

# 이미지 인식을 이용한 비마커 기반 모바일 증강현실 기법 연구

## Non-Marker Based Mobile Augmented Reality Technology Using Image Recognition

조 휘 준\*, 김 대 원\*  
Huijoon Jo\*, Daewon Kim\*

### 요 약

증강현실 기술이 많이 보편화 되고 사용 양태가 다양화됨에 따라 적용되는 분야 및 범위 또한 우리 생활 곳곳에서 쉽게 나타나고 볼 수 있게 되었다. 기존의 카메라 비전 기반 증강현실 기법들은 현실 세계의 실제 정보 이용 보다는 마커를 이용한 기술을 더 많이 사용하였다. 이러한 마커 인식을 통한 증강현실 기법은 그 응용 범위가 제한적이고 사용자가 해당 서비스 응용 프로그램에 몰입하는데 적절한 환경을 제공하는데 한계가 존재한다. 본 논문에서는 스마트 모바일 단말 기반 증강현실 기술 구현을 위해 단말 장치에 내장된 카메라와 이미지 처리 기술을 활용하여 어떠한 마커도 없는 상태에서 사용자가 현실세계의 영상으로부터 객체를 인식하고 해당 객체에 연결된 3D 콘텐츠와 관련 정보를 현실 세계의 영상에 추가되게 함으로써 서비스가 구현되는 증강현실 기술을 제시하였다. 이미지로부터의 객체 인식은 미리 등록되어 있는 참조용 정보와 비교하는 과정을 통해 진행되었으며 이 과정에서 스마트 모바일 장치의 특성을 고려하여 구동 속도 향상을 목표로 유사도 측정을 위한 연산량을 줄이도록 하였다. 또한 3D 콘텐츠가 단말 화면상에 출력된 후 사용자는 스마트 모바일 장치를 이용한 터치 이벤트를 통해 상호작용이 가능하도록 설계되었다. 이 후 사용자의 선택에 따라 웹 브라우저를 통하여 객체와 연관된 정보를 얻을 수 있도록 하였다. 본 논문에 묘사된 시스템을 이용하여 기존 기술과의 객체 인식 및 동작 속도, 정확도, 인식 오류 검출 정도 등의 성능 차이를 비교 분석하였고 그 결과를 제시함으로써 스마트 모바일 환경에 적합한 증강현실 기법을 소개하고 실험을 통해 검증하고자 하였다.

### Abstract

AR(Augmented Reality) technology is now easily shown around us with respect to its applicable areas' being spreaded into various shapes since the usage is simply generalized and many-sided. Currently existing camera vision based AR used marker based methods rather than using real world's informations. For the marker based AR technology, there are limitations on applicable areas and its environmental properties that a user could immerse into the usage of application program. In this paper, we proposed a novel AR method which users could recognize objects from the real world's data and the related 3-dimensional contents are also displayed. Those are done using image processing skills and a smart mobile embedded camera for terminal based AR implementations without any markers. Object recognition is done from the comparison of pre-registered and referenced images. In this process, we tried to minimize the amount of computations of similarity measurements for improving working speed by considering features of smart mobile devices. Additionally, the proposed method is designed to perform reciprocal interactions through touch events using smart mobile devices after the 3-dimensional contents are displayed on the screen. Since then, a user is able to acquire object related informations through a web browser with respect to the user's choice. With the system described in this paper, we analyzed and compared a degree of object recognition, working speed, recognition error for functional differences to the existing AR technologies. The experimental results are presented and verified in smart mobile environments to be considered as an alternate and appropriate AR technology.

**Keywords** : mobile, augmented reality, application, image, recognition, Android, 3D contents

\* 단국대학교

투고 일자 : 2011. 10. 18 수정완료일자 : 2011. 10. 31

계재확정일자 : 2011. 11. 1

### I. 서 론

최근 스마트 단말 장치가 고성능화 되면서 휴대폰 기능 이외에 컴퓨터 지원, 인터넷 기능 등이 추가되어 본래의 음성 통화 기능에 충실하면서도 보다 지능화된 기기로 발전했다. 이렇게 다양한 기능들의 추가는 스마트폰에서 단순한

통신만이 아닌 이메일, 웹 브라우징, 팩스, 모바일 뱅킹, 게임 등 다른 많은 기능들을 수행 가능하도록 하였다. 이렇게 멀티미디어 단말 기기의 고성능화와 다양한 기능들의 추가로 인해 스마트폰에 대한 관심이 증가하였고 이것은 곧바로 일반 사용자의 증가로 이어졌다. 스마트폰 관련 시장의 규모도 커졌는데 그 중 스마트폰 앱 시장은 통신 사업자뿐만 아니라 일반 사용자들까지도 수입 창출이 가능하게 하여 이 분야에 대한 관심이 증폭되었고 앱 스토어 환경 또한 급진적으로 발전하게 되었다. 스마트폰 콘텐츠 시장의 발전은 자연스럽게 스마트폰에서 동작하는 콘텐츠 내용을 복잡하고 다양하게 만들었다. 시간이 흐를수록 스마트폰 관련 앱은 기존의 단순하고 단방향적인 응용 프로그램에서 벗어나 새로운 기술에 기반을 둔 고품질의 양방향 서비스를 요구하고 있다. 이러한 일반 사용자의 요구에 부응할 수 있는 대안으로 증강현실을 들 수 있다. 이는 가상현실의 한 분야로서, 기존의 가상현실이 실세계 전체를 가상으로 대체하는 기술이라고 한다면, 증강현실은 실세계와 가상세계를 융합한 환경을 사용자에게 제공함으로써 더욱 향상된 현실감과 몰입감, 그리고 상호작용성을 제공하는 기술이다. 여기서 상호작용성이란 사용자가 장치를 통하여 현실에 증강된 정보를 이용하여 현실세계와 대화하며 상호작용 할 수 있도록 하는 것을 말한다 [1]. 이동성을 갖고 있는 멀티미디어 단말 기기들의 성능이 향상되면서 증강현실 기술은 스마트폰에 다양하게 활용되고 있다. 증강현실을 구현하는 방법으로는 Optical flow, 마커 기반, NFT (Natural Feature Tracking) 기반, 센서 기반 기술 등이 있다 [2]. 증강현실과 관련하여 스마트폰을 사용할 때 발생하는 이동시의 문제점을 해결하고 사용자에게 효율적인 UI/UX (User Interface/User eXperience) 및 적절한 콘텐츠를 제공하기 위해 해당 기술들이 많이 연구되고 있다. 그 중 현재의 스마트폰에 대부분 활용되고 있는 마커 기반 증강현실 기술은 기존의 데스크탑 환경 상의 연구결과를 그대로 옮겨 놓은 기술이기에 입출력장치, 카메라 활용성, 시스템의 성능 등 여러 가지 측면에서 많은 차이가 있다 [3]. 또한 다양한 분야에서 활용되는 영상처리 기법은 주어진 영상을 2차원 신호로 취급하여 이에 해당하는 표준 신호 처리 기법을 적용하고 있다 [4]. 본 논문에서는 단순한 마커 기반보다는 영상 처리 기법을 활용한 비마커 기반 증강현실 기술 구현을 통하여 스마트 단말 상에서의 증강현실 서비스 기술을 개선하고 일반 사용자에게는 높은 몰입감과 향상된 상호작용성을 제공하고자 하는 목적의 연구를 진행하였다. 영상 인식 및 검출 기술을 바탕으로 한 비마커 기반 증강현실로 구현된 모바일 앱은 구글 안드로이드 플랫폼 기반 스마트 단말에서 콘텐츠의 효율적인 배치와 운용 속도 향상 등을 고려하여 실험이 진행 되었다. 또한 핵심 알고리즘 구현 및 기존 기술과의 성능 비교와 스마트폰의 제한적인 환경 극복을 위해 응용 프로그램의 독립적 화면

구조 및 UI/UX가 설계되었다.

## II. 이미지 객체 검출 및 인식을 통한 증강현실

### 2.1 증강현실

일반적으로 사용자가 가상의 세계에 몰입하게 되어 실제 세계를 볼 수 없게 되는 가상현실과는 달리 증강현실은 완벽한 가상의 공간만을 제공하지는 않는다. 증강현실이란 가상현실의 한 분야로 실제세계에 존재하는 사물이나 장소에 부가적인 정보나 의미를 부여하여 생성된 객체를 실제세계에 존재하는 사물처럼 보이도록 하는 기술이다 [5]. 그림 1은 현실세계와 가상현실, 그리고 증강현실 공간 사이의 관계를 연속선의 측면에서 나타내고 있다.

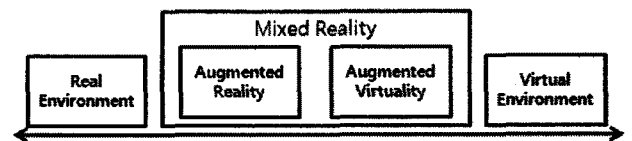


그림 1. 현실세계와 가상현실 및 증강현실 사이의 관계  
Fig. 1. Relationship among reality, virtuality, and augmented reality

일반적인 정의에 따르면 증강현실은 첫째, 실제 영상과 가상 영상을 결합하며 둘째, 실시간으로 상호 작용하고 셋째, 실세계에 가상의 영상을 등록하는 기술로 정의된다. 증강현실 적용이 가능한 분야로는 의료 [6], 군사, 교육 [1], 게임, 의료, 부동산, 유통, 자동차, 관광, 금융, 항공, 영화, 디자인, 건축, SNS (Social Network Service) 등으로서 거의 모든 분야를 망라하고 있다. 증강현실에는 마커 검출, 마커 정합, 트래킹, 3D 오브젝트 렌더링 등 다양한 기법이 존재하는데 핵심적으로 요구되는 기술 중 하나는 현실세계와 3D로 이루어진 객체를 정확하게 일치시키는 것이다. 이를 위해서는 먼저 영상에서 목표를 정확하게 인식하고 추적하는 것이 요구된다. 객체추적은 일반적으로 센서 기반, 컴퓨터 시각 기반, 센서와 컴퓨터를 결합한 하이브리드 시각 기반 기술로써 현실화 된다. 객체추적 기술 중 가장 보편화 된 기술은 인공 마커를 추적하는 것인데 이 기술은 마커를 이용함으로써 객체 탐지를 용이하게 최소한의 오류로 수행하기 때문에 증강현실 구현에 적절했다. 하지만 이러한 마커의 이용은 점차 다양해지고 세분화 되는 응용분야에서 사용하는데 제약사항이 많다는 단점이 있다. 본 연구에서는 이러한 단점의 보완을 위해 비마커 기반 방식을 채택하였고 이를 위해 카메라로부터 들어온 영상을 통해 점, 선, 모서리, 텍스처 등과 같은 특징을 추출하여 증강현실 응용 프로그램에 적용하였다. 해당 객체의 특징을 추출하여 동작하기 때문에 사용자가 사전에 정의한 특정 패턴을 마커와 같이 이용할 수도 있다. 비마커 기반 방식은 증

강현실에 있어 가장 난이도가 높고 중요한 기술로 평가받고 있는데 이는 객체의 특징을 기반으로 회전각, 방향, 조명의 변화, 부분 겹침 등을 인식하고 검출하는데 뛰어난 성능을 보이기 때문이다. 이렇게 확보된 영상정보에 GPS, 관성 센서, 컴퓨터 시각 센서 등의 정보를 이용한 실외 추적 기능을 덧붙인다면 더욱 만족할 만한 객체 인식 및 추적의 결과를 얻을 수 있다 [7][8].

**2.2 이미지 객체 검출 및 인식**

객체 검출이란 여러 가지 모델 물체들에 대한 정보를 갖고 있는 상태에서 임의의 영상이 투입된 경우 입력 영상과 모델 물체들 자료간의 정합과정을 거쳐 입력 영상 안에 포함되어 있는 물체와 가장 유사하다고 판단되는 모델을 검색해내거나 목표 물체의 위치를 찾아내는 일련의 과정을 말한다. 이 동작은 입력 영상의 상태에 따라 민감한 영향을 받는데 그 이유는 입력 영상 안에서 객체의 크기, 방향, 위치 등이 다르게 제공될 수 있고 또, 한 영상 내에 여러 개의 객체가 포함되어 있거나 목표 객체가 다른 물체에 의해 가려져 객체의 일부분만 보이거나 일그러져 보이는 경우도 발생할 수 있기 때문이다. 객체 인식 체계 구축을 위해서는 보통 두 가지 방법이 사용되는데 첫 번째 기법은 외관 기반 기법으로서 입력 영상에서 특정 영역의 영상을 사용하는 것이다. 이 기법은 특정 영역의 밝기 값을 그대로 사용하기 때문에 물체의 모양이나 질감 정보가 함축적으로 표현된다. 또한 특정 영역 내의 모든 화소들의 밝기 값들로 정의된 벡터를 사용하여 해당 객체나 패턴을 표현하는데, 영역의 크기에 따라서 벡터의 크기가 커질 수 있기 때문에 차원 감소 기법이 사용되며 주로 얼굴 인식 영역에서 활발하게 사용되고 있다 [9][10]. 두 번째 기법은 입력 영상에 대한 전처리 과정을 거쳐 추출된 특성을 이용하는 특징 기반 판단 기법이다. 이 기법은 영상의 밝기 값의 변화량을 계산하여 변화가 심한 모서리 부분을 검출하고 이를 바탕으로 객체의 윤곽선이나, 윤곽에 근사하는 선분, 곡선 등을 찾아낸다. 또한 입력 영상 안에 존재하는 모서리 점, 특정 변환에 대한 불변의 점, 그리고 회전변환이나 크기변환 등에 대하여 불변인 특정 점 등을 추출하여 객체 검출에 이용하기도 한다 [11][12]. 객체 검출과 더불어 이미지 인식 기술은 입력 영상을 받아들여 얻어낸 정보를 통해 특징을 추출하여 시스템에 등록된 자료와 비교한 후 해당 객체가 무엇인지를 판단해 내는 기술이다. 이는 의료, 지문, 얼굴, 위패, 교통법규 위반차량, 특정표적 등의 분야에서 응용 가능한 인식 기술로써 최근 다양한 기술과 융합 되어 사용되고 있다 [13][14]. 본 연구에서는 특징점 추출을 통한 이미지 인식 및 검출 기술을 활용하여 객체를 판단한 후 증강현실 기술과 접목, 응용이 가능한 서비스 기술을 다루고 이를 구현하여 기존 기술과의 성능 비교 및 평가를 진행하였다

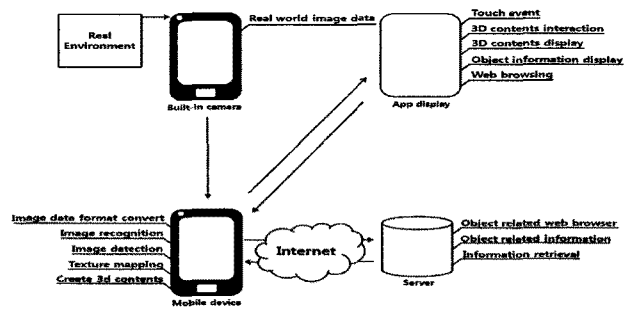


그림 2. 객체인식 및 검출과정과 융합된 증강현실 기술 구조도

Fig. 2. Structure of object recognition, detection and integration of augmented reality technology process

또한 객체 인식 및 검출이 완료되면 사전 등록된 해당 3D 콘텐츠를 불러내 화면에 디스플레이 하도록 하고 지정된 동작을 통해 웹 상에 존재하는 해당 콘텐츠 관련 정보를 검색하고 획득하도록 하였다. 그림 2는 객체 인식 및 검출의 단계로부터 시작하여 해당 특징에 대한 정합 과정을 통해 증강현실 기술과 융합한 후 최종 정보를 디스플레이 하고 온라인 상의 서버에 접속하여 필요 정보를 검색한 후 획득하는 과정을 순차적으로 보이고 있다.

**III. 모바일 환경 상 이미지 객체 인식을 통한 증강현실**

본 연구는 스마트폰의 내장 카메라로부터 입력 받은 영상을 인식하여 비마커 기반의 증강현실 기술을 학습하고 구현하는데 그 목적을 두었다. 해당 모바일 앱은 디스플레이 이용 화면에 내장 카메라로부터 들어온 입력 영상을 띄우고 중간에 입력영상의 특정부분 추출 표시를 위한 컨트롤 박스를 배치하고 있다. 핵심 부분인 이미지 인식 및 검출 관련 기술은 G20 서울 정상회의 회원국인 20개 국가의 국기를 특정 영역을 지정하여 이미지 형성 후 해당 국가 관련 정보를 웹 상에서 검색, 브라우징 하는데 활용 되었다. 이어 증강현실 기술을 이용하여 해당 국가 관련 3D 콘텐츠를 디스플레이하고 인터랙션 동작을 마친 후 그 나라의 대표 웹 사이트로 이동하도록 하였다. 그림 3은 이미지로부터 객체를 인식하고 증강현실을 이용하여 서비스를 제공하는 응용 알고리즘의 실행 절차를 보이고 있다. 실시간으로 현실의 영상을 모바일 카메라로 입력 받은 후 절차에 따라 저장된 영상을 전처리 하고 해당 이미지 인식 및 검출 결과를 3차원 콘텐츠 디스플레이를 위한 모듈에 전송시켜 화면에 출력되게 한다. 이 후 선택 절차에 따라 해당 국가의 정보가 있는 홈페이지로 이동할 수 있도록 하고 있다. 이는 G20 정상회의에 참석한 VIP 및 IT 관계자들에게 모바일 상의 증강현실 기술을 소개하고 해당 국가 담당자들에게 웹 상에서 필요한 정보를 제공하기 위하여 진행된 작업이

다.

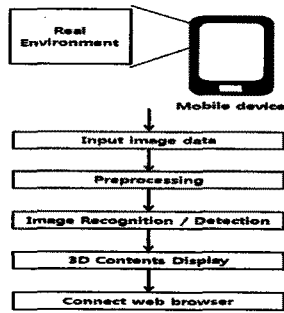


그림 3. 이미지 인식 및 증강현실 적용 절차

Fig. 3. Application process of the image recognition and augmented reality

그림 3에서 저장된 이미지가 인식 및 검출 된 후 해당 국가에 대한 분류가 완성되고 나면 사용자의 선택에 따라 브라우저 기능을 이용하여 해당 국가 홈페이지로 이동할 수 있게 하였다.

3.1 객체 검출 및 인식

이미지 인식 및 검출을 위한 영상처리 기법은 비전 기반 증강현실 분야에서 매우 중요한 부분을 차지하고 있는 기술로써 다양한 연구가 진행되고 있다. 데스크탑 기반의 영상처리 기술은 장치사양과 작업환경의 발전으로 많은 발전이 있어 왔지만 제한적인 리소스 기반 모바일 단말 장치에서의 이미지 윤곽 인식 및 검출 관련 기술은 연구 개발이 더 필요하다. 본 연구에 사용된 모바일 단말 장치에서는 타겟이 되는 객체를 지정된 영역 안에 고정시키고 화면상에 터치 이벤트를 발생 시켜 전처리 과정을 시작하고 스마트 단말 내장 카메라로부터 YUV420 포맷의 영상을 입력 받는다. 이것을 다음 단계에서의 사용을 위해 RGB포맷으로 변환 시키고 인식된 특정 부분만의 신호를 추출해낸다. 입력된 영상은 스마트 단말 카메라의 특성 상 회전 및 축소의 처리 과정을 거치게 되는데 이는 구동 시의 리소스 절약과 처리 속도 향상을 위해서이다. 입력영상에 대한 정보가 정확히 전달되고 전처리 과정이 오류 없이 완료되면 다음 단계인 영상인식 및 추출과정이 진행된다. 이 단계에서는 전처리가 완료된 영상정보와 미리 준비되어 있는 참조영상과의 비교 분석을 통해 유사도를 측정하고 이를 통해 객체를 추출한다.

$$R(x,y) = \sum_{x',y'} (T'(x',y') \cdot I'(x+x',y+y')) \quad (1)$$

$$, x' = 0..w-1, y' = 0..h-1$$

$$T'(x',y') = T(x',y') - \frac{\sum_{x'',y''} T(x'',y'')}{(w \cdot h)} \quad (2)$$

$$I'(x+x',y+y') = I(x+x',y+y') \quad (3)$$

$$= \frac{\sum_{x'',y''} I(x+x'',y+y'')}{(w \cdot h)}$$

수 식 (1)에서 T는 전처리가 완료된 입력 영상, I는 사전에 준비된 참조 영상, 그리고 R은 이미지 인식 및 검출 후의 출력 영상을 의미한다. 또한 w는 영상 T의 가로 길이, h는 세로 길이를 나타내며 이를 통해 두 값의 변동 방향 및 양상을 나타내는 공분산을 구한다. 공분산이 0보다 크면 영상 T의 픽셀 값이 증가할 때 I의 픽셀 값도 증가하고, 공분산이 0보다 작으면 반대로 영상 T의 픽셀 값이 감소할 때 I의 픽셀 값도 감소하는 경향을 보인다. 여기서 영상 T의 크기는 I의 크기보다 작도록 설계되었으며 공분산의 값은 -∞부터 ∞까지의 값을 가질 수 있다. 공분산은 측정 단위에 따라 그 값이 크게 변할 수 있는데 절대값의 크기가 클수록 두 이미지 간의 관련성이 크다고 할 수 있다. 식 (2)와 (3)에서 T'와 I'는 각각 이미지 인식 및 검출 시, R에 해당하는 결과를 얻고자 이전 단계에서 비교 분석을 통해 유사도를 측정하고 이를 통해 객체를 추출하는데 사용되는 중간 단계에서의 입력(T') 및 참조(I') 영상이다. 입력 영상과 참조 영상 사이의 다양한 비교를 위해 공분산의 값을 영상 T의 표준편차와 I의 해당 영역의 표준편차로 나누어 상관계수를 구한 후 비교하도록 하였다. 이러한 방법으로 영상 T가 I의 전체 범위를 탐색하면서 해당 위치마다 공분산을 구하게 된다. 따라서 모바일 단말 장치의 카메라에 입력되는 객체를 추출하기 위해 상관계수 특정 추출 기법을 이용해 이미지 간의 유사도를 계산하였다. 상관계수 특정 추출 과정에서는 두 영상간의 평균값을 계산하여 비교한 후 일치 및 불일치 등의 매칭 여부 판단 작업을 수행하였다. 두 영상이 일치하는 지정 기준값을 만족하면 1을, 연관성이 부족하여 기준값을 만족시키지 못할 경우에는 0을 출력한다.

$$Z(x,y) = \sqrt{\sum_{x',y'} T'(x',y')^2 \cdot \sum_{x',y'} I'(x+x',y+y')^2} \quad (4)$$

$$R'(x,y) = \frac{\sum_{x',y'} (T'(x',y') \cdot I'(x+x',y+y'))}{\sqrt{\sum_{x',y'} T'(x',y')^2 \cdot \sum_{x',y'} I'(x+x',y+y')^2}} \quad (5)$$

수 식 (4)는 공분산을 상관계수로 만들기 위한 영상 T와 I의 해당 영역의 표준편차를 구하는 과정을 나타낸다. 여기서 표준편차는 영상 T와 I의 해당 영역 평균값으로부터 그 영상 데이터의 산술 평균 거리이다. 공분산 값은 같은 데이터를 이용해 측정하더라도 그 단위에 따라 수치에 많은 차이가 날 수 있기 때문에 표준화를 통해 상대적으로 비교하기 수월한 값으로 바꿔서 사용한다. 또한 수 식 (5)는 수 식 (1)을 수 식 (4)로 나눈 것으로서 상관계수를 구하는 과정을 보이고 있다. 수 식 (5)와 같이 연산하는 이유는 값의 비교를 쉽게 하고 영상처리 과정에서 빛의 간섭 영향을 줄이는 효과도 얻을 수 있어서이다. 상관계수는 공분산을 표준화 시킨 것인데 이 과정이 끝나면 결과 영상인 R의 상관

관계 수 맵으로부터 가장 유사한 값이 발생한 위치에 대한 정보를 얻을 수 있다. 이 위치 정보를 이용하여 가장 유사한 객체의 정보를 추출하게 되며 그 후 해당 객체의 3D 콘텐츠를 불러오게 된다. 이 과정에서 입력된 객체의 특징 값과 가장 유사한 참조 객체의 정보를 비교하여 다음 단계로 넘어가게 된다.

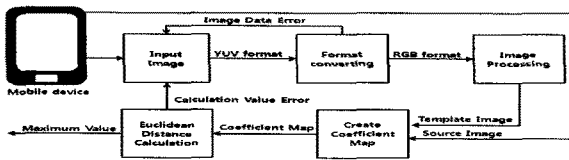


그림 4. 이미지 객체 검출 및 인식 과정도

Fig. 4. Block diagram of the image object detection and recognition process

이 단계에서 곧바로 객체 정보를 전달하게 되면 사용자가 원하지 않는 아이템이 다음 과정으로 전달되는 오류가 발생할 가능성이 있다. 그러한 오류의 발생을 예방하기 위해 유사 값 정보를 미리 지정해 놓은 기준치와 비교하도록 하였다. 만약 유사도를 나타내는 측정값이 기준치를 넘지 못하면 해당 객체는 참조 객체로 사진 등록되어 있지 않다고 판단하게 된다. 이러한 경우 시스템은 사용자에게 경고 메시지를 전달하고 전처리 이전의 과정인 영상입력 과정부터 다시 실행하도록 하였다. 오류 없이 모든 과정이 진행된다면 카메라로부터 입력되어 추출된 정보의 객체와 일치하는 참조 객체의 3D 콘텐츠가 호출된다. 그림 4는 앞서 설명한 과정들을 도식화하여 블록 다이어그램으로 보이고 있다.

### 3.2 모바일 증강현실을 통한 3차원 콘텐츠 인터랙션

여기 묘사된 애플리케이션 실행을 위해서는 현실세계의 영상 획득을 위한 모바일 장치에 내장 카메라가 필요하고 3D 콘텐츠와의 상호작용을 위해 터치스크린이 장착되어 있어야 한다. 또한 해당 정보 검색 및 제공을 위해 웹사이트 연동이 가능한 3G망이나 무선 네트워크 환경이 갖추어져 있어야 한다. 모바일 상에서 증강현실을 구현하는 과정에는 사용자에게 보다 나은 몰입감을 경험할 수 있도록 고려해야 할 제약 사항이 있다. 여기에는 모바일 기기의 소형 화면, 데스크탑 환경에 비해 낮은 하드웨어 사양, 무선의 특성 상 유선에 비해 상대적으로 느린 네트워크 환경 등이 있다. 또한 모바일 단말의 작은 화면으로 인해 3D 콘텐츠의 크기와 사용자 인터페이스를 설계하는 과정에 제약 조건이 따르는데 기기 사용자에게 실행방법에 대해 설명하면서 인터랙션이 가능하도록 하였다.

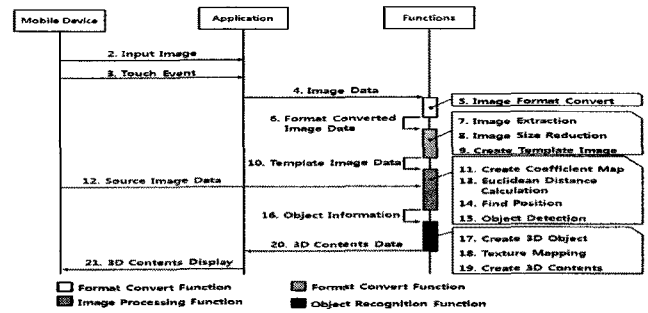


그림 5. 3D 콘텐츠 디스플레이 과정을 포함한 증강현실 신호 흐름도

Fig. 5. Signal flow of the 3D contents display involved augmented reality

제안된 방식에 따라 구현된 응용 프로그램은 인식된 객체가 텍스처 맵핑된 3D 큐브를 콘텐츠로 사용하였는데 이는 데스크탑 환경에 비해 상대적으로 낮은 모바일 단말기기의 하드웨어 사양으로 인해 사용자의 몰입감 유도에 방해가 될 수 있기 때문이다. 또한 구현된 응용 프로그램에는 3D 콘텐츠 디스플레이 이후 인식된 객체에 해당되는 정보를 담은 웹에 접속하는 기능이 있으며 3G나 Wi-Fi 환경이 갖추어져 있다면 연결이 가능하도록 설계 되어 있다. 그림 5는 지금까지 설명된 일련의 과정들을 순서도로 나타내고 있다. 비전 인식을 이용하는 비마커 기반 증강현실 기술은 실제 세계의 특징을 통해 얻은 정보를 이용해 사용자에게 증강된 정보를 제공하고 몰입감을 향상 시킨다. 3.1장에서 서술된 영상인식 및 추출 기법에 의해 객체에 대한 정보를 전달 받게 되면 이에 해당하는 3D 콘텐츠를 디스플레이 하도록 설계 되었다. 앞서 묘사되었듯이 3D 콘텐츠는 3D 큐브로 제작되었으며 사용자의 몰입감을 향상시키기 위해 객체와 연관이 있는 텍스처를 맵핑하여 디스플레이 함과 동시에 해당 관련 음원이 재생 되도록 하였다.



(a) (b) (c)

그림 6. 실시간 모바일 증강현실을 이용한 3차원 콘텐츠 디스플레이: (a) 3D 큐브 콘텐츠 디스플레이, (b) 3D 콘텐츠 터치 인터랙션, (c) 웹 사이트 브라우징

Fig. 6. 3D contents display and control in real time for the mobile augmented reality: (a) 3D cube contents display, (b) 3D contents touch interaction, (c) Web site browsing

그림 6은 스마트 단말 상에 구현된 앱이 실제로 구동되는 모습을 보이고 있다. 그림 6의 화면들은 스마트 단말 상에서 구동되는 응용프로그램의 캡처 화면인데 (a), (b) 각 그림의 배경은 스마트폰에 내장된 카메라를 통해 실시간으로 비춰지는 G20 정상회의 안내용 리플렛에 인쇄된 참가국의 국기들이다. 단말 디스플레이를 통해 단순히 3D 콘텐츠를 보기만 하는 것이 아니라 이를 제어 하면서 증강된 정보와 간단한 터치 이벤트를 통해 인터랙션이 가능하도록 설계되었다. 그림 6의 (a)는 이미지로 인식된 국기에 해당하는 3D 큐브 콘텐츠가 관련 정보와 함께 출력되어 있는 것을 보이고 있다. 스마트폰 화면에 3D 큐브가 디스플레이 되면 플링 이벤트를 통해 이를 원하는 방향으로 회전하는 것이 가능하다. 다음 과정으로 그림 6의 (b)에 보이는 것처럼 3D 콘텐츠에 더블 탭 이벤트를 행하여 이를 사라지게 하고, 마지막으로 터치 이벤트를 통하여 (c)에 보이는 것과 같이 해당 국가의 대표 웹사이트로 이동하게 된다.

IV. 실험결과 및 성능평가

이 논문에 제시된 시스템은 안드로이드 2.2 SDK와 이클립스 3.5 환경에서 개발되었고 실제 타겟 단말을 이용하여 실험이 진행되었다. 시스템 구현에 사용된 단말 장치들은 Samsung SHW-M100S (Galaxy A), SHW-M110S (Galaxy S), SHW-M180S (Galaxy Tab) 등이며 영상인식 및 추출을 위해 각 단말 장치들의 내장 카메라를 사용했다. 각 장치들의 제품 사양이 달라 실행 속도 면에서 차이가 조금씩 있긴 했지만 내장 카메라의 성능에 있어서는 많은 차이가 나타나지 않았다.

	SHW-M180S (Galaxy Tab)	SHW-M110S (Galaxy S)	SHW-M100S (Galaxy A)
CPU	A8 Cortex 1GHz		
OS	Android OS 2.2		
RAM	512 MB		
MEMORY	16 GB		8 GB
화면크기	7 inch LCD	4 inch Super AMOLED	3.7 inch AMOLED Plus
터치스크린	정전식 Multi-Touch		
해상도	1024 x 600	800 x 480	
동영상	Divx, XviD, MKV(1080p)	Divx Xvid(720p)	Divx (720p)
Wi-Fi	802.11 b/g/n type		
배터리용량	4000 mAh	1500 mAh	

표 1. 실험에 사용된 하드웨어 장치 사양 비교  
Table 1. Comparison of the hardware devices' specification used in the experiment

표 1은 실험에 사용된 무선 단말 장치들의 사양을 나타내고 있다. 제안된 시스템은 표 1에 나타난 바와 같이 안드로이드 OS 2.2 기반의 스마트 폰에 설치되어 실험 진행되

었으며 기존 응용 프로그램과의 성능 비교 및 분석을 위해 마커 기반 증강현실 기술을 이용하여 구현된 애플리케이션을 사용하였다. 성능 평가의 비교 대상이 된 기존 기술은 자바용 Ny-AR-Toolkit의 핵심 라이브러리를 사용하는 안드로이드용 Ny-AR-Toolkit을 통해 구현 되었으며 증강현실 클래스 라이브러리로써 AR-Toolkit과 호환이 가능하다. 그림 7은 시뮬레이션 시 응용 프로그램을 구동하는 화면을 보이고 있다. (a)와 (c)를 통해 본 논문에서 제안하는 시스템의 동작을 보이고 있는데 (a)는 시스템을 시작한 후 G20 국가 중 임의의 국기를 인식하는 모습이고 (c)는 (a)에서 화면 터치 후 인식된 나라의 정보와 함께 국기로 텍스트링된 3D 큐브가 화면상에 보이는 장면을 나타내고 있다. (b)와 (d)는 마커 기반의 기존 기술 동작을 보이고 있는데 (b)는 Ny-AR-Toolkit 동작 시 초기 화면이다. 여기서 스마트 단말의 내장 카메라를 통해 입력받은 영상에 마커가 없으면 (b)와 같은 화면이 디스플레이 장치에 보이게 되고 만약 마커가 존재한다면 (d)와 같이 해당 3D 객체가 스마트 단말의 화면에 나타나게 된다.

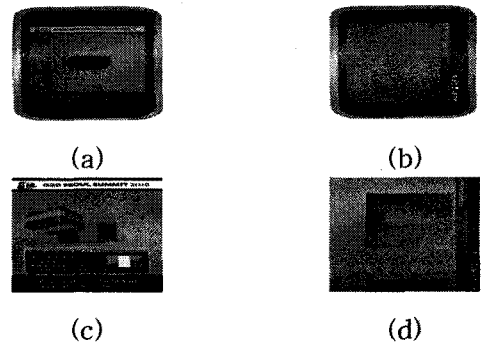


그림 7. 시뮬레이션 결과: (a) 오브젝트 인식, (b) Ny-AR-Toolkit 동작 초기 화면, (c) (a)의 3D 콘텐츠 디스플레이, (d) (b)의 3D 오브젝트

Fig. 7. Simulation results: (a) Object recognition, (b) Initial operation display of the Ny-AR-Toolkit, (c) 3D content's display of (a), (d) 3D object of (b)

실험은 스마트 폰의 내장 카메라를 통해 입력받은 정지 영상을 사용하여 이미지 처리를 위해 포맷 변환을 거친 후 이미지 상의 객체를 정확히 인식하는 과정까지 걸리는 시간과 에러율을 측정하는 방식으로 진행 되었다. 단 한 개의 마커를 인식하도록 구현된 기존 기술과의 성능 비교를 위해 실험 과정에서는 제안된 응용 프로그램 또한 동등한 조건에서 구동되도록 하였다. 따라서 실험 수행을 위해 단 한 개의 객체인 임의의 국가의 국기를 인식하도록 한 후 진행 되었으며 최대한 객관적인 결과를 얻기 위해 각각 100회에 걸쳐 구동 시간을 측정하고 기록하였다. 표 2는 실험 후 측정 결과를 마커 기반 인식 기술과 비교하여 전체 100회 중 매 5회 마다 그 결과를 기록하여 표로 작성한 것이다. 이는

스마트 폰의 내장 카메라를 통한 객체 인식 실험을 진행한 회 수 마다 여기에서 제안된 시스템과 마커 기반의 기존 기술을 이용한 응용 프로그램의 구동 시간을 막대그래프로 비교하여 나타낸 것이다.

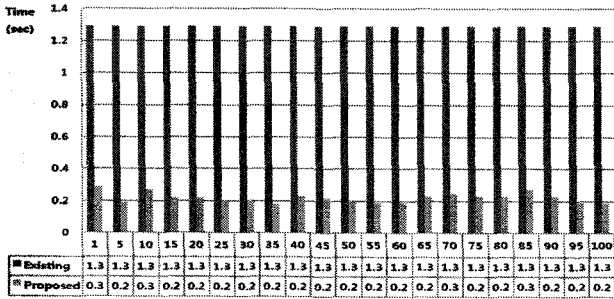


표 2. 객체 인식 속도 측정 결과

Table 2. Results of the object recognition velocity measurement of the proposed method

표 2를 보면 본 연구에서 제시된 기술의 평균 동작 시간은 0.219초이고 마커 기반의 기존 기술의 평균 동작 시간은 1.3초에 달하여 제안된 기술이 기존의 것보다 구동 속도 측면에서 거의 평균 6배 차이가 날 정도로 우월하게 빠른 것을 확인할 수 있다. 여기서 구동 속도 차이의 근본적인 이유는 영상 처리 과정에서의 연산량에서 많은 차이가 나기 때문이다. 마커 기반의 기존 기술은 입력 영상 전체로부터 마커가 존재하는지를 찾아서 모바일 장치의 해당 위치에 3D 콘텐츠를 출력하는 반면 본 논문에 묘사된 기술은 스마트 모바일 환경에의 최적화를 위해 카메라로부터의 입력 영상에서 특정 부분만을 추출하고 그 특징을 분석하여 객체를 인식함으로써 동작 시간을 줄이도록 하였다.

횟수	1	10	20	30	40	50	60	70	80	90	100
기존 기술	o	o	o	o	o	o	o	o	o	o	o
제안된 기술	o	o	o	o	o	o	o	o	o	o	o

표 3. 객체 인식 측정 결과

Table 3. Results of the object recognition measurement

주목할 만 한 점은 마커 기반의 기존 기술은 영상 전체를 이용하여 마커를 인식하며 바이너리 이미지를 사용하는 데 비해 제안된 기술은 YUV420 포맷의 컬러 영상을 입력 받아 특정 추출을 위해 부분적으로 영상처리를 수행한다는 것이다. 흑백 영상에 비해 픽셀 당 정보량이 많은 YUV420 포맷의 컬러 영상으로서는 영상처리를 수행하는 과정에서 연산량이 증가할 수밖에 없는데 그럼에도 불구하고 뛰어난 구동 성능을 보인 제안된 기술은 스마트 단말의 구동 환경에서 동작하기에 매우 효과적인 시스템인 것으로 사료된다. 표 3은 입력 영상의 객체 인식 정확도를 측정한 결과를 나

타내고 있다. 정확도 측정 실험은 제안된 기술은 임의의 국가의 국기를, 기존의 기술은 마커를 인식하는 방식으로 진행 되었다. 여기서 국기의 인식 동작 자체가 처음부터 원활하게 이루어지지 않는 경우가 있었는데 이는 각 나라의 국기를 컬러 출력 후 인식 시 입력 자료로 사용할 때 선명성이 부족해 손실된 정보가 존재하는 경우가 있기 때문이다. 미국, 사우디아라비아 등의 국기가 이에 해당한다. 초기 인식 오류를 일으키는 또 다른 요인으로는 영상 입력 과정에서 그림자가 많이 포함되는 경우이다. 아르헨티나, 일본, 대한민국 등의 국기가 주로 이 경우에 해당 되었다. 실험 결과 표 3에 나타난 바와 같이 총 100회 시도 중 매 10회마다 측정했을 때 두 가지 기술 모두 입력 영상에서 객체가 파악되고 나면 객체 인식 과정에서는 오류가 발생하지 않는 것을 알 수 있다. 다음으로는 오류 측정 실험을 진행하였는데 이는 해당 응용 프로그램이 인식하지 말아야 할 객체를 잘못 인식하는 경우가 있는지 측정하는 실험이다. 마커를 이용한 기존 기술의 경우 마커와 유사한 객체 또는 마커와 전혀 상관이 없는 객체들을 오류를 일으키며 인식하는지 측정 하였다. 또한 제안된 기술은 인식을 하기 위해 미리 등록해놓은 20개 국가의 국기 이외에 다른 나라의 국기를 인식하는지 그리고 국기와 전혀 상관없는 객체들을 인식하는지 측정하였다.

횟수	1	10	20	30	40	50	60	70	80	90	100
기존 기술	o	o	o	o	x	o	x	o	o	o	o
제안된 기술	o	o	o	o	o	o	o	o	o	o	o

표 4. 객체 인식 오류 측정 결과

Table 4. Results of the object recognition error measurement

실험 결과 표 4를 보면 본 논문에서 제안된 기술은 잘못된 객체를 인식하는 경우가 발생하지 않았고 기존의 기술은 약 18%의 확률로 해당 객체를 잘못 인식하는 경우가 발생하였다. 따라서 제안된 기술은 구동 속도 측면에서 뿐만 아니라 객체 인식률에서도 우위를 보여 향후 스마트 단말 환경에서 활용 가치가 크다고 할 수 있다. 비마커 기반 이미지 인식 및 검출 기술을 활용하여 구성된 본 논문에서의 제안 기술은 인식 과정이 완료되면 해당 3D 콘텐츠를 중강현실의 일환으로 디스플레이에 출력하고 지정된 동작을 통해 관련 웹 사이트로 이동하도록 설계되었다. 스마트 폰을 위시하여 태블릿 PC 등 다양한 스마트 단말이 활용되어지는 지금, 단말에 내장된 카메라를 통해 입력된 영상을 활용하여 응용 프로그램 서비스를 제공하는 경우가 많아지고 있으며 본 논문에서 제안된 기술 또한 관련된 응용 분야에서 일정 부분 기여할 수 있는 역할이 있으리라 기대된다.

V. 결론

본 논문에서는 멀티미디어 중심의 스마트 단말에 내장된 카메라를 통해 입력된 영상을 비 마커 기반의 객체 인식 및 검출 알고리즘을 사용하여 판단한 후 해당 3D 콘텐츠를 단말 디스플레이에 출력하는 증강현실 서비스에 관하여 묘사하였다. 실제 세계의 영상으로부터 얻은 정보에 덧붙여 사용자에게 증강된 정보를 제공함에 있어 영상 증첩을 구현하기 위해서는 입력된 영상을 인식한 후 정확한 매칭 여부 판단을 내리는 기술이 필요하다. 기존 마커 기반의 증강현실 기술은 스마트 단말의 작은 디스플레이 화면과 장치의 기본 리소스 및 성능, 배터리 소모량 등 다양한 측면에서 무선 멀티미디어 단말 장치의 사양에 최적화 되지 않은 부분이 존재한다. 또한 마커의 사용으로 인해 응용 범위가 제한적이었고, 사용자의 몰입감이 감소되는 경향이 있다. 따라서 본 연구에서는 카메라로부터 입력된 영상과 사전 자료로 저장되어 있는 참조 영상 간의 공분산 및 상관 계수 등을 계산하여 이미지 간의 유사도를 판단하도록 하였다. 또한 영상인식 및 특징 추출 기술을 통해 구동 속도와 사용자의 몰입감이 향상 되도록 하였고, 스마트 단말의 작은 화면을 고려하여 3D 콘텐츠와의 터치 동작을 통한 수월한 정보 교환을 목표로 하였다. 안드로이드 기반의 스마트 단말을 통한 실험은 구동 속도, 정확도 및 인식 오류를 측정하는 내용으로 진행되었고 이들 결과는 기존 마커 기반의 기술과 비교하여 매우 경쟁력이 있는 것으로 나타났다. 향후 스마트 단말 상에서 동작되는 비전 기반의 증강현실 기술이 많아진다면 의미있는 상대적 성능 평가 또한 가능할 것으로 보인다. 본 논문에서 제안된 비마커 기반 증강현실 기술은 3D 콘텐츠가 디스플레이 될 때 입력 영상이 정지하고, 그 입력 영상에서 객체를 추출하는 과정에서 객체의 크기 조절이나 회전 동작이 부자연스러운 부분이 존재한다. 향후 현재 구현된 기술을 비디오 영상으로 입력받아 실시간으로 처리하고 그 과정 속에서 사용자와 3D 콘텐츠가 실시간 교감이 가능하도록 하는 연구를 진행할 계획이다. 추가적으로 영상 인식 시 특징 파악 후 객체를 추출할 때 그 크기나 회전의 정도에 대해 스마트 단말 환경 상에서 주위 상황에 최소한의 영향을 받으며 콘텐츠 전송 네트워크의 조건에 적절히 대응하여 사용자에게 필요한 정보를 적재적소에 제공해주는 알고리즘 개발이 필요할 것으로 사료된다.

참고 문헌

[1] M. L. Yuan, S. K. Ong, and A. Y. C. Nee, "Registration Using Natural Features for Augmented Reality Systems", *IEEE Transactions on Visualization and Computer Graphics*, Vol. 12, No. 4, pp. 569-580, 2006.

[2] U. Neumann, and S. You, "Natural Feature Tracking for Augmented Reality", *IEEE Transactions on Multimedia*, Vol. 1, No. 1, pp. 53-64, 1999.

[3] S. DiVerdi, and T. Höllerer, "Heads Up and Camera Down: A Vision-Based Tracking Modality for Mobile Mixed Reality", *IEEE Transactions on Visualization and Computer Graphics*, Vol. 14, No. 3, pp. 500-512, 2008.

[4] S. Samavi, S. Shirani, and N. Karimi, "Real-Time Processing and Compression of DNA Micro-array Images", *IEEE Transactions on Image Processing*, Vol. 15, No. 3, pp. 754-766, 2006.

[5] J. B. Kim, and H. S. Jun, "Simulating Low-Cost Cameras for Augmented Reality Compositing", *IEEE Transactions on Visualization and Computer Graphics*, Vol. 16, No. 3, pp. 369-380, 2010.

[6] R. Azuma, Y. Baillo, and R. Behringer, "Recent advances in augmented reality", *IEEE Computer Graphics and Applications, Computers & Graphics*, Vol. 21, Issue 6, pp. 34-47, 2001.

[7] C. Bichlmeier, "The Virtual Mirror: A New Interaction Paradigm for Augmented Reality Environments", *IEEE Transactions on Medical Imaging*, Vol. 28, No. 9, pp. 1498-1510, 2009.

[8] M. Pressigout, and F. Chaumette, "Real-Time Markerless Tracking for Augmented Reality: The Virtual Visual Servoing Framework", *IEEE Transactions on Visualization and Computer Graphics*, Vol. 12, No. 4, pp. 615-628, 2006.

[9] D. Wagner, and G. Reitmayr, "Real-Time Detection and Tracking for Augmented Reality on Mobile Phones", *IEEE Transactions on Visualization and Computer Graphics*, Vol. 16, No. 3, pp. 355-368, 2010.

[10] A. Mohan, C. Papageorgiou, and T. Poggio, "Example-Based Object Detection in Images by Components", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 4, pp. 349-361, 2001.

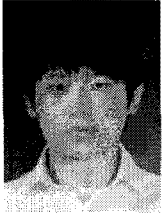
[11] D. J. Lee, and K. T. Choi, "Recognizable-Image Selection for Fingerprint Recognition With a Mobile-Device Camera", *IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics*, Vol. 38, No. 1, pp. 233-243, 2008.

[12] M. A. Gavrielides, E. Šikudová, and I. Pitas, "Color-Based Descriptors for Image Fingerprinting", *IEEE Transactions on Multimedia*, Vol. 8, No. 4, pp. 740-748, 2006.

[13] J. Zhang, "Low-Resolution Gait Recognition", *IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics*, Vol. 40, No. 4, pp. 986-996, 2010.

[14] D. M. Monro, and S. Rakshit, "DCT-Based Iris Recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 29, No. 4, pp. 586-595, 2007.





조 휘 준(Huijoon Jo)

2010년 2월 단국대 멀티미디어공학(학사)

2010년 2월 ~ 현재 단국대 대학원 컴퓨터과학

석사과정 InCross 기술연구소 연구원

※주관심분야 : 모바일 응용, 임베디드 시스템,  
AR



김 대 원(Daewon Kim)

正會員

1993년 2월 중앙대 전자공학과(공학사)

1996년 5월 USC(So. Cal.)-EE (M. S.)

2002년 5월 ISU(Iowa St.)-EE (Ph. D.)

2002년 5월 ~ 2004년 8월 삼성전자 TN총괄 통

신연구소 차세대 단말기술 Lab, 책임 연구원

2004년 9월 ~ 현재 단국대학교천안캠퍼스 공학대학 멀티미디어  
공학과 부교수

※주관심분야 : 디지털신호처리, 모바일 응용, NDE