

스마트폰으로 검출된 객체의 음성합성 시스템

권 순 각[†]

Speech Synthesis System for Detected Objects by Smart Phone

Soon-Kak Kwon[†]

ABSTRACT

This paper designs an application for detecting various objects using a smart phone with camera sensor, then implements the application that detects the number of faces in front of a user by using the Face API provided by android and generates a speech to the user. For implementing the application, the GoF strategy pattern is applied to design the application. It provides some advantages: first, the algorithm development schedule can separate the whole application development schedule; next, it makes easier to add the algorithm. For example, another detecting algorithm for the other objects (character, motion detection) that may be developed in the future, or it may be replaced by a more high-performance algorithm. With the propose method, a general smart phone can make some advantages that can provide information of various objects (such as moving people and objects, and detected character from signboards) to the person who is visually impaired.

Key words: Camera, TTS, Face Detection, Strategy Pattern

1. 서 론

카메라를 사용하여 특정 객체를 인식하는 소프트웨어에서는 기본적으로 카메라로 입력되는 영상을 처리하여 다양한 객체를 추출하는 기술 개발이 중요하다. 입력된 영상으로부터 특정 객체를 추출하기 위한 기존 연구들이 많이 있다. 이러한 연구들의 대부분은 시스템을 테스트하기 위한 환경이 각각 달라서 일반인들이 사용하고 있는 스마트폰에 적용할 경우 기존의 연구 결과와는 다른 성능과 결과를 나타낼 수 있다. 그렇지만 현재 비약적으로 발전하고 있는 스마트폰의 하드웨어 중에서 특히 카메라를 이용한 다양한 객체 인식은 이를 가공하여 충분히 의미 있는 결과를 음성정보로 전달할 수 있다. 또한 이러한 기술

개발은 시각 장애인들에게 많은 도움을 줄 수 있다.

스마트폰을 이용하여 특정한 객체를 인식한 후, 이를 음성으로 제공하는데 사용할 수 있는 주요 기술로 TTS(Text To Speech)가 있다. Fig. 1은 TTS 엔진의 일반적인 구성요소를 보여준다[1].

현재 스마트폰에 설치되어 있는 TTS는 일상적인 정보를 얻기 위해 보조수단이다. TTS는 문자로 표현된 정보를 소리로 들어야만 하는 시각 장애인들이나, 자기의사를 원활하게 표현하기 힘든 언어 장애인, 그리고 지적 장애인에게 더욱 유용하게 활용될 수 있다.

기존의 음성합성 시스템들은 텍스트를 포함한 사진을 결과물로 가공하여 문자를 인식한 후, 음성합성을 구현하는 경우가 대다수이지만, 본 논문에서는 카

* Corresponding Author : Soon-Kak Kwon, Address: (47340) Eomgang-ro 176, Busanjin-gu, Busan, Korea, TEL : +82-51-890-1727, FAX : +82-51-890-2629, E-mail : skkwon@deu.ac.kr

Receipt date : Jan. 20, 2016, Revision date : Feb. 3, 2016
Approval date : Feb. 11, 2016

[†] Dept. of Computer Software Engineering, Donggeui University

* This work (Grants No. C0298779) was supported by Business for Cooperative R&D between Industry, Academy, and Research Institute funded Korea Small and Medium Business Administration in 2015.

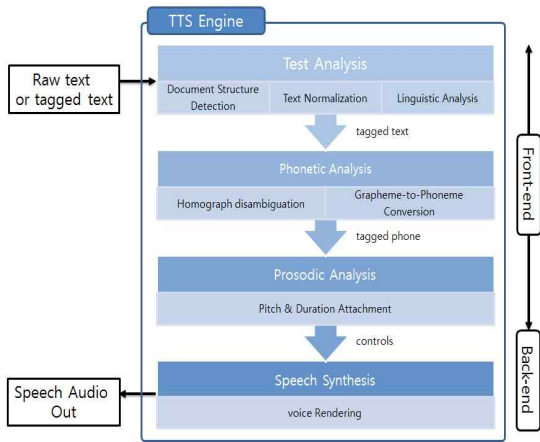


Fig. 1. Main component of the TTS Engine.

메라가 인식할 수 있는 다양한 객체들 중에서 사람의 얼굴을 인식하여 이에 대한 정보를 시각장애인들에게 음성으로 전달하는 방법을 연구하였다.

2. 연구내용과 방법

2.1 관련 연구

카메라로부터 입력된 영상을 사용하여 객체를 인식하기 위한 주요 알고리즘으로 SIFT, SURF 등의 많은 방법들이 제안되고 있다[2-3]. 카메라로 취득되는 일반적인 정보 중에는 주위에 있는 사람들과나 문자들이 있다. 일반적으로 컴퓨터 비전 분야에서 많이 연구되고 있는 광학 문자 인식(Optical character recognition)은 스캐너나 카메라로 획득된 영상을 컴퓨터에서 편집 가능하도록 디지털화 시키는 것이다. 주요 방법으로는 패턴 정합법(pattern matching)과 구조 분석법(structure analysis)이 있다[3-8].

스마트폰을 이용한 실시간 광학문자 인식은 카메라로부터 얻어지는 영상에 포함되어 있는 다양한 문자정보를 추출하여 실시간 음성정보를 제공해 줄 수 있다는 측면에서 시각장애인들에 아주 유용하다. 그러나 현재 사용되고 있는 스마트폰의 하드웨어 사양

과 관련기술이 해결해야하는 문제점들이 많아서 시각장애인들에게 필요한 정보를 제공해 주기 위해서는 추가적인 기술 개발이 필요하다.

얼굴 인식 알고리즘 중에서 가장 흔하게 사용되는 것은 PCA(Principal Component Analysis; 주성분 분석), LDA(Linear discriminant analysis; 선형 판별 분석), 그리고 EBGM(Elastic Bunch Graph Matching) 등이 있다. PCA 알고리즘은 얼굴 사진에서 미세한 오차는 제외하고 주된 특징들만을 분석하는 방법으로 효율성이 높지만 조명이나 표정의 변화가 있을 경우 구분이 힘들다는 단점이 있다. LDA는 통계학과 형태 인식에 쓰이는 방법으로 사물이나 형태를 이루는 선들의 조합과 특징을 분석하는 기법이며, EBGM은 앞의 두 알고리즘이 선형 분석 방법에 의한 것이라는 것과는 달리, 실제 얼굴 이미지가 담고 있는 비선형적인 요소를 비교하는 기법이다[9].

카메라를 이용하여 객체를 인식하기 위한 오픈 소스인 OpenCV는 스마트폰 플랫폼의 한 종류인 안드로이드를 위해 다양한 API와 알고리즘을 제공한다. 이를 잘 활용하면 카메라로 입력되는 영상에서 다양한 정보를 추출할 수 있다. OpenCV에서 제공하는 기술 중에서 전방에 있는 얼굴을 인식하는 알고리즘으로 Multi-Scale LBP(Local Binary Pattern)가 있으며 이를 응용한 다양한 알고리즘이 연구되고 있다[10-12]. 영상의 텍스처(texture)를 분류하기 위한 용도로 개발된 LBP는 얼굴인식과 같은 분야에 응용되고 있다. LBP는 영상의 모든 픽셀에 대해 해당 픽셀의 주변 3×3 영역에서 각 픽셀이 가지는 상대적인 밝기 변화를 2진수로 코딩한 인덱스 값이다. 중앙 픽셀보다 밝으면 1, 어두우면 0으로 코딩한 후, 이 값들을 연결한 이진수를 로컬 텍스처의 인덱스로 활용한다. Fig. 2는 LBP 알고리즘이 어떻게 이미지를 가공하는지를 보여준다.

위의 그림에서 원 영상의 중심 화소(밝기 값: 5)에서 그 주위에 있는 8개의 화소에 대하여 상대적인 밝기 값을 가져와서 비교한 후, 비교되는 픽셀이 더

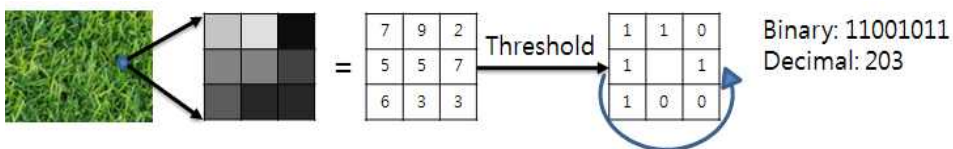


Fig. 2. Example of the image processing using the LBP Algorithm.

밝으면 1, 어두우면 0으로 설정하고 이 값들을 좌측 중앙에서 반 시계방향으로 읽으면서 이진화한다. 이후 각 화소별로 계산된 인덱스 값에 대하여 히스토그램을 구하여 영상 영역에 대한 텍스처 모델로 활용한다. LBP를 얼굴인식에 활용하는 방법을 Fig. 3에 나타냈다.

얼굴인식에 적용되는 LBP는 얼굴영역을 일정한 크기의 블록으로 분할한 후, 각 블록에 대하여 LBP 히스토그램을 구한 후, 구해진 히스토그램을 일렬로 연결한 벡터를 최종 feature로 사용하는 방법과 이를 응용한 방법들이 사용된다.

2.2 스마트폰의 카메라를 이용한 객체 인식과 음성 정보 제공

시각장애인들과 관련된 스마트폰의 주요 기술은 음성 자동 변환 기능인 TTS, 그리고 음성인식 기능이 대표적이며, 이러한 기능을 이용한 추가적인 어플리케이션의 개발이 필요한 실정이다. 또한 스마트폰에 부착되어 있는 센서는 카메라 외에도 GPS, 자이로스코프, 가속도 센서, 기압계, 나침반, 지문 인식, 동작 인식, GPS, A-GPS, GLONASS, 심박측정기, 그리고 근접 센서가 있다. 이러한 센서들로부터 수집되는 데이터를 적절하게 조합하여 처리하면, 시각 장애인들에게 카메라로 인식되는 다양한 객체들의 보다 상세한 정보를 TTS 기술과 조합하여 음성으로 안내해 줄 수 있다.

카메라를 사용한 객체 인식에서 어려운 점은 시맨틱 갭(Semantic gap)으로, 예를 들어 한 사람의 얼굴 인식에 있어서도, 옆모습, 뒷모습, 마스크 착용, 그리고 안경을 착용한 경우 등이 발생할 경우, 컴퓨터는 객체에 대한 본질을 알지 못하기 때문에 같은 사람으로 인식할 수 없다는 것이다.

대중적인 스마트폰을 사용하여 얼굴을 인식하는

어플리케이션 구현에 안드로이드에서 제공하는 Camera.Face 클래스를 이용하면 카메라 정면에 있는 사람의 얼굴을 인식할 수 있다. 이것을 응용하면 정면에서 접근하고 있는 사람의 얼굴을 인식한 후, 이와 관련된 정보를 장애인들에게 음성으로 안내해 줄 수 있다[12-17].

안드로이드의 plug'n play는 최적의 소프트웨어 구현을 위해 다른 구성요소를 조합할 수 있는 기능을 제공한다. 문자인식을 통한 음성합성시스템인 TTS의 구현에 있어서도 다양한 엔진을 사용할 수 있다. TTS 주요엔진으로는 대부분의 안드로이드 장치에 설치되어 있는 "SVOX Pico"가 있으며, 국내의 단말 기에는 일반적으로 "삼성 TTS"도 설치되어 있다.

3. 소프트웨어 설계

3.1 스마트폰의 TTS 설정과 개발

스마트폰의 카메라를 사용하여, 전방에서 인식된 사람들의 얼굴에 대한 정보를 Textview로 출력한 후, 이를 음성합성 TTS 기술과 연동하여 실시간 음성 정보로 제공할 수 있다. Fig. 4에 스마트폰에서 TTS의 설정환경을 확인하는 예를 나타낸다.

안드로이드 기반의 TTS 애플리케이션 개발은 android.speech.tts 패키지를 설치해야하고, TextToSpeech 클래스를 사용한다.

3.2 GoF Strategy Pattern 적용

기존 알고리즘의 성능을 개선하기 위한 제안되는 다양하고 많은 수의 알고리즘들의 성능을 효율적으로 테스트하기 위한 소프트웨어 설계를 위하여 디자인 패턴을 활용할 수 있다. 최적의 성능을 나타내는 알고리즘의 선택을 위해 여러 가지 방법 중에 어떤 것을 선택하는 것이 효과적인지 결정하지 못했을 경

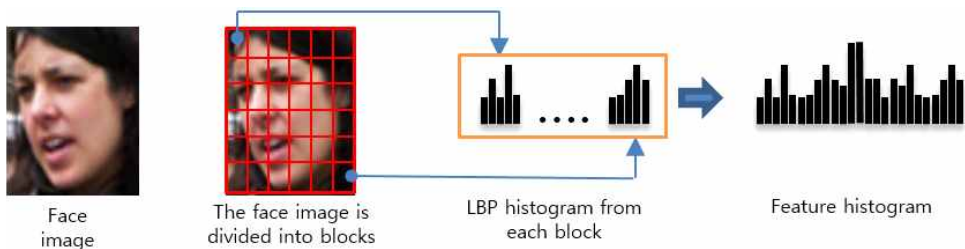


Fig. 3. Face detection method using LBP algorithm.

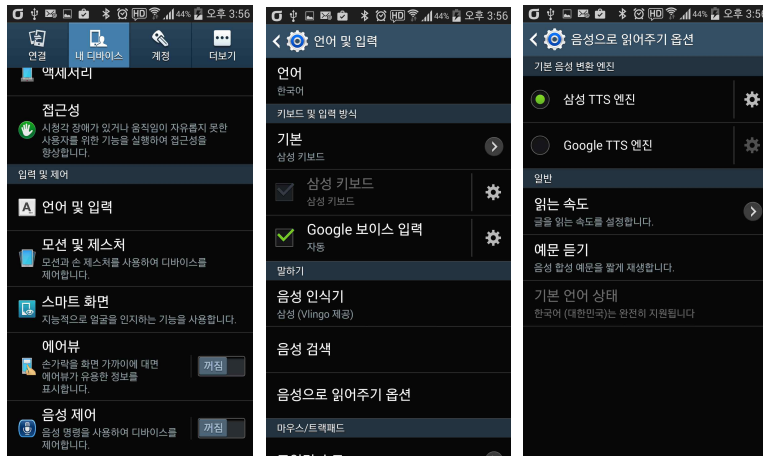


Fig. 4. TTS setup example.

우, 당장은 효율이 떨어져도 개발 비용이 적게 드는 방법으로 문제를 해결한 후, 추후 알고리즘의 성능을 개선할 수 있다. 이러한 경우에 GoF의 전략패턴 (Strategy Pattern)을 설계에 적용하면 현재 개발하고 있는 알고리즘이나 추후 개발될 알고리즘에 대한 추상화를 할 수 있다. 또한 전체 어플리케이션 개발 일정의 큰 지장 없이 독립적으로 추가적인 알고리즘을 개발할 수 있다. 이러한 설계를 통하여 소프트웨어를 분산하여 개발할 수 있다[18]. 전략 패턴을 사용하여 시각장애인을 보조하기 위한 어플리케이션을 개발하기 위한 클래스 다이어그램의 설계 예를 Fig. 5에 나타낸다.

위의 다이어그램에서 일반적으로 사용자 어플리케이션이 가지고 있는 Context 객체는 추상화된 ObjectDetectStrategy 인터페이스를 사용하여, 얼굴을 인식할 수 있는 FaceDetectStrategy, 특정 객체의 특징 추출과 인식을 위한 SiftStrategy, SURF-Strategy, 문자를 추출하는 Pattern Matching Strat-

egy, StructureAnalysisStrategy 등을 동적으로 호출할 수 있게 된다. 각각의 알고리즘을 구현한 전략 (Strategy)들은 해당 결과 값을 사용자 어플리케이션에 보여줄 수 있게 된다.

클래스 다이어그램에 나타난 클래스의 객체가 필요한 기능을 수행하기 위하여 다른 객체들의 기능을 호출하는 순서의 예는 Fig. 6에서 볼 수 있다.

위의 다이어그램에서 어플리케이션 사용자는 카메라로 취득된 객체의 정보를 얻기 위해 objectDetectStart() 메서드로 client 객체를 호출한다. 다음으로 client 객체는 getObject() 메서드를 사용하여 Context 객체를 호출한다. 호출된 Context 객체는 현재 개발되고 있는 FaceDectStrategy 객체의 알고리즘을 호출하여 인식된 얼굴의 개수를 Context 객체로 반환한다. 추가적으로 개발이 진행되고 있거나, 추후 개발될 예정인 SiftStrategy 객체나 SURF-Strategy 객체의 알고리즘도 순환적으로 호출하여 인식된 객체들에 대한 정보를 되돌려 받는다.

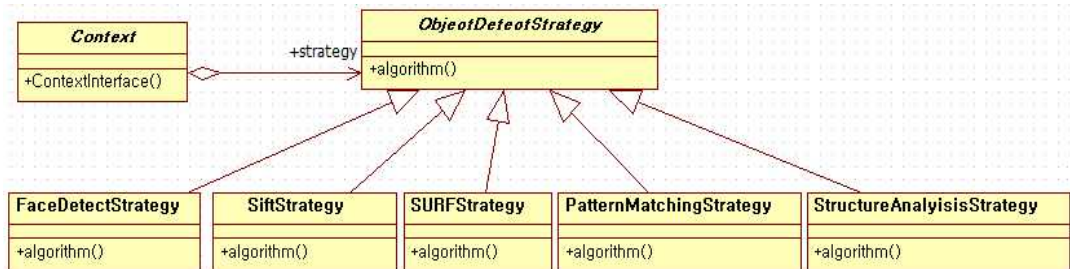


Fig. 5. Example of the object-detection software design using the GoF strategy pattern.

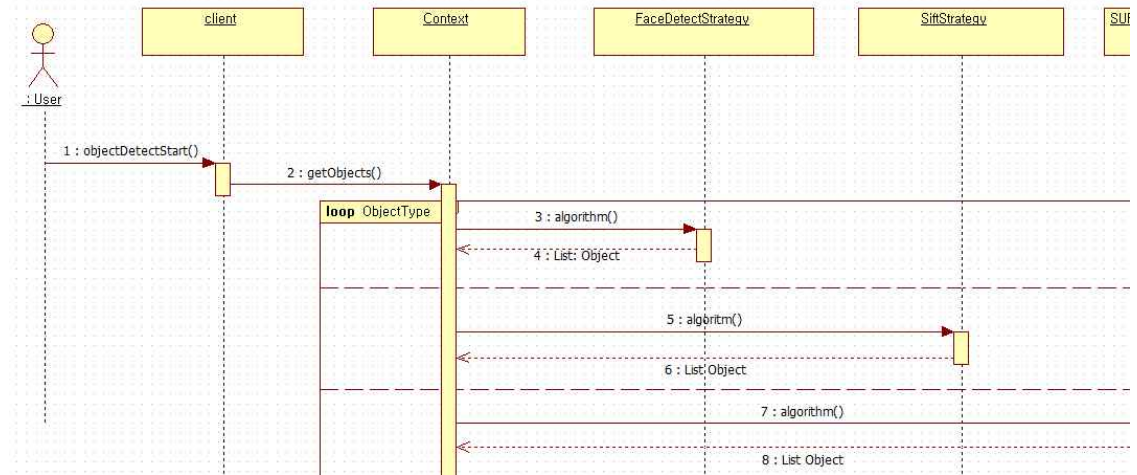


Fig. 6. Example of the sequence diagram using the strategy pattern.

4. 얼굴인식 음성전달 방법의 구현

인식된 객체의 정보를 음성 메시지로 전달하기 위하여 TTS를 이용한다. 안드로이드에서 Main Thread와 별도로 동작하는 TTS쓰레드 구현은 Handler API와 메시지 큐를 이용해야한다. 주 쓰레드와 작업 단위 쓰레드에서 메시지 큐를 이용하여 메시지를 주고받는 방법을 Fig. 7에 나타낸다.

FaceDetectStrategy 클래스를 사용하여 얼굴을 인식하기 위한 FaceDetectionListener의 주요 구현을 Table 1에 나타낸다.

소스 코드에서 굵은 글씨체로 표시된 부분이 Main Thread가 작업 쓰레드에게 사람얼굴을 검출한 결과를 메시지 큐로 전송하는 부분이며, 메시지 전달을 효율을 높이기 위하여 Bundle 클래스를 이용한다.

시작되는 CameraActivity 클래스에서는 surface

의 생성후에 Context 클래스 사용한다. 안드로이드에서 제공되는 Camera.Face 클래스의 주요 메서드는 오른쪽과 왼쪽 눈의 좌표를 얻어오는 rightEye()와 leftEye() 메서드, 입의 중심좌표를 얻어오는 mouth() 메서드, 얼굴의 영역을 얻어오는 rect() 메서드, 그리고 얼굴 검출결과가 어느 정도의 신뢰도를 가지는가를 0~100의 숫자로 나타내는 score() 메서드가 있다.

카메라 뷰의 위쪽에 있는 TextView에 몇 명의 사람얼굴이 인식되었는지를 문자 정보로 나타낸다. 얼굴을 인식하는 FaceOverlayView 클래스는 CameraView 위에서 Camera.Face 클래스의 rect() 메서드를 사용하여 검출된 얼굴 영역을 사각형으로 보여준다. 또한 score() 메서드를 사용하여 검출된 얼굴 영역의 신뢰도를 나타냈는데, score() 메서드는 반환되는 값이 50 이상일 때 사람 얼굴로 인식되도록 처리한다.

구현된 소프트웨어의 테스트에 사용된 스마트폰은 Sony의 XperizZ2(D6503), 삼성의 갤럭시S4(SHV-E330k), 갤럭시S6Edge+를 사용하였다. Table 2에 실험에 사용된 스마트폰의 주요 사양을 정리하였다.

Sony의 XperizZ2(D6503)와 삼성의 갤럭시S4(SHV-E330k)는 4개의 코어를 사용하는 CPU를 채택하였고 해상도는 둘 다 1920x1080인 것을 알 수 있으며, 갤럭시S6Edge+는 8개의 코어를 사용하는 CPU를 채택하였고, 해상도는 2560x1440인 것을 알 수 있다.

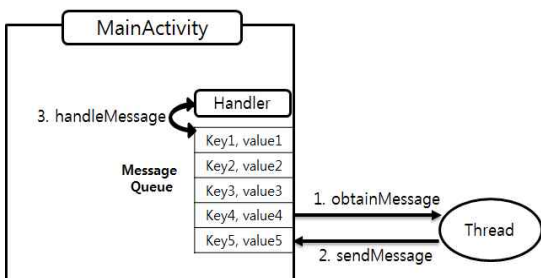


Fig. 7. Message passing sequence between Main Thread and a Thread.

Table 1. Example of the FaceDetectionListener in the FaceDetectStrategy

```
private class FaceDetectStrategy extends ObjectDetectStrategy implements Camera.FaceDetectionListener {
.....
    public int algorithms() { .....
        cameraActivity.getCamera().setFaceDetectionListener(this);
        cameraActivity.getCamera().startFaceDetection(); .....
    }
    public void onFaceDetection(Face[] faces, Camera camera) {
        messageTextView = (TextView)findViewById(R.id.messageText);
        if (faces.length == 0){
            messageTextView.setText( "전방에서 사람 얼굴을 인식할 수 없습니다.");
        }else{
            mFaceView.setFaces(faces);
            messageTextView.setText("전방에서 " + String.valueOf(faces.length) +
                "명의 사람의 얼굴이 인식되었습니다.");
        }
        currentTTSText = messageTextView.getText().toString();
        if(!currentTTSText.equals(savedTTSText)) {
            Message sendMsg = faceDetectMessageHandler.obtainMessage();
            Bundle bundle = new Bundle();
            String msg = new String(currentTTSText);
            bundle.putString("text", msg);
            sendMsg.setData(bundle);
            faceDetectMessageHandler.sendMessage(sendMsg);
            savedTTSText = currentTTSText;
        }
    }
};
```

Table 2. Specifications of smart phones for simulation

Model	Specifications
Sony Xperia Z2 D6503	<ul style="list-style-type: none"> • 32bit OS • CPU: Qualcomm Snapdragon 801 MSM8974-AB(Quad-core, 2.3GHz, 3GB RAM) • GPU: Adreno 330 • Resolution: 1920x1080 Pixels • Aperture: f/2.0
GalaxyS4 LTE-A (SHV-E330k)	<ul style="list-style-type: none"> • 32bit OS • CPU: Qualcomm Snapdragon 800 (Quad-core, 2.3GHz, 2GB RAM) • GPU: Adreno 330 • Resolution: 1920x1080 Pixels • Aperture: f/2.2
GalaxyS6 Edge+	<ul style="list-style-type: none"> • 64bit OS • CPU: ARM Holdings Cortex-A57 2.1GHz Quad-core, ARM Holdings Cortex-A53 1.5GHz Quad-core • GPU: ARM Holdings Mali T760 MP8 700 MHz • Resolution: 2560x1440 Pixels • Aperture: f/1.9

각각의 스마트폰에서 구현된 앱을 실행한 후, 사람의 얼굴이 인식되는 최대 거리를 측정하였고, 다가오는 사람에 대하여 얼굴을 인식할 수 있는 최대 속도를 측정하였다. 3개의 스마트폰에서 사람의 얼굴이 인식되는 최대 거리를 테스트한 영상은 Fig. 8에서 볼 수 있다.

근접거리에서 얼굴을 인식시킨 후, 실험자가 뒤로 물러서면서 얼굴이 인식되는 최대거리에서 정지영상을 포착하였다. 좌측의 영상은 Sony XperiaZ2가 얼굴을 인식하는 최대 거리고, 중앙은 갤럭시S4가

인식하는 최대 거리고, 우측은 갤럭시S6가 얼굴을 인식하는 최대거리다. Sony XperiaZ2는 약 2.5m, 삼성의 갤럭시S4는 약 3.5m, 갤럭시S6는 약 6.0m가 최대 얼굴 인식 거리였다.

스마트폰을 정지시켜놓은 상태에서 정면에서 걸어오는 사람에 대해 어느 정도 속도에서 사람 얼굴을 인식할 수 있는지를 테스트한 결과는 Fig. 9에서 볼 수 있다.

Sony XperiaZ2의 경우 약 0.3m/sec 이하인 경우 얼굴인식이 가능하였고, 비슷한 사양의 삼성 갤럭시

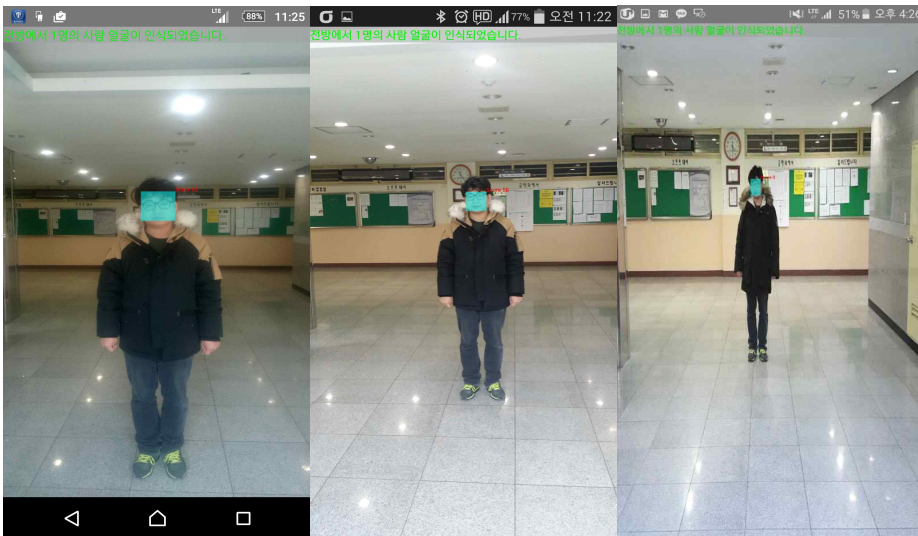


Fig. 8. Maximum distance to detect face(Left: Sony XperiaZ2 D6503, Center: GalaxyS 4 LTE-A(SHV-E330k), Right: GalaxyS6 Edge+).



Fig. 9. Maximum walking speed to detect face (Left: Sony XperiaZ2 D6503, Center: GalaxyS 4 LTE-A(SHV-E330k), Right: GalaxyS6 Edge+).

S4의 경우도 사람의 걸음걸이 속도가 약 0.5m/sec 이하일 때, 얼굴인식이 가능하였다. 갤럭시S6의 경우 보통사람의 걸음걸이(속도: 1m/sec)로 다가오는 사람의 얼굴을 인식하였다. Table 3에 실험 결과를 요약했다.

어느 정도의 거리에서 얼마나 많은 얼굴을 인식하는가를 테스트하기 위해 6명의 실험참가자를 대상으로 인식되는 얼굴의 개수와 거리를 측정하였다. Fig. 10에서 그 결과를 볼 수 있다.

갤럭시S4의 경우 2.5m 내에서 6명의 실험참가자 얼굴을 모두 인식하였고, 갤럭시S6의 경우 6m내에서도 얼굴을 모두 인식하였다. 조명이 밝은 곳에서 얼굴인식이 잘 되었으며, 스마트폰의 떨림도 인식하는 얼굴의 숫자에 큰 영향을 미쳤다. Table 4는 실험 결과를 나타낸다.

실험 결과로 볼 때, 3개의 스마트폰 중에서 Sony Xperia Z2, GalaxyS4의 경우는 전방에서 마주 오는 사람을 제대로 인식하여 음성으로 전달해 주기 힘든 것으로 판정되었다. 반면 삼성의 최신 스마트폰인 갤럭시S6Edge+의 경우 6m에서도 사람의 얼굴을 인식하였고, 어느 정도 정상 걸음걸이로 걸어오는 사람에 대하여 얼굴을 인식하여 음성으로 전달하는 것이 가

능하다는 것을 보여주었다.

우수한 성능을 보인 갤럭시S6Edge+와 다른 2대의 스마트폰과의 가장 큰 차이점은 갤럭시S6에서는 처리속도가 향상된 8개의 코어, 성능이 더 우수한 GPU, 64bit 운영체제를 지원한다는 것이다. 또한 갤럭시S6에 장착된 카메라의 사양에서 조리개는 f/1.9로 다른 스마트폰 보다 밝으며, 또 어두운 환경에서도 밝고 선명한 이미지를 제공하는 기능이 있어서 얼굴인식이 더 뛰어났던 것으로 판단된다.

위의 결과를 볼 때, 대중적인 스마트폰을 사용하여 앞에서 다가오는 다수의 사람들에 대한 정보를 실시간 음성으로 전달하여 시각장애인들의 보행에 도움을 주기위해서는 스마트폰 카메라모듈과 하드웨어의 지속적인 성능향상이 필요하다는 것을 알 수 있다.

5. 결론 및 향후 연구

스마트폰의 카메라로부터 취득되는 영상의 객체 분석 기술과 TTS를 이용한 음성합성 기술은 문자를 읽기 어려운 장애인이나 문자를 읽을 수 없는 사람(아이, 외국인 등), 난독증과 같은 학습장애인, 인지

Table 3. Simulation results according to three smart phones

Model	Item	Measured value
Sony Xperia Z2 D6503	Maximum distance	2.5m
	Maximum speed	0.3m/sec
GalaxyS4 LTE-A (SHV-E330k)	Maximum distance	3.5m
	Maximum speed	0.5m/sec
GalaxyS6 Edge+	Maximum distance	6.0m
	Maximum speed	1m/sec



Fig. 10. Maximum detected faces according to smart phone (Left: GalaxyS 4 LTE-A(SHV-E330k), Right: GalaxyS6 Edge+).

Table 4. Maximum detected faces

Model	Item	Measured value
GalaxyS4 LTE-A (SHV-E330k)	Maximum distance	2.5m
	Maximum faces	6
GalaxyS6 Edge+	Maximum distance	6.0m
	Maximum faces	6

장애인과 노인, 그리고 다문화가족 등과 같은 많은 사람들에게 편의를 제공해 줄 수 있다.

시각장애인에게 카메라로 인식된 객체의 정보를 전달하기 위해 본 논문에서 안드로이드의 Camera, Face 클래스와 TTS를 이용하여 인식된 얼굴을 음성으로 전달하는 어플리케이션을 구현하였다. 효율적인 소프트웨어 설계를 위한 GoF의 Strategy 패턴을 적용하였다. Face 클래스를 이용한 객체 인식 기능을 우선적으로 구현하여 정면에서 다가오는 사람 얼굴을 인식한 후, 해당 정보를 TTS로 전달하였다.

서로 다른 3대의 스마트폰을 이용하여 사람의 얼굴을 인식하는 최대거리와 마주 오는 사람을 인식할 수 있는 속도를 테스트하였다. 테스트 결과에서는 사람의 얼굴인식에 영향을 미치는 주요 요소가 운영체제의 처리 속도, CPU와 GPU의 성능, 그리고 카메라의 성능(조리개, 해상도)가 주요한 요인이었으며, 어두운 환경에서도 밝은 영상을 얻는 기술이 얼굴을 인식하는데 많은 도움이 된다는 것을 알 수 있었다.

컴퓨터 비전 분야에서 객체를 인식하기 위한 다양한 알고리즘들이 현재 지속적으로 성능이 업그레이드되고 있는 스마트폰의 다양한 센서들과 조합되면 시각장애인들에게 유용한 정보를 제공해 줄 수 있다. 효과적인 정보 전달을 위해 기존에 연구되고 있는 컴퓨터 비전의 객체인식 분야 기술을 스마트폰에 적용하여 시각장애인을 위한 객체인식 기술을 지속적으로 향상시킬 필요가 있다. 객체를 탐색하는 다양한 알고리즘의 최적화를 통하여 보다 높은 객체 인식률을 얻을 수 있는 방법들의 연구도 필요하며, 정면에서 다가오는 사람들에 대한 정보를 먼 거리에서 파악하여 여유 있게 대처할 수 있도록 보다 해상도가 높은 스마트폰 카메라 모듈 개발이 필요하다.

REFERENCE

- [1] J. Yang, *A Study on Syllable-level Concatenation for HMM-based Mixed-lingual Text-to-speech*, Master's Thesis of Gwangju Institute of Science and Technology, 2010.
- [2] Demo Software: SIFT Keypoint Detector, <http://www.cs.ubc.ca/~lowe/keypoints/> (accessed Jan., 05, 2016).
- [3] Character recognition, <http://terms.naver.com>, (accessed Jan., 13, 2016).
- [4] character recognition, <https://ko.wikipedia.org>, (accessed Jan., 13, 2016).
- [5] N. Kim, D. Kim, S. Kim, and S. Kwon, "Vocabulary Generation Method by Optical Character Recognition," *Journal of Korea Multimedia Society*, Vol. 18, No. 8, pp. 943-949, 2015.
- [6] E. Go, Y. Ha, S. Choi, K. Lee, and Y. Park, "An Implementation of an Android Mobile System for Extracting and Retrieving Texts from Images," *Journal of Digital Contents Society*, Vol. 12, No. 1, pp. 57-67, 2011.
- [7] M. Kim, "Individual Identification Using Ear Region Based on SIFT," *Journal of Korea Multimedia Society*, Vol. 18, No. 1, pp. 1-8, 2015.
- [8] G. Yoo, K. Jeong, and H. Moon, "Object Recognition Based on Speed Up Robust Feature Algorithm(SURF) for Smartphone Environment," *Proceeding of The Institute of Electronics Engineers of Korea*, pp. 544-545, 2010.
- [9] The development of face recognition technology and content services, <http://www.koc-ca.kr>, (accessed Apr., 10, 2014).
- [10] D. Kim, M. Sohn, and S. Lee, "A Study on Face Recognition Method Based on Binary Pattern Image under Varying Lighting Condition," *Proceeding of The Institute of Electronics Engineers of Korea*, pp. 61-74, 2012.
- [11] J. Kim, "Face Recognition by Fiducial Points Based Gabor and LBP Features," *The Journal*

of the Korea Contents Association, Vol. 13, No. 1, pp. 1-8, 2013.

[12] M. Abdur Rahim, M. Najmul Hossain, T. Wahid, and M. Shafiul Azam, "Face Recognition using Local Binary Patterns (LBP)," *Global Journal of Computer Science and Technology*, Vol. 13, No. 4, pp. 1-8, 2013.

[13] Object Recognition, <http://www.haenaki.com/118>, (accessed Jan., 11, 2016).

[14] The 7 Best Android Text-To-Speech Engines, <http://www.geoffsimons.com/2012/06/7-best-android-text-to-speech-engines.html>, (accessed Jan., 05, 2016).

[15] The Implementation of ths TTS in the Android, <http://www.devblog.kr.>, (accessed Jan., 05, 2016).

[16] Android Developer Text To Speech, <http://developer.android.com/reference/android/speech/tts/TextToSpeech.html>, (accessed Jan., 05, 2016).

[17] Using the Google TTS(TextToSpeech), <http://blog.daum.net/mailss/26>, (accessed Jan., 05, 2016).

[18] E. Gamma, R. Johnson, R. Johnson, and H. Vissides, *Design Patterns, Elements of Reusable Object-Oriented Software*, Addison-Wesley Publishing Company, Boston, 1995.



권 순 각

1990년 2월 경북대학교 전자공학과 졸업
 1992년 2월 KAIST 전기및전자공학과 석사
 1998년 2월 KAIST 전기및전자공학과 박사

1997년 3월~1998년 8월 한국전자통신연구원 연구원
 1998년 9월~2001년 2월 기술보증기금 기술평가센터 팀장
 2003년 9월~2004년 8월 Univ. of Texas at Arlington 방문 교수
 2010년 9월~2011년 8월 Massey University 방문 교수
 2001년 3월~현재 동의대학교 컴퓨터소프트웨어공학과 교수
 관심분야 : 멀티미디어신호처리, 영상통신