning	
	Fixed Method	Automatic Method
Sad	51	62
Surprised	72	81
Joy	88	90
Disgust	84	93
Average	73.75	81.50

(c) Comparison of Automatic and Fixed Audience Area  
(c) 고정과 자동 관객 영역 방식의 비교

표 2의 a는 경우 고정 관객 영역 방식을 이용하였을 때의

실험 결과로 c의 표에서 보여지는 것처럼 평균적으로 73.75의 Hit Ratio의 결과를 보여주었다. 관객이 자세를 바꿀 때 관객 영역에서 벗어나는 경우가 발생하였다.

표 2의 b는 제안하는 자동 관객 영역을 이용하였을 때의 실험 결과로 c의 표에서 보여지는 것처럼 평균적으로 81.50의 Hit Ratio를 보여주었다. 관객이 자세를 바꿀 경우에 자동적으로 관객영역을 적응하기 때문에 Hit Ratio가 증가하는 것으로 판단된다.

표 2와 같이 자동으로 관객 영역을 검출하여 적용할 경우 좌측 기반의 고정된 관객 반응 검출의 성능이 더 우수한 것을 알 수 있다. 평균적으로 10.5 % (7.75 점의 Hit Ratio)에 대한 성능 향상이 나타나는 것으로 파악되었다. 이것은 얼굴 검출을 이용해 관객의 자세변화나 체형에 맞게 관객 반응 영역을 설정하고, 영역 내부의 셀을 적절하게 분할했기 때문으로 판단된다. 하지만 얼굴 검출에 실패하는 경우가 발생하여 관객 반응영역을 엉뚱한 곳에 설정하는 경우가 있었다. 이의 방지를 위해 규칙을 설정할 경우 성능향상을 이룰 것으로 예상된다. 또한 유은정 등의 연구에서 지적인 것처럼 감정별로 반응 시간이 다르기 때문에 감정별로 프레임 수를 차별적으로 선택하면 관객 반응의 성능 향상을 나타낼 것으로 예상된다.

## V. 결론 및 향후 연구

본 논문에서는 관객의 영상 데이터를 신경회로망 기법을 이용하여 관객의 반응을 관측하여 판단하는 방법과 시스템을 제안하였다. 또한 관객 영역을 얼굴 검출 기술을 이용하여 자동으로 할당하는 방법을 제안하고 시스템으로 구현하였다. 제안하는 방법과 시스템의 성능을 평가하기 위해 고정적 형태의 관객 영역 할당 방법과 비교를 진행하였다. 제안하는 방법은 고정적 형태의 관객 영역 할당 방법에 비해 10.5% (7.75 만점의 Hit Ratio)의 성능이 뛰어난 것으로 나타났다. 이것은 제안하는 방법과 시스템이 기존의 방법보다 성능이 우수한 것을 의미한다.

하지만 관객 영역 할당 모듈에서 얼굴 검출 시에 오류가 발생하여 관객 반응 추출 성능을 떨어뜨리는 경우가 발생하였다. 이것은 탐색영역을 관객석 내로 한정하거나 추가적인 제약조건을 지정하여 성능 향상을 이룰 수 있을 것으로 예상된다. 따라서 향후에는 관객 반응 추출 성능 향상을 위해 관객 영역을 좀 더 정확하게 찾아낼 수 있는 방법에 대한 추가적인 연구를 진행할 필요가 있다.

## REFERENCES

- [1] S. G. Barsade, and D. E. Gibson, "Group emotion: A view from top and bottom," *Research on Managing Groups and Teams*, Stanford, CT: JAI Press, Vol.1, pp. 81-102, 1998.
- [2] J. M. Ryu, S.-B. Park, J. K. Kim, "A Study of the Reactive Movement Synchronization for Analysis of Group Flow," *Journal of Intelligent Information System*, Vol. 19, No. 1, pp. 79~94, March 2013.
- [3] N. Ambady, and R. Rosenthal, "Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis," *Psychological Bulletin*, Vol. 111, No. 2, pp. 256-274, 1992.
- [4] M. Coulson, "Attributing emotion to static body postures: recognition accuracy, confusions, and viewpoint dependence," *Journal of Nonverbal Behavior*, Vol. 28, No. 2, pp. 117-139, 1992.
- [5] R. Laban, "*Modern Educational Dance*," Macdonald & Evans, 1968.
- [6] E. C. Ryoo, H. Ahn, J. K. Kim, "The Audience Behavior-based Emotion Prediction Model for Personalized Service," *Journal of Intelligent Information System*, Vol. 19, No. 2, pp. 73-85, June 2013.
- [7] M. K. Leung, and Y. H. Yang, "Human Body Motion Segmentation in A Complex Scene," *Pattern Recognition*, Vol. 20, pp. 55-64, 1987.
- [8] J. S. Lee, and M. G. Lee, "A Study on Pattern Recognition using DCT and Neural Network," *The Korean Institute of Communications and Information Sciences*, Vol. 22, pp. 481-492, 1997.
- [9] J. W. Davis, "Appearance-Based Motion Recognition of Human Actions," *Technical Report*, MIT, 1996.
- [10] A. P. Dempster, 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, pp. 1-38, 1977.

[11] Y. Zhu, 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, pp. 573-576, 1998.

[12] M. Coulson, “Attributing emotion to static body postures: recognition accuracy, confusions, and viewpoint dependence,” Journal of Nonverbal Behavior, Vol. 28, No. 2, pp. 117-139, 1992.

[13] C. Lee, K. Um, and K. Cho, “Design of a Character’s System to Express Mixed-Emotion by Gesture and Posture,” The Proceedings of Spring on Korea Multimedia Society 2009, pp. 191-194, 2009.

[14] W. S. McCulloch, and W. Pitts, “A Logical Calculus of the Ideas Immanent in Nervous Activity,” The Bulletin of Mathematical Biophysics, Vol. 5, No. 4, pp. 15-33, 1943.

[15] S.-B. Park, “The Controller Design for Truck Backer-Upper By ART2 Neural Network,” Master Thesis, Inha University, February 1997.

[16] D. Kim, “Theory and Application of Neural Network,” Hitech Information, 1993.

[17] D. Hong, and W. Woo, “Emotional User Interface of Based on Augmented Reality for I2-NEXT,” The Proceedings of Korea HCI Society 2005, pp. 96-101, 2005.

[18] J.-G. Jeong, S.-S. Park, and D.-S. Jang, “Real-Time face detection using the Skin color and Haar-like feature,” Journal of the Korea society of computer and information, Vol. 10, No. 4, pp. 113-121, September 2005.

[19] Y. T. Baek, S.-B. Park, “Shot Type Detecting System using Face Detection,” Journal of The Korea Society of Computer and Information, Vol. 17 No. 9, pp. 49-56, September 2012.

[20] S.-B. Park, J. M. Ryu, J. K. Kim, “A Group Arousal Analysis Based on the Movement

Synchronization of Audiences,” Multimedia Tools and Applications, Online published, 2014.

## 저 자 소개

### 백 영 태



1989: 인하대학교 전자계산학과 이학사  
 1993: 인하대학교  
 전자계산공학과 공학석사  
 2002: 인하대학교  
 전자계산공학과 공학박사  
 1993-1998 : 대상정보기술(주)  
 정보통신연구소 선임연구원  
 1998-현 재: 김포대학교  
 스마트콘텐츠과 교수  
 관심분야: 멀티미디어 정보검색,  
 웹교육 시스템,  
 모바일 시스템  
 Email : hannaee@kimpo.ac.kr

### 유 은 순



1995: 인하대학교  
 불어불문학과 문학사  
 2000: Franche-Comté 대학교  
 언어학 석사  
 2007: Franche-Comté 대학교  
 언어학 박사  
 현 재: 단국대학교 미디어콘텐츠연구원  
 리서치펠로우  
 관심분야: 스토리텔링, 추천시스템,  
 소셜미디어, 빅데이터, 감성  
 Email : tesniere@naver.com

### 박 승 보



1995: 인하대학교 전기공학과 공학사  
 1997: 인하대학교  
 전기공학과 공학석사  
 2011: 인하대학교  
 정보공학과 공학박사  
 현 재: 인하대학교 IUT 사업단 연구교수  
 관심분야: 멀티미디어 정보검색,  
 스토리 공학, 감성 검색  
 Email : molaal@naver.com