I.J. Image, Graphics and Signal Processing, 2021, 6, 30-44
Published Online December 2021 in MECS (http://www.mecs-press.org/)
DOI: 10.5815/ijigsp.2021.06.04

Sound Source Localization Ability in Hearing Aids: A Survey

Jyoti M. Katagi
Biluru Gurubasava Mahaswamiji Institute of Technology/Electronics & Communication Engineering, Mudhol-587313, Karnataka, India
Email: katagi.jyoti@gmail.com

Pandurangarao N. Kulkarni
Basaveshwar Engineering College (Autonomous)/Electronics & Communication Engineering, Bagalkot-587102, Karnataka, India
Email: pnk_bewoor@yahoo.com

Received: 14 August 2021; Revised: 01 September 2021; Accepted: 19 September 2021; Published: 08 December 2021

Abstract: Ability to locate sound source in human acoustic system is a prime factor. The source of sound has various spectral, temporal and strength characteristics depending on where it is located. To identify the sound location, the listeners analyze these characteristics arising from various directions on the horizontal and the vertical surfaces. In noisy background, it is very difficult to understand the speech for individuals with sensorineural hearing loss. In order to reliably distinguish various sound sources and increase speech intelligibility in noisy conditions, binaural hearing is adopted. Diffraction induced by the pinnae, head, shoulders and torso changes the pressure waveform when sound waves travel from the audio source to the listener's eardrum. Two transfer functions that specify the relation between the sound pressures at the listener's right and left ear drums will catch these propagation effects. These spectral changes are recorded by Head Related Transfer Functions (HRTFs). Different hearing aid algorithms are to be studied to measure their effectiveness in improving speech perception through series of subjective evaluations involving subjects with sensorineural hearing loss with different types of loss characteristics under different listening conditions. We investigated the various proposed approaches, weighed in on their benefits and drawbacks and most importantly, examined whether and how the resulting HRTFs perceptual validity is evaluated. This paper brings out current research efforts on sound source localization ability in hearing aids, which includes use of Head Related Transfer Functions (HRTFs) for generating spatial sounds in elevation and azimuth plane, evaluating the effect of monaural and binaural hearing aid algorithms on source localization under different listening conditions on subjects with different hearing losses and also to assess the effectiveness of localization with type of hearing aids.

Index Terms: Sound source localization, Head related transfer function, Hearing aids, Inter-aural Level Difference, Inter-aural Time Difference.

1. Introduction

We would be able to hear sounds coming from all directions and distances. Individual sounds can be classified based on their volume, pitch, tone, and spatial position. The spatial position of a sound is determined by its three-dimensional state. Localization is approach to interpret the sound source location in three dimensions [1,2]. By locating the sound source, we will have a more relaxed and natural listening experience. In order to improve communication in noisy listening environments, the listener requires some additional visual signals so that the subject can turn towards the sound source and localize it for safety purposes, such as avoiding incoming traffic, approaching cyclists on a running path, or a dropping object. Hearing loss may lead to a loss of localization ability; however, amplification may not always be effective in restoring localization to normal levels. Hearing aids with wireless communication and digital signal processing methods have significantly helped in the preservation of localization cues.

A linear and time invariant acoustic filtering system can be used to filter sound signals from a sound source position to the ears drums. The head related transfer function (HRTF) represents the transformation of incident sound signals by biological structures (such as head, torso and pinnae) in the device transfer function [3]. HRTF varies greatly from subject to subject due to various anatomical characteristics of individuals. Virtual sound reproduction is a common application of HRTF, in which a virtual sound source at a specific spatial location can be created by filtering a mono stimulus with paired HRTFs (for left and right ears) and then recreating through headphones. Many studies indicate that listeners' individual HRTFs are required for accurate virtual sound reproduction; otherwise, localization errors such as degraded elevation localization output and an increase in front-back and up-down confusions can occur. As a result, customizing the listener's HRTFs improves the auditory efficiency of virtual sound reproduction [4,5,6,7,8,9,10,11].

There is an apprehension that the binaural hearing aids may degrade the localization cues thus indirectly affecting the source localization. In the light of this apprehension we have investigate the effectiveness of different Monaural and Binaural hearing aid algorithms in improving speech perception using noise reduction technics and their impact on source localization and we have also studied different types of hearing aids which are suitable for sound source localization. So we wanted to achieve optimized hearing aid algorithm with minimal adverse effect on source localization.

The remaining part of the paper is structured in the following manner. Section II introduces about sound localization system & different auditory localization cues. Section III presents head related transfer function (HRTF) approach, CIPIC database and anthropometric features. Section IV discusses about the effects on localization from different types of hearing aids. The survey is summarized in final section V.

2. Sound Localization System

Humans have the ability to hear and see three dimensionally in space [12]. By listening to a sound, they may calculate the exact sound source location and also the distance from the sound source to the listener. The ability of a listener to identify the source or locate the sound in terms of distance and direction is known as sound localization. Hearing systems measure interaural time differences (ITDs) to determine the horizontal direction of a signal, while hearing strength and the head shadow effect cause interaural level differences (ILDs) [13,14]. ITDs are primarily used for low frequency localization i.e frequencies lower than 800 Hz. For frequencies above 1.5 kHz, ILDs are the primary horizontal localization cues because head shadow effects increase with increasing frequency.

The popularity of 3D sound systems has increased in recent years. For researchers, accurate 3D sound generation is an interesting subject. Virtual reality, gaming, online coverage of activities, music concerts, Augmented reality, human computer interface, entertainment and pilot displays are all built on three dimensional view of audio and video. We need to incorporate three-dimensional sound effects to make these devices more practical for the users. 3D sound systems create sound fields that tend to come from a specific azimuth and elevation in space. These sound systems use auditory cues that specify the sound source position to achieve sound localization in 3D space. Special functions known as Head Related Transfer Functions are used to encode these cues. HRTFs are spectral filters for filtering the sound from the source to the listener’s eardrum. HRTFs are becoming a common way of recreating binaural sounds as a result of major advancements in signal processing hardware [15].

HRTFs are the fundamental functions that enable virtual three-dimensional (3D) auditory systems to be created. Here, we build a Virtual Acoustic Space (VAS) in which sounds heard through headphones tend to come from a specific direction in 3D space. The listener is given the impression of being surrounded by a virtual sound source. Virtual Acoustics is a subfield of the rapidly growing field of ‘Virtual Reality' (VR) [16].

Rendering is the process of creating cues for the various senses (3D picture, 3D audio, etc.). Diffraction, which is one of the most complicated phenomena in general linear acoustics, is often used to transmit sound. Localization and detection of a sound source is a major challenge when creating a practical acoustic device.

In order to locate binaural sound a probabilistic model is proposed in [17]. In this study, cochlear model is used to generate cochleagrams that are used as input to the system to extract binaural cues. This model measures ITD and ILD separately for a variety of frequencies and then processes the data as a whole. This produces activity charts, which are binaural cues in terms of two-dimensional frequency versus time-delay representation. Based on these activity charts, a probabilistic assessment is performed to determine the sound source location over time.

Adaptive blind channel identification for sound source localization is discussed in [18]. This paper assesses the performance of a direction of arrival (DOA) approach for binaural hearing aids that makes use of the difference in time and the level difference of impulses, calculated from the Adaptive Eigen value Decomposition Algorithm (AEDA) or Adaptive Principal Component Algorithm (APCA). Both algorithms are tested for various reverberation times, SNR, and source locations. In comparison with AEDA, APCA is relatively insensitive to noise, but it can tolerate only reasonable reverberation.

A robust model for acoustic localization is presented in [19]. This paper gives the solution for degrading reverberation effect and the existence of various sources on localization performance by using Gaussian mixture models (GMMs) to robustly analyze complex acoustic scenes.

Maximum likelihood (ML) approach to informed sound source localization is used to estimate the Direction of Arrival (DoA) of the target signal in [20]. In comparison to other informed sound source localization (SSL) techniques which adopt binaural microphones for locating the signal, MLSSL uses signals of one or even more microphones mounted on only one ear, thereby reducing the wireless overhead transmission of binaural listening aids. MLSSL discloses an average absolute DoA estimated failure of 5°, when the target position is restricted towards the front-horizontal surface, in a large-scale noise situation with no reverberation.

Time delay estimation is used to localize acoustic events in a noisy environment is discussed in [21]. In this study, three delay estimation approaches are investigated: normalized cross correlation, Least Mean Squares (LMS) adaptive filters, and cross power-spectrum process. Different tests were conducted in different acoustic sources in order to compare these three techniques. The CPS-based technique outperforms LMS by 20% and NCC-based techniques by 40% when a reasonable tolerance of 5° is used. However, accuracy of localization was not satisfied, especially when adopting NCC and LMS-based approaches. From these tests it is concluded that cross power spectrum phase based technique is the most stable both in clean and in non-critical noisy situations than the other two techniques in acoustic event localization.

The efficiency assessment by listeners with medium to extreme sensory loss of hearing for various compression factors to enhance speech perception is discussed in [22]. By adopting multiband frequency compression, the impact of spectral masking can be decreased. In this approach, the range of speech is divided into several analytical bands. The constant compressor factor compresses spectral samples in each of these strips towards the middle of the band. The efficiency of the strategy for various compression factors is examined in order to enhance speech perception. The modified rhyme test revealed that the most significant increase in recognition ratings resulted from a compression factor of 0.6: approximately 17% for normal hearing persons with simulated hearing loss and 6-21% for persons with mild to serious sensorineural hearing loss.

The performance of statistical methods and NMF approaches for the voice enhancement challenge is considered in [23]. Speech and noise spectral coefficients are assumed to be super Gaussian, resulting in enhanced PESQ, Short Time Objective Intelligibility (STOI), and SNR values compared to standard speech enhancement techniques. When it comes to non-stationary signals like speech, a template-based Non-Negative Matrix Factorization (NMF) technique delivers better outcomes than standard speech augmentation approaches. It is possible to considerably enhance the performance measurements of PESQ, SNR and STOI by integrating the advantages of both statistical and NMF methods.

Reducing background music while using frequency domain speech enhancement techniques is presented in [24]. Due to the high variation and poor estimation of noise and speech spectra, musical noise arises in each speech frame. This study utilizes a method based on wavelet thresholding the multitaper spectrum coupled with a noise estimating approach to decrease the musical noise. Soft and hard thresholding with Daubauchies wavelet were employed in this article to test the suggested technique in terms of objective quality metrics under eight distinct real-world disturbances at three distortions in input SNR. Measures like Segmental SNR, Weighted Spectral Slope Distance, Log Likelihood Ratio, Perceptual Evaluation of Speech Quality (PESQ), and composites are compared to wavelet de-noising techniques to predict speech quality in the presence of noise. Spectral Subtraction and Multiband Spectrum Subtraction provide consistent performance for all eight noises in the majority of the cases.

Localization cues are produced by the interaction of the listener's anatomy with the acoustic wave field. As a result of this interaction, wave reflections, diffraction and scattering occur, and the spectral quality of the acoustic signal is altered in a direction-dependent manner, resulting in a range of cues that allow the listener to interpret the position of the sound in 3-Dimension. These cues are discussed below.

A. Binaural Cues

By comparing acoustic input from two separate detectors we can do the binaural localization. When the acoustic signal arrives at both the ears, the two signals are compared to estimate the sound source localization, this is called binaural localization. Two acoustic cues: Inter-aural Level Difference (ILD) and Inter-aural Time Difference (ITD) are used to measure the direction of sound source arrival at both ears. These are the strongest directional cues. When the sound waves arises from a source it reaches the near ear first and then the far ear. The biological binaural cue is given by this split-second delay. This is called as Inter-aural Time Difference (ITD). For human beings maximum ITD is 0.66ms [25].

Whenever the sound source is placed in the horizontal surface, its angle with regard to the head is referred to as its azimuth angle. If the sound source is directly to the right of the listener then it indicate 90° azimuth angle, directly in front indicates 0°, and directly behind the listener indicates 180°. If any sound signal is arising exactly from right side of the subject i.e. at 90° azimuth angle, it will take less time to reach the right ear as compared to the left ear. This shows how long it took for the sound to hit both ears. This time advance or delay is calculated and used to find the sound source location.

The magnitude of a sound signal is high when it reaches the near ear compared to the sound signal that reached the far ear; it is called as the "Inter-aural Level Difference" (ILD). The shadowing effect of the head (above 1.5 kHz) and the difference in wave propagation time (below 1.5 kHz) cause these variations.

Cone of confusion is created when the azimuth angle is kept constant while changing the elevation angle. ILD and ITD remain constant when a sound source shifts its position inside the cone of confusion. The Head Related Transfer Function is a special acoustic process that is used to resolve the cone of confusion.

The main cause of ITD is separation between the two ears. When the source is placed exactly to the right or left of the subject (i.e. at +90° or -90°) the maximum ITD value is about 0.66ms. The shadowing effect of the head causes Inter-aural Level Difference (ILD). The minimum ILD that can be detected is 0.5dB [26].

Binaural three dimensional localization of voice sources using the HRTF's composite function vector is presented in [27]. This article exploits the interaural magnitude and phase characteristics of the HRTF. This article introduces a new feature vector that blends these two characteristics non-linearly and a method for the extraction of this feature vector free of speech spectral distortion. The proposed procedure is tested and examined with a correlation-based HRTF method and a two-step localization technology for different source locations, HRTF (individuals) and voice inputs. For average signal-to-noise ratios, the results indicate that up to a 20% increase in localization efficiency can be achieved.

The effectiveness of HRTF-based SSL methods in terms of spectral similarity among HRTFs and microphone array geometry is discussed in [28,29]. These papers show that the performance of the SSL algorithms based on HRTF is dependent on the DoA of the output signal because of similarity among HRTFs. The studies have shown that the geometry of the microphone array is significant even though the increase in number of microphones in the microphone array enhances SSL efficiency.

Depending on the Time Difference of Arrival (TDoA) of the output sound at two microphones located on the ears of the Hearing Aid System (HAS) user, one on each ear, an informed DoA estimator is presented in [30]. The maximum likelihood method is used to calculate the TDoA and the DoA, which is based on the noise-free target sound and calculation of the ambient noise features. The likelihood function can be accurately estimated using inverse discrete-Fourier-transform techniques for $M = 1$ and $M = 2$ microphones in this paper. In this paper, the proposed algorithm's efficiency in large crowd noise situations is evaluated for different DoAs, Signal to Noise Ratios (SNRs) and lengths.

Smartphone aided Technology for enhancing Sound Localization in Hearing Aid Devices is introduced in [31]. This paper is used to correctly track a single voice source in very low SNR situations with the Non- Uniform Non-linear Microphone (NUNLA) Array. To understand the impact of the approach on real-time implementation, it is checked for various SNRs and data lengths for speech sources. The proposed NUNLA-based algorithm outperforms standard uniform linear arrays for a 360° scan (ULA). This would improve overall SSL efficiency for hearing aid devices (HAD) in noisy situations, giving users more comfort.

Normal and hearing-impaired listeners' speech intelligibility and sound localization abilities with wireless coordination of multi-channel Wide Dynamic Range Compression is discussed in [32]. For 20 participants, the subjective assessment was conducted with eight normal hearing participants and twelve had bilaterally symmetrical sensorineural loss of hearing. Every participant completed the sound localization test and Hearing in Noise Test (HINT) with two kinds of stimuli. Localization results were checked for Front/Back and Left/Right dimensional errors. When the sound source was broadband, enabling wireless synchronization decreased the rate of Front/Back confusions by 10.5% in the hearing disabled population. The wireless condition had no effect on the Left/Right dimension localization performance. For the HINT, there was no clear gain from wireless WDRC synchronization; however, hearing disabled listeners had improved localization with wireless synchronization.

Spectral splitting of the input signal is carried out by using filter bank summation process for binaural dichotic analysis, also dynamic range compression and noise reducing method such as wiener filter is proposed in [33]. Subjective and objective assessments focused on Mean Opinion Score (MOS) and Perceptual Evaluation of Speech Quality (PESQ) results are utilized to determine the perceived efficiency of speech in various SNR situations. The outcomes of the hearing assessments (using MRT) gave the highest increment of (27.299%, 23.95%, 24.503%, 23.602% and 23.498%) in speech identification results at SNR values of (-6dB, -3dB, 0dB, +3dB, +6dB) as examined with unprocessed speech identification results. Response times were lower with lower SNR values than unprocessed speech response times. Response time was decreased by 1.581, 1.41, 1.329, 1.279 and 1.01s respectively at SNR values of -6, -3, 0, +3 and +6 dB. This indicates improved speech intelligibility at lower SNR values.

B. Monaural Cues

External filters such as the shoulders, ears, outer ear or "pinna," and torso aid in monaural sound localization. Sounds are frequency filtered based on the angle from which they impact the different external parts. The notch filtering effect is caused by destructive interference of waves reflected from the outer ear, and this effect of pinna notch serves as the most effective filtering cue for sound localization. The frequency at which the sound is critically notch filtered is determined by the angle from which it enters the outer ear. The monaural cues are good at detecting elevation information (using notch effect) while the binaural cues are efficient in detecting azimuth information (using ITD) [34].

Diffraction of sound waves by the human shoulders, torso, outer ears (pinna), and head will change the range of sound that reaches the eardrums. HRTFs are responsible for encoding these changes. The HRTF varies not only with elevation, azimuth, frequency, and range in a complex way, but also from person to person.

The spectral cues present in the HRTF were extracted using cepstral analysis method [35]. The fine spectral format of the voice and the spectral envelop of the HRTF can be easily distinguished using cepstral analysis. This study shows how binaural cues and HRTF data are truncated in the cepstral field, and it shows that only a few cepstral elements preserve essential information for localization. It is also shown that the changing spectral features of speech can easily be normalized in order to decrease their impact on spectral cue extraction. Finally, the method proposed is compared to the localization method based on convolution. The higher uncertainty with regard to convolution is because of its preference for the higher energy area of the signal spectrum, which is amplified by reducing SNR. Finally, the higher efficiency of the stated approach at low SNR suggests that the spectral positioning of binaural locations in the median plane is more effective.

On 2D sound localization efficiency, the impact of monaural spectral pinna hints and the Head Shadow Effect (HSE) is studied in [36]. For various conditions, a subjective assessment of 9 listeners with chronic unilateral listening problem (18–50 years old) is carried out. The localization test results from arbitrarily interleaved sound levels reveal that every monaural deaf audience rely heavily on HSE, while binaural audiences ignore it. Regardless of the level of sound, certain monaural listeners partially reacted to actual azimuth sound source. These listeners used their pinna cues to get azimuth details. The more monaural listeners can identify the azimuth with spectrum signals, the better they can find the elevation of the sound source. In a successive localization study with one single sound level, monaural listeners quickly followed an approach based on the HSE. These findings suggest that monaural spectral indicators are insufficient for accurate two dimensional sound localization in unfamiliar acoustic environments. As a result, monaural listeners place a high value on the ambiguous HSE, which can help them in dealing with known acoustic conditions.

The head-related transfer function (HRTF) was used to create two sound localization algorithms in [37]. To find the location of a sound source, they use the ITD, ILD, and monaural spectral cues. Since most localization methods would have to operate in the presence of background noise, one of the methods' localization output was evaluated at signal-to-noise ratios (SNRs) ranging from 40 to -40 dB. Ten real-world broadband sounds were used as stimuli, with 5° intervals in azimuth and 0° elevation. The algorithm was used to localize sounds to a 5° accuracy using both two and four microphones. At 20 dB SNR and higher value, the two-microphone edition of the algorithm committed mean localization error of less than 2°, while the four-microphone edition committed mean error of approximately 1° at 10 dB SNR or above.

The ability of humans to detect sound elevation and find whether a sound is originating from the front or back is heavily reliant on the pinnae's monaural spectral features [38]. The pinna observances were collected from the median HRIRs of 45 independent HRIRs in the CIPIC HRTF database, to get the near enough acoustic depiction of HRTF personalization. These are later structured as linear amalgamation of 4 or 5 simple temporal features for every elevation on the median plane by PCA (principal components analysis) in the time domain. Replacement of KEMAR HRIR pinna response at given height is done by customizing the results obtained for calculations done on strength of each basis function at that height. Out performance of KEMAR HRIRs were obtained by developing vertical impacts with minimized front/back doubtfulness in the median plane when four people with normal hearing ability were capable of developing a sequence of HRIRs in response to filtered stimuli over headphones.

A cascaded device of noise control and multiband compression improves the experience of speech for people who use monaural hearing aids for sensorineural hearing loss is presented in [39]. The input signal has a multiband compression scheme dependent on auditory critical bandwidths (ACB) and is split into 18 frequency bands varying from 0 to 5 KHz. All spectral energy is concentrated at the middle of each band to mitigate the spectral masking effect. The background noise is reduced using a noise reduction scheme called the wiener filter. Assessment procedure includes i) the evaluation of voice quality (using PESQ and MOS) on normal listening people ii) A hearing impaired persons intelligibility review (using MRT). 50 sets of consonant-vowel-consonant (CVC) words make up the test material. There are six words in each set. For 1800 presentations each subject is answered (300 CVC words x 6 various SNR values). The obtained MOS, PESQ and MRT results indicate improved quality of speech and intelligibility for SNR values from -6 dB to + 6 dB. Table 1. gives the comparison of different sound localization systems. Different methods are used to get the better sound source localization performance.

Table 1. Comparison of Sound Localization Systems

Sl.No.	Paper Title	Advantage	Remarks
1	Maximum likelihood approach to "informed" sound source localization for hearing aid applications[20]	In comparison to other informed sound source localization (SSL) techniques which adopt binaural microphones for locating the signal, MLSSL uses signals of one or even more microphones mounted on only one ear, thereby reducing the wireless overhead transmission of binaural listening aids. MLSSL discloses an average absolute DoA estimated failure of 5°, when the target position is restricted towards the front-horizontal surface, in a large-scale noise situation with no reverberation.	SSL's resilience to reverberation is a critical concern.
2	Effects of Active and Passive Hearing Protection Devices (HPDs) on Sound Source Localization, Speech Recognition & Tone Detection[40]	A total of four distinct HPDs were examined, including two that had never been tested before. All investigated HPDs considerably reduced performance compared to unoccluded performance; however one active HPD enhanced high-frequency tone detection thresholds and did not affect speech recognition.	According to the data, high-frequency spectral signals, which are crucial for avoiding front-back confusions, are considerably distorted in people with HPDs.

3	Informed TDoA-based Direction of Arrival Estimation for Hearing Aid Applications[30]	The maximum likelihood method is used to calculate the TDoA and the DoA, which is based on the noise-free target sound and calculation of the ambient noise features. The likelihood function can be efficiently estimated using inverse discrete-Fourier-transform techniques for M = 1 and M = 2 microphones in this paper. In this paper, the proposed algorithm's efficiency in large crowd noise situations is evaluated for different DoAs, Signal to Noise Ratios (SNRs) and lengths.	Future research will examine the effects of reverberation and more realistic acoustic settings.
4	On the influence of microphone array geometry on HRTF-Based sound source localization[29]	This paper shows that the performance of the SSL algorithms based on HRTF is dependent on the DoA of the target signal because of similarity among HRTFs. The studies have shown that the geometry of the microphone array is significant even though the increase in number of microphones in the microphone array enhances SSL efficiency.	This research solely looked at horizontal target placements and Behind the Ear (BTE) hearing aids. In addition to azimuth, future study should incorporate elevation and range. The research will also be enriched if other types of hearing aids are considered in addition to the BTE.
5	On the preprocessing & postprocessing of HRTF individualization based on sparse representation of anthropometric features[41]	CIPIC HRTF database experimental findings suggest that the preprocessing methods have a greater impact on HRTF individualization performance, with higher variance in preprocessing approaches. In sparse presentation, adding nonnegative restrictions increases efficiency. Using standard score in anthropometry normalization, log magnitude spectra of HRTFs, and non-negative sparse representation with normalized weights, the best results are achieved. However, this approach has a spectral distortion of 5.86 dB, which is significantly lower (8.11 dB) than the closest HRTF set method, as well as reasonably close to the theoretical bottom bound (5.12 dB) of such linear regression based HRTF individualization methods.	The HRTF individualization procedures will be subjected to subjective review in the future.
6	Binaural localization of speech sources in 3-D using a composite feature vector of the HRTF[27]	This article exploits the interaural magnitude and phase characteristics of the HRTF. This article introduces a new feature vector that blends these two characteristics non-linearly and a method for the extraction of this feature vector free of speech spectral distortion. The proposed procedure is evaluated and compared with a correlation-based HRTF approach and a two-step localization technology for different source locations, HRTF (individuals) and voice inputs. For moderate signal-to-noise ratios, the results indicate that up to a 20% increase in localization efficiency can be achieved.	In the future, this feature vector idea will be extended to multi-source localization.
7	Concurrent Localization of Sound Sources and Dual-Microphone Sub-Arrays Using TOFs[28]	As a result of hearing aid applications, this study proposes a localization technique for dual microphone subarrays. As part of the localization process, hearing aid microphones estimate time-of-flights (TOFs) of the target signals that originate from unknown sound sources. There was no far-field assumption in this article because it was treated as a linear system of equations.	Future studies will incorporate more realistic scenarios that take into account head presence and reverberation.
8	Human Voice Localization in Noisy Environment by SRP-PHAT and MFCC[42]	In comparison with steered response power (SRP), the Mel-frequency cepstral coefficient (MFCC) approach is more effective in localizing human speech. Using a vowel corpus comparison, the suggested technique will identify the speaker's voice and enhance it. By utilizing MFCC feature extraction, human voice may be retrieved from a noisy environment using SRP-PHAT.	Only vowel sounds are taken into account in the suggested technique, thus a consonant-specific expansion might increase the system's accuracy.
9	Window-Dominant Signal Subspace Methods for Multiple Short-Term Speech Source Localization[43]	Besides robustness, the suggested techniques can also count the number of voice sources. Because of their low latency, the suggested techniques are useful for real-time voice source localization. To enhance voice source localization in unfavorable settings, the generalized sparsity assumption can be used.	Frequency adjacent bins can also be regarded as part of the generalized assumption, in addition to time adjacent bins.
10	Binaural localization of speech sources in the median plane using cepstral HRTF extraction[35]	The fine spectral structure of the speech and the spectral envelop of the HRTF can be easily distinguished using cepstral analysis. This study shows how binaural signals and HRTF data are truncated in the cepstral domain, and it shows that only a few cepstral components retain essential information for localization. It is also shown that the variable spectral characteristics of speech can easily be normalized in order to reduce their impact on spectral cue extraction. Finally, the method proposed is compared to the localization method based on convolution. The higher uncertainty with regard to convolution is because of its preference for the higher energy area of the signal spectrum, which is amplified by decreasing SNR. Overall, the higher accuracy of the proposed method at low SNR suggests that the spectral positioning of binaural locations in the median plane is more effective.	Future research will look at expanding the notion of cepstral HRTF extraction to include multiple source localization, as well as azimuth and elevation localization.

11	Sound localization with and without hearing aids[44]	Researchers have studied how hearing challenged people locate sounds in three different planes: the frontal horizontal plane, the horizontal plane, and the elevation. The processing of binaural cues dominates the first dimension, whereas the processing of monaural spectral cues dominates the remaining two dimensions.	In this paper, it is shown that hearing aids might have a negative impact on localization abilities. Using certain hearing aids resulted in lower left-right accuracy and reduced capacity to resolve front-back confusions, according to the study.

3. Head Related Transfer Functions

HRTF can generate 3D digital sound. Depending on the sound source location, the audio wave may undergo many spectral changes when it moves from a sound source to the ear drum and these spectral changes are recorded by HRTFs. The head and pinna filter sounds in a spatial and frequency-dependent manner, resulting in acoustical clues to the direction of the sound source. The spectral and temporal transformations are calculated by the head and pinna's physical dimensions. HRTF is an auditory transfer function which is described as the sound pressure ratio at the listener's eardrum to that at the center of the head when the listener is absent. HRTF's time domain representation is a Head-related Impulse Response (HRIR). HRTF varies depending on frequency, physiological structure, and sound direction. HRTFs may be greatly affected by slight differences in anthropometric size and shape. The source location based pre-filtering effects of these external structures help in the determination of sound source location. The elevation, azimuth, and distance between the sound source and the listener all affect the spectral changes.

In order to quantify the sound pressure of such random audio sources on an ear drum, the pulse response h(t), from the sound source to the ear drum is required. This is referred to as the Head Related Impulse Response (HRIR). When the left and right ear HRTF is available, it is easy to precisely synthesize the binaural sound from a monaural sound source.

HRTF is calculated using 4 variables: frequency and 3 spatial coordinates. When a sound source is more than one meter away in a spherical coordinate system, it is said to be in the far region, and the HRTF value varies inversely with distance. Since HRTFs are usually measured in the far region, they are only a function of elevation, azimuth, and frequency.

HRTF magnitudes differ as a function of the following parameters, according to research findings [45]:

- Frequency: The amplitude of the HRTF varies with frequency. HRTFs have a lot of spectral diversity.
- Source Position: The asymmetric shape of the pinnae contains details about source direction. HRTFs differ vertically and horizontally depending on the source location.
- Subject: Pinnae and cavum conchae are unique to each person.

HRTFs must fit the localization cues that each person needs to extract from real-world sounds because physical elements such as the head, torso, and outer ears are the defining factors. As a result, each HRTF pair should be specified separately for each potential listener. Since HRTF is unique to each subject and changes as the relative source position changes, a wide collection of HRTFs for various elevations and azimuths is needed for realistic implementation of 3D audio.

Some sound specialization systems currently use empirically determined HRTFs for each user. These customized HRTFs, on the other hand, are anthropometrically accurate for each user. The development of 3D acoustic systems for an individual necessitates a huge number of measured HRTFs as well as specialized equipment, facilities, and expertise. It takes a long time to implement and is challenging. This limits their use to high- end sound specialization systems with very unique applications. Generic HRTFs, which are recorded from a KEMAR mannequin as a subject, are one of the most popular alternatives. For this KEMAR mannequin, the average values of anthropometric data are used to represent a group of potential listeners. Asymmetric pinnae structures for different individuals are shown in Fig.1.

Virtual sound sources were simulated with HRTFs at five separate horizontal distances [46]. Every simulated position has an electromagnetic sensor to determine the sound source direction and distance. In a free field situation, listeners were able to accurately localize the sound.

Different methods to obtain the HRTFs for navigation purpose were compared in [47]. HRTF synthesis, HRTF personalization and direct measurements for navigation applications were compared in this paper using basic listening tests. Since only four anthropometric variables (width of torso, width of head, diameter of concha, and depth of concha) must be collected from the user, localization using synthesized HRTFs is more accurate than other methods. Furthermore, the approach is very basic, and HRTF can be incorporated in any direction. The personalization approach necessitates a larger database of high-resolution HRTFs and more anthropometric parameters. The measurement approach is extremely in terms of equipment, time, and post-processing.

Fig. 1. Pinna structures for different persons

Mumble is an open source VoIP conferencing programme that incorporates a HRTF-based 3D sound convolution engine [48]. The listening exercises were done in an office environment. When more than one speaker is present, the speaker's intelligibility improves.

The HRTFs are measured for two dummy heads, KEMAR and BHead210 which represents the average human head and Chinese person head respectively [49]. Then the localization performance for both dummy heads is observed. In general, BHead210 outperforms KEMAR for Chinese listeners, with the exception of locations [45,0], [180,0] and [180,-30]. This is most likely due to the following factors: (1) The torso has a significant impact on the sound localization from lower planes; BHead210 does not have a torso unlike KEMAR. (2) Sound localization is influenced by a number of subtle factors, particularly the size and pattern of the ear pinna. For Chinese subjects BHead210 performs better than KEMAR, particularly in frontal plane.

OpenAL-Soft is a three dimensional audio application interface which uses simple HRTF dataset. To improve the audio quality of the game an improved HRTF dataset which uses asymmetric HRTFs is proposed in [50]. The asymmetric HRTFs on the left and right are sampled in complete space. The spectra of the binaural signals made by OpenAL-Soft API are more precise.

The selection and tuning of HRTFs is done individual for every wanted direction in [51]. Even the left and right sides of a HRTF pair can be tuned. The resulting transfer functions showed a significant reduction in coloration and an increase in global localization, but only small changes in frontal location. This selection and tuning process can be completed by a single tuning expert, and the results of the listening test are verified. As a result, in a device with HRTFs and headphone replication, the perception is less personalized than one would imagine. Sparse representation to acquire personalized magnitude response of HRTFs is presented in [52]. A sparse depiction of the new listener's anthropometric characteristics is found by using the sparse representation of the CIPIC database anthropometric features. The method proposed has an average 5.53dBs spectral distortion value which is extremely low compared to other approaches in [53,54].

Using an individual HRTF is typically impractical since the measurement needs an anechoic chamber, specialized technology to locate the source on a circular grid, and tracking devices to guarantee that the subject remains stationary. Therefore, other methods have been developed to acquire "individual" HRTFs without the need to measure them vocally. For those who don't have access to a personalized collection of HRTFs, a typical option is to find an equivalent in a database. Selection might be based on a variety of factors such as location clues, subjective appraisal or anthropomorphic similarity, among others. A localization experiment is widely used to evaluate HRTF selection methods. As a result, it is usually considered that the ideal HRTF minimizes localization error. It is therefore possible to determine the HRTF that would minimize angular errors or confusion rates. Due to the large variation and multimodal distribution of answers, such approaches generally require specialized hardware, take a lot of time in order to complete, and the results are sometimes difficult to understand. For this reason, subjective techniques have been developed, which are aimed at reducing HRTF assessments' setup and testing durations.

A. CIPIC Database

CIPIC database is a public domain database which contains a set of high spatial resolution HRTFs. HRTFs for 45 people at 50 different elevations and 25 different azimuth angles at relatively 5° angular intervals were determined at the University of California Davis CIPIC Interface Laboratory. This will offer us 1250 distinct points in space around each subject's head. A collection of anthropometric measurements is also available in the CIPIC database, which can be used in scaling studies [55].

There were 45 subjects in all, 43 of whom were humans and two of whom were KEMAR mannequins of various sizes and shapes. Out of 43 human subjects 16 were women and 27 were men. The HRTFs for each ear of the subject were measured by placing the subject at the center of a 1m radius hoop whose axis was aligned with the subject’s inter-aural axis.

Bose Acoustimass loudspeakers were located at different points along the hoop. Special Etymotic research ER7C probe microphones were utilized to collect Golay code signals produced by a modified Snapshot device from Crystal River Engineering. To create a raw HRIR, the microphone outputs were sampled at 44.1 kHz with 16 bit resolution and processed with Snapshot's Oneshot feature. The calculated raw HRIR was windowed with a modified Hanning window to eliminate room reflections, and the outcomes were free field compensated to account for the transducers' spectral characteristics. Each HRIR has a length of 200 samples and is about 4.5 ms.

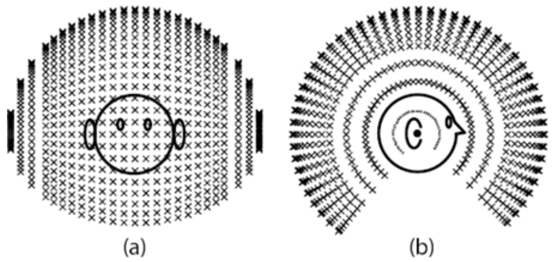

Fig. 2. Sampled points locations in space [55]

In inter-aural polar co-ordinates, the elevation angle ϕ and azimuth angle θ specify the direction of the sound source. Elevation angles are sampled uniformly in 5.625° increments from -45° to +230.625°. At -80°, -65°, -55° from -45° to 45° in stages of 5°, 55°, 65° and 80°, azimuth angles are sampled. In three-dimensional space, this yields 1250 distinct points, as shown in Fig. 2.

The CIPIC HRTF data base uses an inter-aural polar co-ordinate scheme with the subject's head at the center, as shown in Fig. 3. The inter-aural axis is a line that runs across the middle of both the right and left ears. The inter-aural mid-point, which is the exact center of the line connecting the two ears, is the source of the spherical coordinate system. Azimuth is described as the distance between the mid sagittal plane or the vertical median plane and a vector to the sound source, and it is represented by θ. From -90° to +90°, the value of azimuth angle changes. The distance between the horizontal surface and the source's projection into the mid sagittal plane is known as the elevation angle. The symbol for it is ϕ. The elevation angle can be anywhere between -90° and +270°.

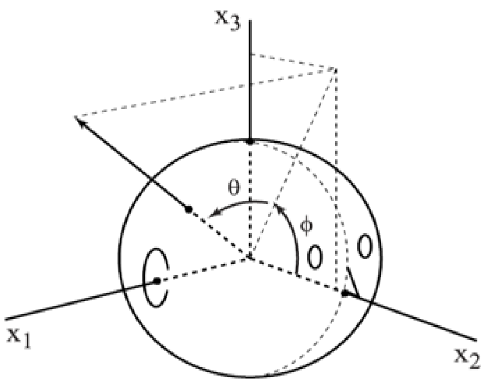

Fig. 3. Inter aural polar coordinate system [56]

A head related impulse response 'h' is a function of elevation, azimuth and time. In the data files HRIR is sampled both spatially and temporally. Azimuth, elevation and time are described by distinct variables naz, nel and nt respectively. A HRIR h (naz, nel, nt) is a three-dimensional array with dimensions of (25*50*200).

Table 2. indicates how azimuth and elevation angle values in inter aural polar coordinate system vary for different points.

Table 2. Direction in 3D Space for different azimuth and elevation angles [56]

Azimuth	Elevation	Direction in three dimensional space
0°	0°	Ahead
0°	90°	Overhead
0°	180°	Behind
0°	270°	Below
90°	0°	To the right
-90°	0°	To the left

B. Anthropometric Features

The CIPIC database contains a collection of anthropometric assessments that are used in scaling analyses. Because even minor differences can cause major variations in the HRTF, particularly for the pinnae measurements, a sufficient collection of well specified and appropriate anthropometric [57] measurements was not specified in a general way.

The database had 27 anthropometric measurements, 10 for the pinna and 17 for the head and torso. The subject's sex, age, and weight are several additional anthropometric parameters.

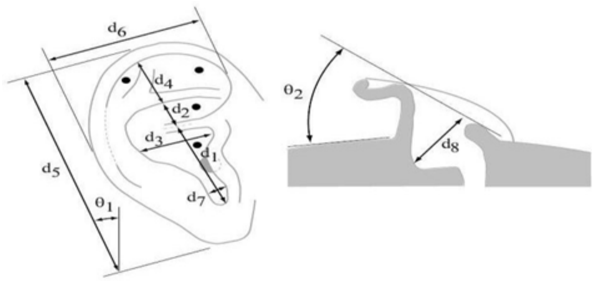

Fig. 4. Pinnae measurements [56]

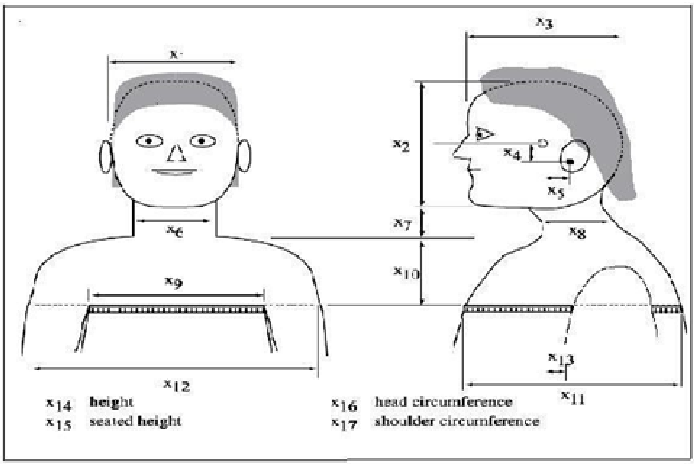

Fig. 5. Head and torso measurements [56]

Table 3. Anthropometric parameters [56]

Var	Measurement
X_1	Head width
X_2	Head height
X_3	Head depth
X_4	Offset down of pinna
X_5	Offset back of pinna

X_6	Width of neck
X_7	Height of neck
X_8	Depth of neck
X_9	Top width of torso
X_{10}	Top height of torso
X_{11}	Top depth of torso
X_{12}	Width of shoulder
X_{13}	Offset forward of head
X_{14}	Height
X_{15}	Seated height
X_{16}	Circumference of head
X_{17}	Circumference of shoulder
d_1	Height of cavum concha
d_2	Height of cymba concha
d_3	Width of cavum concha
d_4	Height of fossa
d_5	Height of pinna
d_6	Width of pinna
d_7	Intertragal incisure width
d_8	Depth of cavum concha
θ_1	Rotation angle of pinna
θ_2	Flare angle of pinna

4. Effects on Localization with Type of Hearing Aids

Speech is one of the most natural ways to communicate with others. When one's ear is damaged, they should use a hearing aid device to communicate with others. In a traditional hearing aid device, the voice signal is only amplified without any processing of background noises. However, when background noises are present, it is extremely difficult and inconvenient for users to localize or understand speech. As a result, removing background sounds and locating sound signals becomes a key consideration when developing a hearing aid device.

Listening techniques and problems with localization must be addressed when instructing hearing disabled person on hearing aids. If the subject has a conductive or combined listening problem, aid fitting would most likely guide them to resolve any difficulties with localization, unless the problem is vertical rather than horizontal. Subjects with sensorineural damage will likely gain location when sounds that are clearly not detectable cannot be located and sounds that are soft will me more difficult to detect than sounds that are pleasant or intense. Certain types of fittings, on the other hand, can make localization more difficult for certain people and the audiologist must take this into account when making a suitable recommendation. For example, closed ear molds that have a bilateral fit can be beneficial at frequencies greater than 4000 Hz for strong hearing people and in certain cases for persons with poor hearing (250-1000 Hz).

The use of open-fit hearing aids with a wide frequency range is discussed in [58,59]. Hearing aids with a wide bandwidth allow for the audibility of high- frequency sounds, which can aid in localization. Even with active feedback cancellation, an open-fit auditory support reduces the peak value of high frequency gain to less than 40 dB. As a result, open-fit wide bandwidth hearing aids are only suitable for people who have a slight to reasonable level of listening impairment in the higher frequencies (around 60-70 dB HL at 4000 Hz).

Bilateral hearing aid impact on frontal horizontal directional hearing is discussed in [60,61]. Bilateral hearing aid participants have under gone localization tests using a variety of stimuli and noise conditions. As a baseline, normal hearing people were used. Hearing aid users were tested for both the traditional omnidirectional microphone system and directional adaptive microphone setup with and without their hearing aids. The findings of this study show that normal hearing people do better in a localization task than people who wear bilateral hearing aids, that bilateral hearing aids do not maintain location indications and that adaptive directional noise reduction may adversely affect localization efficiency.

An effect on aided localization output on horizontal surface for people with hearing loss is investigated by inter-ear coordinated compression (IE) and pinna compensation in [62,63,64]. A group of ten hearing assisted subjects with bilaterally symmetrical sensorineural hearing loss who had already undergone localization training were assessed. On the horizontal surface, 12 loudspeakers placed 30° away were used to measure the effectiveness of localization. Some listeners will be able to better localize sounds coming from the sides if they use IE.

So in this section what we wanted to summarize is that each individual with sensorineural hearing loss has to choice the hearing aid in such a way that it should remove the background noise for good perception and it should not have adverse effect on sound source localization.

5. Conclusion

The difficulty in perceiving speech in noisy environments is the biggest problem for people with sensorineural hearing loss. So an exhaustive literature survey is carried out in this paper to evaluate the optimal signal processing algorithms for monaural and binaural hearing aids to improve speech perception using noise reduction technics and their influence on source localization. Localization is one of the facets of auditory perception that have gained insufficient consideration in hearing aid fitting and understanding hearing impaired people's problems. Traditionally, the primary emphasis has been on determining speech comprehension problems and developing and fitting hearing aids to enhance speech understanding. We have reviewed most of the international journals and conference papers; they involve theoretical computation research, experiment research and implementations of HRTF based three dimensional interactive audio displays. Experiments have shown that using a non-individualized HRTF will result in a high error rate and poor results. A few studies' comparison experiments show that individualizing HRTF is needed for accurate localization and high-quality auditory display. We summed up their basic benefits and drawbacks, as well as taking stock of current developments and finding several areas for change. We were particularly interested in the presence and findings of related perceptual research. Further we wanted to test optimized monaural and binaural hearing aid algorithms on hearing impaired and normal hearing people with different SNR values, azimuth angles, elevation angles and then to correlate the improvement in speech perception and the effect of hearing aid algorithm on source localization. So we wanted to achieve optimal hearing aid algorithm with minimal adverse effect on source localization.

Each technique has its own set of specifications and characteristics. As a result, the study centered on methodology classification based on features, in-depth analysis of a variety of research solutions, and prospective research directions. A comprehensive literature survey on sound source localization ability in hearing aids is presented.

References

[1] Kuk F, Korhonen P. “Localization 101: Hearing aid factors in localization”, Hearing Review, vol.21, no.9, pp.26-33, 2014.

[2] Chantai Laroche, “Review of the literature on sound source Localization and applications to noisy Workplaces”, Canadian Acoustics, vol.22, no.4, pp.13-18, 1994.

[3] M. Usman, K. Kamal, R. Qayyum, S. Akram, S. Mathavan, “3D Sound Generation Using Kinect and HRTF”, IEEE 2nd International Conference on Signal and Image Processing, Singapore, pp. 307-310, Aug. 2017.

[4] Corentin Guezenoc & Renaud Seguier, “HRTF Individualization: A Survey”, Audio Engineering Society Convention 145, Audio Engineering Society, Oct 2018, New York, United States.

[5] Song Xu, Zhizhong Li & Gaviriel Salvendy, “Individualization of Head-Related Transfer Function for Three-Dimensional Virtual Auditory Display: A Review”, International conference on virtual reality, pp. 397-407, July 2007.

[6] Jan-Gerrit Richter &Janina Fels, “On the Influence of Continuous Subject Rotation During High- Resolution Head-Related Transfer Function Measurements”, IEEE/ACM Transactions on audio, speech & language processing, Vol. 27, No. 4, April 2019.

[7] Michele Geronazzo, Simone Spagnol & Federico Avanzini, “Do We Need Individual Head-Related Transfer Functions for Vertical Localization? The Case Study of a Spectral Notch Distance Metric”, IEEE /ACM Transactions on audio, speech and language processing, Vol.26, No.7, July 2018.

[8] Michael Buerger, Stefan Meier, Christian Hofmann, Walter Kellermann, Eghart Fischer & Henning Puder, “Retrieval of Individualized Head-Related Transfer Functions for Hearing Aid Applications”, 25th European Signal Processing Conference (EUSIPCO), 2017.

[9] Hannes Gamper, David lohnston & Ivan J. Tashev, “Interaural time delay personalization using incomplete head scans”, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),2017.

[10] Holzl, J., “A Global Model for HRTF Individualization by Adjustment of Principal Component Weights”, Master Thesis, Graz University of Technology, 2014.

[11] Denk, F., Heeren, J., Ewert, S. D., Kollmeier, B., & Ernst, S. M., “Controlling the Head Position during individual HRTF Measurements and its Effect on Accuracy”, in DAGA, Kiel, 2017.

[12] Wen-Chih Wu, Cheng-Hsun Hsieh ; Hsin-Chieh Huang ; Oscal T. -C. Chen ; Yu-Jen Fang, “Hearing aid system with 3D sound localization”, IEEE Region 10 Conference, pp.1- 4, Taipei, 2007.

[13] M. Raspaud, H. Viste, G. Evangelista, "Binaural Source Localization by Joint Estimation of ILD and ITD", IEEE Transactions on Audio, Speech, and Language Processing, vol.18, Issue 1, pp.68-77, Jan 2010.

[14] W. A. Yost,L. Loiselle, M. Dorman and J. Burns, “Sound source localization of filtered noises by listeners with normal hearing: A statistical analysis”, Journal of the Acoustical Society of America, vol.133, no.5, pp. 2876– 2882, May 2013.

[15] Danfeng Li, S.E.Levinson, “A Bayes-rule based hierarchical system for binaural sound source localization”, Proc.IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) , vol.5, pp v521-v524, April 2003.

[16] Nguyen Duy Hai, Nitesh Kumar Chaudhary, Santi Peksi, Rishabh Ranjan, Jianjun He & Woon-Seng Gan, “Fast HRTF Measurement System with Unconstrained Head Movements for 3D Audio in Virtual and Augmented Reality Applications”, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),2017.

[17] V. Willert , J. Eggert , J. Adamy , R. Stahl and E. Korner, "A probabilistic model for binaural sound localization", IEEE Transactions on Systems, Man, and Cybernetics, Part B, vol. 36, no. 5, pp. 982-994, 2006.

[18] Ivo Merks, Gerald Enzner, Tao Zhang “Sound source localization with binaural hearing aids using adaptive blind channel identification”, Proc. IEEE International Conference on Acoustics, Speech and Signal Processing, pp.438-442, 2013.

[19] T. May, S. van de Par, and A. Kohlrausch, “A probabilistic model for robust localization based on a binaural auditory front-end”, IEEE Transactions on Audio, Speech, and Language Processing, vol.19, no.1, pp.1–13, 2011.
[20] Mojtaba Farmani, Michael Syskind Pedersen; Zheng-Hua Tan; Jesper Jensen, “Maximum likelihood approach to “informed” Sound Source Localization for Hearing Aid applications”, Proc.IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp 16- 20, South Brisbane, QLD, April 2015.
[21] M.Omologo and P. Svaizer, "Acoustic event localization using a crosspower-spectrum phase based technique", Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing, pp. II/273-II/276, Apr 1994.
[22] P. N. Kulkarni, P. C. Pandey, and D. S. Jangamashetti, “Multi-band frequency compression for sensorineural hearing impairment”, Proc.16th IEEE International Conference on Digital Signal Processing (DSP 2009), pp. 322-327, Santorini, Greece, 2009.
[23] Ravi Kumar Kandagatla, P V Subbaiah, “Performance Analysis of Statistical Approaches and NMF Approaches for Speech Enhancement”, International Journal of Image, Graphics and Signal Processing, Vol.11, No.7, pp. 9-38, 2019.
[24] P.Sunitha, K.Satya Prasad, “Speech Enhancement based on Wavelet Thresholding the Multitaper Spectrum Combined with Noise Estimation Algorithm”, International Journal of Image, Graphics and Signal Processing, Vol.11, No.9, pp. 44-55, 2019.
[25] https://www.sciencedirect.com/topics/engineering/interaural- time-difference.
[26] Daniel J. Tollin, Kanthaiah Koka, Jeffrey J. Tsai, “Interaural Level Difference Discrimination Thresholds for Single Neurons in the Lateral Superior Olive”, The Journal of Neuroscience, vol. 28, no.19, pp.4848 – 4860, May 7, 2008.
[27] Xiang Wu, Dumidu S. Talagala, Wen Zhang and Thushara D. Abhayapala, “Binaural localization of speech sources in 3-D using a composite feature vector of the HRTF”, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brisbane, Australia, April 2015.
[28] Mojtaba Farmani, Richard Heusdens, Michael Syskind Pedersen, Zheng-Hua Tan, and Jesper Jensen, “Concurrent Localization of Sound Sources and Dual-Microphone Sub- Arrays Using TOFs”, 19th International Conference on Information Fusion, pp.1931-1936, July 2016, Germany.
[29] Mojtaba Farmani, Michael Syskind Pedersen, Zheng-Hua Tan, Jesper Jensen, “On the influence of microphone array geometry on HRTF-based sound source localization”, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp.439-443, 2015.
[30] Mojtaba Farmani, Michael Syskind Pedersen, Zheng-Hua Tan and Jesper Jensen, “Informed TDoA-based Direction of Arrival Estimation for Hearing Aid Applications”, IEEE Global Conference on Signal and Information Processing (GlobalSIP), pp.953-957, 2015.
[31] Anshuman Ganguly, Chandan Reddy, Yiya Hao, Issa Panahi, “Improving Sound Localization for Hearing Aid Devices Using Smartphone Assisted Technology”, IEEE International Workshop on Signal Processing Systems (SiPS), pp.165-170, Dec-2016.
[32] Ibrahim I, Parsa V, Macpherson E, Cheesman M, “Evaluation of speech intelligibility and sound localization abilities with hearing aids using binaural wireless technology”, Audiology Research, Volume 3, no.1, 2013.
[33] Rajani S. Pujar, Pandurangarao N. Kulkarni, "Wiener Filter Based Noise Reduction Algorithm with Perceptual Post Filtering for Hearing Aids", International Journal of Image, Graphics and Signal Processing, Vol.11, No.7, pp. 69-81, 2019.
[34] Denis Byrne & William Noble, “Optimizing Sound Localization with Hearing Aids”, Trends in Amplification, Vol. 3, No. 2, pp.51-73, 1998.
[35] Dumidu S. Talagala, Xiang Wu, Wen Zhang and Thushara D. Abhayapala, “Binaural Localization of Speech Sources in the median plane using Cepstral HRTF Extraction”, 2014 22nd European Signal Processing Conference (EUSIPCO), Lisbon, Portugal, Sept.2014.
[36] Marc M. Van Wanrooij and A. John Van Opstal, “Contribution of Head Shadow and Pinna Cues to Chronic Monaural Sound Localization”,The Journal of Neuroscience, vol. 24, no.17, pp. 4163– 4171, April 2004.
[37] J. A. Macdonald, "A localization algorithm based on head- related transfer functions", The Journal of the Acoustical Society of America, vol. 123, no. 6, pp. 4290-4296, 2008.
[38] Shin, K. H. and Park, Y., “Enhanced vertical perception through head-related impulse response customization based on pinna response tuning in the median plane”, IEICE Transactions on Fundamentals of Electronics, Communications & Computer Sciences, pp.345–356, 2008.
[39] Rajani S. Pujar, Pandurangarao N. Kulkarni, "Cascaded Structure of Noise Reduction and Multiband Frequency Compression of speech signal to Improve Speech Perception for monaural Hearing Aids”, 16th IEEE India Council International Conference (INDICON 2019), Gujarat, 13th – 15th December 2019.
[40] Andrew D. Brown, Brianne T. Beemer, Nathaniel T. Greene, Theodore Argo, G.Douglas Meegan & Daniel J. Tollin, “Effects of Active and Passive Hearing Protection Devices on Sound Source Localization, Speech Recognition & Tone Detection”, https://doi.org/10.1371/journal.pone.0136568, 27th August 2015.
[41] Jianjun He, Woon-Seng Gan & Ee-Leng Tan, “On the preprocessing & postprocessing of HRTF individualization based on sparse representation of anthropometric features”, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), South Brisbane, QLD, Australia, 19-24 April 2015.
[42] Arjun Das H, Lisha Gopalakrishnan Pillai & Muruganantham C, “Human Voice Localization in Noisy Environment by SRP-PHAT and MFCC”, International Research Journal of Advanced Engineering and Science, Volume 1, Issue 3, pp. 33-37, 2016.
[43] Dongwen Ying, Ruohua Zhou, Junfeng Li & Yonghong Yan, “Window-Dominant Signal Subspace Methods for Multiple Short-Term Speech Source Localization”, IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 25, no. 4, April 2017.
[44] T. Van den Bogaert, E. Carette & J. Wouters, “Sound localization with and without hearing aids”, Belgium, January 2009.
[45] Ngai-Man Cheung Steven Trautmann & Andrew Homer, “Head-Related transfer function modeling in 3-D sound systems with genetic algorithms”, IEEE, pp.3529-3532, 1998.

[46] Douglas S. Brungart, Brian D. Simpson, “Auditory Localization of nearby sources in a Virtual Audio Display”, Proceedings of the 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics (Cat.No.01TH8575), New Platz, NY, USA, Oct.2001.

[47] Frantisek Rund, Filip Saturka, “Alternatives to HRTF Measurement”, 2012 35th International Conference on Telecommunications and Signal Processing (TSP), Prague, Czech Republic, July 2012.

[48] Martin Rothbucher, Tim Habigt, Johannes Feldmaier and Klaus Diepold, “Integrating a HRTF-based Sound Synthesis System into Mumble”, 2010 IEEE International Workshop on Multimedia Signal Processing, Saint Malo, France, pp. 24-28, Oct. 2010.

[49] Xiaohai Tian, Zhonghua Fu and Lei Xie, “An Experimental Comparison on KEMAR and BHead210 Dummy Heads for HRTF-based Virtual Auditory on Chinese Subjects”, IET 3rd International Conference on Wireless, Mobile and Multimedia Networks (ICWMNN 2010), Beijing, China, Sept. 2010.

[50] Ruixing Wu1, Guangzheng Yu, “Improvements in HRTF dataset of 3D Game Audio Application”, 2016 International Conference on Audio, Language and Image Processing (ICALIP), Shanghai, China, July 2016.

[51] Silzle A., “Selection and Tuning of HRTFs”, AES 112th Convention, Munich, Germany, May 10-13, pp.1–14, 2002.

[52] Mo Zhu, Muhammad Shahnawaz, Stefano Tubaro, Augusto Sarti, “HRTF Personalization based On Weighted Sparse Representation of Anthropometric Features”, 2017 International Conference on 3D Immersion (IC3D), Brussels, Belgium, Dec. 2017.

[53] H. Hu, L. Zhou, H. Ma, and Z. Wu, “HRTF personalization based on artificial neural network in individual virtual auditory space”, Applied Acoustics, vol. 69, no.2, pp. 163– 172, 2008.

[54] L. Li and Q. Huang, “HRTF personalization modeling based on RBF neural network”, 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp.3707–3710, 2013.

[55] V. R. Algazi, R. O. Duda, D. M. Thompson and C. Avendano, "Then CIPIC HRTF database", Proceedings of the 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics, New Platz, NY, pp. 99– 102, 2001.

[56] “CIPIC database,” http://interface.cipic.ucdavis.edu.

[57] Kazuhiro Iida, Hikaru Shimazaki & Masato Oota, “Generation of the Individual Head-Related Transfer Functions in the Upper Median Plane Based on the Anthropometry of the Listener's Pinnae”, IEEE 7th Global conference on consumer electronics (GCCE),2018.

[58] William Noble, S. Sinclair, Denis Byrne, “Improvement in Aided Sound localization with Open Ear molds: Observations in People with High-Frequency Hearing Loss”, Journal of the American Academy of Audiology, vol.9, no.1, pp. 25-34, Feb 1998.

[59] Kuk F, Baeksgaard L, “Considerations in fitting hearing aids with extended bandwidth”, Hearing Review, vol.16, no.11, pp.32-38, 2009.

[60] Tim Van den Bogaert, T. J. Klasen, M. Moonen, Lieselot Van Deun, Jan Wouters, “Horizontal localization with bilateral hearing aids: Without is better than with”, Acoustical Society of America, pp.515–526, 2005.

[61] Iman Elsabah Ibrahim, “Effects of Coordinated Bilaterial Hearing Aids and Auditory Training on Sound Localization”, PhD Thesis, University of Western Ontario, April 2013.

[62] Kuk F, Crose B, Kyhn T, Morkebjerg M, Rank M, Norgaard M, Foh H, “Digital wireless hearing aids part-3: audiological benefits”, Hearing Review, vol.18, no.8, pp.48-56, 2011.

[63] Korhonen P, Lau C, Kuk F, Deenan D, Schumacher J. “Effects of coordinated dynamic range compression and pinna compensation algorithm on horizontal localization performance in hearing-aid users”, Journal of the American Academy of Audiology, vol.26, no.1, pp 80-92, Jan 2015.

[64] Keidser G, O’Brien A, Hain J, McLelland M, Yeend I, “The effect of frequency-dependent microphone directionality on horizontal localization performance in hearing-aid users”, International Journal of Audiology, vol.48, no. 11, pp. 789- 803, 2009.

Authors’ Profiles

Smt. Jyoti M. Katagi received her B. E. Degree from Karnataka University, Dharwad and M.Tech. degree in Digital Communication and Networking from Visvesvaraya Technological University, Belgaum. Presently she is pursuing her Ph. D. in Speech Processing at ECE Department, Basaveshwar Engineering College (Autonomous), Bagalkot, Visvesvaraya Technological University, Belgaum. She has teaching experience of 15 years. Her research interests are: Speech Proceesing, Control Systems and Digital Signal Processing.

Dr. Pandurangarao N. Kulkarni has completed his Ph.D from Indian Institute of Technology Bombay, India, in the year 2010. He has thirty years of teaching experience and eleven years of research experience. Currently he is working as a professor, department of Electronics and Communication Engineering, at Basaveshwar Engineering College, Bagalkot, Karnataka, India. His fields of interest include digital communication, digital signal processing and applications. His research area is speech processing to improve speech perception in sensorineural hearing loss.

How to cite this paper: Jyoti M. Katagi, Pandurangarao N. Kulkarni, " Sound Source Localization Ability in Hearing Aids: A Survey", International Journal of Image, Graphics and Signal Processing(IJIGSP), Vol.13, No.6, pp. 30-44, 2021.DOI: 10.5815/ijigsp.2021.06.04