

시각장애인을 위한 촉각 TV 시스템에 관한 연구

류영일*, 노현준*, 정웅준**, 류은석*

*가천대학교 컴퓨터공학과

**세화고등학교

e-mail : wlrmlrm99@gc.gachon.ac.kr

ggyo@gc.gachon.ac.kr

june88823@naver.com

esryu@gachon.ac.kr

A Study on Novel Video Conferencing System for Visually Impaired

Yeongil Ryu*, Hyun-Joon Roh*, Woongjune Chung**, Eun-Seok Ryu *

*Dept. of Computer Engineering, Gachon University

**Sehwa High School

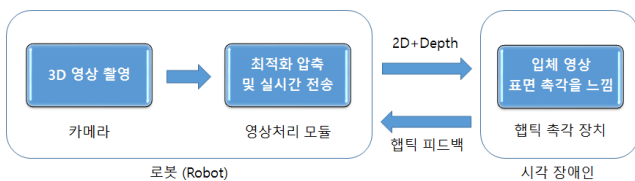
요 약

정보통신 기술의 발달과 스마트폰의 보급으로 일반 사용자들은 언제, 어디서나 가족 및 주변 사용자들의 모습을 바라보며 통화를 할 수 있게 되었고, 원하는 영상이나 사진을 감상할 수 있다. 하지만 시각장애인들은 이들을 위한 연구 및 사회적 인프라의 부족으로 인하여 이와 같은 서비스들의 제공 대상에서 늘 제외되어 왔다. 이러한 사항을 개선하기 위하여 본 연구는 시각장애인을 위한 새로운 방식의 촉각 TV 시스템을 제안한다. 제안하는 시스템은 크게 3D 캡처(Capture) 기술, 실시간 전송/스트리밍 기술, 햅틱(Haptic) 장치 및 액츄에이터(Actuator)제어 기술로 구성된다

1. 서론

세계적으로 약 3 천 9 백만명의 사람들이 법적 시각 장애인 판정을 받고 살아감에도 불구하고, 이들을 위한 TV 기술은 ‘영상 인식에 의한 음성 안내’ 및 ‘길 안내’ 정도의 초보적인 수준에 머무르고 있으며, 그 어떤 시스템도 영상 이미지를 직접 제공하지 못한다. 이에, 본 연구는 시각장애인을 위한 새로운 형태의 촉각 TV 시스템을 제안한다.

제안하는 시스템은 이동이 불편한 시각 장애인이 원격의 가족 얼굴 윤곽을 느끼고 인식하는 약한 수준의 화상통화를 할 수 있고, 그간 시각장애인에게는 불가능하였던 TV 시청 및 사진감상을 제한적 기능으로라도 가능하게 한다.



(그림 1) 제안하는 시스템의 개념적 구성

2. 관련 연구

국내외에서 시각장애인을 위한 정보전달과 관련된

여러 연구가 수행되었으나, 현재까지 영상 이미지를 직접 시각장애인에게 전달하려는 시도는 없었다. 관련된 연구를 아래에 설명하고, 연구내용 및 한계점을 정리하였다.

2.1 Dennis Hong 교수(UCLA)의 시각장애인을 위한 자동차

시각장애인이 운전을 할 수 있도록 카메라로 인식된 장애물 등을 파악하여 핸들을 돌릴 방향을 햅틱적으로 알려주는 시스템이다.

핸들과 장갑에서 압축 공기를 방출하여 사용자에게 정보를 전달하는 촉각 액츄에이터를 설치하여, 영상 인식을 통해 장애물의 위치를 알리는 기술이 주요 특징이다. 하지만, 사용자가 직접 영상을 구체적으로 느끼고 판단하기 보다는 컴퓨터가 영상 처리를 통해 인식한 장애물을 판단해주기 때문에 기존의 시각장애인을 위한 네비게이션 시스템과의 차이점이 적다.

2.2 MIT Media Lab 의 Tangible media 기술

물리적 텔레프레전스(Physical telepresence)라고 이름 붙은 기술로, 수많은 액츄에이터들로 원격을 물체를 시각화 하려는 연구로서 ‘만질 수 있는(Tangible) 정보 처리’ 측면에서 의미가 매우 크다고 볼 수 있다. 하지

만, 개발 자체의 목적이 시각장애인을 위한 연구가 아니므로 그들의 촉각능력 및 판단 능력에 맞게 개발되지 않았고, 또한 색감 및 기타 정보 전달 방법 및 영상처리 기술이 포함되지 않았다.

2.3 카네기멜론 대학교(CMU)의 ‘Assistive Robotic System’

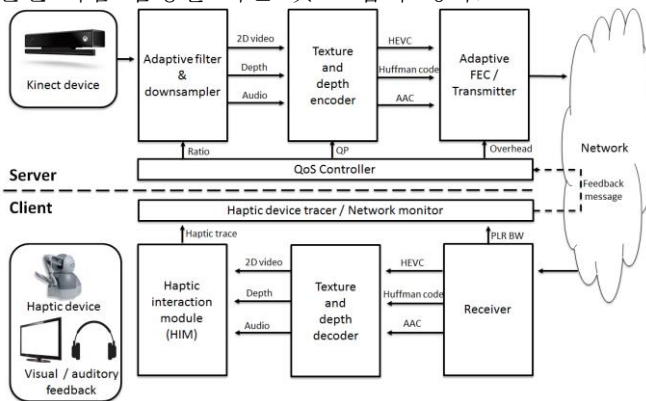
CMU의 로보틱스 랩에서는 시각장애인의 길 안내를 대신할 로봇을 연구하였는데, 초기에는 GPS에 의한 방향 안내를 하는 스마트 지팡이(Smart cane) 수준이었으나, 현재는 시각장애인을 따라오며 길 안내를 돕는 로봇을 연구개발 하였다. 이러한 기술은 시각장애인의 길안내 보조 도구로서 의미가 크지만, 영상 자체를 전달하여 시각장애인이 직접 판단하고 느끼는 영상처리 기술은 포함되지 않았다.

2.4 워싱턴 대학과 조지아공대의 시각장애인을 위한 로봇 및 그 확장 연구

조지워싱턴 대학교(GWU)와 조지아 공과대학교(Georgia Institute of Technology)의 로봇연구실에서는 시각장애인이 로봇에 달린 3D 카메라로부터 받아들인 입체영상을 햅틱 장비로 느끼며 로봇을 조종하던 기술을 연구하였다. 워싱턴 대학에서는 이 기술을 확장하여 3D 영상을 시각장애인의 햅틱 장비로 스트리밍하는 시스템을 연구개발하고 있으며, 가천 대학교에서 3D 영상 스트리밍 기술의 연구개발을 협업 진행한다 [2-4].

3. 제안하는 시각장애인을 위한 촉각 TV 시스템

세계적으로 약 3 천 9 백만명의 사람들이 법적 시각장애인 판정을 받고 살아감에도 불구하고 이들을 위한 TV 기술은 전무하다. 또한, 시각장애인은 그 활동의 제약성으로 인해 멀리 있는 가족의 모습을 보고 느낄 수 없으며, 박물관을 방문하여 유리상자 안에 보관된 작품 감상을 하는 것도 쉽지 않다.



(그림 2) 시각장애인을 위한 촉각 TV 시스템

물론, 시각장애인을 위한 보조 도구와 기술들이 개발되었지만, 대부분의 기술이 정보를 음성으로 변환해서 읽어주거나 점자(Braille)와 유사한 최소한의 촉각정보를 통해 전달되고 있기에 수많은 시각장애인들

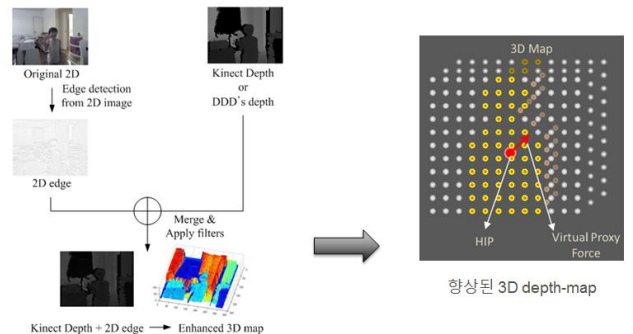
의 불편함은 매우 크다. 따라서, 본 연구는 (1) 3D 캡처 기술, (2) 실시간 전송/스트리밍 기술, (3) 햅틱 장치 및 액추에이터 제어 기술 등이 결합된 다음과 같은 ‘시각장애인을 위한 촉각 TV 시스템’을 제안한다. 제안된 시스템은 시각장애인이 스트리밍된 3D 영상의 윤곽을 햅틱 장비를 통해 느낄 수 있게 돕는다.

시스템 구조는 크게 3D 영상 캡처, 전송, 사용자 인터페이스로 나뉜다. 이 중 3D 영상 캡처는 Microsoft사의 키넥트(Kinect) 장치를 사용하여 캡처하는데, 획득된 영상의 3D 깊이정보(Depth-map)은 각종 필터기술과 다운샘플링(Down sampling) 기술을 통하여 시각장애인에게 최적화 된 형태로 바뀌게 된다. 그 후, 효율적 인코딩 및 실시간 전송을 위해, H.264 및 최신 비디오 코딩 기술인 HEVC 기술을 사용하고, 랩터코드(Raptor Code)로 전송 패킷을 보호한다. 시각장애인과 인터페이스 처리는 햅틱-인터랙션-모듈을 통해 시각장애인에게 최적의 형태로 이미지 영상을 제공한다.

3.1 3D 캡처 기술

3D 캡처 기술은 Kinect Software Development Kit (SDK)을 활용하여 구현된다. 이 때, Kinect SDK는 2D 영상의 크기 및 프레임 비율(Frame rate)을 조절해가며 캡처를 할 수 있는 기능이 있고, 마찬가지로 VGA(640x480) 화면크기의 깊이 정보를 캡처할 수 있다. 따라서 본 연구는 이 SDK를 활용하여 C/C++언어로 2D+Depth 영상 캡처 모듈을 개발한다. 이후, 캡처 처리의 실시간성 향상을 위하여 Kinect SDK의 여러 옵션에 따른 실행 시간을 Visual studio 개발툴의 프로파일링(Profiling) 도구를 활용하여 분석 후, 이에 따라 시각장애인의 요구사항에 맞는 실시간성 향상을 위하여 캡처 모듈을 개선시킨다. 이 때 시각장애인의 햅틱 장비를 이용한 영상 인식 속도 (영상의 윤곽을 느끼는 속도)가 그리 빠르지 않아 충분한 캡처 속도를 보장받을 수 있는데, 이 인식 속도 자체가 하나의 중요한 향후 연구주제이기도 하다.

3.2 Depth-map 처리 기술



(그림 3) Depth-map 향상 기술

원본 2D 이미지로부터 얻은 모서리(Edge) 정보와 깊이(Depth) 정보를 활용하여 시각장애인이 인지하기 좋은 형태로 좀 더 강조된 3D Depth-map을 생성한다. 이 강조의 정도 및 필터기술은 시각장애인의 실질적 의견 반영을 통해 최대한 현실적이면서 상용 가능한

기술로 개발한다.

3.3 영상 서브샘플링 기술

시각장애인이 햅틱 장비를 통해 영상의 윤곽을 촉각으로 느끼고 인식하는 데는 일정 시간이 필요하며, 이는 영상 콘텐츠의 내용 및 복잡도, 시각장애인의 개인차에 따라 다를 수 있다. 따라서 그러한 통계적 정보에 따라 영상 데이터를 시각장애인의 촉각 민감도에 맞게 서브샘플링 할 수 있는 모듈을 개발한다.

3.4 오디오/비디오 실시간 압축기술

2013 년도에 표준화 된 차세대 비디오 압축기술인 HEVC(High Efficiency Video Coding)는 기존 H.264 AVC 영상 압축기술에 비하여 고화질 영상처리 부분에서 약 40%의 성능향상을 가져온다. 또한 이러한 영상 데이터 압축을 스케일러블(Scalable)하게 하기 위한 SHVC (Scalable HEVC) 기술도 2014 년에 국제표준화되었다. 따라서, 이러한 표준기술들의 참조 소프트웨어 (Reference SW)인 HM 과 SHM, 그리고 x265 SW 를 활용하여 실시간 영상 데이터 압축 모듈을 구현하고, 성능을 개선한다.

3.5 실시간 전송 및 재생 기술

주로 시각장애인의 햅틱 장비는 유선으로 시스템에 연결되어 있겠지만, Kinect 를 통한 3D 영상 캡처를 담당하는 부분은 무선으로 캡처된 영상을 전송할 수 있다. 따라서 레이트 컨트롤(Rate control) 기법은 필수적이며, 측정된 네트워크 대역폭에 최적화 하여 전송할 수 있도록 구현 개발한다.

또한, 이를 위해 Rateless codes 의 일종인 Raptor FEC(Forward Error Correction) 기술을 구현하고, 전송할 영상화면 중요도에 따른 차등패킷보호기술(Unequal Error Protection)을 적용하여 에러 내성을 향상시킨다.

3.6 햅틱 장치 제어 향상 기술

햅틱 장치 개발회사에서 제공하는 SDK 와 OpenGL, Scene Graph 등을 활용하여 가상공간에서의 영상을 시각장애인이 느낄 수 있도록 햅틱 장치 제어 향상 기술을 개발한다. 이 때 힘(Force) 디스플레이 기능과 상태 피드백 기능을 잘 활용하여 사용자의 제어 기능을 향상시킨다.

4. 실험

본 논문에서 제안하는 시각장애인을 위한 촉각 TV 시스템은 현재 연구/개발 중이며, 그 중 이미 구현된 일부 기술을 이용하여, 개발하기 전에 주어진 임의의 물체를 햅틱 장치로 얼마나 정확히 인지할 수 있는지 가능성 실험(Feasibility test)을 실시하였다.

실험은 일반인 10 명을 대상으로 진행하였고, 주어진 5 가지 입체도형((1)원기둥, (2)원뿔, (3)정사면체, (4)정육면체, (5)구) 중 제시된 도형을 피실험자가 햅틱 장치를 활용하여 눈으로 도형을 보지 않고 골라내는

방식으로 진행하였다.

실험 결과는 피실험자 10 명 모두 제시된 입체도형을 골라내어 고무적이었으나, 정답을 맞추기까지의 걸린 시간은 편차가 크다는 아쉬움이 있었다. 하지만 이는 실험 전에 햅틱 장치의 특징과 기본적인 사항을 사전교육 후 실험을 진행한다면, 피실험자들이 신속하게 정답을 골라낼 수 있다고 예상된다. 따라서, 실제 이 장비와 실험에 훈련된 시각장애인이라면 짧은 시간 안에 전송된 영상을 알아 맞출 수 있었을 것으로 판단된다.

5. 결론

본 연구는 3D 영상을 시각장애인이 인지하기에 적합한 형태로 처리하여 제공하는 촉각 TV 시스템을 제안한다. 제안하는 시스템은 컴퓨터가 인지하고, 판단하여 그 결과만을 시각장애인에게 전달하는 기존 연구들과는 다르게 실제 영상 이미지와 기타 정보를 시각장애인에게 전달하여, 시각장애인이 직접 느끼고 판단할 수 있도록 한다. 또한, 주어진 서로 다른 종류의 5 개의 입체도형 중 제시된 도형을 맞추는 가능성 테스트 시행 결과, 10 명의 일반인 참가자 모두 정답을 골라내어 본 연구가 실현 가능성을 밝혔다.

본 연구는 향후에 시각장애인용 화상시스템이나 원격수술(Tele-surgical) 로봇 시스템의 햅틱 피드백 요소 기술로 확장이 가능하다.

Acknowledgement

이 논문은 2015 년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(No. NRF-2015R1C1A1A02037743)

참고문헌

- [1] D. De Silva, W. Fernando, H. Kodikaraarachchi, S. Worrall, A. Kondoz, "A depth map post-processing framework for 3d-tv systems based on compression artifact analysis", pp. 1-1 (2011).
- [2] C. H. Park, A. Howard, "Towards real-time haptic exploration using a mobile robot as mediator", in Haptics Symposium, 2010 IEEE, pp. 289-292 (2010).
- [3] C. H. Park, A. M. Howard, "Real-time haptic rendering and haptic telepresence robotic system for the visually impaired", in World Haptics Conference (WHC), 2013, pp. 229-234 (2013)
- [4] C.H. Park, E.-S, Ryu, A.M. Howard, "Telerobotic Haptic Exploration in Art Galleries and Museums for Individuals with Visual Impairments," IEEE Transactions on Haptics, vol.8, no.3, pp.327-338, July-Sept. 1 2015.
- [5] D. Shah, J. Ascenso, C. Brites, F. Pereira, "Evaluating multi-view plus depth coding solutions for 3d video scenarios", in 3DTV-Conference: The True Vision – Capture, Transmission and Display of 3D Video (3DTVCON), 2012, pp. 1-4 (2012).
- [6] C. Tomasi, R. Manduchi, "Bilateral filtering for gray and color images", in Computer Vision, 1998, Sixth International Conference on, pp. 839-846, IEEE (1998).