

가상 인간의 감정 표현 인식을 위한 비언어적 다중모달 영향 분석

Impact Analysis of nonverbal multimodals for recognition of emotion expressed virtual humans

김진옥*
Kim Jin Ok

요약

디지털 콘텐츠에서 HCI로 활용되는 가상 인간은 얼굴 표정과 신체자세와 같은 모달을 이용하여 다양한 감정을 표현하지만 비언어적 다중모달의 조합에 대한 연구는 많지 않다. 감정을 표현하는 가상 인간을 제작하려면 계산 엔진 모델은 얼굴 표정과 신체자세와 같은 비언어적 모달의 조합이 사용자에 의해 어떻게 인식되는지를 고려해야 하기 때문에 본 연구는 가상 인간의 감정 표현 디자인에 필요한 비언어적 다중모달의 영향을 분석하여 제시한다. 먼저 가상 인간에 대한 다중모달 별 감정 인식을 평가하여 다른 모달 간의 상대적 영향성을 분석하였다. 그리고 일치하는 얼굴과 자세 모달을 통해 기본 감정 및 정서와 활성화 인식에 대한 영향을 평가하며 감정이 불일치하는 다중모달을 통해 일상생활에서 빈번하게 드러나는 중첩된 감정의 인식 정도를 관측하였다. 실험 결과, 가상 인간의 얼굴과 신체자세의 표정이 일치하면 감정 인식이 용이하며, 얼굴 표정으로 감정 카테고리를 판별하지만 감정의 활성화 차원 판단에는 자세 모달리티가 선호됨을 확인하였다. 본 연구 결과는 감정을 드러내는 가상 인간의 행동 동기화 및 애니메이션 엔진 시스템 구현에 활용할 수 있다.

ABSTRACT

Virtual human used as HCI in digital contents expresses his various emotions across modalities like facial expression and body posture. However, few studies considered combinations of such nonverbal multimodal in emotion perception. Computational engine models have to consider how a combination of nonverbal modal like facial expression and body posture will be perceived by users to implement emotional virtual human. This paper proposes the impacts of nonverbal multimodal in design of emotion expressed virtual human. First, the relative impacts are analysed between different modals by exploring emotion recognition of modalities for virtual human. Then, experiment evaluates the contribution of the facial and postural congruent expressions to recognize basic emotion categories, as well as the valence and activation dimensions. Measurements are carried out to the impact of incongruent expressions of multimodal on the recognition of superposed emotions which are known to be frequent in everyday life. Experimental results show that the congruence of facial and postural expression of virtual human facilitates perception of emotion categories and categorical recognition is influenced by the facial expression modality, furthermore, postural modality are preferred to establish a judgement about level of activation dimension. These results will be used to implementation of animation engine system and behavior synchronization for emotion expressed virtual human.

☞ keyword : non-verbal multimodal(비언어적 모달), emotion recognition(감정 인식), emotional modal(감정 모달), virtual human(가상 인간)

I. 서론

감정은 신체 기관이 의미 있는 자극에 대응하여 인식과 인식 처리, 주관적 느낌, 행동 경향, 생리적 변화, 운동신경 표현의 5개 관련 요소를 서로 연관시키고 동시에 발생시키는 사건이다[1]. 감정 정보는 얼굴 표정 음성, 언어, 자세, 머리 움직임, 신체 움직임, 자세를 포함한 광범

위한 모달을 통해 전달되므로 감정 관련 다중모달의 표현 동시성이 감정 인식에 중요함에도 감정 인식에 대한 대부분의 연구는 얼굴 표정 및 신체 자세의 개별적인 모달을 통해 감정 인식을 수행하고 있다[2]. 또한 감정 인식에 대한 이전 연구들은 감정 표현을 담당하는 개별 채널에 집중하면서 의도적이고 과장된 표정만을 주로 인식하며 사람들이 일반 환경에서 만나는 복합적인 다중모달 감정 신호는 그다지 고려하지 않는 문제점을 보이고 있다. 실제 환경에서 사람들은 한 가지 이상의 모달을 통해 감정을 표현하며 얼굴과 자세와 같은 비언어적 모달이 중요한 역할을 한다.

* 중신회원 : 대구한대학교 국제문화정보대학 모바일콘텐츠학부 교수 bit@du.ac.kr

[2012/07/25 투고 - 2012/07/30 심사 - 2012/09/ 심사완료]

디지털콘텐츠 분야에서는 가상 인간이 HCI의 새로운 방법으로 대두되면서 감정을 잘 표현하는 가상 인간을 필요로 한다[3]. 가상 인간을 실제 사람의 모습 및 행동과 일치하게 디자인하려면 비언어적 모달을 통해 자연스럽게 감정을 전달하는 방법이 필요하다[4]. 지금까지의 가상 인간은 주로 얼굴 표정 모달을 통해 감정을 표현하고 있으나 활발한 신체 표현으로도 감정을 다양하게 드러내야 한다. 이에 따라 가상 인간의 특징을 포착해 생성하는 컴퓨터 계산 모델은 감정을 드러내는 여러 모달이 조화롭게 인식되는 과정을 파악하고 사용자가 여러 모달을 통합하여 가상 인간의 감정 표현을 이해하도록 해야 한다. 또한 자동화된 가상 인간의 자세 합성 과정에서 감정을 자세로 모핑하는 과정과 신체 자세 등 감정을 나타내는 비언어적 모달 간의 상관관계를 반영할 필요가 있다.

가상 인간의 감정 표현에 대한 이러한 필요성을 바탕으로 본 연구는 가상 인간이 감정을 표현하는 얼굴과 신체 자세를 모달로 이용하여 다양한 감정을 표현하는 방법을 분석한다. 또한 일상생활에서 빈번하게 나타내는 중첩된 감정의 인식 과정에서 서로 일치하는 않는 얼굴과 자세표정이 감정 인식에 미치는 영향을 다룬다. 특히 가상인간의 감정 크기를 인식하는데 영향을 끼치는 자세 모달리티를 새롭게 분석한다.

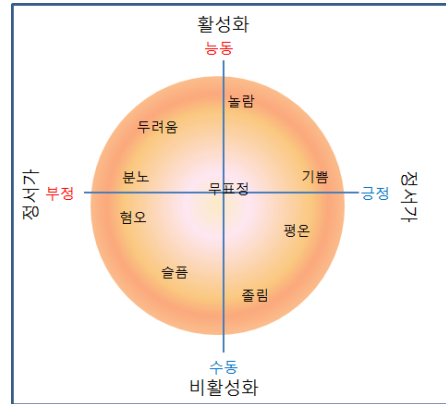
본 연구의 II장 1절에서는 기존 감정 인식에 대한 연구를 조사하고 2절에서는 얼굴, 자세, 얼굴+자세 각 모달이 기본 감정과 감정의 활성화/정서가 인식에 미치는 영향을 실험한다. 3절에서는 감정 표현 시 다중 모달 간 일치와 불일치가 기본 감정과 감정의 활성화/정서가 인식에 미치는 영향을 실험한다. III장에서는 가상 인간 제작에 필요한 비언어적 행동의 모달의 영향력과 중첩된 감정 인식을 대한 실험 결과를 정리하고 향후 연구 방향을 제시한다.

II. 연구 내용

1. 관련 연구

1.1. 다양한 감정 표현 상태 분류

에크만(Ekman)이 감정 현상으로부터 주요 감정 카테고리 구별하는 연구를 제안한 이후 기쁨, 놀람, 두려움, 분노, 슬픔, 혐오 외에 경멸 감정이 기본 감정에 추가되기도 했으며 이자르(Izard)[5]의 연구에서처럼 에크만의 기본 감정에 관심, 부끄러움, 죄책감 같은 감정을 기본 감정으로 추가한 내용도 제시되었다. 최근 감정을 연구



(그림 1) 러셀의 감정 원형 모델

하는 연구자들은 감정의 다양성을 설명하면서 개별 감정 상태를 독립적이지 않고 시스템적으로 서로 관련이 있다고 본다. 이 가운데 러셀(J. Russell)은 즐거움과 고통, 흥분과 침체에 대한 기준을 설정하고 (그림 1)과 같은 2차원 감정 원형 모델(A circumplex model of affect)을 제안하여 감정을 독립적이면서도 차원을 달리한 다양한 감정 상태로 분류하였다[6].

러셀의 감정 원형 모델은 구체적인 감정을 개별적으로 분리하고 개별 감정 하부의 잠재된 상태를 특징별로 설정하여 감정을 2차원적으로 설명하고 있다. 그는 2차원 감정 원형 모델의 세로방향 양극은 활성화(activation, aroused, 능동/수동)상태로, 가로방향 양극은 정서가(valence, pleasant, 긍정/부정 또는 쾌/불쾌)로 나누고 여러 감정을 연속적으로 배치하여 감정을 독립적이면서도 연속적인 상태로 설명하고 있다. 이에 따라 감정의 활성화와 정서가 정도에 따라 다변화되고 중첩되는 감정 상태를 쉽게 이해할 수 있다. 이 외, 클라우스 슈러(Klaus Scherer)는 다양한 실험을 통해 감정의 중첩 상태를 제시하여 일상생활에서는 도드라진 감정 한 두개만 존재하는 것이 아니라 동시에 몇 가지 감정이 여러 표정 모달에 섞여 드러남을 설명하였다[7].

1.2 얼굴과 자세를 통한 감정 표현

실험심리학 연구에서는 감정을 주로 얼굴 표정 집합으로 표현하며 얼굴을 통해 드러나는 표정 수는 감정에 따라 다르다고 본다[8]. 기쁨과 경멸은 표정으로 드러내는 방법이 적고 놀람과 혐오에는 약 10개의 다양한 대표 표정이 존재하며 두려움, 슬픔, 분노에는 20개의 표정이

있다. 이렇게 나타난 표정을 인식하는 인식을 역시 감정에 따라 다르다. 기쁨, 분노, 놀람과 슬픔은 잘 인식하는 반면 두려움과 혐오는 인식이 낮다[9].

감정을 드러내는 다양한 요소를 합성하여 정확히 감정을 인식하는 것에 대한 중요성은 익히 알려져 있고[1] 감정 표현의 확장과 통합 또한 중요하지만 감정 표현 시 얼굴과 신체 요소를 어떻게 조화시키는지, 감정 인식에 여러 모달이 어떤 역할을 하는지에 대한 분석 연구는 많지 않다.

감정을 표현할 때 신체 부위별 위치와 동작은 얼굴 표정과 음성 감정을 더 상세하게 이해하고 해석하는 데 도움이 된다[10]. 신체 자세가 어떻게 감정을 표현하는지를 연구한 기존 연구에 따르면 신체동작은 감정을 느끼는 정도에 따라 다른 정보를 전달하며[11, 12] 각 자세에는 감정을 구분하는 판별 특징이 존재하기 때문에] 월보트(Wallbot)[13]는 자세의 특징을 관측하여 감정 카테고리 간에 판별 가능한 움직임을 정량적으로 측정된 결과를 제시하였다. 얼굴 표정이나 자세 등 한 가지 요소보다 얼굴 표정과 자세의 감정을 같이 이해할 때 감정 상태를 더 잘 인식될 수 있음은 귄(Gunes)과 피카디(Piccardi)[10]가 가장 먼저 설명했으며 이 연구를 바탕으로 쉬러(Scherer)와 엘그링(Ellgring)[14]는 목소리, 얼굴, 신체 모달을 이용해 감정 표현을 비교한 결과를 제시했다. 이들의 연구에서 얼굴 움직임은 FACS로 코딩하고 신체표현의 코딩은 기능과 해부학적 특성을 이용하여 격한 분노, 질책, 극도의 두려움, 염려, 절망, 슬픔, 큰 기쁨, 행복, 관심, 지루함, 수치, 긍지, 혐오와 경멸 등 같은 감정을 다른 강도로 표현하는 감정그룹을 분석한 결과 얼굴과 자세를 이용한 다중모달 반응 패턴은 긍정적 감정일 때 가장 빈번하게 나타난다.

히테넨(Hietanen)과 래파넨(Lappanen)[15]은 정적인 상태의 얼굴 표정과 능동적인 상태의 손 움직임을 통합한 인식과정을 연구했다. 감정을 신체 모달 별로 서로 일치하거나 상반되게 설정하여 분석한 결과 감정 판단은 주로 얼굴 표정에 근거하여 이루어지지만 손의 표현에도 영향을 받았다. 손의 움직임 크기에 얼굴 표정을 연관시켜 관찰한 결과 기쁨 감정을 짓는 얼굴표정에서 손의 움직임이 기쁨이나 무표정을 나타낼 때 기쁨 감정을 인식하는 정도는 비슷했다. 하지만 얼굴 표정을 기쁘게 짓고 손의 움직임이 화난 감정을 표현할 때는 기쁨 감정의 인식이 현저히 저하되었다. 감정의 얼굴표정 인식과정에서 감정 인식이 손 움직임의 표현정도에 영향을 받을 수 있다.

미렌(Meeren)[16]은 사람의 얼굴 표정과 신체 자세를 동

시에 보여주는 정적 이미지의 감정 인식을 연구했다. 두 모달이 일치하거나 서로 다른 감정을 표현하는 실험에서 모달들이 서로 다른 감정을 표현하는 경우 얼굴 표정을 통한 감정의 인식은 자세가 전달하는 감정에 영향을 받는다. 감정 모달이 일치하는 신체 언어에 대해서는 얼굴 표정 인식이 개선됐지만 일치하지 않는 경우에는 신체가 전달하는 감정이 얼굴 표정의 감정 판단을 방해했다.

가상 인간을 활용한 디지털 콘텐츠 제작이 활성화됨에 따라 가상 인간의 감정 표현을 담당하는 신체 움직임과 자세 분석에 대한 관심 역시 높아지고 있다. 감정 모달의 영향을 분석한 기존 연구들은 사람의 신체자세와 얼굴 표정을 대상으로 분석 연구를 진행했으므로 가상 인간과 같은 가상 캐릭터를 대상으로 감정 표현에 영향을 미치는 모달을 분석할 필요가 있으며 본 연구는 가상 인간의 비언어적 모달 특히 자세가 감정 표현에 미치는 영향을 제시하여 생생하고 자연스러운 가상에이전트 제작에 기여한다.

1.3 가상 인간의 비언어적 모달을 통한 감정 표현

얼굴과 자세를 통한 감정 표현에 대한 이론적 배경을 제공한 심리학과 인지과학을 바탕으로 디지털 콘텐츠에서는 가상 인간의 얼굴 변형을 통해 상황에 따라 감정을 드러내도록 디자인하고 있다[17,18]. 가상 인간의 자세는 전체 동작, 신체 부위 이완, 시간 추이에 따른 자세 변화, 우아함, 힘의 강약, 동작의 반복 등을 매개변수로 하여 다양한 움직임을 나타내며[19] 되도록 가상 인간의 자세 표현을 사람과 가장 유사하게 나타내고자 한다[20]. 가상 인간의 외형 디자인에 집중하는 연구 외 가상 인간의 얼굴 표정과 신체적 표현의 조화를 분석하여 사람과 가장 비슷하게 감정을 표현하는 연구는 부족한 상태이다. 다만 부사인(Buisine)[21]이 다중모달 코퍼스(corpus) 연구를 통해 자세와 얼굴 표정에 주석을 표시하여 표시된 내용을 재현하는 것을 정의했을 뿐이다. 그는 대상의 다른 표정 재현을 어떻게 인식하는지에 대한 실험적 연구를 수행했으며, 얼굴과 자세 표정의 개별적 인식과 통합인식의 비교 실험은 제시하지 않았다. 지능형 가상 인간에서 감정을 다중모달로 표현하는 애니메이션 계산 모델 역시 얼굴과 자세를 동시에 고려하는 데는 한계가 있다.

기존 디지털콘텐츠 응용사례를 보면 가상 인간과 인식자간의 거리가 가까우면 인식자는 가상인간의 얼굴 표정을 통해 감정을 인식하지만 거리가 멀어지면 전역적 정보를 활용하려는 경향이 있다. 대상과 가까운 거리에

		모달 개별	다중모달 일치	다중모달 불일치
감정 모달		얼굴 자세 얼굴+자세	얼굴+자세	얼굴+자세
감정 인식	기본 감정	분노 기쁨 슬픔 놀람 두려움	분노 기쁨 슬픔	기쁨+슬픔 기쁨+분노 분노+기쁨 분노+슬픔 슬픔+분노 슬픔+기쁨
	감정 활성화	능동/수동	능동/수동	능동/수동
	감정 정서기	긍정/부정	긍정/부정	긍정/부정

(그림 2) 제안 연구 내용

서는 얼굴이 가장 중요한 감정표현 모달이지만 거리가 멀어질수록 신체 전체의 표현을 통해 대상의 감정을 이해해야 한다. 얼굴 표정이 정확하게 인지되지 않은 때 감정을 전달하는 신체언어가 특히 더 중요한 역할을 하게 되므로 가상 인간의 감정 다중모달은 당연히 고려되어야 할 내용이다. 따라서 본 연구에서는 가상 인간이 감정을 여러 모달을 통합하여 전달하는 데 효과적인 표현 방법을 분석하고 가상 인간이 나타내는 감정 표현을 피실험자가 어떻게 인식하는지를 평가한 내용을 다룬다. 이를 위해 가상 인간을 설정한 후 (그림 2)와 같은 구조로 얼굴, 자세, 얼굴+자세 감정 모달에 대해 기본 감정 인식 정도, 감정 활성화 정도와 감정 정서기를 대상으로 가상 인간의 감정 표현에 필요한 감정 모달 개별로, 다중 모달이 일치했을 때, 다중 모달이 불일치할 때의 경우를 실험하여 평가한다.

디지털 콘텐츠에서 보여지는 가상 인간의 감정 표현 모달은 얼굴과 자세가 대표적이며 가상 인간이 표현하는 대표적 감정 카테고리는 에크만이 제시한 기본 감정을 대상으로 한다. 여기에 러셀의 감정 원형모델을 도입하여 얼굴과 자세 모달을 통해 표현하는 기본 감정 카테고리의 세기 정도를 감정 활성화(능동/수동 상태)와 감정 정서기(쾌/불쾌)를 통해 포착하여 일상 생활에서 중첩하여 자주 나타나는 감정 상태를 쉽게 표현하고 이해하도록 한다. 동일한 감정 카테고리일지라도 얼굴과 자세 모달을 통해 강하게 표현된 정도와 약하게 표현된 정도에 따라 인식자가 감정을 달리 인식할 수 있기 때문이다.

2. 가상 인간의 비언어적 모달 감정 인식

본 연구는 사람이 가상 인간이 표현하는 감정을 인식

할 때 얼굴, 자세, 얼굴+자세와 같은 비언어적 모달이 어떤 영향을 주는지 분석한다. 이를 위해 슬픔, 두려움, 분노, 놀람, 기쁨 과 같은 기본 감정을 얼굴 표정과 자세로 표현한 결과를 피실험자가 인식하고 감정 표현의 능동/수동적 측면을 고려한 활성화 정도, 긍정적 감정인지 부정적 감정인지를 판단하는 정서가 정도를 인식하는 실험을 수행한다. 실험은 두 단계로 나누어 수행한다. 첫 번째 실험에서는 피실험자가 가상 인간의 얼굴과 자세, 얼굴+자세 각 모달 별로 드러난 기본 감정을 인식하고 기본 감정의 정도를 활성화와 정서기로 평가하여 각 모달의 영향을 제시한다. 디지털 콘텐츠에서 가상 인간이 드러내는 다양한 감정은 독립적 사건으로 특정 순간 바로 나타날 수도 있지만 대개는 가상 인간이 처한 상황에 따라 연속적으로 감정이 드러나기 때문에 에크만이 분류한 기본 감정 외 러셀이 정의한 2차원 감정 원형 모델에 나타난 연속적인 감정의 흐름인 활성화와 정서가 차원을 가상 인간의 감정 표현에 반영할 필요가 있기 때문이다.

2.1. 모달 별 감정 표현 인식 실험

가상 인간의 개별 모달로 표현된 기본 감정 인식을 수행하기 위해 기존 연구에서 구축한, 감정을 드러낼 때의 얼굴 표정 구별 단서와 감정별 자세 데이터 특징[9]을 기반으로 비언어적 감정 표현을 위한 가상 인간을 디자인 하였으며 분노, 두려움, 기쁨, 놀람, 슬픔의 기본 감정을 얼굴과 자세를 통해 드러내도록 했다. 감정 모달을 분리하기 위해 감정별로 얼굴 표정과 자세를 분리했다.

얼굴과 자세 모달 별로 감정 표현을 인식하는 실험은 무표정에서 시작해 감정을 표현하고 다시 무표정으로 되돌아가는 3초간의 동영상 애니메이션을 15개 제작하여 이를 피실험자가 평가하는 과정으로 이루어진다. 5개 애니메이션은 (그림 3.a)와 같이 신체의 나머지 부분은 없애고 얼굴만으로 기본 감정을 나타냈다. 두 번째 5개의 애니메이션은 (그림 3.b)와 같이 해당 감정을 자세로만 나타냈다. 나머지 애니메이션은 (그림 3.c)와 같이 해당 감정을 얼굴과 자세를 일치하여 나타냈다. (그림 3.a)와 (그림 3.b)에서 얼굴과 자세 모달을 별도로 구분하고 다른 모달 영역을 지우면 감정 평가에 영향을 미치는 다른 모달의 영향을 줄일 수 있다. 피실험자는 가상 인간의 모달 별 애니메이션을 보고 모달 별 기본감정 카테고리과 감정의 활성화, 정서가 정도를 평가한다. 실험은 무표정에서 시작해 감정을 표현하고 다시 무표정으로 되돌아가는 3초간의 동영상 애니메이션을 15개 제작하여 이를 피



(a) 얼굴 (b)자세 (c)얼굴+자세 조합
(그림 3) 감정을 드러내는 신체 모달의 종류

실험자가 평가하는 과정으로 이루어진다.

42명의 20대 남녀 학생들을 첫 번째 실험에 피실험자로 참여시켜 무작위로 두 그룹으로 나눈 후 각 집단에게 애니메이션의 절반씩을 보게 했다. 첫 번째 가상 인간 애니메이션은 학습 데이터로 설정하여 피실험자들은 가상 인간의 감정을 평가하기 전에 미리 가상 인간의 상태에 익숙해지도록 했다. 그 다음 피실험자가 테스트 애니메이션을 본 다음에는 5점 척도의 리커트 스케일(likert scale)로 각 감정을 얼마나 정확히 인식했는지를 표시하도록 했다. 피실험자는 가상인간의 기본 감정 카테고리를 선택하고 해당 감정의 활성화 및 정서가 정도를 평가했다.

2.2 모달 별 감정 표현 인식 결과

1) 모달 별 기본 감정 평가

첫 번째 실험의 결과는 <표 1>과 같다. 각 모달리티의 의도된 감정인식 수준을 측정하고 모달 간의 감정표현 차이가 동일 모달 내 감정 표현 변동량에 비해 크다는 것을 테스트하기 위해 분산분석(ANOVA)기법으로 반복하여 측정하였다. 실험 결과 분노 감정은 얼굴, 자세, 얼굴+자세 3가지 형태의 모달 모두에서 잘 인식되었다. 슬픔 감정은 얼굴과 얼굴+자세 모달에서는 잘 인식되었으나 자세 모달에서는 인식율이 낮았다. 기쁨 감정은 얼굴 모달과 얼굴+자세 모달에서 잘 인식되었지만 자세 모달만으로는 놀람으로 다수 오인식되었다.

놀람 감정을 드러낸 얼굴 모달은 슬픔, 놀람, 두려움 감정으로 다수 오인식되었지만 자세 모달을 통한 놀람 감정이 가장 높은 인식율을 나타냈다. 놀람 감정을 드러낸 얼굴+자세 모달은 놀람과 두려움으로 다수 오인식되었다.

두려움을 드러낸 얼굴 모달은 슬픔과 무표정으로 다수 인식되었다. 두려움을 드러낸 자세 모달과 얼굴+자세 모달은 놀람으로 다수 오인식되었다. 결과적으로 분노, 기쁨, 슬픔 감정 애니메이션은 얼굴, 자세, 얼굴+자세 세 가지 모달에서 잘 인식되었으나 놀람과 두려움 감정은

(표 1) 다중 모달에서 감정의 평균 인식 정도

(최소값=0, 최대값=4)

<1.1> 얼굴

	분노	기쁨	슬픔	놀람	두려움	무표정	ANOVA $p < .001$
분노	3.21	0.05	0.42	0.63	0.16	0.05	$F(6.90) = 61.79$
기쁨	0.05	2.31	0.10	0.22	0.10	0.79	$F(6.85) = 31.43$
슬픔	0.10	0.16	3.16	0.26	0.68	0.26	$F(6.90) = 61.32$
놀람	0.28	0.12	1.69	1.70	1.53	0.39	$F(6.150) = 18.96$
두려움	0.12	0.21	1.12	0.13	0.39	1.80	$F(6.150) = 19.60$

<1.2> 자세

	분노	기쁨	슬픔	놀람	두려움	무표정	ANOVA $p < .001$
분노	3.63	0.05	0.31	0.05	0.21	0.05	$F(6.90) = 176.44$
기쁨	0.05	3.74	0.00	1.58	0.00	0.50	$F(3.64) = 124.44$
슬픔	0.05	0	3.47	0.37	0.58	0.37	$F(4.72) = 67.66$
놀람	0.26	0.34	0.32	3.02	2.38	0.17	$F(6.165) = 76.01$
두려움	0.41	0.26	0.60	2.79	2.54	0.00	$F(4.132) = 69.69$

<1.3> 얼굴 +자세

	분노	기쁨	슬픔	놀람	두려움	무표정	ANOVA $p < .001$
분노	3.47	0.05	0.28	0.50	0.28	0.00	$F(4.68) = 100.82$
기쁨	0.84	2.47	0.16	2.26	0.26	0.00	$F(4.72) = 14.90$
슬픔	0.28	0.05	1.95	0.44	0.22	0.89	$F(6.85) = 7.92$
놀람	0.32	1.23	0.09	2.89	1.46	0.06	$F(4.132) = 32.28$
두려움	0.17	0.64	0.67	2.63	2.77	0.08	$F(6.170) = 69.69$

다른 감정으로 오인식되는 경우가 많았다.

실험 결과는 한 가지 기본 감정만을 표현한 가상 인간의 감정이 의도된 감정 외 다른 감정으로도 다수 인식됨을 보여 준다. 기쁨, 분노, 슬픔이 놀람과 두려움보다 잘 인식되며 놀람은 세가지 모달에서 모두 두려움으로 오인식되었다. 이는 기존 연구[22, 23]에서 사람의 행동과 자세를 통해 인식한 감정 카테고리 중 슬픔보다는 놀람이 잘 인식되며 기쁨 감정이 분노 감정보다 잘 인식된 것과는 다소 다른 결과로 가상 인간 제작 시 감정을 특징짓는 방법을 더 명확하게 개선해야 함을 의미한다.

2) 모달 별 감정카테고리의 활성화 및 정서가 평가

<표 2>는 감정 카테고리의 활성화 및 정서가 평가 결과로서 피실험자가 다중 모달을 통해 가상 인간의 감정 카테고리를 인식한 후 동일한 감정 카테고리에서 해당 감정의 능동/수동적인지를 판단하는 활성화 정도와 긍정적인지 부정적인지 감정의 정서 차원을 다룬 정서가예를 인식한 결과를 보면 가상인간이 의도한 정서가 정도는 잘 인식하였으나 활성화 차원은 잘 인식하지 못했다. 또한 감정 표현을 얼굴에만 드러낸 애니메이션을 평가할 때 활성화의 능동/수동 차원 인식에는 큰 차이가 없으나 자세 모달 애니메이션에서는 활성화 차원을 더 정확하게 인식하였다. 이를 통해 감정의 활성화 수준을 인식하는 데는 얼굴 표정보다는 자세 모달이 효과적이며 가상 인

〈표 2〉 감정별 모달에 대한 활성화와 정서가의 평균 인식 정도 (최소 값=0, 최대 값=4)

〈2.1〉얼굴

	활성화		정서가	
	능동적	수동적	긍정적	부정적
분노	1.42	1.37	0.31	3.42
기쁨	0.78	2.26	2.94	0.37
슬픔	1.47	1.31	0.53	2.89
놀람	1.42	2.11	0.65	2.64
두려움	0.77	2.25	0.71	1.64

〈2.2〉자세

	활성화		정서가	
	능동적	수동적	긍정적	부정적
분노	3.16	0.21	0.42	3.31
기쁨	3.21	0.10	2.36	1.16
슬픔	1.05	1.84	0.47	2.37
놀람	3.00	0.32	1.85	1.21
두려움	2.78	0.38	1.00	2.15

〈2.3〉얼굴+자세

	활성화		정서가	
	능동적	수동적	긍정적	부정적
분노	3.21	0.47	0.16	3.63
기쁨	3.37	0.21	3.63	0.05
슬픔	1.22	2.05	0.33	3.1
놀람	2.71	0.62	0.82	2.14
두려움	2.82	0.97	0.76	2.26

간의 자세를 통해 활성화 효과를 극대화시킬 수 있음을 예측할 수 있다.

2.3 모달 별 감정 인식 비교 결론

피실험자가 인식한 가상 인간의 감정이 모달에 종속되는지, 감정을 단일 모달로 드러낼 때보다 얼굴과 자세

를 일치시켜 드러낼 때 더 감정을 잘 인식하는지, 얼굴과 자세 각 모달이 감정 인식에 증가의 영향을 미치는지를 분석한 결과 분노 감정의 인식은 세 가지 모달 조건에서 유사하게 높은 결과를 보였다. 기쁨 감정은 얼굴, 자세 모달보다 얼굴+자세 모달에서 더 잘 인식되었다. 슬픔은 자세보다 얼굴, 얼굴+자세에서 더 높은 인식결과를 보였다. 놀람, 두려움은 얼굴 표정보다는 자세, 얼굴+자세 모달을 통해 더 높은 인식결과를 보였다. 또한 분노, 슬픔, 기쁨 감정이 감정의 능동/수동 상태를 담당하는 활성화와 긍정/부정 상태를 담당하는 정서가 인식 정도에서 가장 높은 값을 기록했다.

첫 번째 실험 결과를 통해 감정 카테고리별로 모달의 기여도가 다를 수 있다. 피실험자는 분노 감정은 모달에 관계없이 잘 인식했지만 다른 감정들에 대해서는 모달에 따라 인식에 영향을 받았다. 애니메이션 제작상의 문제때문인지 혹은 다른 인식과정 때문인지 정확히 판별하기 어렵지만 얼굴 및 자세 각각의 모달을 통한 감정 활성화 인식 결과에서 피실험자가 서로 다른 인식 결과를 보였다.

3. 가상 인간의 비언어적 다중 모달 일치/불일치 감정 표현 인식

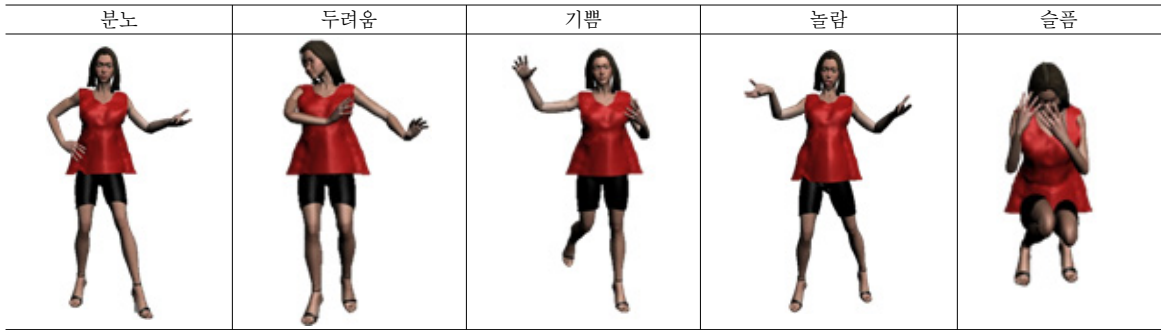
두 번째 실험에서는 기본 표정을 표현하는 다중 모달이 서로 일치할 때와 다중 모달이 각기 다른 감정을 표현할 때 피실험자의 기본 감정 인식과 기본 감정의 활성화와 정서가 정도에 미치는 영향을 다룬다.

3.1 다중 모달 일치/불일치 감정 인식 실험

두 번째 실험은 (그림 4)와 같이 가상 인간의 얼굴과



(그림 4) 가상 인간의 얼굴과 자세 모달 일치



(그림 5) 가상 인간의 얼굴과 자세 모달 불일치

자세가 일치하여 기본 감정을 드러낼 때와 (그림 5)처럼 얼굴과 자세 모달이 불일치하여 자세는 분노를 나타내지만 얼굴은 화난 감정을 표현할 때 가상 인간의 감정이 어떻게 인식되는지에 관한 것으로 가상 인간의 감정 모달의 불일치가 감정 인식에 끼치는 영향을 분석한다. 일상적으로 사람들은 단일 감정을 여러 감정 모달을 통해 동시에 드러내기도 하지만 여러 감정을 여러 모달을 통해 같이 드러내기도 한다. 가상 인간이 디지털콘텐츠에서 HCI로 사용될 때는 단일 감정을 여러 모달을 통해 동시에 드러내는 것이 확실한 감정 인식이 효과적이지만 여러 감정 차원을 나타낼 때는 얼굴과 자세 모달에서 조화롭게 드러나는 것이 자연스럽다.

첫 번째 실험에서와 같이 지시된 감정을 표현한 첫 번째 애니메이션의 캐릭터를 학습 애니메이션으로 하여 피실험자가 테스트 전에 미리 가상 인간의 표정에 익숙하게 한다. (그림 6)과 같은 애니메이션을 본 후 피실험자는 5점의 리커트 스케일을 이용하여 애니메이션에서 각 감정을 얼마나 잘 인식했는지 기본 감정을 선택하고 감정의 활성화와 정서가 정도에 점수를 매긴다.

3.2 다중 모달 감정 표현 인식 평가 결과

1) 다중 모달 일치 기본 감정 인식

두 번째 실험의 평가 결과, (표 3)과 같이 모달이 일치하는 애니메이션에 대해 피실험자는 가상인간이 의도한 감정을 잘 인식함을 알 수 있다.

2) 다중 모달 불일치 시 기본 감정 인식

(표 4)는 다중 모달이 불일치할 때 슬픔 얼굴표정에 분노자세를 조합한 애니메이션에서 자세에 드러난 분노 감정을 잘 인식한 경우를 제외하고는 대부분 자세보다는

(표 3) 얼굴+자세 모달 일치 시 기본감정 인식 정도

모달	감정	분노	기쁨	슬픔	놀람	두려움	무표정	ANOVA p < .001
얼굴+ 자세	분노	3.52	0.00	0.10	0.19	0.05		F(4.76) = 233.20
얼굴+ 자세	기쁨	0.00	3.67	0.00	1.24	0.00		F(1.20) = 86.00
얼굴+ 자세	슬픔	0.14	0.10	2.19	0.33	0.19		F(6.100) = 34.76

(표 4) 모달 불일치 시 기본감정 평균 인식 정도

얼굴	자세	분노	기쁨	슬픔	놀람	두려움	무표정	ANOVA p < .001
기쁨	슬픔	0.10	2.05	0.10	0.76	0.05	0.29	F(5.100) = 17.36
기쁨	분노	0.10	2.43	0.00	0.67	0.00	0.19	F(3.60) = 41.62
분노	기쁨	3.00	0.00	0.14	0.48	0.19	0.00	F(3.60) = 126.76
분노	슬픔	2.81	0.00	0.24	0.29	0.10	0.19	F(4.80) = 64.86
슬픔	분노	1.52	0.14	1.29	0.33	0.62	0.14	F(6.100) = 9.21
슬픔	기쁨	0.48	0.29	1.95	1.00	0.71	0.33	F(6.100) = 8.10

얼굴 표정에서 드러난 감정 카테고리를 중점적으로 인식함을 보여 준다.

3) 다중 모달 불일치 시 활성화와 정서가 인식

모달이 불일치하는 애니메이션의 감정 카테고리에 대한 활성화와 정서가 정도를 비교하기 위해 슬픔 얼굴표정과 분노 자세를 조합한 애니메이션, 기쁨 얼굴 표정과 분노 자세를 조합한 애니메이션, 분노 얼굴 표정과 슬픔 자세를 조합한 애니메이션, 분노 얼굴 표정과 기쁨 자세를 조합한 애니메이션을 설정하여 평가하였다. 분노 얼굴 표정을 기쁨 자세와 조합한 애니메이션만이 예외를 보였는데 분노와 기쁨이 완전히 다른 정서가 차원임에도 활성화 차원은 같은 감정이기 때문에 감정 인식률이 높은 것으로 판단된다.

4) 동일 자세의 얼굴 표정 불일치 시 인식 결과

자세에 드러난 감정은 같고 얼굴의 감정은 다른 애니메이션들을 비교했을 때 피실험자는 얼굴 표정에 드러난

(표 5) 모달 불일치 시 활성화와 정서가 인식

얼굴	자세	활성화		정서가	
		능동	수동	긍정	부정
기쁨	슬픔	1.57	1.33	2.62	0.62
기쁨	분노	2.52	0.52	2.95	0.33
분노	기쁨	2.95	0.29	0.33	2.95
분노	슬픔	2.05	1.24	0.24	3.14
슬픔	분노	2.38	0.65	0.71	2.52
슬픔	기쁨	2.67	0.48	0.67	2.52



(그림 6) 애니메이션 사례

감정을 먼저 인식한다. 슬픔의 인식은 분노 얼굴 표정에 슬픔 자세를 조합한 애니메이션보다 슬픈 얼굴 표정에 분노 자세표정을 조합한 애니메이션이 더 높은 인식률을 보인다. (표 5)의 정서가의 경우, 자세 감정 카테고리의 정서가가 부정적일 때 얼굴 표정 역시 부정 정서가라면 인식 결과는 높은 부정 정서가를 보인다. 따라서 피실험자가 얼굴 표정으로 전달되는 감정 카테고리와 정서가를 같이 인식함을 알 수 있다.

5) 동일 얼굴 표정의 자세 모달 불일치 시 인식 결과 피실험자가 감정을 심도있게 판단하기 위해서는 자세

에 드러난 감정을 고려하지 않기 때문에 자세 모달을 통해 드러난 감정을 먼저 선택하지 않는다. 예를 들어 분노 인식의 경우 슬픔 얼굴 표정에 분노 자세를 조합한 애니메이션의 분노 인식 값이 슬픔 얼굴 표정에 기쁨 자세를 조합한 애니메이션의 슬픔 인식 값보다 낮다. 이는 자세를 심도있게 인식하지 않는다는 의미이다. 이와 달리 활성화 차원의 경우 (표 5)와 같이 얼굴과 자세가 기쁨과 슬픔 감정을 달리 드러낸 애니메이션보다 얼굴과 자세 표정이 같은 활성화 차원을 보이는 분노와 기쁨 애니메이션이 더 높은 인식률을 보였다.

3.4 다중 모달 일치 및 불일치 시 감정 인식 결론

실험을 통해 기본 감정 카테고리 인식에는 자세보다 얼굴 표정이 더 영향을 미친다는 것과 감정 카테고리의 활성화 차원을 판단하기 위해서는 자세 모달을 활용함을 알 수 있다. 이는 슬픔 감정에 대한 활성화와 정서가 정도는 기쁨 얼굴 표정에 슬픔 자세를 조합한 애니메이션보다 슬픔 얼굴 표정에 기쁨 자세를 조합한 애니메이션이 더 높은 인식률을 보였고 기쁨의 인식은 슬픔 얼굴 표정에 기쁨 자세를 조합한 애니메이션보다 기쁨 얼굴 표정에 슬픔 자세를 조합한 애니메이션이 더 높은 인식 결과를 통해 확인하였다. 감정 카테고리 중 기쁨과 슬픔이 특히 얼굴 표정을 통해 잘 인식되었다.

종합적으로 감정은 자세보다는 얼굴 표정을 통해 잘 인식되지만 감정의 능동성/수동성을 나타내는 활성화 정도는 자세 모달을 참조하였다.쾌/불쾌를 설명하는 정서가 정도는 자세보다 얼굴 표정을 통해 더 잘 전달되었다. 하지만 (그림 7)처럼 사람이 감정을 표현할 때 모달들의 감정이 서로 불일치하면 주로 얼굴 표정을 통해 주요 감

얼굴	기쁨	기쁨	기쁨	놀람	슬픔	슬픔
자세	화남	화남	놀람	슬픔	화남	화남

(그림 7) 사람의 감정 다중모달 불일치 사례

정을 전달하고 자세는 감정의 활성화 정도를 전달할 수 있어 감정 인식이 수월하다. 이 달리 가상 인간은 디지털 콘텐츠 디자인 시 얼굴 표정을 통해 상세하고 집중적으로 감정을 전달하기 어렵기 때문에 얼굴 외 자세 모달에 감정 활성화와 정서가 효과를 극대화시킴으로써 복합적이고 연속적 감정 표현이 바람직하므로 가상 인간의 표정 및 자세 설계에 이를 고려할 필요가 있다.

III. 결 론

감정을 표현해야 하는 가상 인간 제작 시 비언어적 모달을 어떻게 조화시키는 지가 감정 인식에 중요한 문제이기 때문에 구현을 담당하는 계산 모델은 감정관련 모달 간의 관계성을 분석하고 이를 바탕으로 콘텐츠를 설계해야 한다. 또한 기본 감정 발현은 독립적 사건이지만 디지털 콘텐츠의 상황에 따라 연속적으로 다양한 감정이 드러나기도 하며 여러 감정이 동시에 신체 모달을 통해 드러날 수 있다는 점을 감안해야 한다.

본 연구는 가상 인간의 자세와 얼굴의 모달 조합을 통해 감정을 표현하는 감정 모달 간의 상관성을 분석하고 얼굴과 자세 모달을 통해 다양한 감정이 제시될 때 모달이 감정 인식에 어떤 영향을 미치는지 확인하기 위해 두 가지 실험을 수행했다. 이를 위해 디지털콘텐츠의 가상 인간과 유사한 캐릭터를 설계하고 이 가상 인간이 드러내는 비언어적 감정 표현을 피실험자가 인식하여 감정 카테고리 상태를 판단하도록 하였다.

가상인간의 감정 표현 시 얼굴 모달만을 이용해도 감정 카테고리 인식이 가능하지만 신체 자세와 얼굴 표정이 일치하는 감정을 표현하면 기본 감정 인식이 더 정확하게 이뤄짐을 알 수 있고, 얼굴과 자세 모달이 동일한 감정을 드러내면 활성화와 정서가 차원이 확대되어 연속적 감정 인식이 가능함을 확인하였다. 두 번째 실험에서는 감정 표현이 일치하거나 불일치하는 모달들이 감정 인식에 미치는 영향을 파악했다. 감정 인식은 주로 얼굴 표정을 통해 이루어지지만 자세에서도 영향을 받는다. 감정 표현 모달 실험은 단일 감정 인식과정에서 얼굴과 자세 그리고 얼굴+자세 모달이 감정 인식에 미치는 영향을 파악하는 것을 목적으로 했다. 실험 결과, 자세와 얼굴 모달이 일치하지 않아 여러 감정이 섞여 나타나는 경우 자세가 감정의 활성화 정도를 평가하는데 유용하였다. 이는 감정을 드러내는 자세에 불일치하는 얼굴표정의 감정을 의도적으로 제시했을 때 즉, 일상적인 생활에서 자주 나타내는 기본 감정들이 서로 다른 모달을 통해

다양하게 표현될 때 피실험자는 얼굴 표정으로 감정을 판별하면서 감정의 활성화 정도를 판단하기 위해서는 자세를 참조함을 알 수 있다.

결과적으로 가상 인간을 통해 감정을 잘 전달하기 위해서는 분노, 기쁨과 같은 감정의 경우 얼굴 표정을 통해 집중적으로 강조하고 동시에 감정과 일치하는 여러 자세 모달을 제작해야 하며 여러 감정이 모달에서 섞여 나타날 때 얼굴 표정 감정을 우선적으로 판별하기 때문에 가상 캐릭터 인간의 감정 표현 시 이를 감안해야 한다. 이와 함께, 디지털콘텐츠에 등장하는 가상 인간은 상세한 얼굴 표정 대신 많은 부분 자세를 통해서 감정의 활성화 정도를 전달하기 때문에 자세를 통한 감정 표현의 중요함을 간과하지 않아야 한다.

가상 인간의 신체 움직임에 대한 평가는 감정연구 전문가 또는 전문 피실험자보다는 온전히 피실험자의 주관적 판단에 따라 이루어지기 때문에 실험환경, 데이터 기록상태, 개인의 감정상태가 평가에 영향을 미친다. 또한 디지털 콘텐츠의 이용 대상에 따라 평가 결과는 달라질 수 있다. 따라서 가상 인간의 감정 평가를 위한 신뢰도를 위해 디지털 콘텐츠의 사용자 특성을 고려하여 어린이, 고령자 등 다양한 성별과 연령대의 피실험자를 통한 가상 인간의 감정과 행동 평가를 할 필요가 있으며 또한 모달을 통해 표현하는 감정의 다양성을 모색할 필요도 있다. 성별, 연령별로 관찰차를 다양화시켜 가상 인간을 평가하는 연구는 향후 계속 진행할 계획이다.

가상인간의 감정 모달 불일치는 부자연스러운 상태의 동작으로 제시되어 디지털콘텐츠 이용의 몰입감, 집중력을 저하시키는 요소로 작용한다. 디지털콘텐츠 제작 시 주로 얼굴에 집중된 감정표현 외 감정의 정도를 자연스럽게 드러내는 자세 설계의 중요성을 확인한 연구 결과를 토대로 후속 연구에서는 감정 모달 불일치를 해결하는 방안으로 일반화된 감정 표현을 위해 모달별 표현방법의 범위를 정량화시켜 데이터 코퍼스로 구축하고 이를 가상 인간 설계 시 그래픽과 동기화시키는 다중 모달 애니메이션 엔진 시스템 제작에 적용하여 아바타와 같은 가상 인간에게 감정에 따라 최적의 행동과 동작을 자동적으로 생성하게 하는 음성, 언어시각화 기술 구현에 기여할 것이다.

참 고 문 헌

- [1] K.R. Scherer, "Emotion, Introduction to Social

- Psychology: A European perspective“, Oxford: Blackwell, pp.151-191, 2000.
- [2] J. A. Harrigen, “The new handbook of methods in nonverbal behavior research”, Series in Affective Science, Oxford University Press, pp. 137-198, 2005.
- [3] T. Partala, V. Suraka, “The effects of affective interventions in human-computer interaction“. *Interacting with Computers*, Vol. 16, pp. 205-309, 2004.
- [4] H. Gunes, M. Piccardi, M. Pantic, “From the lab to the Real World : Affect Recognition Using Multiple Cues and Modalities,” *Affective Computing, Focus on Emotion Expression, Synthesis and Recognition, Tech Education and Publishing*, pp.184-218, 2008.
- [5] C. Izard, “Human Emotions”, Plenum, New York, 1977.
- [6] J. A. Russell, “A circumplex model of affect“, *Journal of Personality and Social Psychology*, Vol. 39, No. 6, pp 1161-1178, 1980.
- [7] K. R. Scherer, “Analyzing Emotion Blends“, *Proceedings of the Xth Conference of the International Society for Research on Emotions*, Fischer, pp.142-148, 1998.
- [8] P. Ekman, W. Friesen, “A new pan-cultural facial expression of emotion“, *Motivation and Emotion*, Vol. 10, pp.159-168, 1986.
- [9] “표정 강도에 강건한 얼굴 표정 인식”, *정보처리학회 논문지 B*, 제 16-B권 5호, pp. 395-402, 2009.
- [10] H. Gunes, M. Piccardi, “Bi-modal emotion recognition from expressive face and body gestures“, *Journal of Network & Computer Applications*, Vol. 30, No. 4, pp.1334-1345, 2007.
- [11] A.T. Dittman, “The role of body movement in communication“, *Nonverbal behavior and communication*, Lawrence Erlbaum, Hillsdale, 1987.
- [12] 김진욱. “감정 자세 인식을 위한 자세특징과 감정 예측 모델“, *한국인터넷정보학회 논문지*, 제 12권 6호, pp. 83-94, 2011.
- [13] H.G. Wallbott, “Bodily expression of emotion“, *European Journal of Social Psychology*, Vol. 28, pp.879-896, 1998.
- [14] K.R. Scherer, H. Ellgring, “Multimodal Expression of Emotion: Affect Programs or Componential Appraisal Patterns“, *Emotion*, Vol. 7, No. 1, pp.158-171, 2007.
- [15] J. K. Hietanen, J. M. Leppänen, “Judgment of other people’s facial expressions of emotions is influenced by their concurrent affective hand movements“, *Scandinavian Journal of Psychology*, Vol. 49, pp.221-230, 2008.
- [16] H.K.M. Meeren, C. van Heijnsbergen, B. de Gelder, “Rapid perceptual integration of facial expression and emotional body language“, *Proceedings of the National Academy of Sciences of the USA*, pp.16518-16523, 2005.
- [17] R. Niewiadomski, “A model of complex facial expressions in interpersonal relations for animated agent“, PhD Thesis, University of Perugia, 2007.
- [18] T. Ziemke, R. Lowe, “On the role of emotion in embodied cognitive architectures: From organisms to robots, *Cognitive Computation*, Vol. 1, pp. 104-117, 2009.
- [19] C. Pelachaud, “Multimodal expressive embodied nonverbal virtuals“, *Proceedings of the 13th annual ACM international conference on Multimedia*, pp.632-689, 2005.
- [20] D. Ballin, M. Gillies, B. Crabtree, “A Framework For Interpersonal Attitude And Non-Verbal communication in Improvisational Visual Media Production“, *1st European Conference on Visual Media Production IEE*, 2004.
- [21] S. Buisine, S. Abrilian, R. Niewiadomski, J. C. Martin, L. Devillers, C. Pelachaud, “recognition of blended emotions: From video corpus to expressive virtual“, *International Conference on intelligent Virtuals*, Springer-Verlag, pp.93-106, 2006.
- [22] J. A. Hall, D. Matsumoto, “Gender differences in judgments of multiple emotions from facial expressions“, *Emotion*, Vol. 4, No. 2, pp. 201-206, 2004.
- [23] 김진욱. “새로운 얼굴 특징공간을 이용한 모델 기반 얼굴 표정 인식, *한국정보처리학회 논문지 B*, 제 17-B권 4호, pp.309-316, 2010.

● 저 자 소 개 ●



김 진 옥 (Kim Jin Ok)

1989년 성균관대학교 졸업(학사)

1998년 성균관대학교 대학원 정보통신공학과 졸업(석사)

2002년 성균관대학교 대학원 전기전자 및 컴퓨터공학과 졸업(박사)

2004년~현재 대구한의대학교 국제문화정보대학 모바일콘텐츠학부 교수

관심분야 : 멀티미디어공학, 패턴 인식, 영상처리, 유비쿼터스 컴퓨팅

E-mail : bit@dhu.ac.kr