Research Article

On Translation of Sonar Stimulation to Proper Visual Representation Using a Photometrically Precise Method of Audible Waves Transformation

S. M. M. Mousavi [a], and S. Mohajer Mazandarani [a & *]

a Biophotonics Laboratory, Department of Physics, Kharazmi University, Karaj, Iran

*Corresponding Author Email: mohajer@khu.ac.ir

ABSTRACT

Received: Nov. 19, 2024, Revised: Feb. 11, 2025, Accepted: Feb. 26, 2025, Available Online: Mar. 18, 2025

This research aims to assign time spectrum and color frame visual representations to acoustic data of audible range. The method of color generation is to assign visible spectrum frequency v, to the audible frequency f resulted from a transformation, defined as the main mapping function $T(f)$. The viability of $T(f)$ is ensured by satisfying four main criteria imposed by human sonar and visual perception. Afterward, a second function is proposed to represent color expressed by the output of $T(f)$ inside sRGB color space. This second function is defined as the rendering function and is obtained by the development of a matrix transformation acting on normalized human tristimulus values to yield sRGB values. Finally, the result is extended to properly translate a sound clip into color frames, changing in response to changes inside the sound clip. Such an algorithm for processing sound enables a person the deduct sounds from different sources through color vision with reasonable quality of details. The main application of such translation of sounds is to aid the Auditory Impaired. The other applications of such a method include simulating a certain type of Synesthesia and determining the cause of seismic waves by the color resulting from its visualization by this method.

I. INTRODUCTION

The plot of normal hearing threshold versus frequency, which also includes thresholds of discomfort and pain, is shown in Fig. **1**. A person with a hearing threshold higher than the blue curve specified in Fig. **1** on any audible frequency domain, is diagnosed with the condition of Hearing Loss [1], [2].

Regardless of the cause of the condition, Hearing Loss has been categorized into seven subclasses, called types, which are classified by

Fig. **1** Average person's thresholds of hearing, discomfort, and pain, are shown by blue, orange, and

red curves respectively; the volume and frequency range of some common phenomena is also shown [1], [3]

the average degree to which a person is incapable of hearing sound of low volumes over all frequencies [2]. Earlier classifications stated that Hearing Loss may be mild, moderate, severe, or profound [1], while current thresholds are presented in Table **1**. Hearing Loss can affect one or both ears and leads to difficulty in hearing conventional speech or soft sounds [2].

TABLE **1** THRESHOLD AND NOTICEABILITY OF EACH TYPE OF HEARING LOSS (CLASSIFIED BY SEVERITY); THE THRESHOLDS ARE AVERAGE HEARING THRESHOLDS OF ONLY 0.5, 1, 2, AND 4 KHZ [2], [4], [5]

Type/Category (by Severity)	Threshold Range (dB)	Noticeability (Effect on Daily Life)
Unilateral	<20 in the better ear >35 in the worst ear	Unnoticeable, Difficulty only when the sound source is close to the worst ear
Mild	20 – 34.9	Nearly Unnoticeable
Moderate	35 – 49.9	Difficulty in Quiet Conversations
Moderately Severe	50 – 64.9	Difficulty in Medium Conversations
Severe	65 – 79.9	In Need of Hearing Aid
Profound	80 – 94.9	Hearing Only Loud Indicatory Sounds
Complete/Total/Full	>95	Unable to Handle Everyday Life

The last three types of Table **1** are named Severe, Profound, and Complete Types, which together are referred to as Disabling Hearing Loss [1] or Severe-to-Profound-Hearing-Loss [6], [7] or simply STPHL [7]; and in this research, people diagnosed with STPHL will be referred to as the Auditory Impaired.

A. *Importance*

Since hearing is one of the most crucial senses of humans, as it imparts tremendous survival advantages and is of great help in indicating the happenstance of something without the need for it to be in any focus [8], therefore, as also specified in Table **1**, the Auditory Impaired are unable to handle aspects of daily life in which hearing is crucial [9]. Therefore challenges arise for the Auditory Impaired in everyday activities including crossing the streets, learning to read and communicate [10]—if the

condition of Hearing Loss existed before learning starts—responding to auditory stimulation outside of focused attention—that is responding to others when they are called upon or feeling the presence of indicatory sounds. Also, unaddressed Hearing Loss and especially STPHL can lead to reduced quality of life (QoL), isolation, dependence, lack of energy, frustration, and even depression [11], [12], [13], [14]. All these challenges may arise for the Auditory Impaired if they are not provided with proper treatment or aid (be the aid of a human, a hearing aid device, etc.) [7]. People with profound Hearing Loss, which implies the existence of very little or no hearing, often use sign language to communicate [10], which therefore implies there is no widespread or useful aid designed for them, to be able to communicate in a manner that could also be conventional for people

capable of proper hearing [1], [15]. This will be discussed in detail in I.B.

STPHL also, is gradually becoming a more common phenomenon in society [6]. Multiple studies on the prevalence and YLD[1] of Hearing Loss have confirmed that both measures are increasing. Reference [16] shows in a clinical database of 32,781 cases, 6.7% of the local clinical population[2] and 0.7% of the general population of the UK, were identified with STPHL by the year 2012; the estimations are certainly higher due to unsafe listening practices [1], [7]. Also, among those older than 60 years, over 25% are affected by STPHL [9], [17]. Reference [7], also, shows by collecting a database containing over 15 million audiograms obtained from regions covering more than 99% of the Swedish population, that 0.28% of the Swedish population are diagnosed by STPHL by the year 2022. The study also declares the global prevalence of STPHL is currently increasing; though another study [18] on a nationwide population-based database from Korean National Health Insurance Service on the South Korean population shows a decreasing percentage in the prevalence of Hearing Loss from 2006 to 2015 after the peak of 0.25 million cases—amounting to 0.5% of South Korean population—in 2010. The most recent study [17] on global Hearing Loss, which investigates the prevalence and YLD rates from the year 1990 until 2019, shows global prevalence and YLD rates are both increasing, with the total number of people identified with STPHL having a 79.1% increase from 225.3 million in 1990 to 403.3 million in 2019. The study also forecasts the number of people diagnosed with Hearing Loss to be 2.45 billion by 2050, amounting to 1 in 10 people alive in 2050 [1]. One of the main causes of this recent growth is unpracticed listening, and preference for buying higher power, and louder audio-playing devices such as party boxes and car audio systems; These cases and other causes shall be discussed in detail in another paper.

B. *Related Works*

First, it shall be stated that STPHL can be divided into purely sensorineural Hearing Loss (SNHL), where the origin of the Hearing Loss can be found in the cochlea or the vestibulocochlear nerve, and mixed Hearing Loss (MHL), which is a combination of SNHL and conductive Hearing Loss caused by damage to the outer and/or middle ear [11]. Because there are more than 300 congenital syndromes related to hearing loss [6], the differential diagnosis for hearing loss is very broad; therefore various methods are used by professionals for the treatment of Hearing Loss. We first discuss conventional methods.

1) Conventional Medical Treatments

The two major causes of conductive Hearing Loss are Otosclerosis, being abnormal bone remodeling in the middle ear [19] and Cholesteatoma, an abnormal collection of skin cells inside the middle ear, which creates benign tumors [20]. These two conditions and other causes of conductive Hearing Loss are usually treated by Surgery, which although widely available, is often hardly affordable[3] [6], [21], [22], [23]. The option is widely available for different types of purely conductive Hearing Loss and it can restore normal hearing in a satisfying number of cases [11], but still, there are limitations as to the cost and effectivity of the procedure—e.g. when the Hearing Loss is not purely result from structural damage and is also partly due to sensorineural causes [11].

For the cases dealing with SNHL, there is a type of implant called a cochlear implant which allows the reception of an 8-channel digital sound, which is hard to understand by an average person [24], and the person will need rehabilitation procedure after the surgery [14], [25], to be trained to recognize everyday sounds

---

[1] Years Lived with Disability
[2] Clinical Population of a certain medical condition refers to the population of all the people identified by all types of the condition under investigation, in this case the condition is hearing loss of all types.

[3] In some countries, the fees are paid mainly by insurance companies, but in some other countries, the terms of insurance policies imply that the fees will only be paid by the company if some specific conditions are met, an example would be the constraints on hospitals, only in those the payment will be done by the insurance company [21].

and speech. Even after rehabilitation, deduction of mood based on the tone of voice would still be far more difficult than normal hearing for cases who previously had normal hearing [24]—So much so that even a device was patented [26] to address the issue by assigning one static predetermined color to predefined moods the person who is speaking might have. In addition, cochlear implants are shown to be most effective on infants [6], [27] for they are not used to normal hearing and their brains will adapt to that limited version of hearing. These facts decrease the benefit-to-cost ratio of the surgery for adult patients due to the low quality compared to high fees [23].

2) Conservative Treatments
There also had been attempts to aid the Auditory Impaired, by the use of Assistive Listening Devices and amplification [6]. Depending on type, some devices have variable amplification ratios, meaning the amount of amplification can be adjusted using a potentiometer, other types have a fixed amplification ratio for every frequency, which is set by examinations of a medical professional for each individual's unique amplification needs [14]. However, for STPHL cases, the threshold is usually so high that amplifying sound to the threshold level would cause pain or further damage to the structure of the ear [28]. Therefore, aside from the expensiveness of some models [6], such hearing aids cannot resolve the issue [5], [11], [13], [28]. This leads us to investigate two other ways, Sign Language and Visualization.

3) Various Sign Languages
Sign language of different kinds [15], [29], although at best only enabling verbal communication between the members of the Auditory Impaired and not a replacement for the sense of hearing, is usually of great help to the Auditory Impaired. Sign Language had long been one of the crucial tools to help educate the Auditory Impaired [29], although its different structure from the verbal language has made a significant difference in the learning of people with and without hearing [10], [15]. This difference also manifests in the communication between people with and without hearing, as the sign language used between the two differs significantly from the one used between the members of the auditory-impaired [30].

Because many people with hearing do not know sign language of any kind, normally a sign language translator will crucially be needed in case of communication between people with hearing and the Auditory Impaired [31]. However, manual translation often lags behind spoken communication. The efforts on non-manual translation have also been proved to be difficult as they are made out of two distinct problems, the first is the pattern recognition of the initial language, and the second is matching the order and grammar to the destination language [10].

4) Visualizations
Methods other than the ones discussed just above have also been proposed. Initial attempts [32] had mainly focused on building an intuition on what sound is through other senses presenting an impression of sound through feeling and seeing sound vibrations. Some more advanced methods, from which the Auditory Impaired would benefit in building further intuition on what sound is and how it propagates, were also invented; one method [33] is to monochromatically[4] visualize sound pressure at each point inside a given space (commonly known as a sound field). However, this invention was mainly intended to provide information on the acoustic properties of various materials built in various geometries [34] and it can convey neither linguistic nor indicatory information that is carried by the sound which is in fact, the most important to be retrieved by the Auditory Impaired.

Other Visualizers, which also do not aim to be hearing aids, made to visualize sounds (e.g. music) for concerts [35], [36], and are widely used for entertainment and their visual output does correspond to the input sound being

---

[4] Monochromatic, in color theory and physics, is referred to light of one static color, represented only by one single electromagnetic wave of one single frequency.

played. However, their algorithms do not exhibit a clear deducible connection with the input sound and input could not be inferred by viewing the output alone. This is because the visualizations usually implement random algorithms in their designs [35], [36].

Another method, which comes with no random algorithms implemented and aims for visualization of musical sound [37], although not claiming to be a hearing aid, works by assigning an EM wave frequency to each musical note; in other words, an EM wave frequency is assigned to each key on the piano. Such a method and the device made based on such method [38], comes with three issues. The first is that the output may extend beyond the visible spectrum and could not be seen or detected by normal human vision in everyday conditions [39]. The second is that even the difference in harmonics [5] different musical instruments produce, (which is the main distinction between them if viewed inside sound space) will not be translated into the output light, resulting in loss of information through the transformation by this method. The third is that everyday sounds cannot be translated by such a method unless only the parts of input sound matching the frequency of musical notes are used from the whole sample.

The succeeding method which was explicitly intended for the Auditory impaired, was mainly focused on exchanging the information sound carries rather than giving insight. This method [26] which previously discussed briefly in I.B.1), was to analyze the fluctuation inside the input sound and to visualize the mood of the speaker, who is reading a paragraph or conversating, by assigning one static, predetermined color of predetermined amplitude to each different mood (e.g. anger, tenderness, joy, etc.).

C. *Objective*

Issues and complications discussed in I.B makes it apparent that, regarding the ways one perceives and processes visual stimulation (all that is perceived by sight e.g. chromatic vision, pattern recognition, and motion detection) and the ways one perceives sonar stimulation (all that is perceived by hearing such as tone recognition, motion detection[6] and the notion of sound texture[7], taking sonar stimulation to be a live recording or previously recorded sound clip in each moment[8]) and their primary properties, there exists an ideal translation that not only exhibits an optimal amount of computation but also maximally represents the input sound inside its output. We are in search of this ideal translation. However, such translation is not apparent to us; therefore, to pinpoint it, we shall consider the most possible basic sound under the action of such ideal translation, that is the translation of a sound clip, exhibiting only three sonar quantities; namely, the amplitude, the frequency and playing duration, all of which can be assigned to the whole sound clip. Calling the action of the unknown ideal translation on this basic sound clip, transformation of the input sound, reveals to us that, if we are to transform the input sound clip properly, we must map each of said quantities to a quantity with the same dimension. This principle comes from the logic of transformation we call consistency of the transformation. Therefore, consistency implies that sonar quantities of the input sound clip shall be mapped to the ideal visual representation's quantities as in Table **2**.

TABLE **2** THE CLOSEST MATCHING QUANTITIES OF COLOR VISION, TO THAT OF SOUND PERCEPTION

Sound Quantity	Reminiscent Light Quantity
The Frequency f of the sound clip	Visible EM wave[9] frequency ν (perceived as a monochromatic color)
Amplitude of f	Light Intensity (Light Amplitude)

---

[5] Hereafter, harmonic, is used in its musical sense, not the physical one, which is defined as a note produced on a musical instrument as an overtone.

[6] more primitive than of the motion detection experienced by sight

[7] the perceived sound texture is only the result of continuous tone detection through time, in other words, sound texture is nothing but how the amplitude of a sequence of tones are changing with respect to time.

[8] As the stimulation of a sound clip with sufficient quality, when played on a device which also shall come with sufficient playing quality, could replicate the sonar stimulation generated by the original sound, which was picked up by the microphone to generate the sound clip itself.

[9] Electro-Magnetic wave

Duration of f	Duration of the color being Displayed

As it is apparent from the mapping, the color perception of the eye is sufficient for mapping the entirety of the investigated sound clip[10]. Other aspects of vision, such as spatial pattern recognition, and three-dimensional depth perception, would not be needed.

Now, for the output to represent the input as closely as possible, we shall consider equal input and output durations, also considering the output intensity to be that of the sound clip amplitude only scaled inside the intensity levels of the screen. Now, there remains the task to properly mapping audible frequencies to the visible portion of EM waves.

II. THE METHOD

A. *Constructing The Main Mapping Function*

To find the specific transformation T, which acts on sound frequency f and outputs visible EM wave frequency v , we shall first investigate the necessary criteria, a general transformation between these two spaces shall have, in order for it to be considered a proper mapping function, which could serve as a useful aid for the auditory impaired. We will see afterwards, that, by consideration of these four crucial criteria, there remains one and only one specific transformation, with specific values for all its necessary parameters. We begin by stating every criterion in detail.

1) Inversibility

T shall act upon every sound frequency such that it would correspond to one and only one definite visible EM wave frequency, for one to be able to deduce the sound frequency played, by the color of the light resulting from the output of T; this can be stated mathematically as:

$$\forall \; T(f) = v : \exists \; T^{-1}(v) = f \qquad (1)$$

2) Conservation of Perception

Every considerable transformation between any input space and an output space other than the input space will add some amount of distortion to the perceived change inside the input space when the input space is viewed through the output space. For the viable transformation, no such distortion shall be present and only the perceived change inside the input space shall be represented inside the output space; in other words, the representation of the input space must remain unchanged during transformation. This narrative implies the perception shall remain conserved during the act of transformation. Therefore, viable is the transformation, which conserves the perception.

To obtain the form of the viable transformation for every possible input and output space, the perceived difference of the input frequency space shall be proportional to the perceived difference in the output frequency space; as for our case, we know from acoustics [40] that the ear is sensitive to the change in the frequency concerning the base frequency—a relative change that is—despite the eye which we know from photometry [41], is sensitive to the change in color regardless of the base color[11]; therefore relative change in the input frequency space, will be proportional to the absolute change in the output frequency space, mathematically stated as:

$$\forall \; T(f) = v : \frac{df}{f} \propto dv$$
$$\rightarrow T(f) = k \ln\left(1 + \frac{f - f_i}{f_0}\right) + c \qquad (2)$$

Where in Eq. (2), f_i would be the starting input frequency, f_0 is a frequency constant that by convenience is chosen to be the peak frequency of human hearing and k and c are constants

---

[10] Motion detection by the sound, which manifests as changing of amplitude through respective recording channels, can also be mapped to the motion of a color gradient inside the output display, but because the exact algorithm as to the gradient properties shall be implemented inside the source code of the computing unit, the transformation of one

channel of the input audio is discussed and the procedure will be the same for all other channels.

[11] The color is perceived as the activation of the three cone cells in photopic vision, and the change in the frequency of the visible EM wave results in a precise amount of change of the perceived color regardless of the base color.

defined by physical boundary conditions and other mathematical constraints.

3) Completeness of The Transformation

For one to recognize all possible input frequencies, all perceivable input frequencies shall be mapped onto all perceivable output frequencies; that is the domain and range of the transformation function shall be all audible frequencies and the visible spectrum respectively; mathematically stated as:

$$\forall \ T(f) = v :$$
$$20\text{Hz} \leq f \leq 20\text{kHz} \tag{3}$$
$$\wedge \ 450\text{THz} \leq T(f) \leq 710\text{THz}$$

Here shall be stated that because neither the visible spectrum nor audible range has a definite dividing line, the visibility of different radiation on the IR and UV range close to the vague ends of visible spectrum depends on the amplitude and the sensitivity of the individual [39] (same is true for the audibility of the ends of hearing range [42]), the standard, agreed upon visible ranges [43] and audible [44] to average person, in normal conditions, are used, respectively.

4) Ascending Output and Peak Mapping

The two other constraints required for convergence of the transformation in future updates are that the lowest, highest and most perceivable frequencies in the input space shall be mapped to their respective counterparts in the output space; these requirements impose two mathematical constraints on the transformation, which are stated as:

$$\forall \ T(f) = v : \frac{dT(f)}{df} > 0 \tag{4}$$

$$\forall \ T(f) = v : T(f = 3.2\text{kHz}) = 540\text{THz} \tag{5}$$

Imposing all these criteria together, the general transformation function will be reduced to one specific form with all the parameters

determined; which after some computation, simplifies to:

$$T(f) = (130.43\text{THz})x + 450.00\text{THz} \tag{6}$$

Where in Eq. (6), x is defined to be:

$$x \equiv \ln\left(1 + \frac{f - 20\text{Hz}}{3.2\text{kHz}}\right) \tag{7}$$

Equations (6) and (7) make up the entirety of the proper transformation from a sound frequency f to an EM wave with a frequency v within the visible portion of the EM radiation, which represents a light of monochromatic color, most closely representing the input audio frequency inside the color vision.

B. *Constructing The Rendering Function*

Although the transformation expressed in Eq. (6) and (7) will yield the output EM wave frequency, such output cannot be directly used to generate the color the output frequency represents; to properly show the corresponding output color inside a display, the output shall again be rendered in a color space [12] and because the visualization is going to occur inside a computing unit with an RGB display, the color space used, will be sRGB color space. Therefore, we shall proceed by derivation of the rendered version of the previous transformation in the form:

$$\mathbf{T}(f) \equiv T_R(f)\hat{\mathbf{R}} + T_G(f)\hat{\mathbf{G}} + T_B(f)\hat{\mathbf{B}} \tag{8}$$

Noting that hatted vectors are respective base color vectors of the sRGB space, and $T_R(f), T_G(f)$ and $T_B(f)$ are respective output brightness values needed to produce the monochromatic color the output EM wave frequency of the previous transformation represents when acted upon sound frequency f.

To obtain $\mathbf{T}(f)$, based on the output values of the previous transformation $T(f)$, we shall

---

[12] A color space is a vector space, in which a set of discrete or continuous colors can be represented by addition of a finite number of base color vectors of corresponding scalar coefficients, therefore, every color inside a color space is represented by a vector of base color coefficients.

assume, there exists at least one other transformation, which we will call the rendering function, and define as:

$$\boldsymbol{\sigma}(\nu) \equiv \sigma_R(\nu)\hat{\mathbf{R}} + \sigma_G(\nu)\hat{\mathbf{G}} + \sigma_B(\nu)\hat{\mathbf{B}} \qquad (9)$$

which takes as input, EM wave frequency ν of a monochromatic color, and outputs the sRGB brightness values needed to represent that same color inside an RGB display. Stating that assumption, the exact definition of $\mathbf{T}(f)$, will be:

$$\begin{aligned} \mathbf{T}(f) \equiv \boldsymbol{\sigma}(T(f)) = \ & \sigma_R(T(f))\hat{\mathbf{R}} \\ & + \sigma_G(T(f))\hat{\mathbf{G}} \\ & + \sigma_B(T(f))\hat{\mathbf{B}} \end{aligned} \qquad (10)$$

Therefore, our task was reduced to obtaining $\boldsymbol{\sigma}(\nu)$. Because such a task relates to how we shall reproduce a certain color in the eye, we will start from the human eye spectral tristimulus function.

1) Human Spectral Tristimulus Function
Humans perceive colors, trichromatically. Meaning, that every perceived color will be decomposed into three activation values for three primary contributions to the perception of three distinct color-like pigments. The reason for expressing the three receptors as perceiving three color-like pigments and not colors themselves is the fact that the data obtained from human tristimulus sensitivity functions [45], which is shown in Fig. **2**, it can be inferred that none of the visible colors inside the visible spectrum are the result of activation in only one single type of cone cell. This inference can be made from the simple observation that there is no color at which only one type of cone cell is active [39], [46][13]. The only case in which the actual pigment associated with one type of cone cells (the L-cone) could be detected by the eye was at high-intensity infrared shining to the eye at an experiment mentioned in [39]. Also, the L-cone is the only type of cone cell for which the pigment can be seen by the naked eye, as

there is nowhere on the visible spectrum that the M-cone or the S-cone are solemnly activated. Therefore, the individual color of the cone cells cannot be represented inside an external device. That is the reason the name of the cone cells is changed to L, M, and S-cones, instead of red, green, and blue cones.
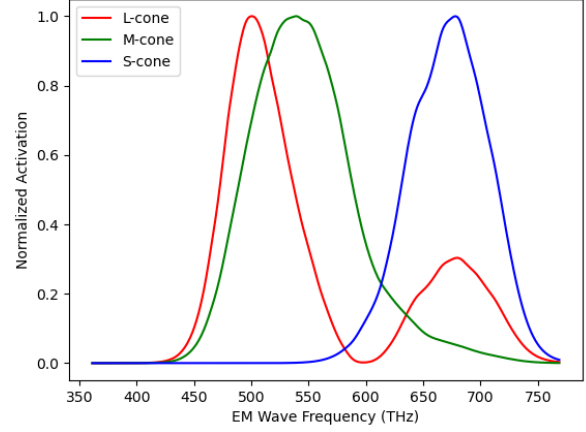

Fig. **2** Normalized Activation of Cone Cells in Photopic Vision concerning D65 White Point

As plotted in Fig. **2**, the tristimulus sensitivity functions are assumed to yield the three activations of each type of cone cell in terms of visible EM wave frequency or wavelength[14]. If we define a numerical visual stimulus function, based on the data plotted in Fig. **2**, in the form:

$$\mathbf{S}(\nu) \equiv a_L(\nu)\hat{\mathbf{L}} + a_M(\nu)\hat{\mathbf{M}} + a_S(\nu)\hat{\mathbf{S}} \qquad (11)$$

Where $a_L(\nu)$, $a_M(\nu)$ and $a_S(\nu)$ are normalized activation values of L, M, and S-cones respectively. Also, the hatted vectors are defined as base activation unit vectors, for respective cone cells[15].

2) Obtaining The Display Matrix
To obtain from the human tristimulus function $\mathbf{S}(\nu)$, the desired rendering function $\boldsymbol{\sigma}(\nu)$, we shall develop the matrix transformation by which we could transfer the vector of a certain color, from the activation space to the sRGB space; such task is done by treating the sRGB

---

[13] In order for the color seen at that imaginary point to be assumed as the color of the cone cell which is solemnly activated at that color.

[14] The original data from CIE determined the activation values with respect to wavelength; but since we have mapped sound frequencies to EM wave frequencies, we plotted the data with respect to frequency.

[15] Note that, $\mathbf{S}(\nu)$ shall not be confused with $\hat{\mathbf{s}}$.

base vectors as normalized color vectors of the three primary colors of the screen inside the tristimulus activation space, assuming v_R, v_G and v_B as their EM wave frequencies respectively. By such assumption, we have:

$$\hat{\mathbf{R}} = \frac{\mathbf{S}(v_R)}{|\mathbf{S}(v_R)|}$$
$$= \frac{a_L(v_R)\hat{\mathbf{L}} + a_M(v_R)\hat{\mathbf{M}} + a_S(v_R)\hat{\mathbf{S}}}{\sqrt{a_L^2(v_R) + a_M^2(v_R) + a_S^2(v_R)}} \quad (12)$$

$$\hat{\mathbf{G}} = \frac{\mathbf{S}(v_G)}{|\mathbf{S}(v_G)|}$$
$$= \frac{a_L(v_G)\hat{\mathbf{L}} + a_M(v_G)\hat{\mathbf{M}} + a_S(v_G)\hat{\mathbf{S}}}{\sqrt{a_L^2(v_G) + a_M^2(v_G) + a_S^2(v_G)}} \quad (13)$$

$$\hat{\mathbf{B}} = \frac{\mathbf{S}(v_B)}{|\mathbf{S}(v_B)|}$$
$$= \frac{a_L(v_B)\hat{\mathbf{L}} + a_M(v_B)\hat{\mathbf{M}} + a_S(v_B)\hat{\mathbf{S}}}{\sqrt{a_L^2(v_B) + a_M^2(v_B) + a_S^2(v_B)}} \quad (14)$$

Substituting the expressions in Eq. (12), (13) and (14), into the definition of the rendering function and rearranging the terms, we will get:

$$\frac{\begin{pmatrix} a_L(v_R) & a_L(v_G) & a_L(v_B) \\ a_M(v_R) & a_M(v_G) & a_M(v_B) \\ a_S(v_R) & a_S(v_G) & a_S(v_B) \end{pmatrix}}{\sqrt[3]{|\mathbf{S}(v_R)||\mathbf{S}(v_G)||\mathbf{S}(v_B)|}} \, \boldsymbol{\sigma}(v) = \mathbf{S}(v) \quad (15)$$

Now, defining M to be:

$$M \equiv \frac{\begin{pmatrix} a_L(v_R) & a_L(v_G) & a_L(v_B) \\ a_M(v_R) & a_M(v_G) & a_M(v_B) \\ a_S(v_R) & a_S(v_G) & a_S(v_B) \end{pmatrix}}{\sqrt[3]{|\mathbf{S}(v_R)||\mathbf{S}(v_G)||\mathbf{S}(v_B)|}} \quad (16)$$

The rendering function will be[16]:

$$\boldsymbol{\sigma}(v) \equiv M^{-1}\mathbf{S}(v) \quad (17)$$

3) Correction for Invalid RGB Values

The definition provided for $\boldsymbol{\sigma}(v)$, works for satisfying cases of color representations. However, it fails to yield proper sRGB values for a significant number of cases. To acquire proper values in all its domains, an error correction algorithm shall map every color outside the color gamut of sRGB color space, to a nearest point inside the color gamut of sRGB. This algorithm shall be implemented inside the source code of the computing unit. However, we are going to briefly introduce the algorithm. For every output of the rendering function, the sRGB values are converted to CIE xy-coordinates; then the algorithm checks whether the coordinate point the color had been transformed into, lies inside sRGB boundaries. If the point was outside the sRGB color gamut (in which case the output sRGB values cannot be displayed), as in Fig. **3**, the coordinates of the intersection point of the line drawn from the point to the white point and the closest sRGB boundary will be converted to sRGB and displayed instead of the original point.

Conversions and computations are carried out in [47].

---

[16] Some additional scaling is needed for the final rendering function to yield 8-bit RGB values in the code implementation.
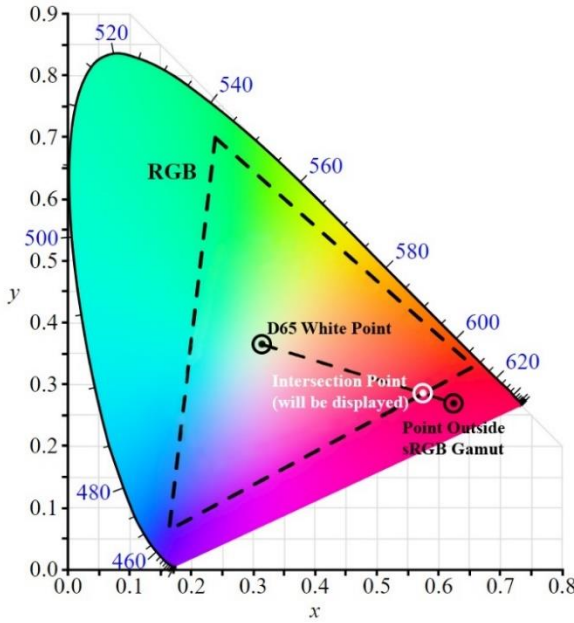
Fig. **3** Human Color Gamut and sRGB Color Gamut; The intersection point is going to be used instead of the original point when the point is detected to be outside the sRGB color gamut.

Applying this algorithm concludes the sRGB numerical rendering function, although, for better results, an iterative smoothing will also be carried. The rendering is shown in Fig. **4**.
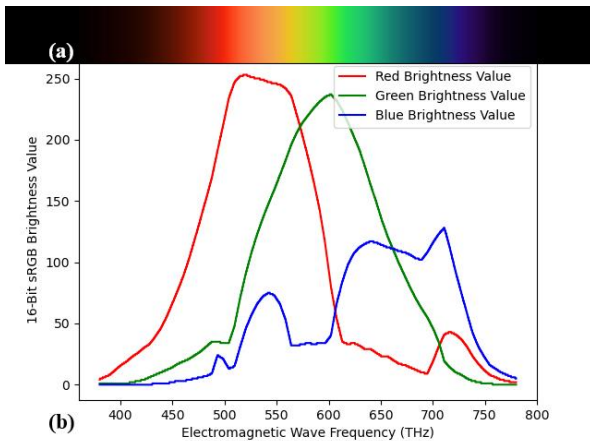

Fig. **4** The visible spectrum (a) and brightness values (b) generated by the obtained sRGB rendering function vs. EM wave Frequency for every color

C. *Applying the Transformation to an Actual Sound Clip*

So far, the method as to the transformation and correlation of sound of one single definite frequency to light of a single monochromatic color has been discussed. However, the correlation of real-world soundtracks is not as simple. First, we consider continuous sound recording and then we cover the digital sound recording procedure in a successive section.

1) Continuous (Analog) Sound Tracks

Even sounds and voices recorded from ordinary phenomena, usually consist of multiple frequencies with different amplitudes. In that case, the above-resulted transformation shall act upon each different frequency of the track as stated below:

$$\mathbf{v}_{color}(t) = \int_{f_i}^{f_f} \frac{a(f,t)\mathbf{T}(f)}{A_{max}} df \qquad (18)$$

Where $\mathbf{v}_{color}(t)$ is the result of applying the transformation on every constituent frequency inside the sound clip at every moment and is therefore the translation of the sound clip to a color frame at time t, f_i and f_f are the beginning and end of the input frequency domain (which in this case is the audible range), and A_{max} is the maximum amplitude of the recorded sound clip[17]. For the above matter, our task will be reduced to finding the constituent amplitudes $a(f,t)$ of every frequency f; the matter is achieved, by applying Fourier Transformation on the recorded sound clip, as below:

$$A(f) = \int_{-\infty}^{\infty} \bar{A}(t)e^{-2\pi ift} dt \qquad (19)$$

Where $A(f)$ is the amplitude of each frequency f, $\bar{A}(t)$ is the normalized recorded pressure difference at each time t. The limits of the integral reduce to the start and end of the sound clip, in the process of integration. While the above expression only gives the sum of all amplitudes over all times, we want to take the amplitude of every frequency f at each different time t, that is:

---

[17] For a live recording of sound, a constant reference shall be considered instead.

$$\int_{-\infty}^{\infty} \bar{A}(t')e^{-2\pi i f t'}\, dt' = \int_{-\infty}^{\infty} a(f,t')dt'$$

$$\rightarrow a(f,t) = \int_{t-\varepsilon}^{t+\varepsilon} \bar{A}(t')e^{-2\pi i f t'}\, dt' \quad (20)$$

2) Finite Sample Rate (Digital) Sound Tracks

In practice, however, the sound cannot be recorded continuously, and there only is, a finite time series of samples, with a definite sampling rate, of the recorded sound; in this case, the above equations transform into discrete sums, also, the color vector, formerly at every moment, shall now be computed at a definite frame rate, which by standard conventions, will be taken to be 24 frames per second. Now, in the case of finite sample rate sound clips, Eq. (20) becomes:

$$a_i = \sum_{j=1}^{N} A_j e^{-2\pi i f_i t_j} \quad (21)$$

Where in Eq. (21), N is defined to be:

$$N \equiv \frac{R_{\text{Sound Sampling}}}{R_{\text{Frame}}} \quad (22)$$

Where the numerator is the sound sampling rate (in Hertz), and the denominator is the frame rate (in fps, which is equivalent to Hertz). Also, because of the discreteness possessed by the current case, the frequency spectrum in Eq. (18) and (19) will reduce to a set of allowed discrete frequencies, with their value being:

$$f_i \equiv i \frac{R_{\text{Sound Sampling}}}{N} = i R_{\text{Frame}} \quad (23)$$

$$i = 1, 2, 3, \cdots, 833$$

The reason for the upper boundary is defined for i is to restrict the computations only to audible frequencies. And Therefore, the integral for obtaining the color frame at each moment, in Eq. (18) changes to a summation and we will have:

$$\mathbf{v}_{\text{frame}} = \frac{\sum_{i=1}^{833} a_i \mathbf{T}(f_i)}{A_{\max}} \quad (24)$$

Where here $\mathbf{v}_{\text{frame}}$ is the color vector of one frame of the generated output[18].

III. CONCLUSION

Using the cascaded rendered transformation function which is in the form of Eq. (8), and applying the Fast Fourier Transform (FFT) Algorithm [48], [49] successively on each one over twenty forth of a second of the sound clip to find a_i defined in Eq. (21), and then substituting the obtained values in Eq. (24), we can generate a proper, consistent translation of any recorded or real-time sound clip to a visual output; Because of the consistency of the procedure, sound texture, which is the change inside the amplitude of a sequence of tones, is also translated as a pattern inside the output time spectrum as shown in Fig. **5**.
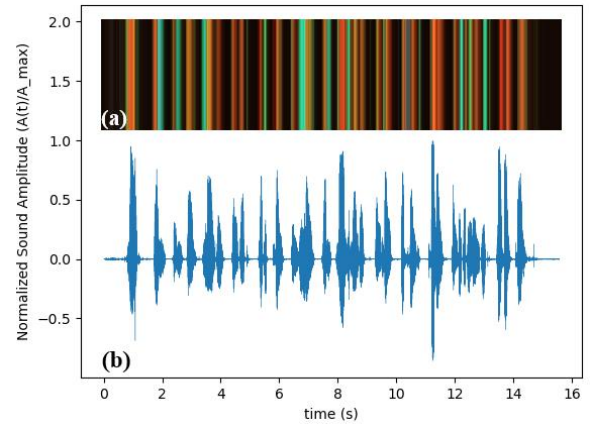

Fig. **5** Time spectrum visual translation (a) of the voice of the first author (b) stating: "Hi, this is a sample input which we can see is translated into different color constituents over time". As apparent, different tones and vocals have distinct colors

Because the formulated translation practically enables perceiving sound through the sense of sight, not only does it provide the Auditory Impaired, the ability to infer and learn vocal structure, and grammar, but also enables them

---

[18] Some additional scaling will be needed in the source code of the computing unit in order for the result of the summation to fit inside sRGB values range.

to detect the presence and the details of other sounds in general (the sound of cars approaching, things dropping as in Fig. **6**, and other indicatory sounds) being the only method to enable this level of sound deduction.
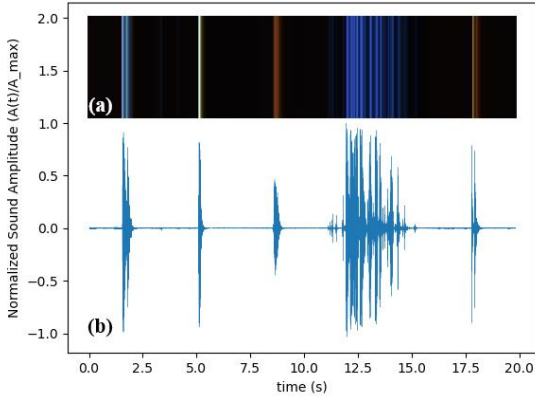

Fig. **6** Time spectrum visual representation (a) of some sounds (b); in order of occurrence sound from dropping of keys, falling of a book, saying 'hey' to call someone, shaking of keys, and putting a glass on a small plate; microphone was fixed at normal height of the ear of a person when sitting, colors are notably different for each, enabling deductions possible

Because the presented algorithm, is a way to translate and visualize any set of mechanical waves of any frequency range and duration, the other important application for such an algorithm is visualizing seismic waves inside the visible spectrum; by applying the discussed translation on a dataset recorded from a seismographer, the cause of every seismic movement can be inferred as each cause generates a distinct set of seismic waves which correspond to a distinct color representation for each different cause of seismic vibrations. This algorithm also presents a proper, consistent, and convenient way to visualize the data obtained from seismographers which are placed in various locations. The data can then be visualized on the map of the area and can then be used to process the propagation of earthquakes or other geological phenomena.

The other application of such a device is that it can reproduce a certain type of condition known as Synesthesia, the type in which the person having the condition, associates colors to sound. Although the association different people experience may differ in the exact

sound-to-color correspondence—which may not even be consistent enough to be considered a correspondence —the presented algorithm can reproduce the same feeling and also can be modified to replicate the association that a certain person reports as having experienced.

Overall, not only can the translation witnessed in Fig. **5** and Fig. **6** aid the Auditory Impaired to understand sounds and vocal languages, or graph any acoustic data of any range by considering the frequency range of data under question, but it is also hoped to open up a window of new possibilities for exploration and research on the physics and the science of human perception, and respective translation between different senses in a precise, quantitative manner.

APPENDIX

A. *Obtaining The Parameters of The Logarithmic Transformations*

As discussed, the frequency transformation function was found by the statement of the **second mapping criterion** to be in the specific form:

$$y(f) = k \ln(f - f_0) + c \qquad (25)$$

Our general problem will be to obtain the values of the 3 parameters k, f_0 and c, in terms of the coordinates of three points:

$$(f_1, y_1), (f_2, y_2), (f_3, y_3)$$

But constructing the three equations with three parameters of the function as the unknowns will result in a non-linear system of equations therefore making the direct solving of the equations an elaborate mathematical problem, even in the numerical solution cases. So, here is explained instead, another method to tackle the problem, which will result in a far more efficient numerical calculation and also yields a direct solution to two of the three parameters and a numerical, but arbitrarily exact solution to the third.

First, we consider the inverse of the function in the form:

$$Y = as^x + b \qquad (26)$$

Which we will call from now on, 'the inverse function'. The parameters of this function are related to the original parameters in the following way:

$$Y \equiv f$$
$$x \equiv y$$
$$a = e^{\frac{-c}{k}} \rightarrow c = -\log_s(a) \qquad (27)$$
$$b = f_0$$
$$s = e^{\frac{1}{k}} \rightarrow k = \frac{1}{\ln(s)}$$

with the coordinate of the points initially considered, getting flipped to account for the inversion of the function and the new points considered will be:

$$(x_1 \equiv y_1, Y_1 \equiv f_1)$$
$$(x_2 \equiv y_2, Y_2 \equiv f_2) \qquad (28)$$
$$(x_3 \equiv y_3, Y_3 \equiv f_3)$$

The three points chosen such that:

$$x_1 < x_3 < x_2$$
$$y_1 < y_3 < y_2$$

Now, if we define,

$$X = s^x \qquad (29)$$

as the new input, our inverse function in the Eq. (26) becomes:

$$Y = aX + b \qquad (30)$$

Which is the definition of a line, the parameters of which can be pinpointed to exact values by considering two of the three points of Eq. (28); so for a and b we will have:

$$\begin{cases} a = \dfrac{\Delta Y_{21}}{\Delta X_{21}} \\ b = \dfrac{X_2 Y_1 - X_1 Y_2}{\Delta X_{21}} \end{cases} \qquad (31)$$

Where ΔA_{ij} is defined to be[19]:

$$\Delta A_{ij} \equiv A_i - A_j \qquad (32)$$

Where A can be any of the above variables (e.g. x, X, Y, etc.). Because of the change of variables considered in Eq. (29), a and b will both depend on the parameters of Eq. (26); however, both parameters are tuned by Eq. (31) in such a way that:

$$\begin{cases} as^{x_1} + b = Y_1 \\ as^{x_2} + b = Y_2 \end{cases} \qquad (33)$$
$$\forall s \in \mathrm{R}^+ - \{0,1\}$$

As mentioned in Eq. (33), both parameters are set in such a way that the inverse function will contain the first and the second points of (28) for all positive values of s except for one and zero; this means we can tune the value of s in order for the inverse function to also contain the third point without losing the two other points.

For the inverse function to contain the third point (x_3, Y_3), the distance of the output of the function on $X_3 = s^{x_3}$ from the y-coordinate of the point shall be zero in other words:

$$| Y_3 - as^{x_3} + b | = 0 \qquad (34)$$

But as said before, considering the variable change of Eq. (29) for X, a, and b are themselves in terms of s; by substituting respective change of variables inside Eq. (31) and substituting a and b inside (34) from (31) we will have:

---

[19] This will be the conventional notation wherever we are faced with incremental differences of values with various indices; which often happens in case of the computations of the parameters of interpolation functions.

$$Y_3 - \frac{\Delta Y_{21} s^{x_3} + s^{x_2} Y_1 - s^{x_1} Y_2}{s^{x_2} - s^{x_1}} = 0 \qquad (35)$$

Which after an amount of simplification yields:

$$s^{x_3} \Delta Y_{21} - s^{x_2} \Delta Y_{31} + s^{x_1} \Delta Y_{32} = 0 \qquad (36)$$

Which is a polynomial equation of degree x_2 on[20] s. Equation (36) is of course a problem solvable only by numerical methods as the values of powers of s can be any real number. Attempting to solve Eq. (36) by Newton's Method will yield three answers, two of which (zero and one) are the apparent solutions to the equation and also obsolete in the sense that they are forbidden base values of exponentials, therefore they are outside the range of possible values of s.

To remove the obsolete solution $s = 0$ from Eq. (36), we can divide both sides by s^{x_1} to obtain:

$$s^{\Delta x_{31}} \Delta Y_{21} - s^{\Delta x_{21}} \Delta Y_{31} + \Delta Y_{32} = 0 \qquad (37)$$

Which will yield a polynomial of degree Δx_{21}. Aside from removing zero out of solutions, such polynomial has the advantage that its powers are usually far smaller than the one in (36); this fact and the fact that this polynomial is a better-behaving curve in the range the solutions exist, will mean reaching the desired solution would need far less iterations, resulting in Eq. (37) being a far better and more efficient choice for numeric computations [21]. As is apparent from Eq. (36) and (37), the problem of solving a system of equations was reduced to one root-finding problem for only one of the parameters and the precision of all three parameters can be controlled to match any arbitrary amount by the number of iterations considered.

By obtaining the value of s, a and b will also be found by substituting the obtained value of s inside Equations (31) and (29); and finally the

parameters of Eq. (25) can also be found using the relations of (27).

As is apparent from our narrative, the parameter s was the parameter that tuned the curvature of our function, to the amount needed for the function to account for the third point in between the two initial points; and as we know from relations of (27):

$$s^x = e^{\frac{x}{k}} \qquad (38)$$

and by taking the limit of s to one—when a and b are evaluated from (31)—we will see that:

$$\lim_{s \to 1} s^x = \lim_{k \to \infty} e^{\frac{x}{k}} \approx 1 + \frac{x}{k} \qquad (39)$$

and the Eq. (26), when a and b are substituted from (31) becomes:

$$Y = \frac{\Delta Y_{21}}{1 + \frac{x_2}{k} - (1 + \frac{x_1}{k})} (1 + \frac{x}{k})$$
$$+ \frac{(1 + \frac{x_2}{k}) Y_1 - (1 + \frac{x_1}{k}) Y_2}{(1 + \frac{x_2}{k}) - (1 + \frac{x_1}{k})} \qquad (40)$$

Which after some amount of simplification becomes:

$$y = \frac{\Delta Y_{21}}{\Delta x_{21}} x + \frac{x_2 Y_1 - x_1 Y_2}{\Delta x_{21}} \qquad (41)$$

Which is the definition of a line passing through the first and second points defined in (28). This case will correspond to the absence of curvature in the inverse function; therefore, we can infer, that the parameter s indicates how *curved* is the inverse function and also our main transformation function. By further examination of Eq. (38), we can see that $s = 1$, corresponds to no curvature and $s = e$ will

---

[20] As is apparent from the constraints applied on the points after the definitions of new points in (28);

[21] It shall be noted that this will only be true when x_1 and x_2 are of the same sign, otherwise (36) is the better choice; in our case (37) is preferred.

result in a curvature amount of one exponential unit; therefore we define:

$$C \equiv \ln(s) = \frac{1}{k} \qquad (42)$$

As the transformation's 'Curvature Constant'. This constant is the exact indicator of how much exactly an exponential or a logarithm is curved away from the equivalent line passing through the endpoints of the interval under investigation; and for the case of transformation from average hearing to average sight, Curvature Constant will be:

$$C_{HS} \approx \frac{1}{132.640964097 \text{THz}}$$
$$\approx 7.53914906158 \text{fs} \qquad (43)$$

As it is apparent, the Curvature of the transformation from Hearing to Sight is very small, but the tiniest change inside its value will vastly change output color mapping[22]; for this also, the achievement of its precise value is crucial to the validity of the transformation. The Curvature Constant declared in Eq. (43) is a fundamental constant of human perception[23] and will be of great importance to the current and future work in the field of sense translation.

ACKNOWLEDGMENT

REFERENCES

[1] World Health Organization, "Deafness and Hearing Loss," Accessed: May 02, 2024. https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss

[2] B. O. Olusanya, A. C. Davis, and H. J. Hoffman, "Hearing loss grades and the international classification of functioning, disability and health," World Health Organization, vol. 97, pp 725–728, 2019.

[3] Informed Health, "Can noise damage your hearing?" Institute for Quality and Efficiency in Health Care. Accessed: Feb. 07, 2025. https://www.informedhealth.org/can-noise-damage-your-hearing.html

[4] G. Stevens, S. Flaxman, E. Brunskill, M. Mascarenhas, C. D. Mathers, and M. Finucane, "Global and regional hearing impairment prevalence: an analysis of 42 studies in 29 countries," Acad. Stevens, S Flaxman, E Brunskill, M Mascarenhas, CD Mathers, M FinucaneThe Eur. J. Public Heal. 2013•academic.oup.com, vol. 23, pp. 146–152, 2013.

[5] E. Holmes and T. D. Griffiths, "'Normal' hearing thresholds and fundamental auditory grouping processes predict difficulties with speech-in-noise perception," Sci. Rep. vol. 9, pp. 16771, 2019.

[6] S. Anastasiadou and Y. Al Khalili, "Hearing Loss," StatPearls Publishing, vol. 22, pp. 1–7, 2023. https://www.ncbi.nlm.nih.gov/books/NBK542323/

[7] C. Löfvenberg *et al.*, "Prevalence of severe-to-Profound hearing loss in the adult Swedish population and comparison with cochlear implantation rate," *Acta Otolaryngol.*, vol. 142, no. 5, pp. 410–414, 2022.

[8] W. E. Brownell and B. R. Alford, "HOW THE EAR WORKS - NATURE'S SOLUTIONS FOR LISTENING," Volta Rev. vol. 99, pp. 9, 1997, 2025. [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC2888317/

---

[22] To express briefly, the reason as to the constant exhibiting dimension, is that it is referring to a delay in perception of the ear relative to the eye, a detailed investigation for the validity of this claim is needed, and will be done.

[23] The exact value of the constant depends on the exact values of parameters of Hearing and Sight of the individual; however, the value declared in Eq. (43) was found by taking the conventional range of audible acoustic waves and visible spectrum which was declared in the main body text.

[9] Y. C. Tseng, S. H. Y. Liu, B. S. Gau, T. C. Liu, N. T. Chang, and M. F. Lou, "Lived experiences and illness perceptions of older adults with age-related hearing loss before the use of hearing aids: An interpretative phenomenological study," Geriatr. Nurs. (Minneap). vol. 61, pp. 231–239, 2025.

[10] N. C. Camgoz, S. Hadfield, O. Koller, H. Ney, and R. Bowden, "Neural Sign Language Translation," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. pp. 7784–7793, 2018.

[11] C. Löfvenberg, S. Turunen-Taheri, P. I. Carlsson, and Å. Skagerstrand, "Rehabilitation of Severe-to-Profound Hearing Loss in Adults in Sweden," Audiol. Res. 2022, Vol. 12, Pages 433-444, vol. 12, pp. 433–444, 2022.

[12] A. Ringdahl and A. Grimby, "Severe-profound hearing impairment and health-related quality of life among post-lingual deafened Swedish adults," Scand. Audiol. vol. 29, pp. 266–275, 2000.

[13] M. A. Ferguson, P. T. Kitterick, L. Y. Chong, M. Edmondson-Jones, F. Barker, and D. J. Hoare, "Hearing aids for mild to moderate hearing loss in adults," Cochrane Database Syst. Rev. vol. 2017, pp. CD012023, 2017.

[14] L. Turton et al., "Guidelines for best practice in the audiological management of adults with severe and profound hearing loss," Seminars in Hearing, vol. 41, pp. 141–245, 2020.

[15] W. C. Stokoe, "Sign Language Structure," Annu. Rev. Anthropol., vol. 9, no. 1, pp. 365–390, 1980.

[16] L. Turton and P. Smith, "Prevalence & characteristics of severe and profound hearing loss in adults in a UK National Health Service clinic," Int. J. Audiol. vol. 52, no. 2, pp. 92–97, 2013.

[17] W. Li, Z. Zhao, Z. Lu, W. Ruan, M. Yang, and D. Wang, "The prevalence and global burden of hearing loss in 204 countries and territories, 1990–2019," Environ. Sci. Pollut. Res. vol. 29, pp. 12009–12016, 2022.

[18] G. J. Im et al., "Prevalence of severe-profound hearing loss in South Korea: a nationwide population-based study to analyse a 10-year trend (2006–2015)," Sci. Reports 2018 81, vol. 8, pp. 1–9, Jul. 2018.

[19] National Institute on Deafness and Other Communication Disorders, "What Is Otosclerosis? Symptoms & Diagnosis." Accessed:, 2025. [Online]. Available: https://www.nidcd.nih.gov/health/otosclerosis

[20] National Health Service UK, "Cholesteatoma." Accessed: Feb. 07, 2025. [Online]. Available: https://www.nhs.uk/conditions/cholesteatoma/

[21] Tehran Ear Clinic, "The Cost of Cochlear Implants, Do the insurance companies cover the fees?" Accessed: 2024. [Online]. Available: https://tehranearclinic.com/blog/cochlear-implant-cost

[22] TriHealth, "Surgical Treatment for Hearing Loss." Accessed: 2024. [Online]. Available: https://www.trihealth.com/services/ear-nose-and-throat/ent-treatments-and-services/surgical-treatment-for-hearing-loss

[23] Forbes Health and Duke Department of Head and Neck Surgery & Communication Sciences, "How Much Do Cochlear Implants Cost?" Accessed: 2024. [Online]. Available: https://headnecksurgery.duke.edu/news/forbes-health-how-much-do-cochlear-implants-cost

[24] Cochlear Implant Brain and Behavior Lab, "What does the world sound like through a cochlear implant?" Accessed: May 04, 2024. [Online]. Available: https://cochlearimplant.lab.uconn.edu/cochlear-implant-information/sounds/

[25] A. Ciorba et al. "Rehabilitation of Severe to Profound Sensorineural Hearing Loss in Adults: Audiological Outcomes," Ear, Nose Throat J., vol. 100, pp. 215S-219S, Jun. 2021.

[26] C. Lee, "Sound to light converter and its method," KR20050089440A, 2004 Accessed: 2024. [Online]. Available: https://patents.google.com/patent/KR20050089440A/en?oq=KR20050089440

[27] F. G. Zeng, S. Rebscher, W. Harrison, X. Sun, and H. Feng, "Cochlear Implants: System Design, Integration, and Evaluation," IEEE Rev. Biomed. Eng. vol. 1, pp. 115–142, 2008.

[28] E. W. Johnson, "Hearing Aids And Otosclerosis," Otolaryngol. Clin. North Am., vol. 26, pp. 491–502, 1993.

[29] C. Valli, C. L. Washington, and D. C. Gallaudet, Linguistics of American Sign Language: An Introduction (3rd Ed.), vol. 1. 2024. [Online]. Available: https://books.google.com/books?hl=en&lr=&id=mfS3GlTLAUMC&oi=fnd&pg=PP13&dq=sign+language&ots=QuNlND-

bBu&sig=mtjMWqhiz-UcPQYy87QvUwcW1fs#v=onepage&q=sign language&f=false

[30] E. Domagała-Zyśk and A. Podlewska, "Strategies of oral communication of deaf and hard-of-hearing (D/HH) non-native English users," *Eur. J. Spec. Needs Educ.*, vol. 34, no. 2, pp. 156–171, 2019.

[31] O. Koller, J. Forster, and H. Ney, "Continuous sign language recognition: Towards large vocabulary statistical recognition systems handling multiple signers," Comput. Vis. Image Underst. vol. 141, pp. 108–125, 2015.

[32] W. C. Hayes, "Sound to light visual vocalization system," US3572919A, 1968 Accessed:, 2024. [Online]. Available: https://patents.google.com/patent/US3572919A/en?oq=US3572919+(A)

[33] M. Kurihara and J. Fujimori, "Sound to light converter and sound field visualizing system," US20120097012A1, 2012 Accessed: May 19, 2024.

[34] H. Lim, M. Imran, and J. Y. Jeon, "A new approach for acoustic visualization using directional impulse response in room acoustics," Build. Environ. vol. 98, pp. 150–157, 2016,

[35] A. Blok, "Sound-to-light graphics system," WO1994022128A1, 1994 Accessed: May 19, 2024. [Online]. Available: https://patents.google.com/patent/WO1994022128A1/en?oq=WO9422128

[36] L. Ok-kyung, "Speaker generating a light corresponding to a sound source," KR20050062950A, 2003 Accessed: 2024. [Online]. Available: https://patents.google.com/patent/KR20050062950A/en?oq=KR20050062950

[37] K. Gil-ho, "Method for transforming sound to color and a light emitting speaker employing the sound to color transformation function," KR20080021201A, 2006 Accessed: 2024. [Online]. Available: https://patents.google.com/patent/KR20080021201A/en?oq=KR20080021201+(A)

[38] K. Gil-ho, "Apparatus for emiting light using led according to sound pitch, sound volume or tone color," KR20050034772A, 2003 Accessed: 2024. [Online]. Available: https://patents.google.com/patent/KR20050034772A/en?oq=KR20050034772+(A)

[39] D. H. Sliney, "What is light? The visible spectrum and beyond," Eye, vol. 30, pp. 222, 2016.

[40] Bruno A. Olshausen, "logs-and-music (Psych 129 - Sensory Processes)." Accessed: 2024. [Online]. Available: http://www.rctn.org/bruno/psc129/handouts/logs-and-music/logs-and-music.html

[41] F. Viénot, D. MacLeod, and J. D. Mollon, "CIE 170-2:2015 Fundamental Chromaticity Diagram with Physiological Axes – Part 2: Spectral Luminous Efficiency Functions and Chromaticity Diagrams," Color Res. Appl. vol. 41, pp. 216–216, 2016.

[42] D. Purves, G. J. Augustine, D. Fitzpatrick, and et al., Neuroscience - The Audible Spectrum. Sinauer Associates, 2001. Accessed: May 24, 2024. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK10924/

[43] T. J. Bruno and P. D. N. Svoronos, "CRC handbook of fundamental spectroscopic correlation charts," CRC Handb. Fundam. Spectrosc. Correl. Charts, pp. 1–226, 2005.

[44] Amplifon, "What is the range of human hearing? | Amplifon UK." Accessed: May 24, 2024. [Online]. Available: https://www.amplifon.com/uk/audiology-magazine/human-hearing-range

[45] International Commission on Illumination, "CIE cone-fundamental-based spectral tristimulus values for 2 degree field size | CIE." Accessed: 2024. [Online]. Available: https://cie.co.at/datatable/cie-cone-fundamental-based-spectral-tristimulus-values-2-degree-field-size

[46] Y. Ohno et al. "The Basis of Physical Photometry, 3rd Edition, CIE 018:2019," 2019.

[47] M. Anderson, R. Motta, S. Chandrasekar, and M. Stokes, "Proposal for a standard default color space for the Internet - sRGB," in Final Program and Proceedings - IS and T/SID Color Imaging Conference, 1996, pp. 238–246. Accessed: 2024. [Online]. Available: https://www.color.org/sRGB.xalter

[48] J. W. Cooley and J. W. Tukey, "An Algorithm for the Machine Calculation of Complex Fourier Series," Math. Comput. vol. 19, pp. 297, 1965.

[49] L. I. Bluestein, "A Linear Filtering Approach to the Computation of Discrete Fourier Transform," IEEE Trans. Audio Electroacoust. vol. 18, pp. 451–455, 1970.