

---

# 영상통신 감성융합 서비스를 위한 실시간 아바타 정합기술

오동식\*, 강준규\*\*, 신민호\*\*\*

## Emotion fusion video communication services for real-time avatar matching technology

Oh Dong Sik\*, Kang Jun Ku\*\*, Sin Min Ho\*\*\*

**요약** 현재 전 세계적으로 미래 수익 사업의 일환으로 각광받고 있는 분야 중 하나는 3D이다. 3D는 현실세계와 가상현실 세계의 차원을 함께 공존시켜주며 형태 및 질감 등을 느낄 수 있게 해주어 기존의 평면적인 2D에서 입체적인 3D로 변화하고 공존의 시대적 현실을 잘 보여주고 있기 때문이다. 3D에 대한 사람들의 관심은 3D아바타를 바탕으로 하는 영화를 통하여 확산되었다. 또한 현재 대기업들의 3D TV 시장으로의 시장 변화 역시 3D시장의 개척에서 도약의 시대로 한층 더 업그레이드 시키게 되었다. 또한 이와 동시에 세계의 현대인이라면 필수품이 되어가고 있는 스마트폰의 열풍 또한 새로운 핸드폰 시장과 IT 시장의 혁신을 이루었으며, 스마트폰은 작은 컴퓨터라 불릴 정도로, 그 파급속도와 여파는 전화, 인터넷의 혁신만큼 많은 이슈를 남기고 있다. 스마트폰은 말 그대로 똑똑한 폰으로 여러 가지 기능을 할 수 있는 핸드폰이다. 현재 iPhone, Android, Windows Phone외에 다수의 스마트폰이 출시되어 있다. 위의 전체적인 향후의 전망과 비즈니스 서비스 모델을 위하여 스마트폰에 가상의 3D캐릭터를 카메라로 입력받아 3D 표정아바타를 사람의 얼굴에 합성할 수 있고 카메라로 사용자의 감성 표정을 인식하여 합성된 아바타를 실시간 다른 휴대전화 사용자에게 정합, 전송, 통신 할 수 있는 실시간 감성융합 영상통신 서비스 어플리케이션의 개발하고자 한다.

**주제어** : 스마트폰, 3D 표정아바타, 어플리케이션, 비즈니스 모델, 영상통신

**Abstract** 3D is the one of the current world in the spotlight as part of the future earnings of the business sector. Existing flat 2D and stereoscopic 3D to change the 3D shape and texture make walking along the dimension of the real world and the virtual reality world by making it feel contemporary reality of coexistence good show. 3D for the interest of the people has been spreading throughout the movie which is based on a 3D Avata. 3D TV market of the current conglomerate of changes in the market pioneer in the 3D market, further leap into the era of the upgrade was. At the same time, however, the modern man of the world, if becoming a necessity in the smartphone craze new innovation in the IT market mobile phone market and also has made. A small computer called a smartphone enough, the ripple velocity and the aftermath of the innovation of the telephone, the Internet as much as to leave many issues. Smartphone smart phone is a mobile phone that can be several functions. The current iPhone, Android. In addition to a large number of Windows Phone smartphones are released. Above the overall prospects of the future and a business service model for 3D facial expression as input avatar virtual 3D character on camera on your smartphone camera to recognize a user's emotional expressions on the face of the person is able to synthetic synthesized avatars in real-time to other mobile phone users matching, transmission, and be able to communicate in real-time sensibility fused video communication services to the development of applications.

**Key Words** : Smartphones, 3D avatar facial expressions, applications, business models, visual communications

---

\* (주)누리플렉시아이티 대표(교신저자)

\*\* (주)아이디피쉬 대표

\*\*\* (주)지오위즈넷 대표

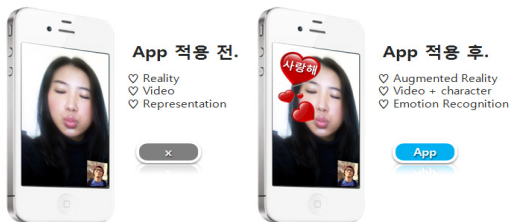
논문접수: 2012년 10월 16일, 1차 수정을 거쳐, 심사완료: 2012년 11월 15일

## 1. 서론

최근 스마트폰의 빠른 확산 및 3D 시장의 발전으로 인해, 3D를 대상으로 하는 콘텐츠 및 H/W, S/W 등 다양한 상품이 출시되고 있다. 국내 및 해외 대기업들이 이러한 차세대 기술 과 디자인에 전력하는 이유는 정보와 함께 흘러가는 정보 소비자의 욕구를 충족시키기 위한 수단으로 그 기업 기술전략의 급성장을 이루고 있다. 이러한 신 기술 사업의 성장은 단순기술이 아닌 정합 기술 사업으로 진출 하였으며 이에 맞는 스마트폰의 영상통화와 3D 아바타의 접목으로 국내 및 해외 시장에 맞는 영상통신 감성융합 서비스를 소비자에게 제공 수단으로 이용한다면 그에 대한 수익도 기대할 수 있다. 예로 네이트온 이모티콘은 일본을 시점으로 전 세계적으로 엄청난 수익 효과를 가져다주고 있는데 국내 카카오톡 역시 이를 응용한 다양한 이모티콘 콘텐츠로 다양한 광고 효과 및 수익 상품을 돌출하고 있다. 이에 본 개발 콘텐츠 역시 발신자 스마트폰 단말(디바이스)의 영상 통신 정합 서비스 또한 향후 막대한 기대 수익을 창출 할 수 있다.



[그림 1] 아바타 정합 개념  
( 화면 전체의 교체와 얼굴 부위 교체로 구분 )



[그림 2] 아바타 정합 개념  
( 표정에 대한 전달 효과를 증대하기 위한 부가 기능 )

[그림 1], [그림 2]와 같이 영상통신 감성융합 서비스는 발신자, 수신자 스마트폰 단말 화면에 가상의 3D객체가 실시간 혼합되어 감정이나 의사전달을 효과적으로 전달할 수 있도록 구현된 아바타 영상통신 서비스로서 감

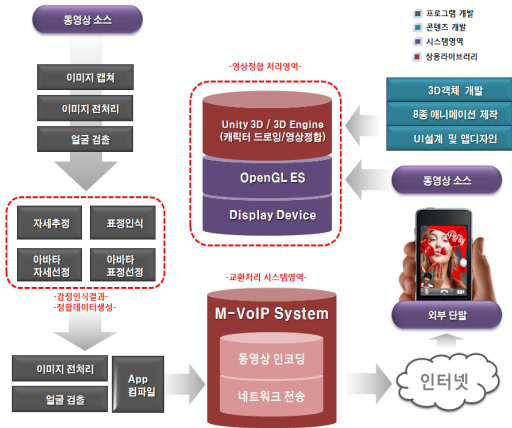
정을 분석하여 특정 감정에 대해 아바타가 대신 표현하도록 하는 아바타 정합과 표정을 분석하여 특정 표정에 대한 효과를 첨가하는 이모티콘 정합의 특징을 표현하는 콘텐츠이며 감정의 정보 전달에 대한 접근성 및 정확도가 뛰어나고, 반응속도도 빨라, 보다 쉽고 편리하게 재미 있고 임팩트적인 효과를 얻을 수 있다. [1][3].

## 2. 영상통신 감성융합 서비스 시스템

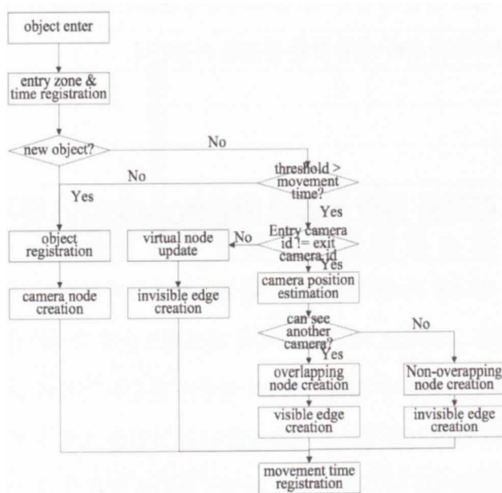
영상통신 감성융합 서비스는 카메라 모듈로부터 동영상 데이터 소스에서 분석할 프레임 이미지를 캡처, 캡처한 이미지를 분석하기 쉬운 상태로 전 처리를 하는 단계를 거친다. 이미지 분석 단계인 엣지 검출 알고리즘 등을 통한 이미지내의 객체 간 경계 파악, 전 처리된 이미지에서 얼굴을 검출하며, 학습 알고리즘 등을 적용하여 얼굴 특징을 분석하고 구성 요소의 이미지상 위치와 관계 데이터를 추출, 얼굴검출 단계를 통해 추출된 인식 정보를 바탕으로 자세추정과 표정인식을 한다. 얼굴이 취하는 자세와 표정에 대한 아바타의 자세와 표정을 선정, 분석된 정보를 통해서 3D 공간상의 아바타(얼굴) 위치 좌표를 결정하고 해당 표정과 감정에 대한 아바타 애니메이션을 선정하고, 3D엔진에 제어 신호(메시지)를 전달 아바타 영상과 동영상상이 표현될 3D공간을 구성하고 분석된 아바타를 해당위치에 놓는 기능을 수행(3D엔진 API를 통한 3D객체 및 3D공간제어) 현재 3D공간을 통해 표현되는 아바타와 동영상소스를 단일 이미지 소스로 정합, 정합된 이미지를 음성소스와 함께 동영상 인코딩을 수행. 음성소스는 동영상 소스로부터 추출하여 처리한다. 화상통화를 위해 구성된 상대단말기로의 네트워크 전송단계로 SIP를 통해 세션을 구성하며, RTP/RTCP를 통해 인터넷으로 전송하게 된다 [그림 3]은 일련의 과정에 대한 최종 시스템 구성도를 보여 준다.[2].

이러한 정합 시스템을 개발, 구현하기 위해서는 겹쳐진 물체와 위치, 전이시간, 얼굴영역 검출, 감정인지등 다량의 주요핵심기술이 필요하게 된다. 사람의 경우 움직임이 자유롭고 식별할 수 있는 뚜렷한 특성이 없고, 공공장소에서의 사람 추적 시 사람과 사람이 서로 겹치는 문제를 고려하여 개발하여야 하며 병합분리(merge-split)와 straight-through 방법을 활용하여 객체 간 겹침을 분리하여야 한다. 이러한 객체의 등장, 퇴장, 연속, 병합, 분

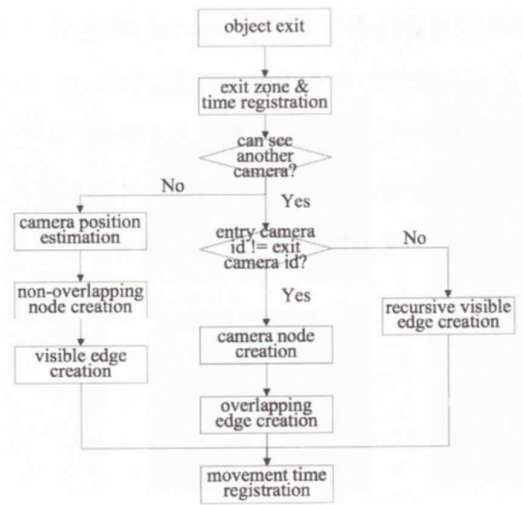
리 이벤트등 에 대한 People Tracking 기술의 개발 후 객체의 크기 카메라의 초점거리와 위치를 좌표로 표현 하여야 한다 객체의 크기는 카메라의 초점거리(focal length)와 같은 조정 정보를 이용하여 계산하고 실제계의 좌표로 위치를 표시 하게 되며, 이러한 일정 시간의 분석된 데이터를 이용하여 자율적으로 학습하는 방식을 사용 중점 또는 비중점 카메라에서 입력받은 영상으로부터 이미지간의 객체를 매칭하여 효과적인 객체 추출 인식 효과를 나타낸다. 이러한 인식은 카메라와 움직이는 사람간의 거리에 따라 피부 영역을 그리드 기반의 추출 방법을 사용하여 개발하게 된다.[3][4].



[그림 3] 최종 시스템 구성도



[그림 4] 객체가 등장할 때 처리하는 순서



[그림 5] 객체가 퇴장할 때 처리하는 순서

[그림 4][그림 5]의 알고리즘을 통하여 객체추적 및 좌표 인지후 객체인지의 얼굴인지 기술(Face Recognition)과 감정인지 기술을 적용하게 된다. 얼굴인지 기술중 얼굴영역검출기술 (Face Detection)은 전체 이미지로부터 얼굴영역을 검출하고 눈, 코, 입 영역을 찾는 검출기술 및 Face Detection 알고리즘을 이용하여 초기 입력이미지인 전체 이미지로부터 얼굴 영역만을 검출하는 기술 skin color 영역 검출 등으로 Binary morphological Dilation 연산(팽창연산)을 통해 노이즈를 제거 라벨링을 통한 각 영역의 중심점을 찾고 각각의 영역을 구현하게 된다.[3].



[그림 6] 얼굴영역 검출 순서

얼굴 기술기의 정도, 정명 각도, 표정 등과 같은 다양한 변화로 인한 얼굴검출이 어려우나 다양한 변화에 대한 규칙을 확장 또는 정규화 하는 지식기반방법(knowledge-based methods)을 적용 이를 해결 할 수 있다.

웃음, 미소, 울음, 찡그림, 놀람 등 얼굴의 표정에 따른 감정인지 기술은 표정분석 관계연삼함수를 통하여 형태와 서로간의 관계로 행복, 슬픔, 기쁨 등의 감정 추론이 가능하다. 이러한 감정객체추출의 단계는 화소, 블랍, 객

체등의 레벨 검출이 구성 적용 되어야만 하는데 배경의 정보를 학습한 배경 모델이 구성되면 배경으로부터 전경 객체를 추출하기 위하여 우선 화소 레벨에서 전경화소와 배경화소를 검출하는 단계를 수행하며 인접 전경화소들을 연결하여 의미 있는 블랍을 검출하는 과정을 거치고 추출된 블랍 중 관심 객체를 구분할 수 있는 블랍을 관심 객체로 정의하는 객체레벨 검출 단계를 수행한다. 균일 분포 배경 모델에 의한 배경차감은 전처리 과정에서 HSI 컬러 공간에서의 배경모델 영상을 획득하기 위해 일정시간의 훈련과정을 거쳐 배경영역의 컬러 값에 대한 균일 배경모델을 구성하는 화소레벨검출을 통해 격자이미지 기법을 사용하여 낮은 해상도로 레이블링 하는 블랍레벨 검출 방법을 통한다 이러한 레벨검출 방법은 노이즈 픽셀 등의 예기치 않은 인위적 결함이 있을 경우에도 좋은 성능을 발휘 하게 된다 블랍레벨검출에서 추출한 집합에서 관심객체인 얼굴과 손 영역을 검출하기 위하여 [그림 7]의 에이다부스트 알고리즘을 적용하여 객체레벨을 검출 한다.[2][3].

$$C(x) = \begin{cases} 1 & \sum_{i=1}^T \alpha_i h_i(x) \geq \frac{1}{2} \sum_{i=1}^T \alpha_i \\ 0 & \text{otherwise} \end{cases}$$

이때,  $\alpha_i = \log \frac{1}{\beta_i}$

[그림 7] 에이다부스트 알고리즘



[그림 8] 객체추출의 얼굴 검출화면 ( 에이다부스트 알고리즘 적용 )

사람은 두 개의 눈, 한 개의 코, 입으로 구성되어 있고 각 요소들은 기하학적 위치 관계로 구성되어 있으며 이러한 객체 추출의 얼굴검출은 지식기반방법의 이미지내 히스트그램을 이용하는 특징기반방법과 외형기반 방법을 통하여 검출 가능하다.

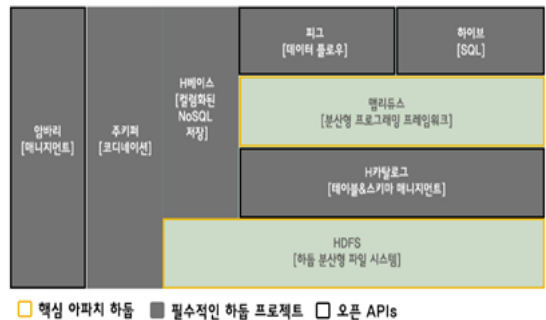
이러한 객체 검출의 전 처리과정후 표정과 제스처 감정을 표현 할 3D객체를 사전에 미리 생성 저장후 3D객체 서버를 활용하여 실시간 정합하게 된다.[3][4][5].



[그림 9] 실시간 정합과정 시스템 순서도

### 3. 교환서버의 Hadoop 분산파일 시스템

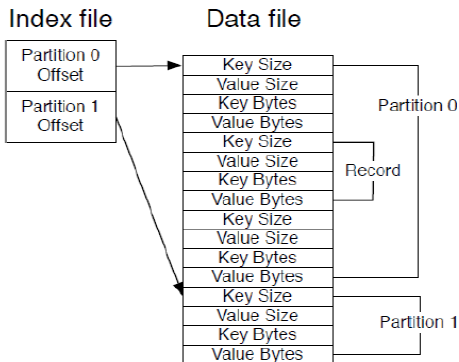
2008년 이래로 데이터량이 44배에 이를 것으로 예측, 실제 이를 증가하고 있다, 이는 단순한 데이터량의 증가가 아니라 데이터의 다양성, 속도등을 함께 포함, 모바일 기기나 소셜 미디어 등에서 생산되는 비정형 데이터의 증가는 소비자의 행태를 바꾸고 있으며 이는 산업계에서 비즈니스 경쟁의 규칙조차 바꾸고 있을 정도의 기술적인 큰 부가가치이다. Hadoop은 HDFS(Hadoop distributed File System)와 분산처리 시스템인 맵리듀스(Map Reduce)를 구현한 것으로 대용량 데이터 처리 분석을 위한 대규모 분산 컴퓨팅 지원 프레임워크이다 (핵심 구성요소 : 저장과 분산처리GDFS를 통해 분산 저장하고, 맵리듀스를 통해 분산 처리 하는 프레임워크).[3].



[그림 10] Hadoop의 구성 요소

※ 2008년 2월(뉴욕타임스) : 1851년부터 1980년 12월까지의 130년 분량의 신문기사 1,100만매를 ‘아마존 S3’에 저장하고, 하둠을 이용해 약 4테라바이트 크기의 데이터를 24시간 만에 변환했고, 이는 일반 서버로 대략 14년 걸리는 작업량임

Hadoop은 Hadoop 맵리듀스(MapReduce)와 대형 클러스터를 위한 맵리듀스 구현, 그리고 HDFS로 구성되어 있다. HDFS란 맵리듀스와 같은 일괄작업 중심의 작업부하를 위해서 설계된 파일 시스템으로 Hadoop의 설치에는 하나의 마스터 node와 많은 작업자 node로 구성되어 있으며 작업(job)-tracker라고 불리는 마스터 node는 클라이언트들로부터 작업을 받아들이고 그들을 더 작은 작업으로 나누고 작업자 node에게 그것들을 분배하는 역할을 한다. 이러한 Hadoop 분산파일은 하드웨어의 오동작을 최소화 하며 간단한 결합 모델로 한번 쓰고 여러 번 읽는 모델에 적합한 구조이다. 파일이 한번 작성되고 단 한번 바뀔 필요가 없는 경우를 위한 것으로 이렇게 함으로써 데이터 처리량을 극대화하여 네트워크 혼잡으로 인하여 발생되어 지는 전체 처리량을 감소화 시키며 이중의 하드웨어와 소프트웨어 플랫폼과의 호환성 및 범용 파일 시스템과 달리 반응 속도보다는 시간당 처리량에 최적화하여 스트리밍 자료 접근을 용이하게 한다.[1]



[그림 11] Hadoop의 구조

Hadoop을 단순히 구축하는 것으로 끝나는 것이 아닌, 다중 작업(job)에 대하여 다중 질의가 발생할 수 있으며, 그 효과적인 질의와 구성 및 응답의 정확도 증가에 대한 계획과 설계 그리고 구축이 필요하다. 이러한 클라우드 클러스터 기술은 OS-level과 High level로 구분하여 언어, 서버 인프라 구성, 시스템 커널드까지 광범위하게 컨트롤 할 수 있는 기술과 노하우가 필요하다.

#### 4. Hadoop 서버 기반의 모바일 웹 플랫폼 구현

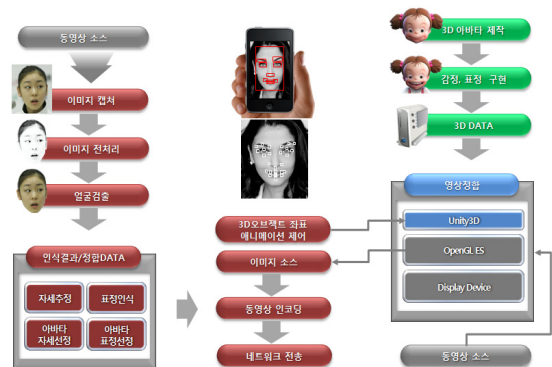
영상통신 감성융합 이란 스마트폰을 통한 화상통신 서비스 기술 개발로 사용자 간에 영상 카메라를 통하여 표정변화를 인지하고 각 표정에 맞는 해당콘텐츠(아바타)를 사람의 실제 얼굴위에 정합하는 것이다. 사용자의 감정과 표정을 인식하기 위한 얼굴검출, 얼굴인식, 감정인식의 핵심기술을 개발 인식된 감정을 대행하는 아바타 및 표정을 극대화 하는 이모티콘의 정합기술 개발로 스마트폰 화상통신 사용자가 사용자간의 표정인식을 통해 현실세계에 불가능한 감정표현을 가능하게 하여 의사전달의 효과를 증대 하는 것이다 <표 1>은 다음의 핵심 기술을 보여 준다.

##### - 주요핵심기술

<표 1>

기술개발 내용	기술개발 목표
얼굴인식 기술	눈, 코, 입 영역 검출
감정인식 기술	웃음, 미소, 놀람 등의 감정 추론
영상정합 기술	검출부위에 가상의 영상 실시간 정합
감성융합 시스템	표정 및 감정에 따른 가상의 오브젝트를 실시간 정합해주는 영상통화 애플리케이션 개발

각 감정과 표정에 따라 제작된 아바타의 소스는 다음의 인식과 정합 단계를 통해서 사용자 애플리케이션에 반영 된다.[1][3].



상용라이브러리 콘텐츠 구현 시스템 영역 기술 구현

[그림 12] 3D아바타의 감정 및 표정 인식과 정합 절차

[그림 12] 3D아바타의 감정 및 표정 인식과 정합 절차  
본 애플리케이션은 화상통화시 사용자의 감정과 표정을 인식하여 해당 감정을 대행하는 아바타와 해당 표정에 효과를 더하는 이모티콘이 정합되는 콘텐츠로서 감정 및 표정 정합지원 화상통화 서비스를 위해 필요한 얼굴 인식, 감정인식, 영상정합, 화상통신 등의 핵심 기술을 개발하고 애플리케이션 콘텐츠와 연동서버를 개발한다.

## 5. 결론

2012년 1월을 기점으로 스마트폰 2258만명, 태블릿 100만명을 돌파할 정도로 단기간에 빠른 성장세를 보이고 있는 스마트폰의 시장은 국내외 IT핵심키워드 1위에 오를 만큼 그 추세는 가히 기하급수적이다. 또한 3D를 포함한 증강현실 역시 미래를 이룰 10대 혁신기술로 꼽히며 향후 미래시장의 선도에서 전 세계를 주도할 것으로 여겨진다. 이와 더불어 국내 해외 대기업간의 콘텐츠를 겸비한 상품화 전쟁이 나날이 지속되고 있는 가운데 3D 감성 융합 사업은 종합적 IT시장의 발전과 기업에 많은 이익창출과 고객들의 가상세계에 대한 인식을 바꿀 수 있는 좋은 기회가 될 것이며, 표정인식 기술에서 한발짝 나아가 표준 동작 데이터에 매칭이 가능한 GTT(Gesture-To-Text)엔진 개발 및 수화 인식 기술에 대한 표준 형상 모형 데이터 기반 구축에 이바지 할 것이다.

## 참 고 문 헌

- [1] 장용석(2008), "3D 형태 모델 기반 3D 얼굴 모델링", 숭실대학교 대학원
- [2] 강병진, 김창현(1998), "얼굴 메쉬와 이미지의 동적 매칭을 이용한 개인 아바타의 3차원 얼굴 합성", 한국정보과학회, Vol.25 No.2II
- [3] 장용석, 정선태, 김부근, 조성원(1998), "3D 변형가능형상 모델 기반 3D 얼굴 모델링", 한국콘텐츠학회논문지, Vol.8 No.1
- [4] 박문호, 교회동, 변혜란(2006), "실제와 합성영상의 비교에 의한 비디오 아바타의 정합", 정보과학회논문지, Vol.33 No.8
- [5] 이웅주(2006), "얼굴특징자 정보를 이용한 인터넷 기반 얼굴관상 해석 및 얼굴아바타 자동생성시스템",

멀티미디어학회논문지, Vol.9 No.8

- [6] 김혁진(2005), "상호대화식 Web3D 콘텐츠의 구현 및 응용", 컴퓨터산업교육학회논문지, 6(2) 317-324 1229-9650
- [7] 배운진(2012), "깊이 정보를 이용한 고속 얼굴 검출 및 추적 알고리즘", 광운대학교

## 오 동 식



- 1999년 2월 : 홍익대학교 기계공학과(공학사)
- 2006년 2월 : 청주대학교 컴퓨터 정보통신과(공학석사)
- 2012년 2월 : 한남대학교 컴퓨터공학과(공학박사)
- 2001년 3월 ~ 현재 : (주)누리뜰희망아이티 대표
- 관심분야 : 네트워크보안, 가상현실, 모바일 소프트웨어
- E-Mail : ocs716@hanmail.net

## 강 준 규



- 2001년 2월 : 우송대학교 컴퓨터디자인과(학사)
- 2012년 8월 : 중앙대학교 창업경영과(석사과정)
- 2001년 8월 ~ 현재 : (주)아이디피쉬 대표
- 관심분야 : 앱콘텐츠, 증강현실, 모바일 소프트웨어
- E-Mail : idfish3@gmail.com

## 신 민 호



- 2009년 3월 : 주성대학 컴퓨터하드웨어응용학과(전문학사)
- 2002년 2월 : 충주대학교 전자계산학과(공학사)
- 2009년 3월 ~ 현재 : (주)지오위즈넷 대표
- 관심분야 : 네트워크보안, 비전인식, 모바일S/W, 3D비전
- E-Mail : help@geowiznet.com