

---

# 시각 정보의 청각 변환을 위한 시각장애인의 정보 선호도 분석

김무섭<sup>1</sup>, 정치윤<sup>1,\*</sup>, 문경덕<sup>1</sup>, 박윤경<sup>1</sup>, 박종미<sup>2</sup>, 이준우<sup>3</sup>

## Analysis of Preference Information on Visual-to-Audio Sensory Substitution for the Visually Impaired

Mooseop Kim<sup>1</sup>, Chi Yoon Jeong<sup>1,\*</sup>, Kyeong-Deok Moon<sup>1</sup>, YunKyung Park<sup>1</sup>, Jong-Mie Bach<sup>2</sup>, Jun-Woo Lee<sup>3</sup>

### 요약

감각 치환 기술은 뇌 가소성에 기반하여 손상되거나 저하된 감각의 정보를 다른 형태의 감각으로 전환하여 전달함으로써 감각을 지각할 수 있도록 하는 기술이다. 본 논문에서는 감각 치환 기술 중 가장 대표적인 시각-청각 변환기술을 묵시적 방법과 명시적 방법으로 구분하여 기술적 특징을 분석하였다. 또한, 166명의 시각장애인을 대상으로 설문조사를 수행하여 시각-청각 변환기술을 통하여 전달받기 원하는 시각 정보들의 선호도를 분석하였다. 기존 기술 분석 및 설문조사 결과를 종합하면 기존 기술은 주로 모양에 관련된 정보를 전달하지만, 실제 시각장애인들은 보행을 위한 거리 및 공간 정보를 요구하여 현재 기술과 사용자의 요구사항에 차이가 있음을 확인하였다. 본 논문의 분석 결과는 감각 치환 장치들이 향후 실생활에 사용되기 위해 극복하여야 할 여러 가지 기술적 이슈들을 해결하기 위한 기본 자료로 활용될 수 있다.

### Abstract

Sensory substitution generally refers to converting or transmitting information from damaged or deteriorated sense to other types of sensory modality based on brain plasticity. In this paper, we review the technical features of the visual-to-auditory sensory substitution that could be classified into implicit and explicit methods. In addition, we analyzed the preference for visual information that the visually impaired want to receive through a questionnaire survey that was conducted with 166 visually impaired participants. We can conclude from the analysis that there are considerable difference between the functional feature of the conventional methods and the preference information of the visually impaired. The analysis results of this paper

---

<sup>1</sup>한국전자통신연구원 인공지능연구소 휴먼증강연구실(책임연구원)

<sup>2</sup>강남대학교 한국CISM연구소(책임연구원)

<sup>3</sup>강남대학교 사회복지학부(교수)

\*Corresponding Author : iamready@etri.re.kr

접수일자 : 2020. 06. 05.

1차 심사 : 2020. 06. 05.

2차 심사 : 2020. 06. 08.

게재확정 : 2020. 06. 09.

DOI : <http://data.doi.or.kr/10.22733/JITAE.2020.10.01.007>



will provide the foundation for the research to solve technical issues that need to be overcome for sensory substitution devices to be used in real life in the future.

Keywords : Sensory substitution, visual-to-auditory neural plasticity, blindness, preference information

## 1. 서론

보건복지부에서 발표한 장애인 현황조사에 따르면, 2018년 12월 기준으로 우리나라의 등록 장애인 2,585,876명의 10%에 해당하는 252,957명이 시각장애를 갖고 생활하고 있다 [1]. 이러한 수치는 시간이 지남에 따라 지속해서 증가하고 있으며, 장애 유형은 중도(severe)·중복(multiple)화 되고 있는 것으로 나타났다 [2]. 시각장애는 인간의 감각 중 가장 심각한 상실이며, 시각장애를 갖고 생활한다는 것은 일상생활의 영위를 위해서 필요한 감각 인지 능력의 90%를 상실하는 것과 같다고 발표되고 있다 [3]. 시각장애인의 증가와 함께 일상생활에서 겪는 어려움의 양상이 복잡하고 다양해짐에 따라 시각장애인의 지원을 위한 기술적, 정서적 측면의 다양한 기술들의 필요성과 중요성이 높아지고 있다.

현재 손실된 시각의 제한적인 기능 회복을 위하여 인공 망막 보철을 사용하거나 광 수용체의 이식 등의 방식이 활용되고 있다. 이러한 방식은 시각장애인에게 제한적이지만 시각 기능을 대체하거나 물리적으로 기능을 복원하여 시력의 직접적인 회복 가능성을 제시하였다. 그러나 개인별로 다르게 나타나는 시각 손실의 원인을 분석하여 개별적으로 대응하여야 하므로 향후 해결해야 할 많은 기술적인 도전 과제를 가지고 있다. 무엇보다 이러한 침습적 방식은 신체에 삽입하거나 부착이 필요하므로 이질감과 함께 수술에 대한 거부감이 크게 작용한다. 또한, 높은 비용에 비하여 낮은 품질, 재활 및 적응에 많은 시간이 필요하다는 점도 역시 극복해야 한다 [4].

최근 들어 이러한 침습적 방법에 대한 기술적 대안으로 장치나 기구의 인체 삽입이 필요

없는 비침습적 방법을 활용하여 시각 기능을 복원하거나 보완하는 연구가 제안되고 있다. 특히, 감각 치환 기술은 경제적 부담이 상대적으로 적고, 수술에 따른 의학적 위험에 대한 사용자의 거부감을 줄일 수 있다는 장점으로 최근 관심이 증가하고 있다.

일반적으로 감각 치환 기술은 기능이 저하되거나 손상된 감각 신호를 다른 감각 기관으로 전달 또는 대체하여 사용하는 것을 의미한다. 이러한 감각 치환 기술은 사람의 뇌가 새로운 환경에 적응하기 위하여 구조, 기능적으로 변화하고 재조직되는 뇌 가소성(brain plasticity)에 기반하는 것으로 알려져 있다 [4].

뇌 과학자들에 따르면 뇌는 다양한 입력 채널에 대응하여 전달되는 신호들을 받아들이고 처리할 내용만 분석하고 판단하므로, 주변 정보나 외부 장치로 감지한 신호를 감각 기관을 통해 뇌로 전달해주면 뇌가 인지할 수 있다고 한다.

이러한 가정은 1969년에 뇌 과학자였던 바키리타(Bach-y-Rita) 교수가 카메라에서 입력되는 영상을 실험자의 등에 매트릭스 형태로 장착된 진동 장치를 통하여 전달함으로써 선천적 시각장애인이 사람의 얼굴이나 물체를 인지할 수 있는 최초의 감각 치환 장치를 개발함으로써 실험적으로 가능성을 증명하였다 [5]. 이러한 가설의 검증 이후 시각 정보를 청각 신호로 변환하거나 시각 정보를 촉각으로 전달하는 등 다양한 감각 치환 방법들이 일부 선진 연구 기관들을 중심으로 제한적으로 진행되고 있다.

현재 진행되는 대부분의 감각 치환 기술들은 감각 정보를 단순한 변환 과정을 활용하여 사용자 시험을 시행하여 전달된 감각 정보의 인지 여부를 확인하는 기술적 개념의 검증 및 가능성을 확인하는 수준의 연구가 단편적으로 진행되고 있다. 따라서 실제 기술들의 활용이 예

상되는 시각장애인들의 기술적 요구에 대한 분석이나 감각 치환 기술을 통하여 전달받기를 원하는 관심 정보에 관한 체계적인 연구나 조사는 아직 진행되지 않고 있다. 본 논문에서는 시각 정보의 청각 신호 변환기술의 개발에 필요한 시각장애인의 시각 정보 인지 및 행동 특성에 대한 분석과 이를 기반으로 166명의 시각 장애인에 대한 조사를 진행하여 시각장애인이 시각-청각 변환기술을 통하여 전달받기를 원하는 상황과 시각 정보 및 선호하는 장치의 형태들에 대하여 분석하였다.

본 논문은 구성은 다음과 같다. 2장에서는 기존 시각 정보의 청각 변환기술을 소개하고, 3절에서는 시각장애인의 정보 선호도를 분석한 결과를 살펴본다. 마지막으로 4절에서는 결론 및 향후 연구 방향을 제시한다.

## 2. 시각 정보의 청각 변환 기술

### 2.1 시각-청각 변환 기술 개요

정보나 데이터를 소리로 변환하여 나타내는 방식을 소니피케이션(Sonification)이라 정의하며 꽤 오랫동안 과학계에서 연구되어왔다. 소니피케이션은 정보를 시각에 의존하지 않고 청각을 이용하여 데이터의 흐름을 새로운 감각의 경험으로 변환함으로써 보다 효과적인 형태로 정보를 접할 수 있는 방식을 제공한다. 최근에는 데이터가 주는 다양성과 예측 불가능한 성질을 예술적 형태로 접목하려는 시도가 있었으며, 최근에는 유사한 형태로 데이터를 시각화하여 표시하는 미디어아트가 주목을 받기도 하였다. 따라서 감각 치환 기술은 크게 소니피케이션의 일부로 파악할 수 있다. 감각 치환 기술의 관점에서 시각 정보의 청각 신호 변환은 이미지나 카메라에서 입력받은 영상 신호를 청각 신호로 변환하여 전달함으로써 예술적 표현이 아니라 청각을 통하여 시각 정보를 인지하는 것을 목표로 한다. 사람의 뇌는 초당 200회 이상 변화하는 소리를 감지할 수 있다고 알려져 있다. 또한, 청각은 유일하게 사람이 잠들었을 때도 감지할 수 있는 감각이며 동적인 활동 없

이도 소리를 들을 수 있다는 장점이 있다. 따라서 이러한 청각의 특징으로 감각 치환 기술 중 시각-청각 변환기술에 관한 연구가 많이 진행되고 있다.

청각 신호로 변환된 시각 정보의 대뇌 지각 여부의 검증은 시각장애인에게 일정 기간에 걸쳐 청각을 통하여 시각 정보를 제공하여 청각 정보를 해석하는 훈련을 진행하면서 대뇌 피질 영역의 반응을 기능적 자기공명 영상(fMRI)으로 관찰함으로써 검증하였다. fMRI를 통한 대뇌 피질의 변화를 관찰한 결과 청각 신호를 활용한 시각 정보의 인지 훈련이 진행됨에 따라 대뇌 피질의 청각 영역뿐만 아니라 시각에 관여하는 시각 피질 영역도 활성화되는 것으로 보고되었다[6]. 이러한 시험 결과는 기능적으로 시각 정보를 상실한 경우라도 청각으로 변환한 시각 정보를 전달해주면 일부분 시각적 기능을 대체할 수 있음을 의미한다.

시각 정보의 청각 신호 변환에 필요한 기술적 구성과 절차는 그림 1과 같이 정리할 수 있다. 먼저 전달하려는 시각 정보를 이미지나 영상 데이터 형태로 수집한다. 다음으로 수집된 입력 데이터에서 청각 신호로 전달하기 위한 대상이 되는 시각 정보를 추출한다. 추출된 정보는 청각 신호로 변환하거나 특정 청각 정보들로 매핑하고 필요한 경우에는 청각 기관에서 수용할 수 있는 내용으로 데이터를 함축하는 과정이 필요하다. 마지막으로 변환하려는 데이터를 사운드 생성 모델을 사용하여 음향 정보로 생성하고 전달하는 과정이 필요하다.

본 연구에서는 그림 1의 시각-청각 변환 과정 중 시각적 특징을 추출하고 시각-청각 변환 과정에서 기존 연구들과 실제 시각장애인이 원하는 시각정보의 특징과 변환 방법을 분석하였다.

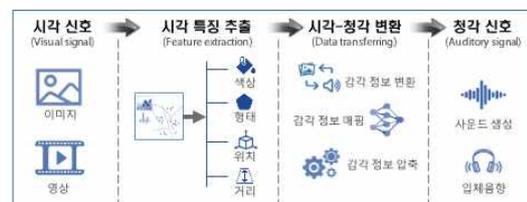


그림 1. 시각 정보의 청각 신호 변환.  
Fig. 1. Visual to auditory sensory substitution.

## 2.2 시각-청각 변환 기술 관련 연구

시각 정보의 청각 신호 변환기술은 감각 간의 정보를 매핑하는 방법에 따라서 명시적 방법들과 묵시적 방법으로 구분할 수 있다. 명시적 방법은 영상의 색상, 위치, 윤곽선 등의 정보를 소리의 주파수, 크기, 리듬 등으로 변환해주는 규칙을 정의한 후, 정의된 규칙에 따라서 특정 영상 정보를 소리로 변환하는 방법이다. 반면, 묵시적 방법은 전문가가 영상 정보를 소리 정보로 변환하는 규칙을 사전에 정의하지 않고 기계가 학습을 통하여 자동으로 입력 영상에 맞는 소리 정보를 생성한다.

명시적 방법의 시각-청각 변환기술 중 초기에 제안되었으나 현재까지도 가장 광범위하게 사용되는 방법이 vOICe이다. vOICe는 네덜란드 델프트공대 피터 메이저(Peter Meijer) 박사에 의해 1992년에 제안된 감각 치환 장치이다[7]. vOICe는 64x64 픽셀 해상도의 2차원 그레이 스케일 이미지를 입력받아 소리로 변환하여 사용자에게 전달함으로써 제한적이지만 소리를 통해 주변을 볼 수 있는 기능을 제공한다. 최근 영국 바스대 심리학과 연구팀이 vOICe 장치의 성능 측정을 위해 수행한 실험 결과에 따르면, vOICe 장치를 활용하면 시각 장애인인 최대 0.05 정도의 시력을 가질 수 있다고 발표하였다[8]. 0.05의 시력은 아주 희미하게 시각적으로 물체를 구분할 수 있는 수준의 시력을 의미한다. vOICe는 단일 화면의 프레임을 청각 신호로 변환하여 표현하는 소리의 단위를 사운드스케이프(sound scape)로 정의하고 변환된 소리를 생성하기 위해 이미지를 구성하는 픽셀의 수평 및 수직 위치, 밝기 정보를 사용한다.

vOICe는 그림 2와 같이 시간 축의 방향으로 이미지 프레임의 열 단위로 각 픽셀의 수직 위치와 픽셀의 밝기 정보를 이용하여 소리로 변환한다. 변환 과정에서 각 열의 수직 좌표에 해당하는 픽셀마다 각각 다른 소리의 주파수가 할당되고 해당 픽셀의 밝기 정보는 해당 주파수를 갖는 소리의 크기로 변환한다. 세로축에 사용하는 주파수는 픽셀 위치마다 최저 500Hz에서 최고 5,000Hz까지 지수적으로 변화하도

록 설정하였다.

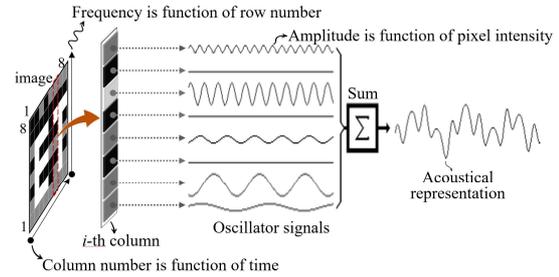


그림 2. vOICe 변환 개념.  
Fig. 2. Concept of vOICe conversion.

이렇게 프레임의 열을 구성하는 개별 픽셀의 음향 파형들은 모두 합하여 그 열을 대표하는 소리 정보로 변환되고 사용자에게 전달되는 소리는 44.1 Khz로 샘플링된 음향으로 생성된다. 따라서 이미지에서 열의 수평 위치는 사운드가 플레이되는 시간의 위치를 통해 표시된다. vOICe는 이러한 작업을 이미지 프레임을 구성하는 모든 열에 대하여 시간 축의 방향, 즉 왼쪽에서 오른쪽으로 진행하면서 동일한 변환작업을 반복해서 수행하여 1개의 이미지 프레임을 1초 단위의 소리 신호로 변환하여 전달한다. vOICe의 이러한 변환 방법은 직관적이면서 단순하여 학습을 수행하는 시각장애인이 학습 방법을 이해하기 쉽고 신호의 변환에 많은 컴퓨팅 자원이 필요하지 않으므로 최근에는 스마트폰 카메라를 이용하는 안드로이드 버전까지 사용되고 있다.

영상에서 색상 정보는 시각 데이터 처리에 중요한 특징 정보이다. 그러나 vOICe는 원래 영상을 그레이 영상으로 변환하여 데이터를 처리함으로써 영상에서 중요한 색상 정보의 손실이 발생한다. 또한, 시각-청각 변환을 위해 vOICe가 생성하는 소리는 사용자가 지속적으로 듣기에는 거부감이 있어 사용자 편의성 측면에서 단점이 있다.

vOICe의 이러한 문제점을 해결하기 위하여 이스라엘 히브리대 연구진은 카메라로 촬영한 시각정보를 '소리'로 전환하여 헤드폰을 통해 전달하여 시각장애인들도 사운드스케이프를 듣고 주변 상황을 인식할 수 있는 EyeMusic을 개

발하였다[9]. EyeMusic은 vOICe와 같이 사운드스케이프를 기반하지만 영상의 색상 정보를 빨강, 파랑, 녹색, 노랑, 흰색, 검정색의 6개 색상 정보로 군집화하고, 분류된 색상별로 다른 악기를 사용하여 사운드 정보를 표현하였다. 사운드 생성에 사용하는 주파수 대역은 최대 1,568Hz로 설정하고 사운드 변환으로 생성되는 소리를 5음계에 해당하는 소리로 매핑하여 주파수 대역이 5,000Hz인 vOICe에 비하여 사용자가 듣기 친숙한 오디오 신호로 전달하여 사용자의 거부감을 줄였다. EyeMusic을 사용한 시험 결과, 선천적 시각장애인도 눈앞에 놓인 물체의 모양과 크기, 색깔까지 구분할 수 있으며, 지속적인 반복훈련을 거치면 방 전체의 풍경을 구별하고 묘사할 수 있으며, 글자를 인식하거나 단어를 읽을 수 있다고 발표하였다[10]. EyeMusic은 입력 영상을 매우 작은 크기로 줄인 후 소리 신호로 변환하여 사용자에게 전체적인 정보를 제공해 줄 수 있지만, 세부적인 정보를 제공해주는 데 어려움이 있다.

명시적 방법의 시각-청각 변환기술로 사운드스케이프를 사용하는 방법과 달리 Vibe에서는 사람의 망막을 모사하기 위하여 센서와 수용체로 정의되는 2개의 레벨로 구성되는 셀(cell)을 가진 가상의 망막을 사용하였다[11]. Vibe에서 개별 센서들은 이미지의 특정 위치의 픽셀들에 대응하며, 사전에 일정하게 분포하도록 정의된 수용체(RF: receptive field)는 정의된 위치 주변에서 랜덤하게 선택된 10개의 센서로 구성되고 이미지의 특정 영역을 대상으로 한다. Vibe는 320x240 크기의 그레이 입력 영상에 20개의 수용체를 사용하여 44.1 KHz의 샘플링으로 사운드를 생성한다. 수용체의 동작은 해당 수용체를 구성하는 10개 센서들의 상태에 대한 동적 함수로 정의되므로 센서들에 대응하는 픽셀들의 값에 따라 수용체의 진폭 크기가 결정되는 사운드 신호를 생성한다. 따라서 Vibe 방법은 수용체를 구성하는 센서들의 분포를 효율적으로 설정하는 방법이 중요하다. 이러한 구성적 특징을 가지는 Vibe는 1초에 1개의 이미지 프레임에 해당하는 음향 신호를 생성하는 사운드 스케이프 방식과 달리 전체 수용체

의 음향 신호가 동시에 변환될 수 있으므로 빠른 시각-청각 신호 변환이 가능한 장점이 있다. 하지만 이미지 프레임을 변환한 음향 신호의 길이가 46ms 정도로 매우 짧아서 사용자가 시각 정보와 청각 변환 신호를 이해하는 데 어려움이 있을 수 있다.

최근 들어 컴퓨터 비전, 음성인식, 자연어 처리, 음성/신호 처리 등의 분야에서 다량의 데이터나 복잡한 자료들 속에서 핵심적인 내용 또는 기능을 요약하는 추상화 작업을 위하여 다양한 비선형 변환기법들의 조합을 활용하여 사람의 추론 방법과 유사한 형태의 추론을 학습하는 딥러닝(심층학습)이 주목할만한 연구 결과들을 보여주고 있다. 기존의 명시적 방법과 달리 사전 정의된 규칙이 없이 딥러닝 알고리즘을 적용하여 기계가 학습을 통하여 입력 감각을 다른 감각으로 변환하는 묵시적 방법의 시각-청각 변환기술로 오토 인코더 기반의 청각 신호 변환 방법(AEV2A)이 2019년에 발표되었다[12].

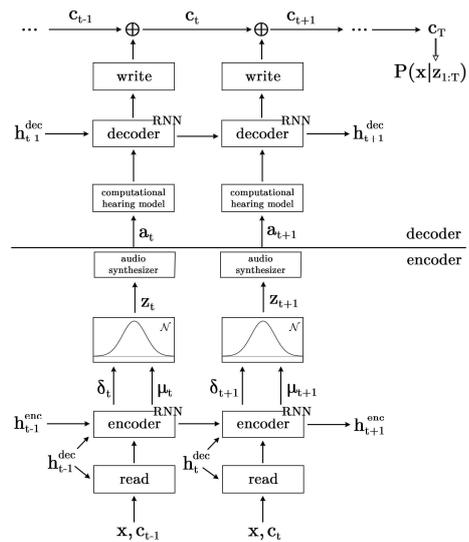


그림 3. AEV2A의 네트워크 구조도.  
Fig. 3. Network structure of AEV2A.

AEV2A는 DRAW(Deep Recurrent Attentive Writer) 네트워크와 음성합성 모델로 구성되며, 네트워크 구조는 그림 3과 같다. DRAW는 가변 오토 인코더(VAE: Variational AutoEncoder)를 기반으로 사람의 시각 인지 메커니즘을 모

방하여 순차적으로 영상을 생성하는 네트워크 구조이며, 공간적 주의 집중(attention) 메커니즘과 LSTM (Long Short-Term Memory) 구조로 되어있다. DRAW의 동적 주의 집중 메커니즘을 통해 각 타임 스텝마다 어떤 부분을 집중해서 읽고 쓸 것인지 결정하며, LSTM 네트워크를 사용하여 순차적으로 주의 집중 영역의 정보를 갱신한다. 따라서 AEV2A 방식은 입력 이미지를 160 x 120 크기로 줄인 후 에지 영상을 검출하여 딥러닝 네트워크의 입력으로 사용하며, DRAW 네트워크에서 잠재변수(latent variable)로부터 음성합성 모델을 사용하여 소리 신호를 생성하고, 청각 모델을 거쳐 청각적 특징을 추출하는 과정을 추가하여 시각 정보를 청각 신호로 변환한다.

AEV2A 방법의 경우, 기계가 학습을 통하여 입력 영상에 맞는 소리 정보로 변환하므로 사용자가 변환 규칙을 이해하기 어려워 학습 가능성이 작을 수 있다. 실제 AEV2A 방법의 사용자 학습 가능성을 검증하기 위하여, 입력되는 정보의 변화에 따라 생성되는 소리에 대하여 동적 시간 왜곡(DTW: Dynamic Time Warping)을 이용하여 생성되는 소리의 유사성을 비교한 결과 입력 이미지의 조그만 변화에도 전혀 다른 소리가 생성되어 사용자 학습이 어려운 것으로 나타났다[14].

지금까지 간략하게 살펴보았듯이, 초기 단계의 연구들은 저해상도의 흑백 영상 정보를 음향 특징(주파수, 음의 높이, 소리 크기 등)으로 단순히 매핑하는 방법을 사용하여 청각 신호로 변환했지만, 최근 연구에서는 상대적으로 고해상도의 컬러 영상 및 거리(Depth) 정보를 청각 신호로 변환하여 전달하기 위한 시도를 하고 있다. 또한, 시각 정보를 음향 정보로 전달하는 단계에서는 사용자가 듣기 좋은 악기 등의 음색을 사용하거나 입체적인 음향 신호를 생성하는 방향으로 발전하고 있다.

현재까지 개발된 시각-청각 변환기술의 경우, 표 1에서 정리한 것과 같이 주로 이미지에서 특정 시각 정보를 1:1 형태의 청각 신호로 변환하였다.

표 1. 기존 시각-청각 변환 방법의 특징.  
Table 1. Comparison of conventional V2A methods.

기술	변환방법	입력정보	제공정보	학습정도
vOICe	명시적	흑백이미지	모양	쉬움
Vibe	명시적	흑백이미지	모양	어려움
EyeMusic	명시적	컬러이미지	모양, 색상	보통
AEV2A	묵시적	흑백이미지	모양	어려움

대다수 감각 정보의 경우 감각 기관의 특성과 감각 수용체가 표현할 수 있는 정보의 양이 다르므로 단일 정보를 활용하여 전체 영상이나 이미지의 정보를 표시하기에는 어려움이 있다. 따라서 개별 감각 수용체의 한계와 특성을 고려하여 적절하게 데이터를 변환하고 정보를 전달하는 방법이 필요하다. 이는 시각-청각 변환 기술의 잠재적 사용자인 시각장애인들의 인지 및 행동 특성에 대한 정확한 이해와 함께 실제로 시각장애인이 시각-청각 변환기술을 통하여 전달받기를 원하는 정보에 대한 체계적인 분석이 선행되어야만 해결할 수 있다.

### 3. 시각-청각 변환을 위한 정보 선호도 분석

#### 3.1. 조사 대상 선정 및 방법

시각의 기능에는 시력, 시야, 대비감도, 광감도, 색각의 다섯 가지 기능이 있다. 이러한 시각의 기능 중 한 가지, 혹은 여러 가지 기능이 복합적으로 손상되었을 때의 상태를 일반적으로 시각장애라고 한다. 시각장애를 장애 정도에 따라 분류하면 전맹(盲)과 저시력(약시) 및 시야결손으로 분류할 수 있다[15]. 본 연구에서는 시각-청각 변환기술의 잠재적 사용자인 시각장애인의 선호 시각 정보를 장애 정도에 따라 조사, 분석하였다.

시각-청각 변환기술의 전달 대상이 되는 시각 정보의 선호도를 조사하기 위하여 2019년 9월 기준으로 수도권에 거주하는 만 19세 이상 시각장애인을 대상으로 하였다. 수도권의 경우 우리나라 전체인구의 50% 이상이 수도권에 집중되어 있으므로 지방과 비교할 때 상대적으로 다양한 연령층과 다수의 장애인 시설이 분포하

므로 조사 대상 지역으로 선정하였다. 조사 대상의 선정은 선호하는 시각 정보를 장애등급(1급~3급)과 성별(남, 여), 나이(20대~70대 이상)를 고려하여 균형 있는 조사가 진행될 수 있도록 하였다.

시각장애인을 대상으로 선호하는 시각 정보의 조사는 수도권에 있는 시각장애인복지관과 시각장애인 단체, 장애인자립생활센터를 중심으로 공문을 발송한 후 협조를 얻어 면접 조사, 전화 설문, 온라인 설문(이메일 & 모바일)을 병행하여 진행하였다. 설문지는 일반 설문지와 점자 설문지를 동시에 활용하였으며, 온라인 접근이 가능한 대상자들의 참여를 높이기 위해 구글 온라인 설문조사 양식을 활용하였다.

본 연구는 설문조사의 전 과정에서 윤리적 이슈를 고려하기 위해 노력하였다. 공용기관생명윤리위원회 지침을 준수하고자 노력하였으며, 조사 진행에 앞서 연구 참여자들에게 사전에 연구 설명문과 협조 요청 공문을 통해 연구의 취지 및 목적을 소개하였다. 설문조사 당일에는 시작에 앞서 연구 참여자들에게 연구의 목적과 취지 및 활용 방법에 대하여 충분한 사전 설명을 진행한 후, 연구 참여에 대한 동의(비밀보장, 익명성, 조사 답변에 대한 거절 가능)를 구하는 서면동의를 얻은 후에 설문조사를 진행하였다.

시각 정보 선호도 조사에 참여 의사를 나타낸 시각장애인을 대상으로 2019년 9월부터 10월에 걸쳐 총 193부의 설문지를 배포하여 190부를 회수하였으며, 회수된 응답에서 중복 응답이나 무응답, 4~6급의 경증 시각장애인의 응답 등 본 조사에서 목표로 하는 대상이 아니거나 분석에 사용할 수 없는 24부의 설문지를 제외한 총 166부의 응답 자료를 분석에 활용하였다.

최종적으로 본 조사에 참여한 시각장애인의 성별 분포는 여성 응답자가 85명(51.2%)으로 남성(81명, 48.8%)의 응답 참여자와 비슷한 비율을 나타내었으며, 조사에 참여한 응답자들의 연령대 분포는 표 2에서 정리한 것과 같이 30세~59세에 해당하는 중장년층이 109명으로 전체 응답자의 65.7%를 차지하였다. 다음

으로 60세 이상에 해당하는 노년층이 43명(25.9%), 20대(20세~29세)가 14명(8.4%) 순으로 조사되었다.

표 2. 참여자 연령별 분포.  
Table 2. Age distribution of participants.

연령		인원(명)	비율(%)
청년	20-29세	14	8.4
장년	30-39세	45	27.1
중년	40-49세	33	19.9
	50-59세	31	18.7
노년	60-69세	27	16.3
	70세 이상	16	9.6
합계		166	100.0

표 3. 참여자 장애 유형.  
Table 3. Type of participants.

장애유형	인원(명)	비율(%)
전맹	89	53.6
저시력	74	44.6
시야결손	3	1.8
합계	166	100.0

본 연구를 위해 수집된 자료들은 실증적 통계분석을 위해 통계프로그램인 IBM SPSS Statistics 21을 이용하여 통계 처리를 진행하였으며, 구체적인 분석 방법은 다음의 세부 과정으로 수행하였다. 첫째, 조사응답자별 특징은 빈도 분석과 기술통계 분석을 시행하였다. 둘째, 조사 문항의 영역별 차이에 대한 비교는 독립표본 T검정(T-Test)과 일원 배치 분산분석(ANOVA)을 활용하여 집단 간 평균 비교 검증을 하였다. 셋째, 복수 응답의 경우, 다중응답 분석을 시행하여 응답의 대표성을 추출하였다. 마지막으로, 주요 항목별 차이 비교는 성별, 장애 유형, 직업 유무로 교차분석을 시행하여 분석 오류를 낮추도록 하였다.

표 4. 주요 조사 내용.  
Table 4. Questionnaire contents.

영역	조사 항목	세부항목	문항 수
A	일반 사항	성별, 연령, 최종학력, 교육환경, 장애 유형, 장애등급, 장애 발생 시기, 장애 발생 원인, 보행 훈련 기관, 직업 유무, 사회활동 참여 현황	11

영역	조사 항목	세부항목	문항 수
B	인식과 경험	- 익숙한 공간에서 혼자 이동 선호 - 타인 시선 의식 - 새로운 정보 장애인 간 공유 - 경험에 따른 정보 인지 차이 정도 - 낮선 환경의 공간 구조 어려움 - 새로운 정보의 이미지화 - 청각을 통한 정보의 이미지화 - 촉각을 통한 정보의 이미지화 - 낮선 실내에서 목적지 찾는 어려움 - 익숙한 공간에서 장애의 불편함 인지 정도 - 보행 시 흰 지팡이 사용 정도 - 낮선 넓은 공간에서 위치·방향 파악 어려움 - 물체나 공간 입체적 인지 정도	13
C	행동 특성	- 보행에 영향을 주는 요인 - 행동을 제약하는 장애물 - 익숙한 공간에서의 행동의 어려움	3
D	시각 정보	- 가장 절실한 필요 정보 - 상황 및 분야별 보고 싶은 욕구 - 선호하는 청각 유형 - 선호하는 감각 유형	4
E	보조공학 기술	- 사용하고 있는 보조공학 기술의 만족도 - 보조공학 기기의 불편함 등의 정도	2
F	보조기술의 개발 방향	- 희망하는 보조공학기기의 유형 - 희망하는 보조공학기기의 기능 - 기술개발 시 필요한 전제조건 - 꼭 얻기를 희망하는 정보 - 기술적인 도움이 있는 경우 비장애인보다 자신감 있는 부분	5
조사 문항 합계			38

### 3.2 조사 내용

시각-청각 변환기술의 전달 대상이 되는 정보의 선호도를 조사하기 위한 주요 문항은 시각장애인의 특성을 고려하여 표 4와 같이 여섯 가지 영역에서 전체 38개 문항으로 구성하였다. 먼저 설문조사 참여자에 대한 일반사항, 선호하는 시각 정보의 기반이 되는 시각장애인의 인식과 경험 그리고 시각장애인의 행동 특성을 포함하였다. 이러한 시각장애인의 행동 특성에 따라 전달받기를 원하는 시각 정보를 구체화하였으며, 시각-청각 변환기능을 제공하는 기기 및 기술에 관한 사항들을 조사 내용으로 포함하였다. 마지막으로 시각-청각 변환기기 및 기술의 개발에 대하여 사용자의 관점에서 고려해야 할 사항들을 조사하였다.

### 3.3 시각 정보의 선호도

시각-청각 변환기술을 활용하여 정보를 전달

하는 경우, 시각장애인들이 평소 가장 절실하게 필요로 하는 시각 정보를 조사한 결과는 표 4와 같다. 조사 결과 시각장애인은 앞에 있는 사람이나 사물과의 거리나 간격을 포함하는 공간 정보를 가장 절실하게 전달받기를 원하는 것으로 나타났으며, 앞에 있는 사람이나 사물의 모양 또는 형태 정보를 다음 순위로 원하였다. 표 5의 시각 정보에 대한 선호도는 남녀 성별과 장애 유형 및 직업 유무에 따른 교차분석의 결과에서도 약간의 차이는 있지만, 전반적으로 같은 양태의 선호도를 보였다.

표 5. 시각 정보의 선호도.  
Table 5. Preference for visual information.

정보의 형태	인원(명)	비율(%)
앞의 사람 또는 사물의 모양·형태	65	39.4
앞의 사람이나 사물의 거리·간격	80	48.5
앞에 있는 사람이나 물체의 색깔	12	7.3
기타	8	4.8
합계	165	100
무응답	1	-

또한, 시각장애인이 시각 정보를 가장 필요로 하는 상황에 대하여 복수의 응답을 선택하도록 조사한 결과는 표 6과 같다. 조사 결과 영상, TV, 도서, 잡지 등의 미디어 정보를 접하거나 병원에서 진료내용이나 검사 결과를 접할 때와 같은 정보접근 분야에 대한 욕구가 케이스 퍼센트 68.5%로 가장 높게 나타났다. 성별에 따라 가장 절실하게 보고 싶어 하는 상황 정보에 대하여 1순위로 응답한 내용과 교차분석을 시행한 결과, 남성은 가장 보고 싶은 욕구를 느끼는 분야로 일상생활 분야와 정보 접근 분야를 비슷하게 응답하였으나, 여성의 경우는 일상생활 분야를 압도적으로 가장 보고 싶은 분야라고 응답하였다. 추가적으로, 장애 유형에 따라 시각 정보를 필요로 하는 상황의 조사를 위하여 동일한 방법으로 1순위로 응답한 내용과 교차분석을 실시한 결과, 전맹의 경우는 이동 편의 분야를 가장 보고 싶은 분야로 응답하였으나 일상생활과 정보 접근 분야와 크게 차이가 없었다. 반면 저시력 시각장애인의 경우는 일상생활 분야를 압도적으로 가장 보고 싶은 욕구를 느끼는 분야라고 응답하였다.

표 6. 시각 정보가 필요한 상황.

Table 6. Situations where the visual information required.

구분	1순위 응답		2순위 응답		3순위 응답		케이스 퍼센트 (%)
	인원 (명)	비율 (%)	인원 (명)	비율 (%)	인원 (명)	비율 (%)	
일상생활	60	36.6	15	9.1	24	14.6	60.0
정보접근	41	25.0	50	30.3	22	13.4	<b>68.5</b>
여가생활	10	6.1	38	23.0	30	18.3	47.3
사회생활	14	8.5	25	15.2	34	20.7	44.2
이동편의	31	18.9	26	15.8	38	23.2	57.6
개인분야	8	4.9	11	6.7	16	9.8	21.2
합계	164	100.0	165	100.0	164	100.0	298.8
무응답	2		1		2		

### 3.4 정보 전달 장치 및 기능의 선호도

시각-청각 변환기술에서 정보의 전달을 위하여 시각장애인들이 주로 활용하기를 희망하는 기기의 형태는 표 7과 같이 다양한 기능이 담긴 휴대폰을 가장 선호하였으며, 손목시계 형태, 목걸이 형태 순으로 선호도가 나타났다. 특히, 시각장애인들은 휴대폰의 편리성을 위해 휴대폰 기기 하나에 다양한 기능을 구현할 수 있는 복합 기능을 선호하는 것으로 나타났다.

조사에 참여한 응답자들이 희망하는 시각-청각 변환 기기의 기능을 분석한 결과는 표 8과 같다. 보행을 지원하는 기능을 1위로 꼽아 시각장애인들이 가장 희망하는 기능으로 나타났으며, 생활에서 필요한 정보들에 대한 정보의 제공, 공간이나 구조의 파악 등의 순서로 기본적인 생활에 필요한 이동과 관련 정보의 제공을 위한 시각적 기능을 원하는 것으로 조사되었다.

표 7. 선호 장치.

Table 7. Preferred device.

구분	1순위 응답		2순위 응답		3순위 응답		케이스 퍼센트 (%)
	인원 (명)	비율 (%)	인원 (명)	비율 (%)	인원 (명)	비율 (%)	
안경	33	20.4	20	12.2	21	13.3	45.1
목걸이	20	12.3	27	16.5	35	22.2	50.0
손목시계	27	16.7	52	31.7	39	24.7	72.0
휴대폰	71	43.8	41	25.0	28	17.7	<b>85.4</b>
3D장치	9	5.6	23	14.0	32	20.3	39.0
기타	2	1.2	1	.6	3	1.9	3.7
합계	162	100.0	164	100.0	158	100.0	
무응답	4		2		8		

표 8. 희망하는 기능.

Table 8. Desired function of the device.

구분	1순위 응답		2순위 응답		3순위 응답		케이스 퍼센트 (%)
	인원 (명)	비율 (%)	인원 (명)	비율 (%)	인원 (명)	비율 (%)	
공간구조 습득	28	17.3	29	18.0	32	20.1	54.6
보행지원	64	39.5	41	25.5	20	12.6	<b>76.7</b>
빛의유무 확인	6	3.7	12	7.5	4	2.5	13.5
여러기능 복합지원	36	22.2	33	20.5	19	11.9	54.0
사전등록 정보제공	14	8.6	32	19.9	52	32.7	60.1
외부소음 차단	14	8.6	14	8.7	31	19.5	36.2
기타					1	0.6	0.6
합계	162	100.0	161	100.0	159	100.0	
무응답	4		5		7		

추가로, 시각장애인을 위한 시각-청각 변환기술 개발에 반드시 선행되어야 한다고 생각하는 전제조건에 대해 질문에는 가장 많은 응답을 보인 문항은 시각장애 유형과 정도에 따른 차별화된 접근이 필요하다고 응답하였다.

### 3.5 시각 정보 선호도 결과 분석

본 연구에 참여한 시각장애인들은 앞에 있는 사람이나 사물과의 거리 또는 간격에 대한 정보를 가장 절실하게 필요로 하는 것으로 나타났으며, 다음으로 앞에 있는 사람이나 사물의 모양 또는 형태의 정보를 필요로 하였다. 이러한 결과는 시각장애로 인해 공간을 지각하는데 어려움을 느끼게 되면서 공간 정보를 획득하는 것과 각 공간에서의 자유로운 활동을 가장 절실하게 필요로 하는 것으로 판단된다.

시각장애인들이 선호하는 시각 정보는 향후 개발되는 시각-청각 변환기술의 희망 기능과 연결되며, 조사 결과 참여자들은 보행을 지원하는 기능을 희망 기능 1위로 선택하여 선호하는 정보와 희망 기능이 일치하는 결과를 보여 주었다.

표 9. 기존 방법과 사용자의 정보 선호도 비교.  
Table 9. Comparison with conventional methods and user preference.

기술	분석 정보	시각정보 [표4]	활용상황 [표5]	제공기능 [표7]	선호장치 [표6]
vOICe	모양, 흑백	모양	일상 생활	복합 기능	안경 /휴대폰
Vibe	모양, 흑백	모양	일상 생활	복합 기능	PC
EyeMusic	모양, 컬러	모양, 색상	일상 생활	복합 기능	안경
AEV2A	형태, 흑백	모양	일상 생활	복합 기능	PC

기존 시각-청각 변환기술에 관하여 본 연구를 통해 조사한 내용인 시각 정보, 시각 정보가 필요한 상황, 희망 기능 및 선호 장치에 대응하여 비교해보면 표 9와 같다. 시각장애인들은 보행 지원을 위하여 거리나 공간적인 정보를 선호하지만, 대부분의 선행 기술들은 흑백 영상을 분석하여 영상에 존재하는 객체의 모양 정보를 전달하는 것을 주요 목적으로 하고 있다. 따라서 사용자의 요구사항을 만족시키기 위해서는 깊이(Depth) 정보를 전달할 수 있는 방법에 관한 연구가 필요할 것으로 생각된다.

시각장애인들은 일상생활과 더불어 스마트폰, PC, 책 등 정보 접근 시 시각 정보가 가장 필요하다고 느끼고 있지만, 현재 개발되는 기술들은 주로 일상생활 지원을 목적으로 하고 있다. 따라서 향후 시각-청각 변환기술을 개발할 때 스크린 정보를 효과적으로 사용자에게 전달하는 방법에 관한 연구가 필요할 것으로 생각된다.

시각장애인이 정보 전달 목적으로 사용하고 있는 기기들은 사용 목적에 따라서 각기 다른 정보를 제공해주며, 사용자가 필요로 하는 모든 정보를 단일 기기가 제공해주지 못하고 있다. 따라서 시각장애인들은 가능한 사용하는 보조 기기의 수를 줄이기를 원하며, 그 결과 다양한 기능들이 탑재될 수 있는 스마트폰을 가장 선호하는 기기로 선택하였다. 기존 기술 중 vOICe는 스마트폰, 안경 등 다양한 형태에 탑재될 수 있는 형태로 개발되었지만, Vibe, AEV2A의 경우 데스크 탑에서 개발되어 사용자 선호 장치와는 차이가 있다. 향후 개발되는

시각-청각 변환기술은 스마트폰에 탑재될 수 있는 형태로 개발되어야 활용성이 높을 것으로 판단되며, 이를 위해서는 변환 알고리즘의 경량화가 필요할 것으로 판단된다. 또한, 새로운 형태의 신호에 대하여 시험자에 따라 효율적인 학습이 가능한 플랫폼의 구축이나 학습 방안이 필요하다[13].

#### 4. 결론

감각 치환 기술은 일반적으로 손상되거나 저하된 감각의 정보를 다른 형태의 감각으로 전환하여 전달 또는 사용하는 것을 의미한다. 본 논문에서는 감각 치환 기술 중 우리나라 장애 인구의 가장 높은 비율을 차지하는 시각장애인을 대상으로 적용할 수 있는 시각-청각 변환기술의 기술적 특징들을 살펴보고 앞으로 해결해야 할 도전적 이슈들을 분석하였다. 또한, 이러한 기술들이 현재 해결해야 할 문제들의 효율적인 분석을 위한 근원적인 접근을 위하여 시각장애인들을 대상으로 실제로 전달받기를 원하는 시각 정보들을 조사하고, 장애 정도와 상황에 적합한 정보들의 선호도를 분석하였다.

본 논문에서의 분석 결과, 시각장애인들은 실생활을 위한 필수적인 시각 정보들을 선호하는 것으로 나타났다. 특히 인간 본연의 욕구인 자의적인 이동에 필요한 공간 정보와 디지털 시대를 살아감에 있어 정보 접근을 위한 시각 정보에 대한 선호도가 극명하게 나타났다. 그러나, 현재 개발된 대부분 기술은 실제 시각장애인이 원하는 공간 정보나 정보 접근을 제공하는 기술과는 차이가 있음을 확인할 수 있었다. 따라서 본 분석 결과는 감각 치환 장치들이 향후 실생활에 사용되기 위해 앞으로 극복하여야 할 여러 가지 기술적 이슈들을 해결하기 위한 출발점을 제시할 수 있다. 또한, 최근 장애인들의 사회 참여 욕구의 증대와 함께 과학 기술의 사회적 책임이라는 관점에서 향후 감각 치환 기술 분야에는 더욱 다양한 기술적 접근을 위한 기초자료로 활용될 것으로 기대된다.

---

## 감사의 글

본 연구는 한국전자통신연구원(ETRI) 연구 운영비지원사업의 일환으로 수행되었음. [20Z S1260, 인간의 감각·지각 능력을 증강하는 다중감각 융합기술 개발 사업].

## 참고 문헌

- [1] 보건복지부, “장애인현황”, 2018.
- [2] 배병덕, 홍재영, “시각중복장애 관련 연구의 동향과 과제: 최근 10년간을 중심으로”, *The Korean Journal of Visual Impairment*, 제34권, 제3호, pp. 95-121, 2018.
- [3] A. Colenbrander, “Measuring vision and vision loss”, *Duane’s clinical ophthalmology*, Vol. 5, pp. 1-39, 2001.
- [4] 문경덕 외, 감각 치환 기술 동향, 전자통신동향분석, 제34권, 제4호, pp. 65-75, 2019.
- [5] P. Bach-y-Rita, C. Collins, F. Saunders, and B. White, “Vision substitution by tactile image projection”, *Nature*, Vol. 221, No. 1, pp. 963 - 964, 1969.
- [6] C. Poirier, A. De Volder, D. Tranduy, and C. Scheiber, “Pattern recognition using a device substituting audition for vision in blindfolded sighted subjects”, *Neuropsychologia*, Vol. 45, No. 5, pp. 1108-1121, 2007.
- [7] P. B. Meijer, “An experimental system for auditory image representations”, *IEEE Transactions on Biomedical Engineering*, Vol. 39, No. 2, pp. 112-121, 1992.
- [8] A. Haigh, D. J. Brown, P. Meijer, and M. J. Proulx, “How well do you see what you hear? The acuity of visual-to-auditory sensory substitution”, *Front. Psychol.* 4:330. doi: 10.3389/fpsyg.2013.00330, 2013.
- [9] S. Abboud, S. Hanassy, S. Levy-Tzedek, S. Maidenbaum, A. Amedi, “EyeMusic:

Introducing a “visual” colorful experience for the blind using auditory sensory substitution”, *Restor Neurol Neurosci*, Vol. 32, No. 2, 2014.

- [10] L. Reich, A. Amedi “Visual parsing can be taught quickly without visual experience during critical periods”, *Scientific reports*, Vol. 5, 2015.
- [11] S. Hannelton, M. Auvray, and B. Durette, “The Vibe: a versatile vision-to-audition sensory substitution device”, *Applied Bionics and Biomechanics*, Vol. 7, No. 4, pp. 269-276, 2010.
- [12] V. Tóth, and L. Parkkonen, “Autoencoding sensory substitution”, *arXiv preprint arXiv:1907.06286*, 2019.
- [13] S. Maidenbaum, S. Abboud, and A. Amedi, “Sensory substitution: closing the gap between basic research and widespread practical visual rehabilitation”, *Neuroscience & Biobehavioral Reviews*, Vol. 41, pp. 3-15, 2014.
- [14] 이수비, 정치윤, “동적 시간 왜곡(DTW)을 이용한 오토인코더 기반 청각치환 방법의 사용자 학습 가능성 검증”, 한국멀티미디어학회 추계 학술대회, 2019.
- [15] 한국장애인고용공단, “2018년 직장 내 장애인 인식개선교육”, 강사양성교재, 2018.