

감각치환 기술 동향

Technological Trends in Sensory Substitution

문경덕 (K.D. Moon, kdmooon@etri.re.kr)	휴먼증강연구실 책임연구원
김무섭 (M.S. Kim, gomskim@etri.re.kr)	휴먼증강연구실 책임연구원
정치윤 (C.Y. Jeong, iamready@etri.re.kr)	휴먼증강연구실 책임연구원
박윤경 (Y.K. Park, parkyk@etri.re.kr)	휴먼증강연구실 책임연구원
신승용 (S.Y. Shin, seneru@etri.re.kr)	휴먼증강연구실 선임연구원
오창목 (C.M. Oh, chmooh@etri.re.kr)	휴먼증강연구실 선임연구원
박준석 (J.S. Park, parkjs@etri.re.kr)	휴먼증강연구실 책임연구원
신형철 (H.C. Shin, shin@etri.re.kr)	휴먼증강연구실 책임연구원/실장

ABSTRACT

Sensory substitution involves mapping the characteristics of one sensory modality to the stimuli of another sensory modality. In general, when a person is visually impaired or deaf, they do not actually lose their ability to see or hear completely; however, they only lose their ability to transmit sensory signals from the periphery to the brain. It has been experimentally proven that a person who has lost the ability to retrieve data from the retina can still visualize subjective images by using data transferred from other sensory modalities such as tactile or auditory modalities. This is because vision processing pathways are still intact in most cases. Therefore, sensory substitution uses human perception and the plasticity of the human brain to transmit sensory signals through pathways that have not been lost. In this study, we analyze the characteristics and problems of various devices used for sensory substitution and summarize the recent technological trends in these devices.

KEYWORDS 감각치환, 감각지각 증강, 휴먼증강, 다중감각 융합, Sensory Substitution Device

1. 서론

최근 통계청의 고령자 통계에 따르면 2018년 65세 이상의 고령자는 738만 명으로 전체 인구의 14.3%이며, 2060년에는 전체 인구의 41%에 이를

것으로 전망하고 있다[1]. 또한 우리나라의 시각, 청각 등 감각장애 인구는 57만 명이며, 세계보건 기구(WHO)에서 시력 손상 및 상실을 포함하여 시각장애를 겪는 전 세계 인구는 13억 명, 청력 손상 및 상실을 포함하여 청력 장애를 겪는 전 세계 인

* DOI: <https://doi.org/10.22648/ETRI.2019.J.340407>

* 본 연구는 한국전자통신연구원(ETRI) 연구운영비지원사업의 일환으로 수행되었음[19ZS1500, 인간의 감각·지각 능력을 증강하는 다중 감각 융합 기술 개발 사업].



본 저작물은 공공누리 제4유형

출처표시+상업적이용금지+변경금지 조건에 따라 이용할 수 있습니다.

©2019 한국전자통신연구원

구는 4억 6천만 명에 이른다고 추정하고 있다[2]. 이와 같이 시각, 청각 등의 감각이 손상되거나 기능이 저하된 고령자와 장애인이 증가하면서 이들의 손상된 감각 및 지각 능력을 향상시켜 지속적인 경제 활동과 삶의 질을 향상 시킬 수 있는 기술적 대안으로 감각치환 기술에 대한 관심이 증가하고 있다.

감각치환은 일반적으로 손상되거나 기능이 저하된 감각의 정보를 다른 형태의 감각으로 전환하여 전달 또는 사용하는 것을 의미한다. 감각치환 기술은 사람의 뇌가 새로운 환경에 적응하기 위하여 구조, 기능적으로 변화하고 재조직되는 뇌가소성(Brain plasticity)에 기반하는 것으로 알려져 있다. 감각 능력이 저하되거나 상실하였을 때 두뇌 피질의 감각 정보들을 담당하는 영역들은 재배치가 일어난다. 이러한 발달 현상은 인체의 특정 감각이 기능을 상실하였을 때 다른 감각을 사용하여 기능을 대체하고 적응하는 것을 가능하게 해 준다.

1960년대 위스콘신 대학의 뇌과학자인 바크-이-리타(Bach-y-Rita) 교수는 지각(Perception)은 주로 신경계에서 일어나기 때문에 눈으로 볼 수 없는 사람도 다른 감각을 통하여 신경계로 정보가 전달되지만 하면 뇌가소성으로 인하여 뇌로 볼 수 있다고 가정하였다. 그는 1969년 카메라 영상을 피험자의 등에 장착된 바이브레이터 매트릭스를 통하여 진동으로 변환하여 전달함으로써 선천적 시각장애인이 사람의 얼굴, 물체, 그림자를 인지할 수 있는 최초의 감각치환 장치를 만들었다[3]. 이후 시각 정보를 청각으로 치환하는 방법, 청각 정보를 촉각으로 전달하는 방법 등 뇌가소성을 이용한 다양한 감각치환 방법들이 선진 연구기관 중심으로 진행되고 있다.

현재 감각의 회복을 위하여 인공 망막 보철, 유전자 치료, 광 수용체 이식 등 물리적으로 기능을

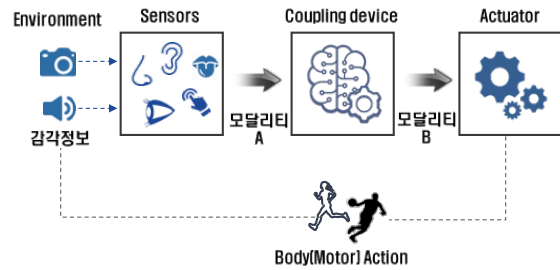


그림 1 감각치환 시스템 구성

대체하거나 복원하는 침습적 방법이 활용되고 있다. 이러한 방법은 시각장애인에게 제한적이지만 환경을 해석할 수 있는 시력 회복을 보임으로써 향후 큰 가능성을 지니고 있지만, 개인별로 다양한 감각 소실의 특정 병인(감각 악화의 유형과 심도, 경로의 병변 부위)을 발견하고 해결하기 위하여 기술적으로 해결해야 할 많은 난제를 가지고 있으며, 고비용, 낮은 품질, 긴 재활과정 등으로 인한 어려움 역시 극복하여야 할 과제이다.

이러한 침습적 방법과 비교할 때 감각치환 장치는 많은 장점을 가지고 있다. 예를 들면, 대표적인 침습 방법인 망막 임플란트에 비해 시각 정보를 청각신호로 변환하여 주는 비침습 방법은 비용이 저렴하여 경제적 부담이 없고 신체의 물리적 변화를 주지 않아서 의학적 위험에 따른 사용자의 거부감을 줄일 수 있다.

감각치환 장치는 그림 1에서와 같이 일반적으로 주위 정보를 수집하는 **센서**, 센서에서 입력 받은 신호를 해석하고 변환해서 전달하는 **커플링 장치** 또는 **시스템**, 변환된 정보를 사람이 인식할 수 있는 형태로 전달하는 자극 장치인 액추에이터로 구성된다.

감각치환에 관한 연구는 초기에는 주로 촉각이나 청각을 통해 시각을 대체하고자 하는 연구가 이루어졌으나, 최근에는 신체 균형 능력 보완과 같이 다양한 분야에서 감각치환 기술에 대한 연구들이

시도되고 있다.

이론적으로 감각치환 기술은 사람의 모든 감각을 대상으로 적용할 수 있다고 알려져 있지만, 본고에서는 우리나라 감각장애 인구 중 가장 높은 비율을 차지하는 시각장애인과 청각장애인을 대상으로 적용할 수 있는 청각치환 기술, 촉각치환 기술의 전반적 동향과 기술적 이슈에 초점을 두고 기술한다.

II. 감각 신호의 청각치환 기술 동향

1. 기술 개요

감각 정보의 청각치환은 사용자 주변의 신호 자극들을 청각 신호로 변환하여 전달함으로써 청각을 통하여 다양한 감각 정보를 인지할 수 있게 하는 기술이다. 사람의 한쪽 귀에는 약 3만 개 이상의 청신경 섬유가 존재하고, 대뇌 일차 청각 피질에는 6천만 개 이상의 뉴런이 존재하여 소리 정보의 경우 뇌에서는 초당 200회 이상의 변화를 감지할 수 있다. 또한 청각은 잠들었을 때도 유일하게 동작하는 감각이며, 움직이지 않고 사방의 소리를 들을 수 있는 장점이 있다. 따라서 이러한 청각의 특징을 활용하여 감각 신호의 청각치환 기술에 대한 연구가 많이 진행되고 있다.

감각 신호의 청각치환 기술은 그림 2에서와 같이 외부 환경의 감각 정보를 센서를 통하여 수집하는 단계, 수집된 감각 정보를 청각 신호로 전달하기 위하여 입력 정보들을 변환하거나 특정 정보들

을 매핑하고 필요한 경우 청각 기관에서 수용할 수 있는 내용으로 데이터를 함축하는 단계, 변환된 데이터를 사운드 생성 모델을 사용하여 음향 정보로 전달하는 단계로 구성된다.

감각 정보의 청각치환 기술 개념은 시각장애인에게 청각을 통하여 시각 정보를 제공하여 청각 정보를 해석하는 훈련을 진행함에 따라 대뇌 피질 영역을 fMRI로 관찰함으로써 검증하였다. fMRI를 통한 대뇌 피질의 관찰 결과, 훈련이 진행됨에 따라서 대뇌 피질의 청각 영역뿐만 아니라 시각에 관여하는 부분도 역시 활성화되는 것을 확인하였다[4]. 이러한 실험 결과는 시각 정보를 상실한 경우 부분적이지만 청각 정보를 통해 시각 정보를 대체하였음을 의미한다.

사람의 청각 기관은 잡음이 많은 환경에서도 음성과 같이 복잡하고 빠르게 변하는 소리 패턴을 처리할 수 있으며[5], 정보 전송량이 높기 때문에 청각치환 기술은 타 감각을 활용한 기술보다 많은 정보를 전달할 수 있다는 장점이 있다. 또한 헤드폰, 카메라, 모바일 컴퓨팅 장치와 같은 인체 착용이 가능한 소형 장치들로만 구성되어 휴대성이 높고 사용자 편의성이 뛰어나다[6]. 마지막으로 감각 자극을 생성함에 있어서 기계적인 장치를 필요로 하지 않기 때문에 에너지 소모량이 적어 청각치환 기술을 탑재한 장치의 사용 시간을 늘릴 수 있는 장점이 있다.

2. 기술 동향

감각 정보를 청각 정보로 변환하여 전달하는 기술에 대한 연구는 전달하려는 감각 정보의 대상과 변환된 감각 정보를 소리로 변환하는 방법에 따라서 구분 지을 수 있다. 본 절에서는 감각 정보를 시각 정보로 한정하여 기술 동향을 살펴보고자 한다.

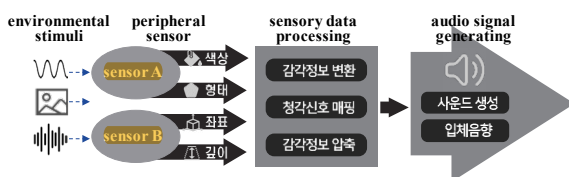


그림 2 감각 신호의 청각치환 기술

시각 정보의 청각치환 기술은 사용자에게 전달하는 시각 정보의 양을 늘리기 위한 방향으로 발전하고 있다. 초기 단계의 연구들은 저해상도의 흑백 영상 정보를 청각 신호로 변환했지만, 최근 연구에서는 상대적으로 고해상도의 컬러 영상 및 거리(Depth) 정보를 청각 신호로 변환하여 전달하기 위한 시도를 하고 있다. 또한 변환된 감각 정보를 음향 정보로 전달하는 단계에서는 사용자에게 친화적이고 입체적인 음향 신호를 생성하는 방향으로 발전하고 있다. 초기 연구들은 변환된 감각 정보를 음향 특징(주파수, 음의 높이, 소리 크기 등)으로 단순히 매핑하는 방법을 사용하였지만, 최근에는 사용자에게 듣기 좋은 악기 등의 음색으로 변환하고 입체음향 효과를 사용하여 공간감을 주기 위한 방법들이 사용되고 있다.

가. vOICe

vOICe는 1992년 네덜란드 델프트공대 피터 메이저(Peter Meijer) 박사에 의해 제안된 감각치환 장치[7]로서, 2차원의 그레이 스케일 이미지를 소리로 변환하여 사용자가 “소리로 볼 수 있는” 기능을 제공하며 현재 시각장애인을 대상으로 광범위하게 사용되는 장치이다.

vOICe는 시각 정보를 표현하는 전체 사운드를 사운드 스케이프스로 정의하고 사운드를 생성하기 위해 이미지 픽셀의 수평 위치, 수직 위치 및 밝

기를 사용한다. 이미지는 가로로 1초 단위로 코드화되므로 이미지에서 픽셀 열의 수평 위치는 사운드 스트림이 스위칭되는 시간축의 위치를 통해 표시된다. vOICe는 그림 3과 같이 사운드 이미지의 수직 좌표에는 각각 다른 사운드 주파수가 할당되고 개별 주파수를 갖는 사운드의 진폭은 해당하는 이미지 픽셀의 밝기에 매핑된다.

vOICe의 이러한 변환 알고리즘의 단순성으로 신호 변환에 높은 컴퓨팅 파워를 필요로 하지 않으므로 이미 휴대폰용 JAVA 앱 형태로 제공되어 사용하고 있다. vOICe 장치를 오랜 기간 사용하고 올바른 방법으로 훈련하면 사용자의 두뇌가 실제로 새로운 지각 방법에 적응하는 것이 실험적으로 입증되었다. 초기 교육 후에 사용자는 대조되는 배경에서 다양한 인공 이미지 또는 단순한 개체를 구별할 수 있다. 그러나 실제로 실생활에 적용하려면 최소 수개월 동안의 지속적인 연습이 필요하다.

또한 vOICe를 사용하기 위해 신호가 사용자의 많은 집중을 필요로 하며 많은 주변 환경 소리를 차단하는 단점이 있다. 무엇보다 vOICe에서 생성되는 신호가 사용자가 듣기에 거부감이 있어 시간이 지남에 따라 매우 자극적일 수 있으므로 사용자의 편의성 측면에서 치명적일 수 있다.

나. EyeMusic

vOICe는 원래 영상을 그레이 영상으로 변환하여 데이터를 처리함으로써 원래 영상에서의 일부 정보 손실이 발생한다. 특히 영상에서 색상 정보는 시각 데이터 처리에 중요한 특징 정보이다. 이러한 문제를 해결하기 위하여 이스라엘 히브리대학 연구진은 카메라로 촬영한 시각 정보를 ‘소리’로 전환하여 헤드폰을 통해 전달하여 시각장애인도 ‘소리 풍경(soundscape)’을 듣고 주변 상황을 인식하는 등 선천적인 장애로 인해 시각피질이 활성화된 적

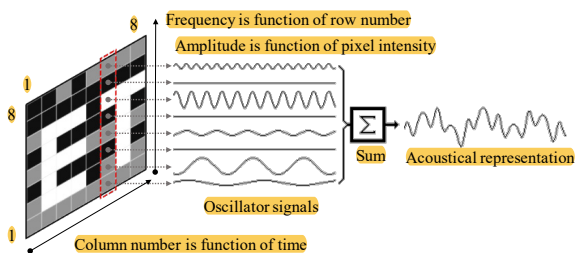


그림 3 vOICe의 음향신호 생성 개념도

없는 중증 시각장애인도 물체를 인식하고 글자까지 읽는 것이 가능해지는 EyeMusic 개발하였다[8].

EyeMusic은 기존의 사운드 스케이프를 기반으로 그림 4에서와 같이 영상의 색깔 정보를 6개의 색상 정보(빨간색, 파란색, 초록색, 노란색, 흰색, 검은색)로 클러스터링을 수행하여 각 색깔별로 다른 악기를 사용하여 사운드 정보를 표현하였다. EyeMusic을 사용한 시험 결과, 선천적 시각장애인도 훈련을 통하여 앞에 놓인 물체의 모양과 크기, 색깔까지 구분할 수 있고, 지속적인 반복훈련을 거치면 방 전체의 풍경을 묘사하거나 구별할 수 있으며, 글자를 인식하거나 단어를 읽을 수 있다고 발표하였다[9].

또한 EyeMusic은 사운드 생성에 사용하는 주파수 대역을 1,568Hz로 설정하고 사운드 생성을 5음계에 매핑하여 주파수 대역이 5,000Hz인 vOICe에 비하여 사용자가 듣기 친숙한 오디오 신호를 생성하여 전달하여 사용자의 거부감을 줄이도록 하였다.



출처 Reprint from S. Abboud et al., "EyeMusic: Introducing a visual colorful experience," *Restor Neurol Neurosci*. 2014; 32(2): 247-257, CC BY-NC

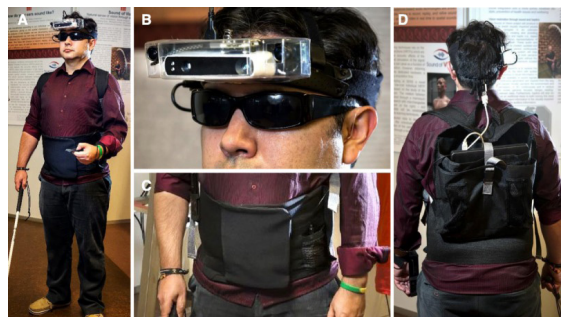
그림 4 EyeMusic의 개요 및 착용 예

다. Sound of Vision

Sound of Vision(SoV)은 유럽연합의 호라이즌 2020(Horizon 2020) 프로그램 중 하나로 2015년부터 3년간 진행된 프로젝트이다. 이 프로젝트의 목적은 시각장애인이 실생활에서 겪는 어려움을 완화시킬 수 있는 웨어러블 기기를 만드는 것이며, 이를 위하여 시각 정보를 음향 정보로 변환하여 전달할 수 있는 그림 5와 같은 웨어러블 기기의 프로토타입을 개발하였다[10].

SoV 시스템은 실내외 등의 다양한 환경 및 조명 조건에서도 신뢰성 있게 동작하기 위해서 실내에서는 ToF(Time of Flight) 카메라, 실외에서는 스테레오 카메라를 기본적으로 사용하면서 조명 조건에 따라서 색깔 등의 정보를 선택적으로 결합하여 깊이맵(Depth map)을 생성할 수 있도록 설계하였다. SoV 시스템에서는 깊이맵을 수평 방향으로 10개의 영역으로 구분한 후, 각 영역의 깊이 정보에 대한 특성값을 추출하여 음향 신호를 생성하였다. 음향 신호를 생성할 때에는 기존의 방법과 달리 유체 흐름 소리 모델(Fluid flow sound model)을 사용하였으며, 10개 영역을 왼쪽에서부터 20°씩 매핑하여 입체 음향효과를 주어 생성하였다.

기존 vOICe 방법에서 입력값을 깊이맵으로 변



출처 Reprint from R. Hoffmann et al., "Evaluation of an Audio-haptic Sensory Substitution Device for Enhancing Spatial Awareness for the Visually Impaired," *Optom Vis Sci*. 2018 Sep; 95(9): 757-765, CC BY-NC-ND.

그림 5 Sound of Vision 시스템

환한 방법과 사용자 설문을 통하여 비교하였을 때 정보 전달의 정확도는 유사한 수준이었지만, 변환된 소리에 대해서 느끼는 만족도는 SoV 시스템이 vOICe보다 매우 높게 나타났다. SoV 시스템은 깊이맵을 10개 영역으로 구분하여 10개 영역의 특성 정보를 제공하기 때문에 사용자가 물체의 형태를 구분하기 어려운 단점이 있다.

라. Vibe

기존의 사운드 스케이프 방법과 달리 Vibe에서는 센서와 수용체로 구성되는 2개 레벨의 셀(cell)을 가진 가상의 망막을 사용하여 사람의 망막을 모사하는 방법을 사용하였다[11]. 개별 센서들은 각각 이미지의 특정 픽셀들에 대응하며, 10개의 센서들로 구성된 수용체(RF: receptive field)는 이미지의 특정 영역을 대상으로 한다. Vibe의 동작 원리는 320×240 grey level의 입력 영상에 수용체 구성에 10개의 센서들을 사용하고 전체 20개의 수용체를 사용하여 44.1KHz의 샘플링으로 사운드를 생성하였다. 수용체의 행동은 수용체를 구성하는 개별 센서들의 행동들에 대한 동적 함수로 정의되고, 센서들의 상태에 따라 수용체의 진폭 크기가 조절되는 주기적인 사운드 신호를 생성한다. 따라서 Vibe 방법은 수용체를 구성하는 센서들의 분포를 설정하는 방법이 중요하다.

참고문헌 [12]에서는 Vibe의 가상 망막 개념을 활용하여 RF의 배치를 균일하게 구성하고 각 수용체의 소리를 멜로디로 튜닝하여 실내 보행에 적용하였다. 입력 정보로 Kinect sensor를 사용한 깊이맵(Depth map) 정보와 깊이 정보를 해석하기 위한 망막 인코더(Retinal encoder) 및 수용체의 정보를 음향 신호로 생성하는 스테레오 톤 생성기(Stereo tone generator)로 구성되었다. 가상의 망막 기능을 위하여 이미지 픽셀마다 8×8 로 균일하게 수용체를

배열하고 수용체를 중심으로 주위의 10개의 픽셀을 무작위로 선택하여 depth 정보 추출하여 각 RF의 수직 위치에 따라 사운드의 높이(Pitch)를 옥타브 스케일의 C4~C5까지 결정하였다. 좌, 우 음향 신호 생성을 위해 수용체의 activity 값을 이용한 음원 소스의 강도로 소리의 레벨 차이(ILD: Interaural Level Difference)를 계산하여 스테레오 음향을 생성하였다. 개발된 방법을 사전 학습 없이 시험하여도 장애물이 없는 실내 보행에 활용할 수 있는 가능성을 확인하였다.

III. 감각 신호의 촉각치환 기술 동향

1. 기술 개요 및 동향

1960년대 인간 두뇌의 가소성을 실험으로 증명한 바크-이-리타 교수는 한 인터뷰에서 “우리는 눈으로 보는 것이 아니라 두뇌로 본다.”라고 말하였다. 촉각치환 기술은 이러한 원칙을 이용하여 다양한 감각 정보를 피부의 진동이나 촉각 정보로 변환하는 감각 스와핑 기술을 의미한다.

감각 정보를 다른 감각 정보로 변환하여 전달하는 최초의 연구는 1960년대에 시각 정보를 촉각 정보로 변환하여 전달하려는 시도에서 시작하였다. TVSS(Tactile-Visual sensory substitution)으로 불리는 이 장치는 치과 진료 의자에 솔레노이드 자극기(Stimulator)를 20×20 로 배열하였으며, 이 의자에 앉은 사람 앞의 시각 정보는 가운데 카메라를 통해 등받이에 매트릭스 형태로 배치된 자극기로 전달하고, 400개의 작은 터치패드가 사람의 등을 누르는 방식으로 영상을 진동으로 변환시켜 전달하였다. 6명의 시험자를 대상으로 20~40시간 동안 훈련을 수행한 결과 25개의 자주 쓰이는 물건의 명사를 알아맞힐 수 있었다. TVSS는 감각치환을 적용하여 제한적이지만 선천적 시각장애인에게 열

물, 물체 및 그림자를 감지할 수 있는 가능성을 검증하였다는 의미가 있으나, 감각치환 장치의 원시적인 형태의 시각 대체용 촉각자극 장치의 초기 버전의 장치로 실용성은 고려되지 않았다.

가. BrainPort

감각치환 개념을 최초 제안한 위스콘신 대학의 마크-이-리타 교수팀은 마이크로 프로세서의 개발과 함께 TVSS에 사용한 진동장치를 고도화하여 전기 자극 장치를 사람의 혀에 탑재할 수 있도록 TDU(Tongue Display Unit)라는 어레이형 형태로 소형화하여 카메라 영상 정보와 전정감각 정보를 촉각 자극 정보로 전달하는 휴대형 장치인 BrainPort를 개발하였다[13].

BrainPort는 영상 정보를 획득하는 카메라가 장착된 헤드셋과 이미지의 픽셀 정보를 촉각 정보로 전달하는 촉각 자극 장치(Intra Oral Device) 및 제어 장치로 구성되어 있다. 롤리팝(Lollipop) 형태의 촉각 자극 장치(Intra Oral Device)는 BrainPort의 핵심요소이며, 이미지를 처리하는 base-unit으로부터 전달된 신호를 0~60볼트 사이의 단상파동(Monophasic pulse) 자극으로 바꾸어 144~400개 그리드 배열의 전극으로 전기 자극(Impulse)을 생성하여 촉각 정보를 생성한다. 제어 장치는 컬러 영상을 저해상도의 그레이 스케일 영상으로 변환하고, White 픽셀은 강한 펄스로, Black 픽셀은 제로 펄스로, Grey 레벨에 따라 적절한 전기 펄스를 생성한다. 입에 BrainPort 진동장치를 넣은 시각장애인이 사진을 400픽셀의 자극으로 매핑하여 자극을 전달하면 유명인의 이름을 맞추었으며, 몇 주간의 적응 기간을 거치면 객체의 위치, 이동, 원근과 깊이 등의 인식이 가능하여 물체를 식별하고, 날아오는 물체를 피할 수 있는 것으로 보고되고 있다. 현재 Brainport 기술은 마크-이-

리타 교수가 설립한 Wicab 회사에서 제품화하여 2015년에 FDA의 승인을 받아 뇌가소성 연구에 활용되고 있다[14].

BrainPort는 촉각의 민감도를 고려하여 구강내부 장치(Intra Oral Device)를 혀에 탑재하여 자극의 효율적인 전달이 가능하지만 항상 입에 자극 장치를 물고 있어야 하는 사용성의 문제와 장기간 사용 시 미각에 부작용을 줄 가능성이 존재한다. 또한 촉각으로 전달되는 제한된 정보를 활용하여 시각 정보를 지각할 수 있는 체계화된 학습 방안에 대한 연구가 추가적으로 필요하다.

나. VEST

베일러 의과대학의 신경과학자인 데이비드 이글만(David Eagleman)은 쉽게 들리지 않는 소리, 가시광선 밖의 빛 등의 정보를 신체에 연결된 장치의 진동 패턴으로 변환하여 뇌로 전달함으로써 새로운 감각의 전달을 통한 인간의 지각 확장을 목표로 하는 감각대체 신경과학(Sensory Substitution Neuroscience) 프로젝트를 2014년부터 수행하고 있다. 프로젝트를 수행하면서 청각장애인에게 들을 수 있는 새로운 감각을 제공하기 위하여 외부 사운드를 진동으로 변환하여 전달하는 착용형 장치인 VEST(Versatile Extra-Sensory Transducer)를 개발하고 있다.

VEST[15]는 26개의 센서를 조끼에 부착하여 6~9m 가량의 주위 소리를 컴퓨터나 스마트폰 마

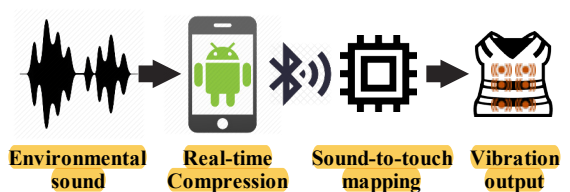


그림 6 VEST 동작 개요

이크로 수집한다. 수집된 소리들은 그림 6에서와 같이 특정 주파수 신호로 분해하고 블루투스를 통해 VEST에 전송하면 VEST에 부착된 센서가 전송받은 주파수 신호에 따라 진동하여 정보를 전달한다.

12일에 걸쳐 48개의 녹음된 단어에 대하여 사용자가 인지하도록 학습하였으며, 시험 결과 학습하지 않은 새로운 음성 단어에 대한 인식률이 학습 전보다 30% 가량 개선된 것으로 알려져 있다.

VEST는 주로 청각장애인에게 들을 수 있는 새로운 감각을 제공하는 것을 목적으로 개발되었지만, 최근에는 오픈 API를 활용하여 트위터 피드나 주식시장의 데이터 등의 다양한 정보를 전달하려는 시도를 하고 있다.

VEST는 사운드 정보를 촉각 정보로 전달하는 새로운 가능성을 제시하였으나, 제안한 조끼 구조의 시스템으로는 사용자가 일상생활에서 사용하기 힘들며, 단순한 변환 외 실생활에서 생길 수 있는 다양한 문제점(이동, 노이즈, 착용성 등)들에 대한 분석과 해결이 필요하다. 마지막으로 학습의 난이도가 높은 만큼 학습하는 방법에 대한 연구 또한 추가적으로 이루어져야 실제 청각장애인이 사용할 수 있을 것으로 예상된다.

다. Facebook 진동촉각밴드

Facebook에서는 사운드 정보를 촉각 장치로 전달하기 위하여 음소(phoneme) 단위로 패턴을 생성하여 사용자가 인지하게 하는 방법을 사용하였다[16]. 구강 내에서 발성하는 위치가 각기 다르다는 사실을 이용하여 발성 위치를 팔에서 진동되는 위치로 매핑하였다. 일반적으로 영어에서 모음이 자음보다 발음이 길다는 특징을 이용하여, 발음을 자극으로 변환하기 위하여 모음은 2개, 자음은 1개의 액추에이터를 이용한 진동패턴으로 표현하

였다.

음소 단위로 생성된 패턴에 대한 학습의 효율성 검증에는 자음과 모음을 무작위로 배치하여 테스트한 사용자와 연관성을 이용한 사용자 사이의 학습 시간을 비교하여 연관성을 통한 학습이 더 용이하다는 사실을 증명하였다. 제안된 기술은 9개 패턴의 생성을 위해 5개의 영어 자음(m, d, dh, w, k)과 4개의 모음(A, AI, E, OO)을 6개의 액추에이터를 이용하여 음소로 매핑한 후, 음소를 먼저 배우고 음소 및 단어를 테스트(Bottom-Up)하는 방법과 단어를 먼저 배우고 음소 및 단어를 테스트(Top-Down)하는 2가지 방법으로 학습하였다. 4명의 시험자를 대상으로 8개의 2음절 단어와 12개의 3음절 단어를 사용하여 2일에 걸쳐 학습한 결과 단어 단위의 인식률은 BU가 88%, TD가 76%, 새로운 단어에 대한 인식률은 BU가 72%, TD가 87%의 인식률을 나타내었으며, 평균 88%의 인식률을 가지는 것으로 발표되었다.

사운드 정보의 학습을 연관 관계를 이용하여 쉽게 할 수 있는 방법은 정확한 음소를 기반으로 패턴을 생성하여 문자를 촉각으로 변환하는 데 유용하지만 다양한 소리가 존재하는 실생활의 음성 대화에서는 사용이 어려울 수 있으며, 언어 이외의 소리는 변환이 불가하다.

지금까지 살펴본 진동촉각(Vibrotactile)을 이용한 감각치환 기술의 효율성은 미 해군에 의해 평가되기도 하였다. 이들이 개발한 장치는 시각에 의지하는 항공기 계기판의 정보를 촉각 정보로 변환하는 시도으로써, 실제 항공기의 움직임을 측정한 데이터를 항공기 조종사가 입고 있는 진동조끼로 전달하여 3차원 운동에서의 이상 기동을 경고하는 데 사용되었다. 실험자들은 1시간가량의 훈련을 통하여 실제 눈을 가린 상태로 곡예 비행을 한 후 항공기를 안정화시키는 데 성공하였다[17]. 이와 같은 시

험들은 감각치환 기술을 이용하여 추가로 제공되는 감각 정보를 활용하여 지각 능력이 향상될 수 있음을 보여준다.

IV. 감각치환 장치의 기술 이슈

감각치환 기술은 의학적 문제의 해결 방법으로 시각적 영상을 ‘듣거나’ 소리를 ‘느끼도록’ 하는 신경 연결 통로의 재건이라는 단순한 목적의 감각의 전달 경로 변경에 대한 가능성 검증을 넘어서 과학과 기술을 이용하여 현재 인간이 가지고 있는 신체적·정신적 한계를 극복하기 위한 시도로 확장하고 있다. 본 절에서는 이러한 관점에서 앞에서 분석하였던 기존 감각치환 장치들이 극복하여야 할 문제들에 대하여 정리한다.

1. 청각치환 장치의 기술 이슈

가. 학습의 어려움

청각치환 장치를 활용하여 새로운 체성 감각의 형태로 감각 정보를 전달하는 경우, 새롭게 변환된 신호를 훈련을 통하여 학습하는 과정이 필요하다. 기존 청각치환 기술의 학습에는 시험자의 지속적인 학습이 필요하고, 무엇보다 제대로 된 지각을 위해서는 전체 학습 과정에서 가이드의 역할이 중요하다. 또한 학습 방법이 시험자 개인의 특성이 아닌 기술을 개발한 공급자 측면에서 제시하는 성향이 강하여 사용자 맞춤형 학습이 어렵다. 따라서 시험자에 따라 효율적인 학습이 가능한 플랫폼의 구축이나 학습 방안의 개발이 필요하다.

나. 사용자의 거부감

감각 정보를 청각 신호로 변환하는 과정에서 사용하는 음향 신호의 피치(Pitch)나 주파수가 사용

자의 측면에서는 거부감을 느끼거나 장시간 집중하여 사용하기에 어려움이 있다.

또한 일상생활을 통해 듣게 되는 소리와 청각치환을 통해 입력되는 소리가 동시에 들리는 경우 뇌에서 어떤 과정을 통하여 통합되는지, 어떤 소리가 선택되는지에 대한 부분도 연구가 되어있지 않은 부분으로 감각 연합에 관련한 기전분석 및 선행 연구도 필요하다.

다. 단일 감각변환의 한계

기존 청각치환 연구들의 경우 주로 단일 감각 정보를 1:1 형태의 청각 신호로 변환하였다. 대다수 감각 정보의 경우 감각 정보가 가지는 특성과 정보의 양이 다르므로 단일 정보로 표시하기에 어려움이 있다. 따라서 감각 정보가 가지는 다양한 정보들을 청각 신호로 복합적으로 변환하기 위한 방안이 필요하다. 이러한 변환 방법은 지나치게 높은 인지적 부하를 초래하지 않고 직관적·자동적으로 변환 정보를 해석할 수 있는 자극을 생성하도록 고도화된 감각치환 알고리즘 개발이 필수적이다.

라. 감각 수용체의 한계

일반적으로 사람의 가청주파수는 20~20,000Hz 정도이며, 예민한 사람은 30,000Hz 소리까지도 들을 수 있다고 한다. 그러나 나이가 들수록 사람의 가청주파수 영역도 줄어든다고 알려져 있다. 10세 정도의 어린이는 20,000Hz까지 듣지만 20대에는 보통 17,000Hz, 40대가 되면 12,000~15,000Hz, 60대에는 10,000Hz까지 줄어들기도 한다. 가청주파수의 제한은 기본적으로 청각으로 변환할 수 있는 데이터에 한계가 있음을 의미한다. 따라서 제한된 감각 수용체의 한계에 맞게 적절한 데이터를 변환하고 전달하기 위한 방법이 필요하다.

2. 촉각치환 장치의 기술 이슈

가. 오랜 학습 시간

청각과 달리 촉각 치환 장치의 경우 데이터의 전달을 위하여 신체 피부에 직·간접적으로 자극을 준다. 이러한 자극으로 전달되는 신호에서 정보를 인지하기까지 상당한 시간이 걸리며 지속적인 학습을 필요로 한다. 또한 시각이나 청각과 달리 촉각은 정보와 직관적인 연결성이 적어 사용자가 학습하기 쉬운 패턴을 설계하고 효과적으로 학습시키는 방법론 개발이 필요하다.

나. 촉각 수용체의 한계

기본적으로 사람의 피부는 부위에 따라 차이는 있겠지만 인가되는 자극을 인지할 수 있는 한계가 있으므로 자극을 인가하는 강도와 인접하여 자극을 줄 수 있는 거리의 제약이 존재한다. 신체 부위 중 등 부분은 1cm 이상의 촉각 공간 분해능을 갖는 반면에 손가락 끝(fingertip), 혀 부분은 1~2mm의 민감한 촉각 공간 분해능을 갖는다. 따라서 많은 정보를 촉각으로 전달하기 위해서는 넓은 자극 면적이 필요하게 되며, 신체에서 자극 면적이 넓은 부분은 한정적이다. 이를 해결하기 위해 변환하려는 정보를 촉각에 맞게 함축/압축하는 과정이 다른 감각기관에 비해 중요하다.

또한 감각 정보를 촉각 신호로 전달하는 경우, 자극의 강도에 따라 사용자가 거부감을 느끼거나 장시간 지속적으로 자극을 인가하는 경우 사용자의 피로감이 발생할 수 있다. 특히 민감성 피부의 사용자의 경우 쉽게 통증을 느끼거나 짜증을 느낄 수 있으므로 자극을 전달할 수 있는 시간이나 자극 부위 및 강도가 제한적이다.

다. 촉각 전달 장치

촉각 정보를 전달하기 위해서는 높은 공간 해상도와 촉각 출력을 제공하는 고집적 촉각 자극 장치(Tactile stimulator)와 제어 기술이 필요하다. 사용자에게 촉각으로 정보를 전달하는 다양한 방법 중 현재 가장 많이 사용되고 있는 진동 액추에이터의 경우 크고 무거워서 착용형 장치에 많은 액추에이터를 사용하기 어렵다. 또한 한 번 설계된 액추에이터로 다양한 촉감을 표현하는 데 제한적이며, 이를 제어하기 위한 하드웨어 또한 복잡해진다. 따라서 향후 촉각변환을 통한 감각치환 장치의 상용화를 위해서는 피부의 촉각 분해능을 수용할 수 있는 고집적 촉각 자극 장치의 발전과 촉각을 자극할 수 있는 다양한 방식의 센서에 관한 연구와 함께 자극을 효율적으로 생성하기 위한 알고리즘의 개발이 필요하다.

V. 결론

감각치환 기술은 일반적으로 손상되거나 저하된 감각의 정보를 다른 형태의 감각으로 전환하여 전달 또는 사용하는 것을 의미한다. 최근 시각, 청각 등의 감각이 손상되거나 저하된 고령자와 장애인 이 증가하면서 이들의 손상된 감각 및 지각 능력을 향상시켜 지속적인 경제 활동과 삶의 질을 향상시킬 수 있는 기술적 대안으로 감각치환 기술 및 장치에 대한 관심이 증가하고 있다.

본 문서에서는 감각치환 기술 중 우리나라 장애인구의 가장 높은 비율을 차지하는 시각장애인과 청각장애인을 대상으로 적용할 수 있는 청각치환 기술과 촉각치환 기술의 전반적인 동향과 기술적 특징들을 분석하였다.

감각치환 기술은 촉각이나 청각을 통해 시각을 대체하고자 하는 개념적인 연구에서 시작되었으며, 최근에는 실생활에 적용을 목적으로 다양한 분야로 연구 영역이 확장되고 있다. 이러한 기술 개발 동향을 고려하여 본 고에서는 감각치환 장치들이 실험실 수준을 벗어나 향후 실생활에 활용되기 위해 극복하여야 할 다양한 기술적 이슈들을 분석하였다.

최근 장애인들의 사회 참여 욕구의 증대와 함께 과학 기술의 사회적 책임이라는 관점에서 향후 감각치환 기술 분야에는 보다 다양한 기술적 접근이 필요할 것으로 예상된다.

용어해설

Sensation(감각) 주변 환경으로부터 눈, 코, 귀, 혀, 피부 등을 통해 입력되는 다양한 물리적 자극을 전기적 형태로 바꾸어 신경계를 통해 전달하는 것

Perception(지각) 자극에 의해 발생한 감각으로부터 의미 있는 패턴을 생성하여 다른 감각과 비교하거나 과거의 기억을 기초로 감지한 것에 의미를 부여하는 과정

Recognition(인지) 다양한 사물을 알아보고 그것을 기억하며, 추리해서 결론을 도출하고, 그로인해 생긴 문제를 해결하는 등의 정신적인 과정

약어 정리

API	Application Programing Interface
FDA	Food and Drug Administration
fMRI	Functional Magnetic Resonance Imaging
ILD	Interaural Level Difference
RF	Receptive Field
SoV	Sound of Vision
SSD	Sensory Substitution Devices
TDU	Tongue Display Unit
ToF	Time of Flight
TVSS	Tactile—Visual Sensory Substitution
VEST	Versatile Extra—Sensory Transducer
WHO	World Health Organization

참고문헌

- [1] 통계청, “2018 고령자 통계,” 2018.
- [2] 한국장애인재단, “WHO세계장애보고서,” 2012.
- [3] P. Bach-Y-Rita et al., “Vision substitution by tactile image projection,” *Nature*, vol. 221, no. 1, 1969, pp. 963-964.
- [4] C. Poirier et al., “Pattern recognition using a device substituting audition for vision in blindfolded sighted subjects,” *Neuropsychologia*, vol. 45, no. 5, 2007, pp. 1108-1121.
- [5] I. J. Hirsh, “Auditory perception and speech,” in R. C. Atkinson et al. (ed.), *Handbook of experimental psychology*, John Wiley: New York, USA, 1988, pp. 377-408.
- [6] C. Capelle et al., “A real-time experimental prototype for enhancement of vision rehabilitation using auditory substitution,” *IEEE Trans. Biomed. Eng.*, vol. 45, no. 10, 1998, pp. 1279-1293.
- [7] P. B. L. Meijer, “An experimental system for auditory image representations,” *IEEE Trans. Biomed. Eng.*, vol. 39, no. 2, 1992, pp. 112-121.
- [8] S. Abboud et al., “EyeMusic: Introducing a “visual” colorful experience for the blind using auditory sensory substitution,” *Restor. Neurol. Neurosci.*, vol. 32, no. 2, 2014, pp. 247-257.
- [9] L. Reich, A. Amedi, “Visual parsing can be taught quickly without visual experience during critical periods,” *Scientific Reports*, vol. 5, 2015, pp. 15359:1-12.
- [10] S. Caraiman et al., “Computer vision for the visually impaired: the sound of vision system,” in *Proc. IEEE Int. Conf. Comput. Vision.*, Venice, Italy, Oct. 2017, pp. 1480-1489.
- [11] S. Hanneton, M. Auvray, B. Durette, “The Vibe: a versatile vision-to-audition sensory substitution device,” *Appl. Bionics Biomechanics*, vol. 7, no. 4, 2010, pp. 269-276.
- [12] V. Fristot et al., “Depth—Melody substitution,” in *Proc. Eur. Signal Process. Conf. (EUSIPCO)*, Bucharest, Romania, Aug. 27-31, 2012, pp. 1990-1994.
- [13] F. Ashwin et al., “Brainport Vision Technology” *Int. Research J. Eng. Technol.*, vol. 3, no. 03, 2016, pp. 475-483.
- [14] BrainPort Technologies, “Brainport Vision Pro,” <https://www.wicab.com/brainport-vision-pro>
- [15] S. D. Novich, D. M. Eagleman, “Using space and time to encode vibrotactile information: toward an estimate of the skin’s achievable throughput,” *Experimental Brain Research*, vol. 233, no. 10, 2015, pp. 2777-2788.
- [16] S. Zhao et al., F. “Coding tactile symbols for phonemic communication,” in *Proc. CHI Conf. Human Factors Comput. Syst.*, Montreal, Canada, Apr. 2018, pp. 392:1-12.
- [17] A. Rupert, “Tactile interface to improve situation awareness,” Aerospace Medical Association, 1996.