APSIPA ASC 2015 **Tutorial 2**
9:00am – 12:30 pm

Assisted Listening for Headphones and Hearing Aids

Signal Processing Techniques

Woon-Seng Gan and Jianjun He

Digital Signal Processing Laboratory
School of Electrical and Electronic Engineering
Nanyang Technological University, Singapore

16th Dec, 2015

Outline of Tutorial : (9am – 10:30am)

I. 3D Sound and Headphones

II. Natural Sound Rendering for Headphone

- Virtualization
- Sound scene decomposition
- Individualization
- Equalization
- Head tracking
- Integration
- 3D Audio Headphones
- Demo

Outline of Tutorial : (11am – 12:30pm)

III. An Overview and Applications of ANC Headsets

IV. Natural Augmented Reality Audio in Headsets

- Signal processing and practical challenges
- ARA headset for mobile and wearable devices
- Multichannel headphone sound reproduction
- Natural augmented reality (NAR) headset using adaptive techniques

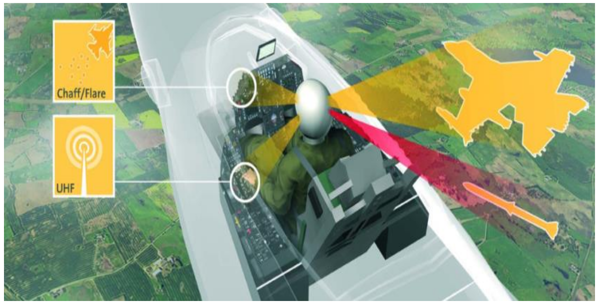
V. Assisted Listening in Hearing aids

- Hearing loss and hearing aids
- Noise reduction and speech enhancement
- Integration of ANC and noise reduction in hearing aids

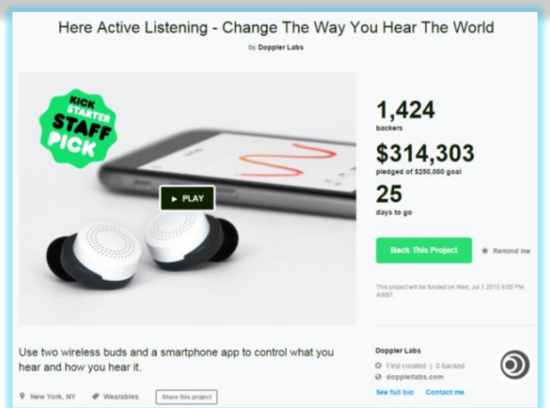
VI. Summary and Demo

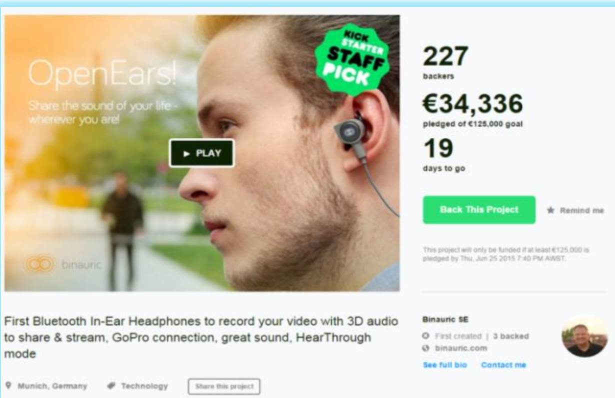
Module I
3D Sound and Headphones

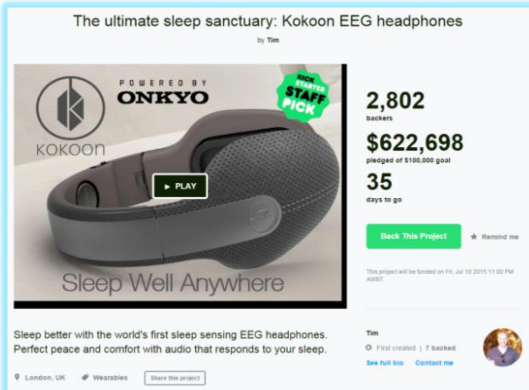
(Assisted) Binaural Listening

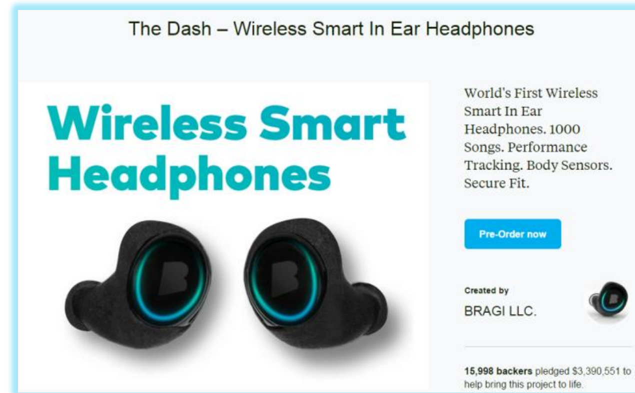


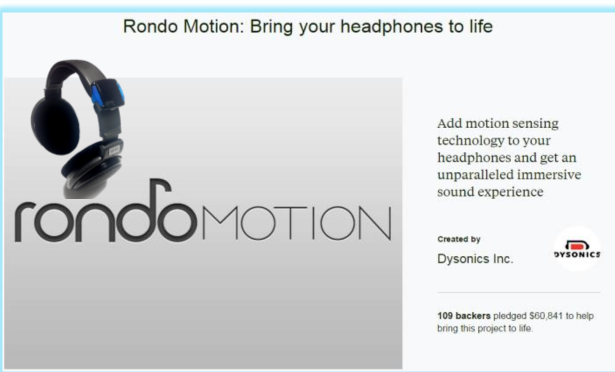
Selected Kickstarter projects on assisted listening

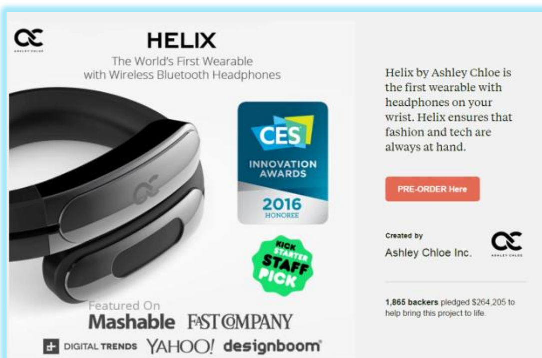


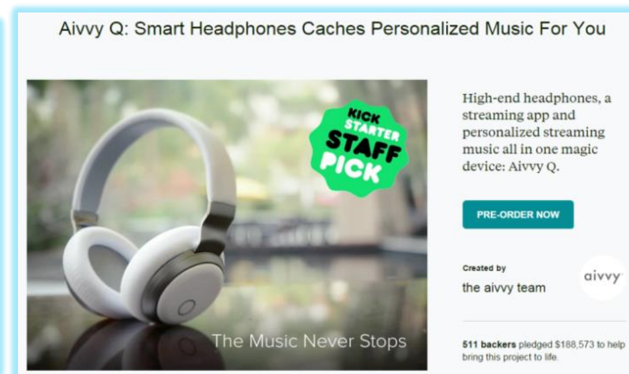












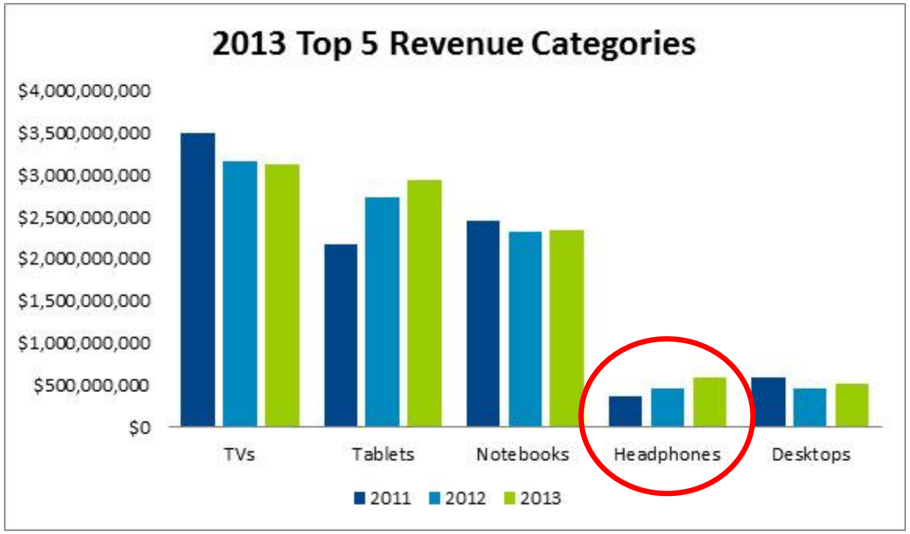
A realistic and engaging experience

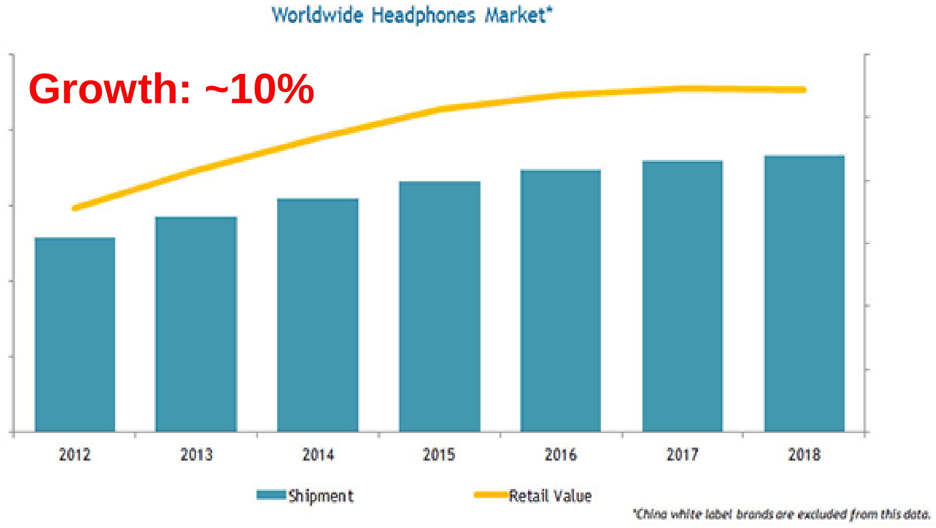
Speakers vs Headphones

Speaker systems

Headphones

Strong headphone market

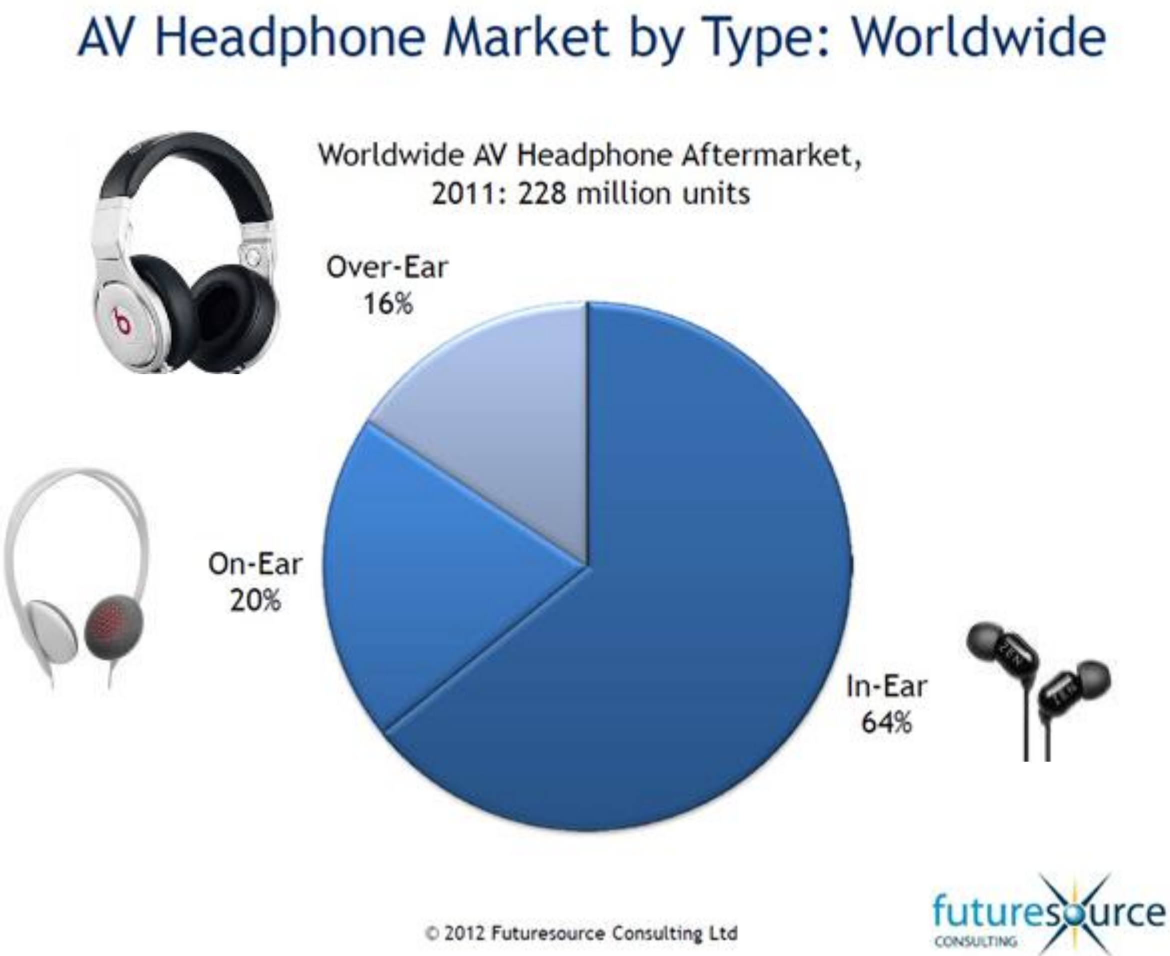




Source from
https://www.npd.com/wps/portal/npd/us/news/press-releases/key-ce-categories-deliver-positive-2013-holiday-results-according-to-npd/

Source from
http://www.ceatec.com/report_analysis/en/ra_150512_2.html

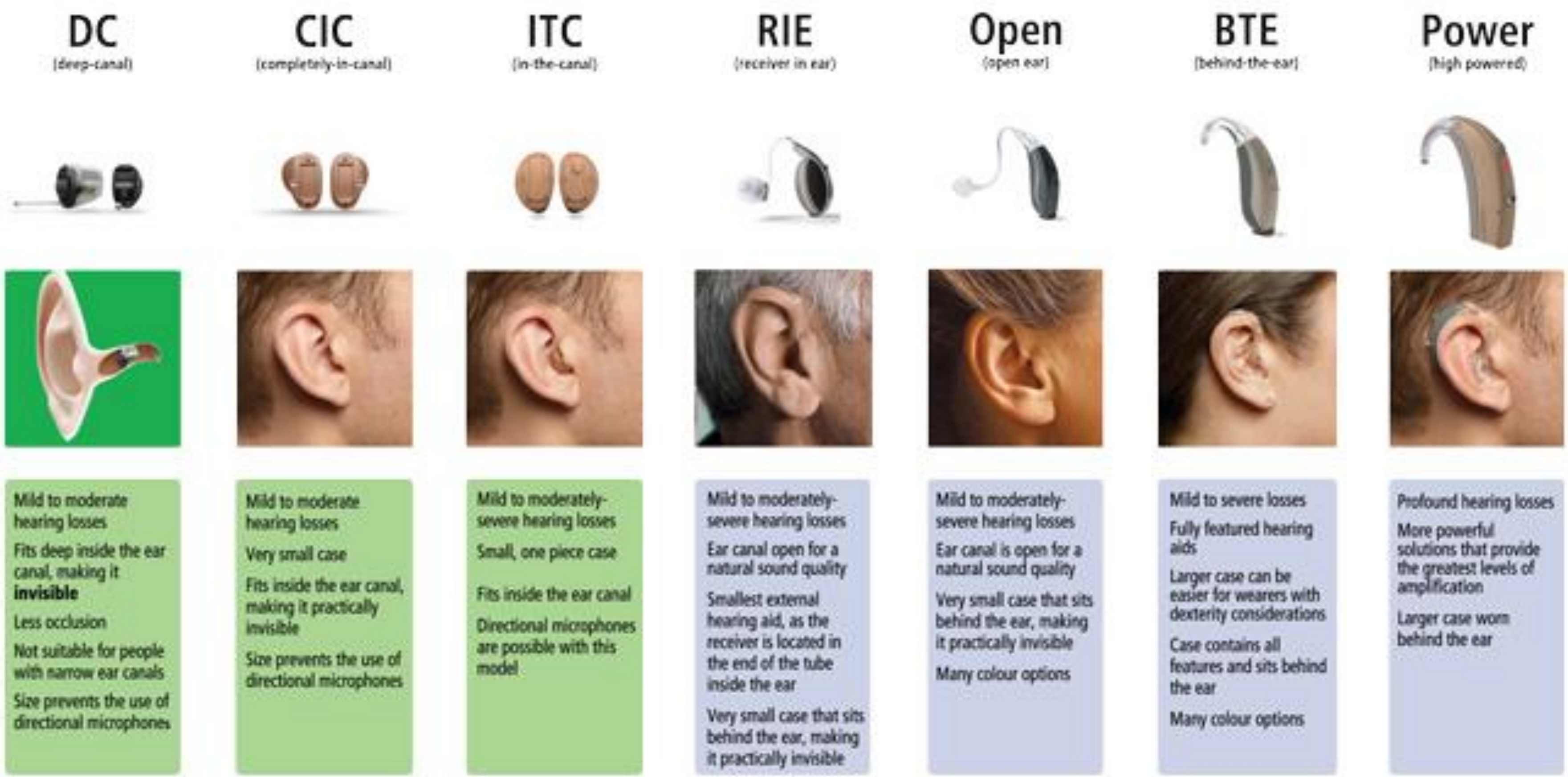
Market Share



Prices of Over-Ear Headphones

Headphone Type	Target Consumers	Headphones Brands	Features and comments	Price
Stereo	General Pro-sumer	AKG, BeyeDynamic, Denon, German Maestro, Grado, Koss, Ultrasone	High fidelity stereo, mostly requires power amplification for optimal sound	**Hundreds to Thousands**
	High-End Consumer	Bose, Fanny Wang, Jays, Monster (beats), Sony, Shure	Stereo, simple to drive, can feature good fidelity stereo	**Hundreds (< $500)**
	General Consumer	Beats, Creative, Eskuche, Hed Kandi, Goldring, iFrogz, Jays, Marshall, Ministry of Sound, Philips, Pioneer, Shure, SkullCandy, SonicGear, Sumajin, TDK	Low cost, stereo, meant for general usage, styling of headphone is general more critical than sound quality	**Hundreds (< $200)**
Stereo with Virtual Surround	General Consumer/Gamers	Acoustic Research, Creative, Corsair, Philips, Logitech, Pioneer, Razer, SteelSeries, Sony, Turtle Beach, Tritton, Ultrasone, Yamaha, Zulman	Virtual surround achieved with HRTF processing, Dolby headphone, S-Logic (ultrasone), CMSS-3D (creative), Dolby Ex (turtle beach)	**Hundreds (< $200)**
Discrete Surround	**Consumer for gaming, general entertainment**	**Creative, Logitech, Mentor, Razer, Psyko, Razor, Turtle Beach, Tritton, Zalman**		**Hundreds (< $400)**

Hearings Aids

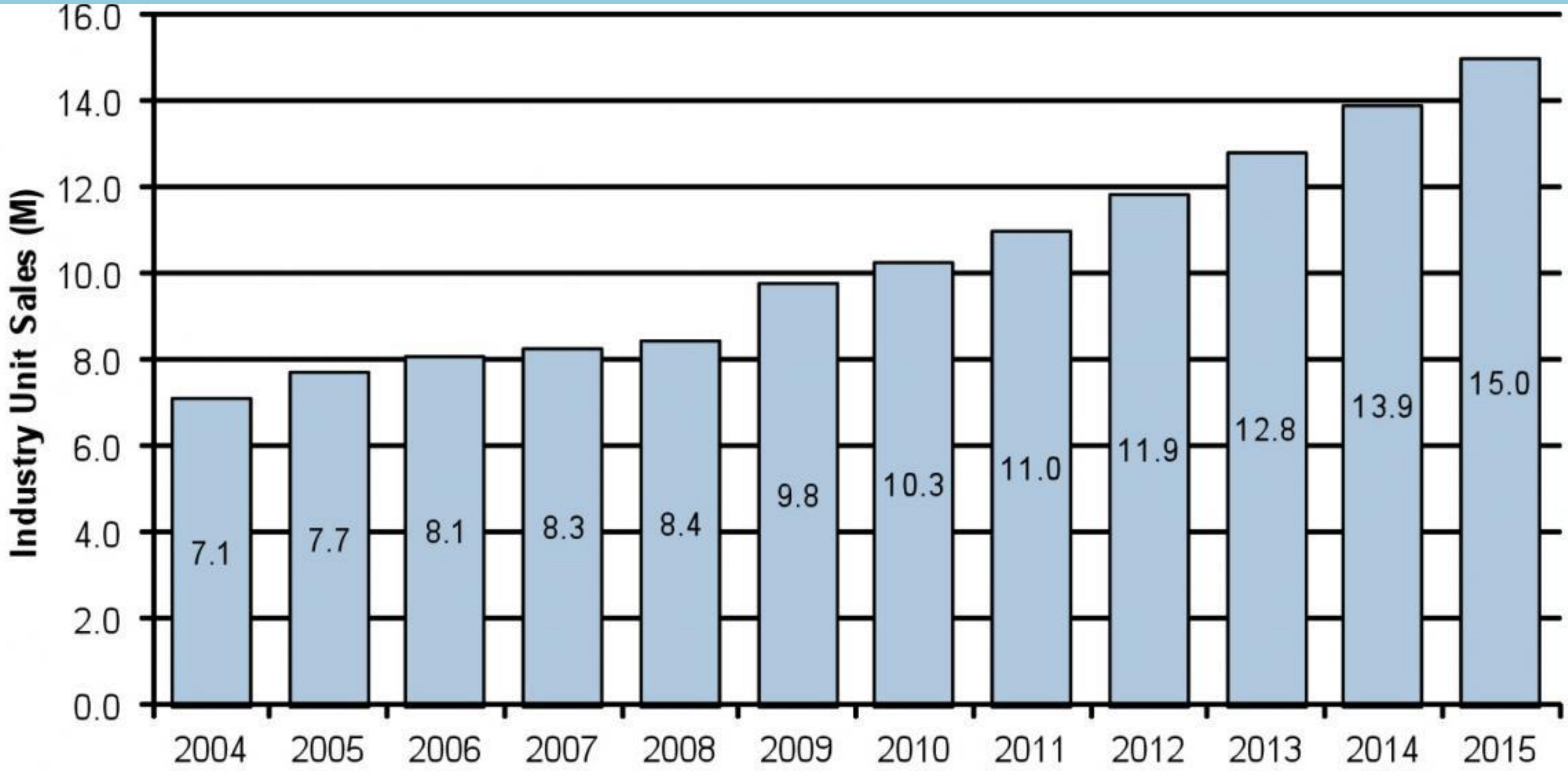


Source from
http://www.globalhearing.com.au/models&sizes.html

Hearing Aids market

"The hearing aids market is expected to reach USD 8,374 million by 2020 from USD 6,183 million in 2015, at a CAGR of 6.3%." - Marketsandmarkets reports

"Hearing Aid Sales Increase by 8.8% in First Half of 2015."- Hearing Industries Association (HIA)



Source from
http://www.bdti.com/InsideDSP/2014/10/16/ONSemi

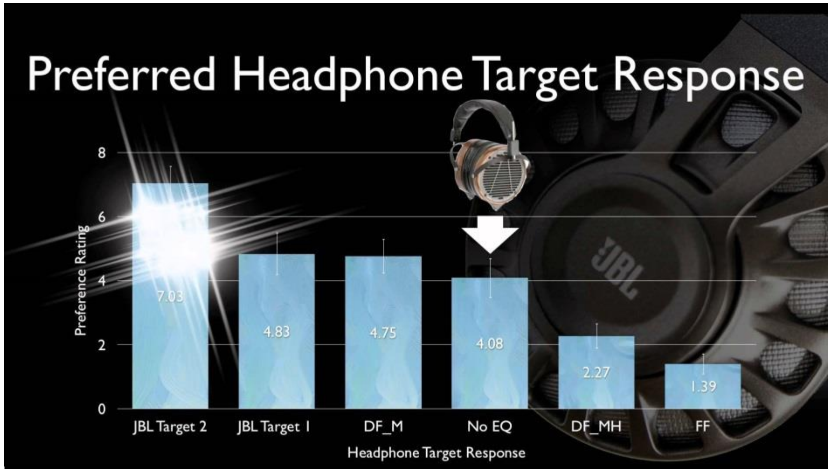
Pursuing natural 3D sound in Headphones Industry

Source from http://www.canadacomputers.com/product_info.php?cPath=1215&item_id=033846

Source from http://www.razerzone.com/referral/invitesurround/?ref=Adrian%20Wong&ref_email=awsh%40techarp.com

Source from http://3dsoundlabs.com/en/

Source from https://www.youtube.com/watch?v=RW-JDy2uSyY



Source from http://seanolive.blogspot.sg/2014/01/the-perception-and-measurement-of.html

Experimenting 3D audio

- Designing an experimentation platform using headphones
  - With user interaction (touch and spatial audio)
  - Sonification / Auralization
  - User preference / hearing profile
  - **Personalized HRTF / anthropometry**
  - Computation and resources demand in portable devices
- Creating new mobile and wearable apps
  - Assistive Applications
  - Enhanced Telepresence
  - Remote Monitoring
- **Goal: design a headphone system which is perceptually indistinguishable from real listening.**

Existing Apps: Binaural Recordings

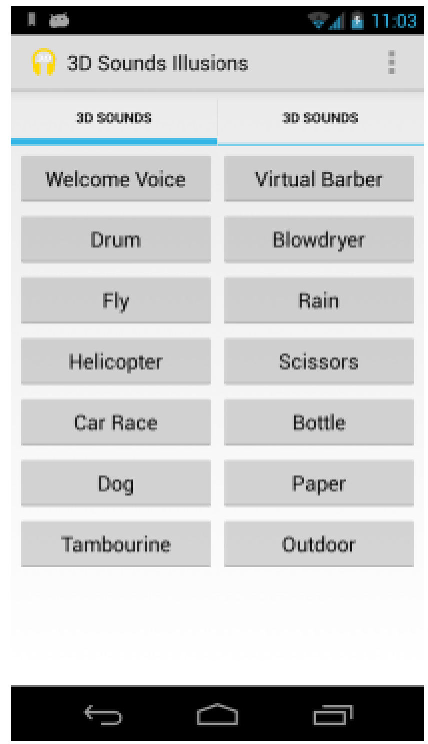
3D Sounds Illusions

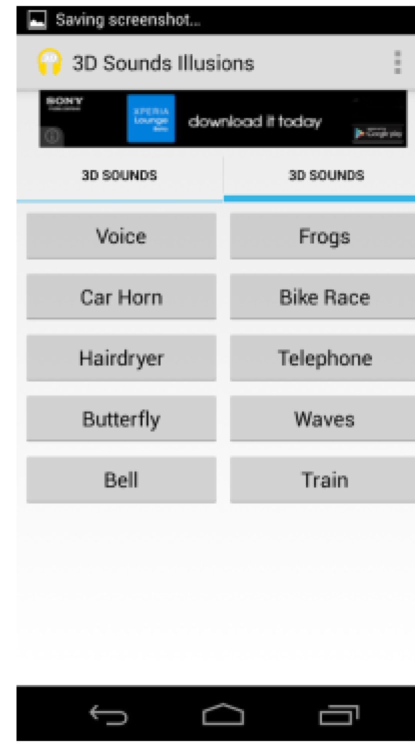
- Package name: sounds3d.soundboard
- https://play.google.com/store/apps/details?id=sounds3d.soundboard

Retail: free (with ads)

Real-time 3D Audio Processing: ✗

HRTF Processing: ✗

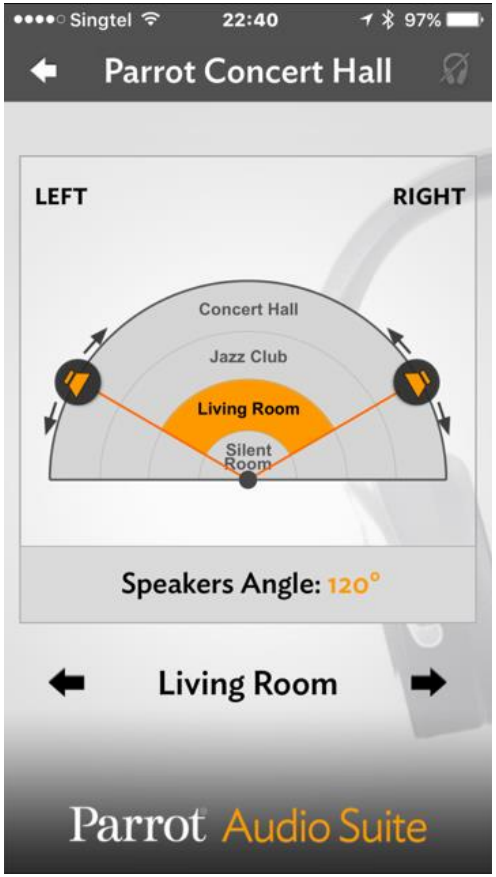




Features:

- A collection of binaural recording files
- Playback binaural recordings.

Existing 3D audio Apps in the Market



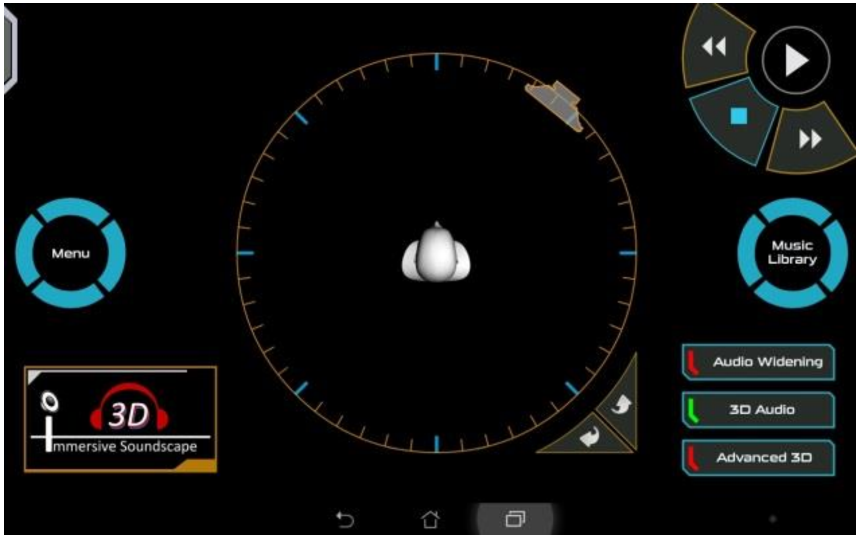
Parrot headphones apps

Headquakes apps

Audio-3D Player Headphones HD 7.1

They are mainly stereo extender and virtual downmix apps

3DA^3



Download here: http://tinyurl.com/kboq2g7

App Statistics

CURRENT / TOTAL INSTALLS	AVG. RATING / TOTAL #
1,048 / 15,676	★ 4.11 / 132

Rating	Count
★★★★★	79
★★★★	22
★★★	14
★★	1
★	16

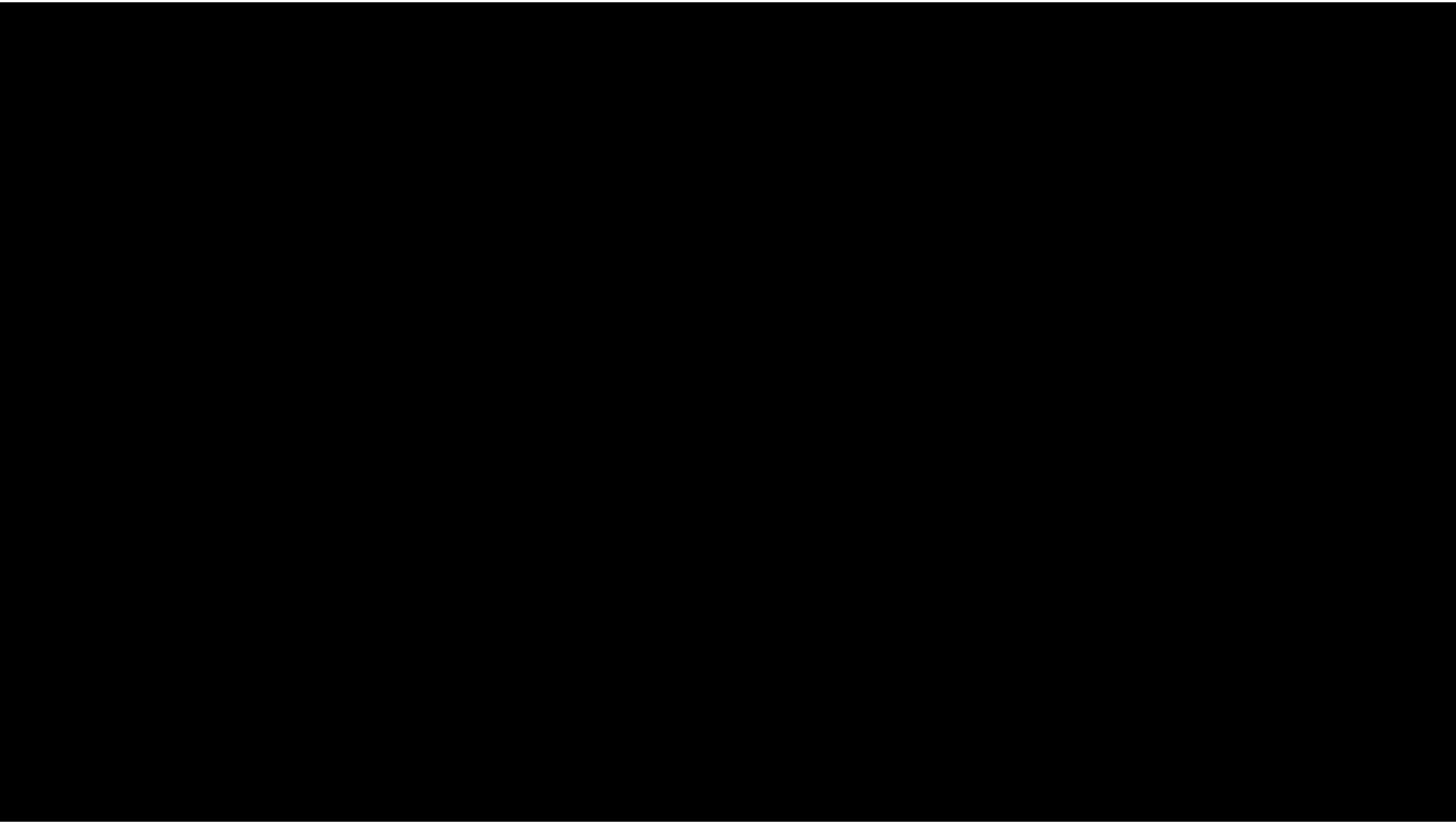
Main Features

1. **Real-time DSP on Android platform**
2. **Fully customized user interface and design.**
3. **Audio Widening (Externalization)**
4. **3D Audio using HRTF filtering**
5. **Near Field 3D (audio depth) with recording**
6. **Elevation function**
7. **Auto-rotate function**
8. **Up to four channels simultaneous processing**
9. **Virtual Bass System integration**
10. **Ffmpeg Audio Decoding**

Technical Specifications

Supported Platform	Android
Supported OS Versions	4.0 and up
Minimum CPU requirement (Chipset)	Snapdragon 400
	Nvidia Tegra 3 T30L
	Mediatek MT6589T
	Exynos 4210
Minimum RAM requirement	512 MB
App Size	6.84 MB
Audio Decoder	FFMPEG library
Supported formats	MP3, WAV
Supported Sampling rate	44100Hz
Frame Size	Stereo: 1152 (Mp3), 1024 (WAV)
HRTF Database	CIPIC HRTF Database
HRTF taps	200
Data type	double-precision 64-bit IEEE 754 floating point
HRTF sets (48 points/set)	Azimuth: 6 sets
	Eelvation: 2 sets
	Near Field: 10 sets
Memory usage (HRTF)	6.75 KB
Azimuth HRTF Resolution	7.5°
Elevation HRTF Resolution	7.5°
Near Field depth	25cm, 50cm, 75 cm, 100cm, 125cm
Devcies tested	Asus Transformer Pad TF300T
	Xiaomi Redmi
	Samsung Galaxy S2, S3, S4, S5
	Asus Padfone Infinity

Demo video on $3DA^3$



Quick Overview

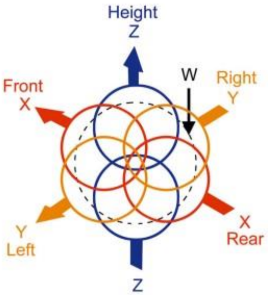
- Techniques to capture 3D sound (Spatial Audio Recording)?
- **Types of processing required**
- Rendering over **headphones** and loudspeakers.

Spatial Audio Recording

- Binaural recording
  - Dummy head
  - Human subjects

- Stereo/5.1

- B-format



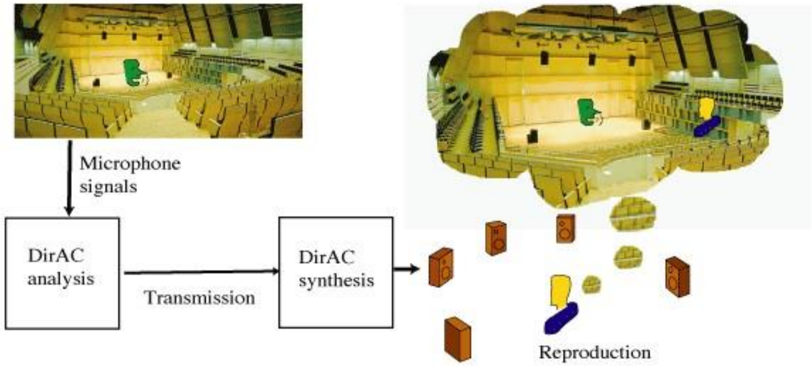
(a)

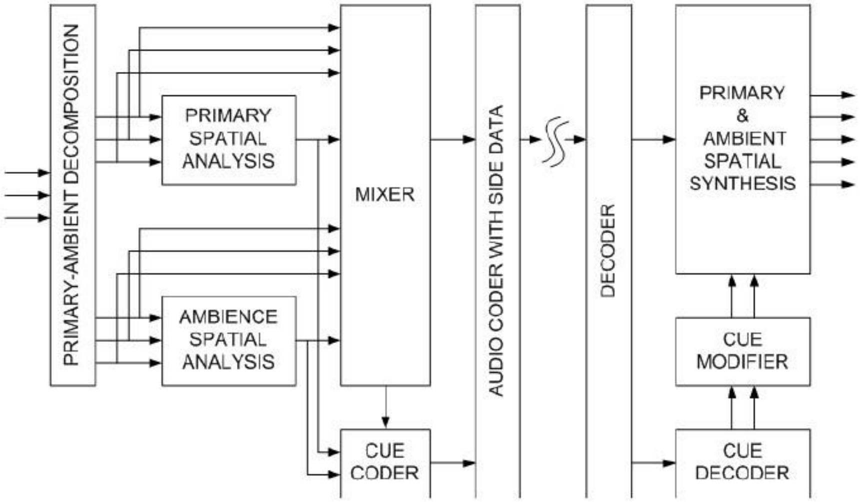

(b)

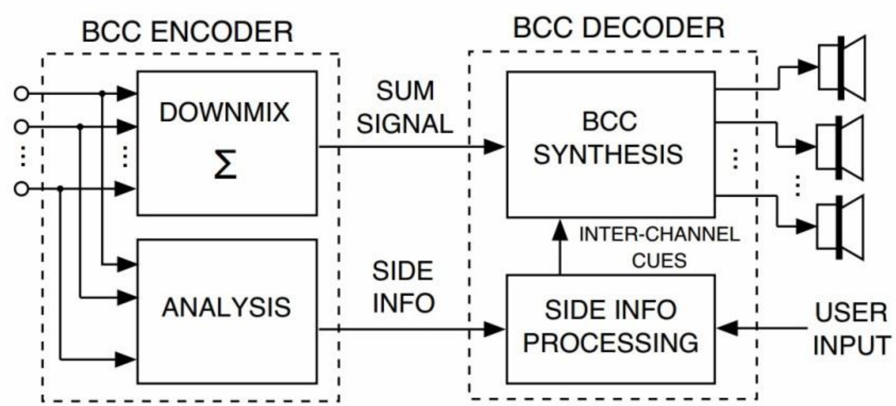
- Microphone arrays

Spatial Audio Processing

- Spatial audio coding
  - Directional audio coding (DirAC)
  - Spatial audio scene coding (SASC)
  - Binaural cue coding (BCC)



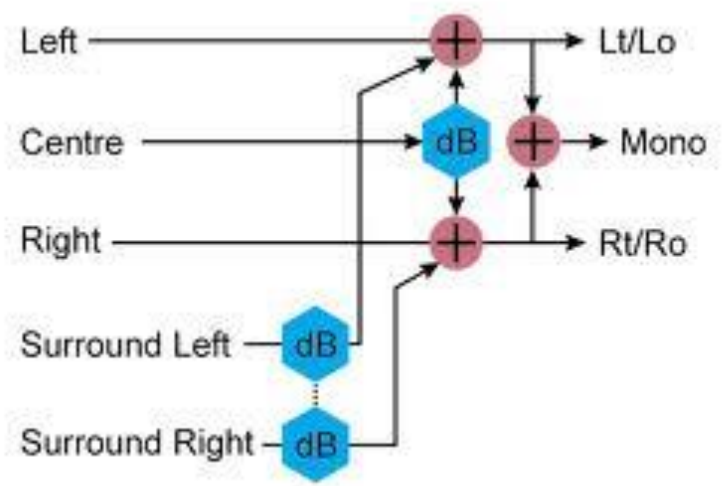


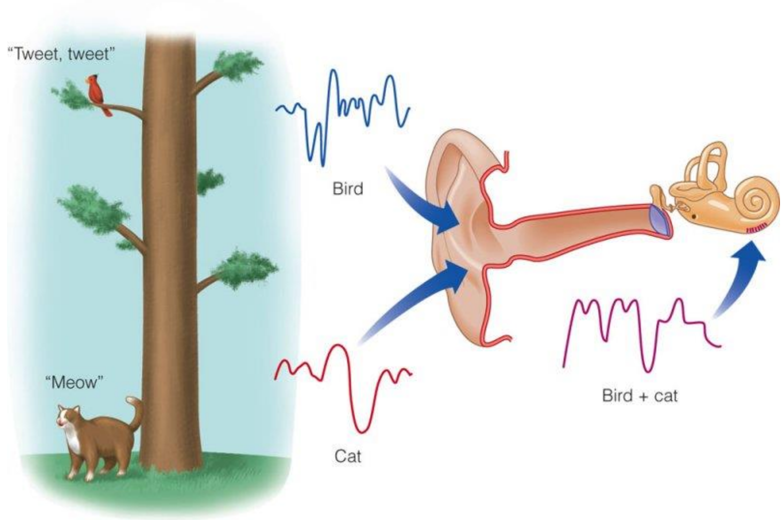


Spatial Audio Processing

- Spatial audio coding
  - Directional audio coding (DirAC)
  - Spatial audio scene coding (SASC)
  - Binaural cue coding (BCC)

- Audio mixing (Down-mix, Up-mix)
- Sound Scene Decomposition

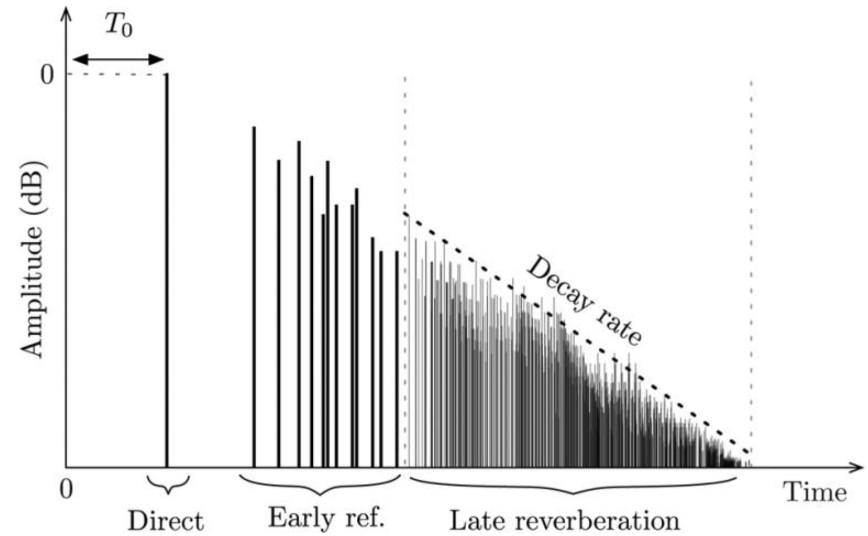


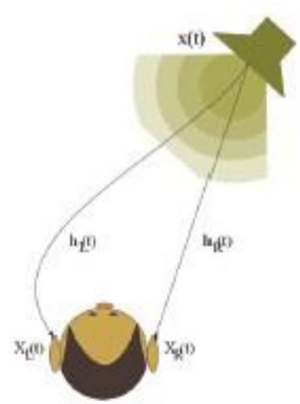


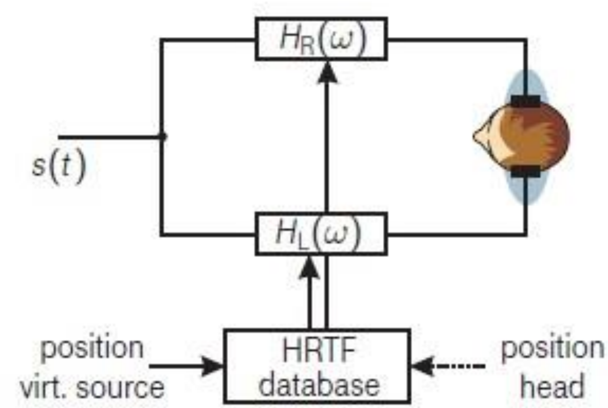
Spatial Audio Processing

- Spatial audio coding
  - Directional audio coding (DirAC)
  - Spatial audio scene coding (SASC)
  - Binaural cue coding (BCC)

- Audio mixing (Down-mix, Up-mix)
- Sound Scene Decomposition
- Binaural synthesis
- Artificial reverberation
- Equalization, Decorrelation, Crosstalk cancellation etc.

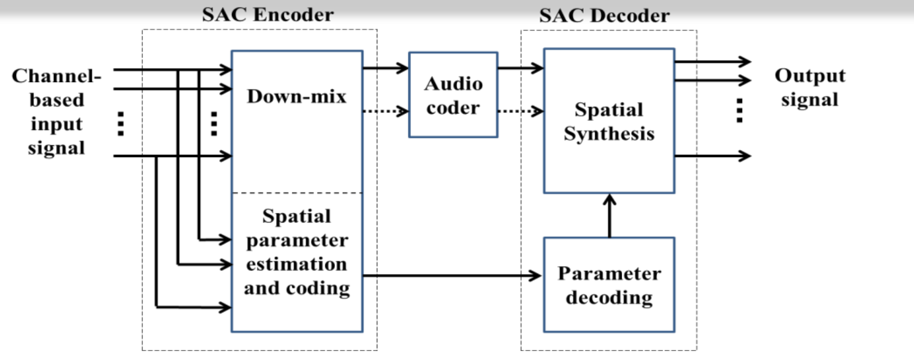




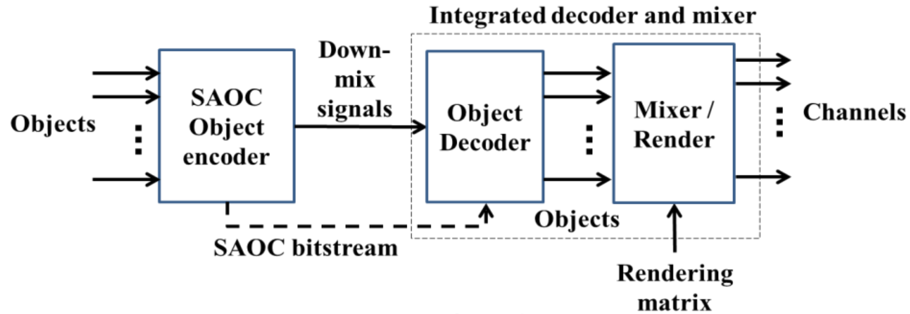


Spatial Audio Standards

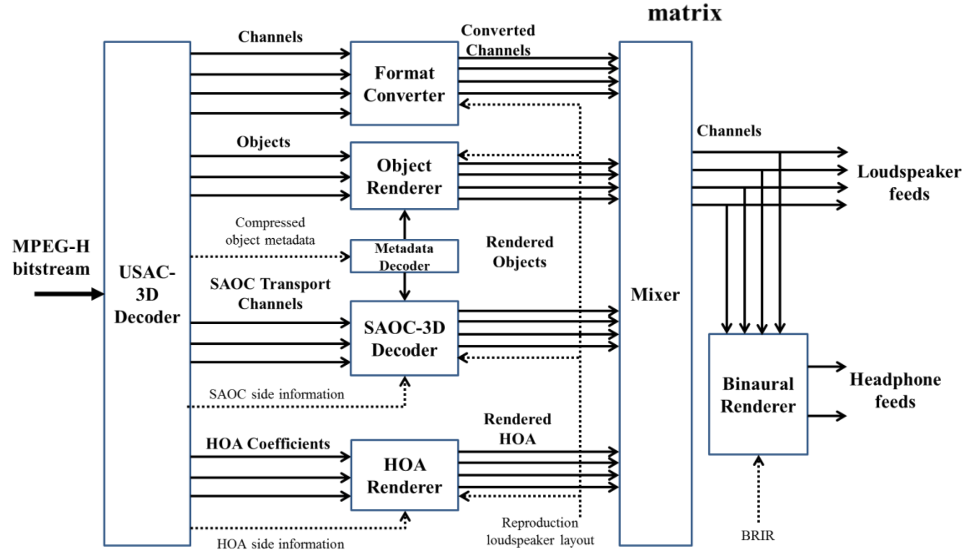
- ➢ MPEG Surround



- ➢ MPEG SAOC



- ➢ MPEG-H 3D Audio



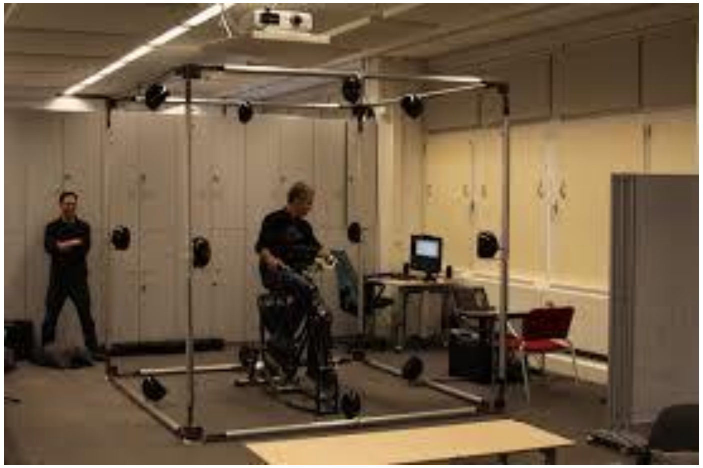
Spatial Audio Reproduction

- Stereo/ Multichannel Loudspeakers (5.1, 7.1, 22.2)

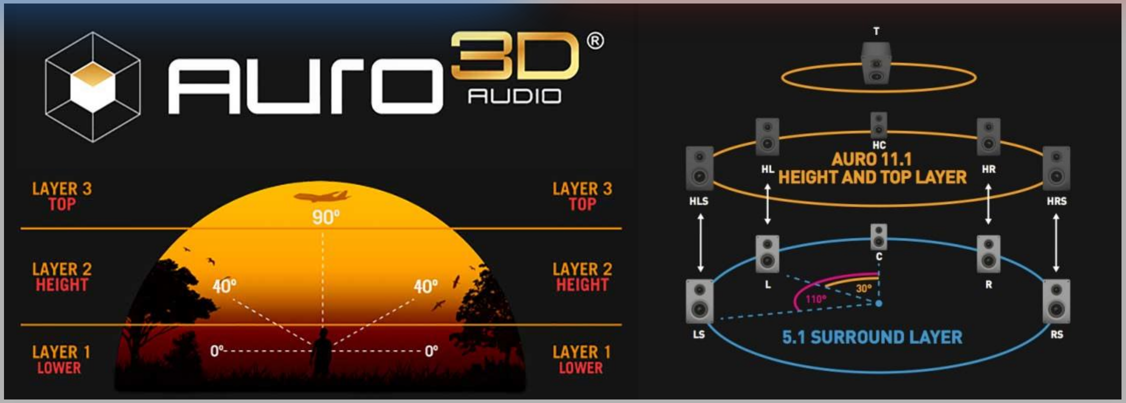
- Stereo Headphone
  - Binaural Headphone

- Wave Field Synthesis (WFS)

- Ambisonics



Emerging Spatial Audio Reproduction



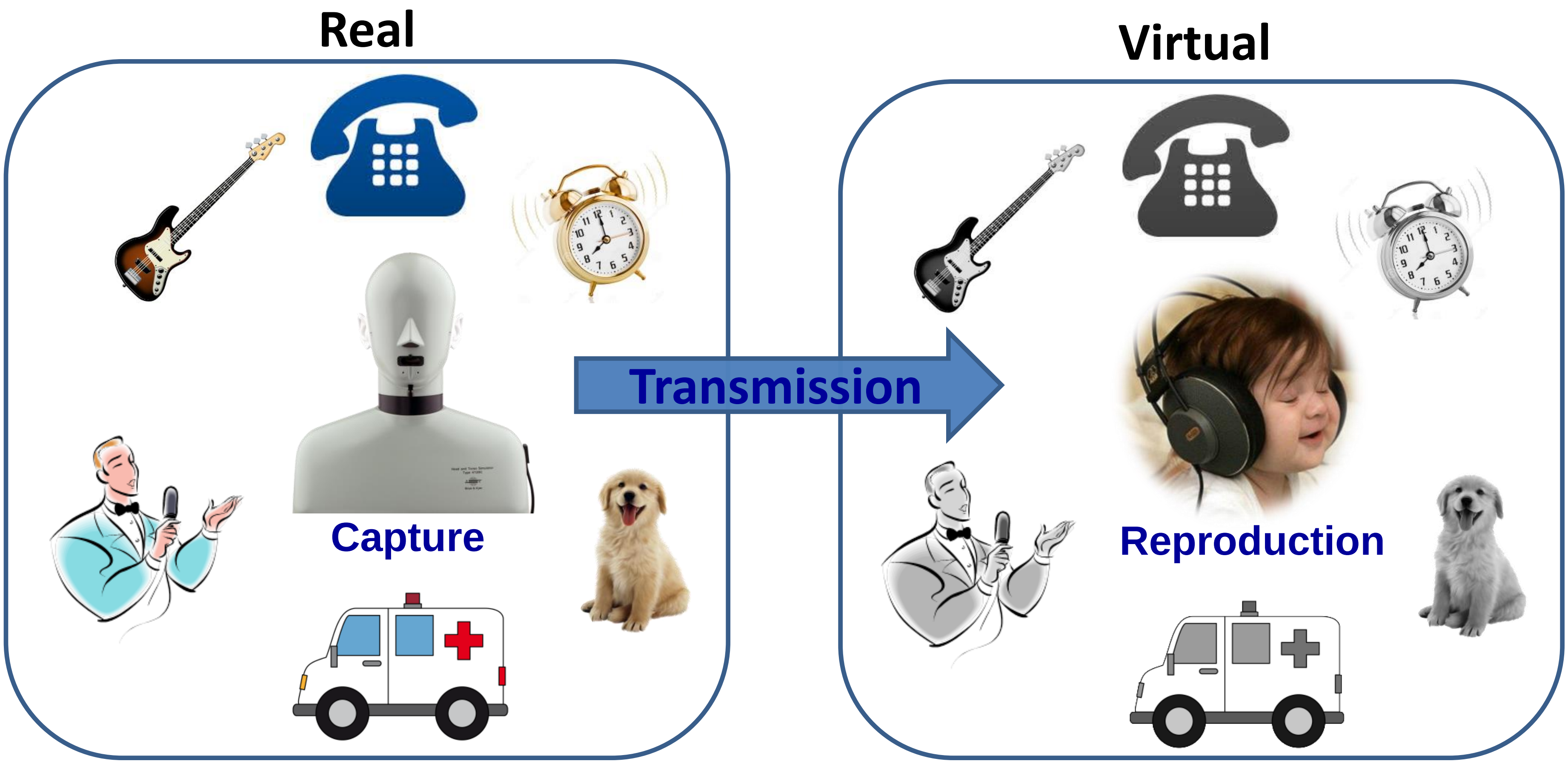
Binaural Technology

* Picture from logicomusic.com.br

- Encompasses a set of tools for recording or synthesizing and rendering binaural signals at the listener's ear

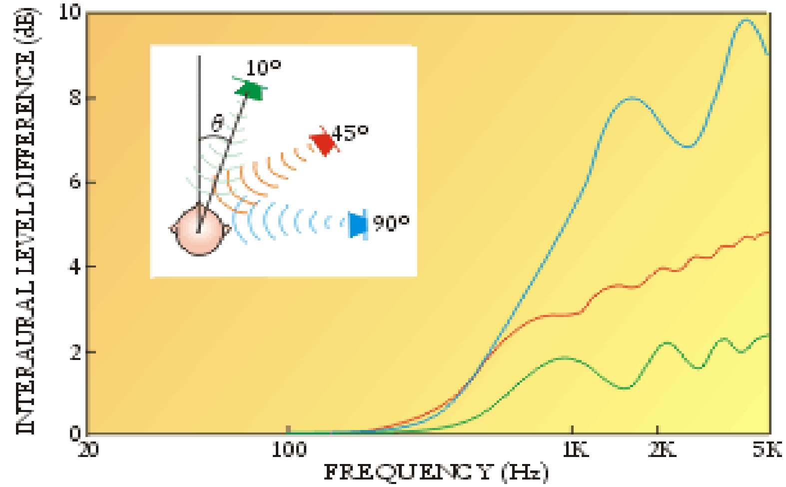
- Deals with the natural cues of auditory localization which results from the reflection and diffraction of the acoustical waves with the human torso, shoulders, and the external ears

From Real to Virtual Reality

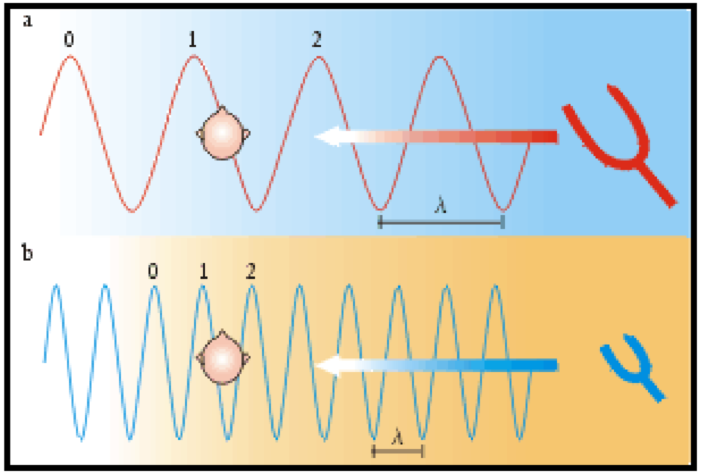


How do we perceive sound location?

- Compare sound received at two ears
  - **Interaural Level Differences (ILD)**
    - Effective for high frequencies above 1.5 kHz
    - Head size > wavelength
    - Smallest detectable ILD = 0.5 dB



  - **Interaural Time Differences (ITD)**
    - Effective for low frequencies below 1.5 kHz
    - **Rayleigh's duplex theory of ILD and ITD**
    - Smallest detectable ITD = 13 μs

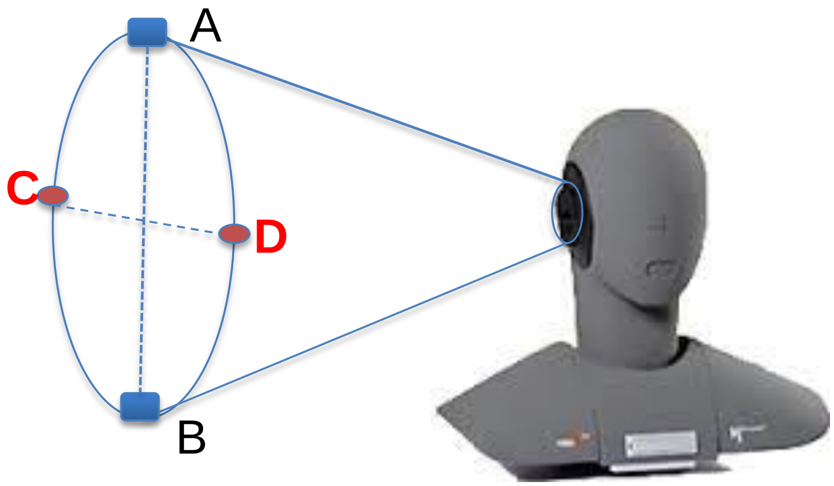


Pictures from W. M. Hartmann, "How we localize sound," Physics Today, 52(11), 24(1999)

Inadequacy of Interaural Difference Cues

- **Similar ITD and ILD due to:**
  - **Cone of confusion**

  Sources A & B; Sources C & D have identical ITD and ILD

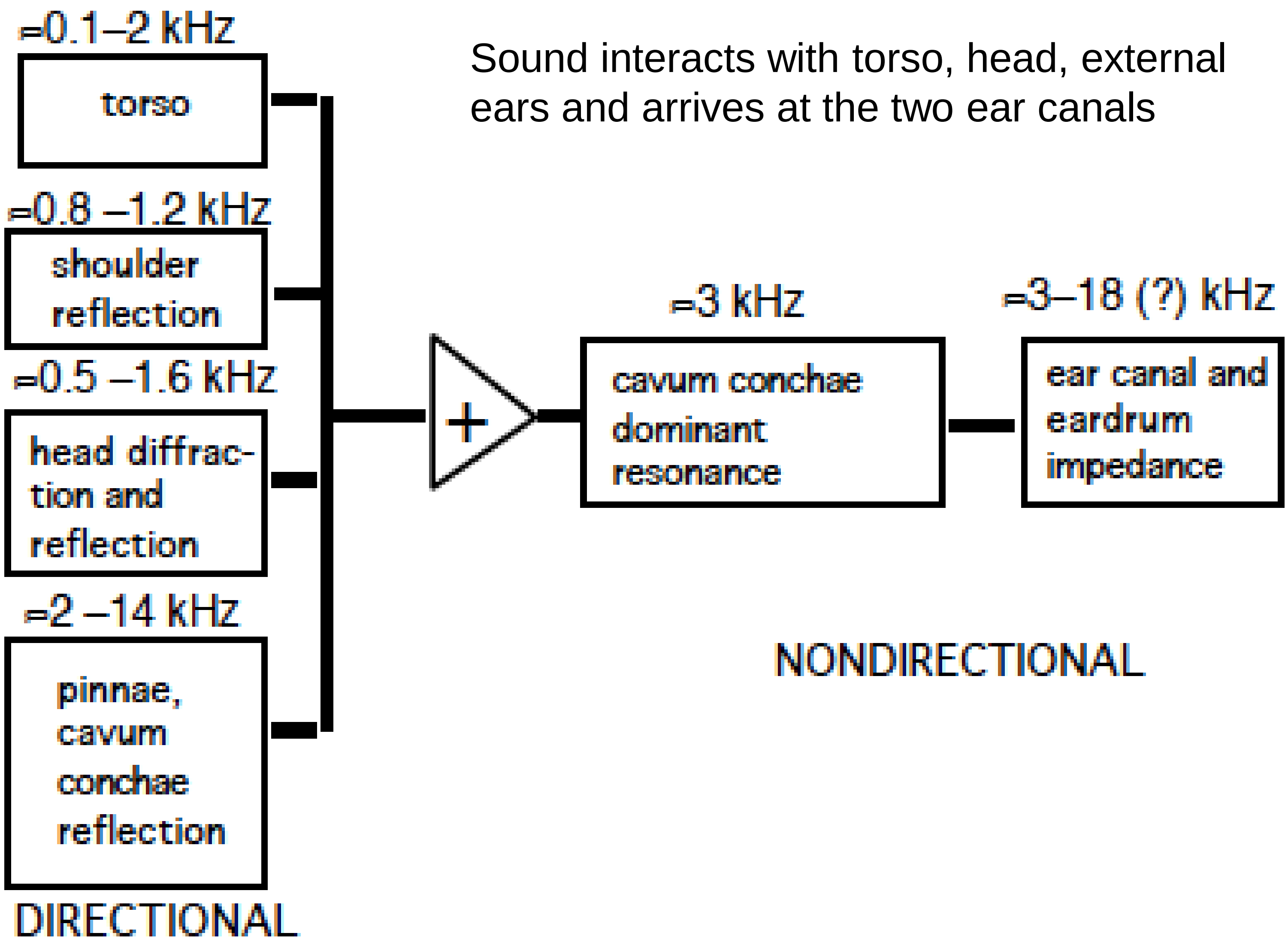


  - **Media Plane (extreme case of cone of confusion)**

- **We need another sound localization cue!**

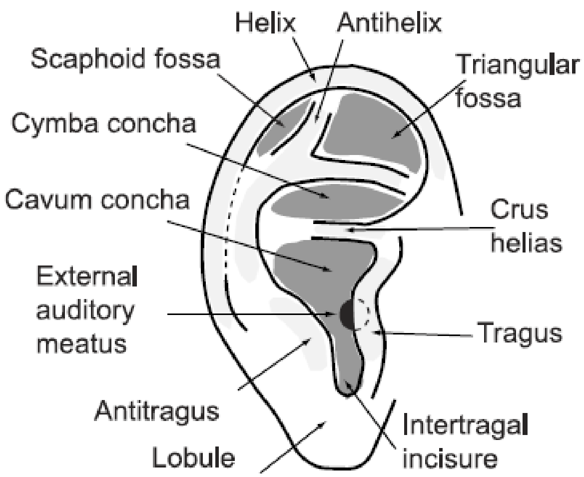
Modeling of Sound Scattering (Human body & ears)

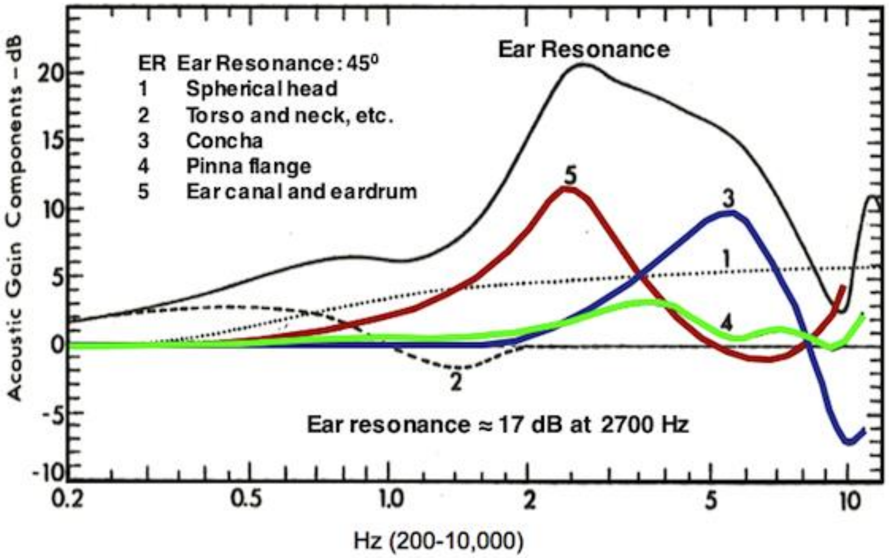
Sound interacts with torso, head, external ears and arrives at the two ear canals



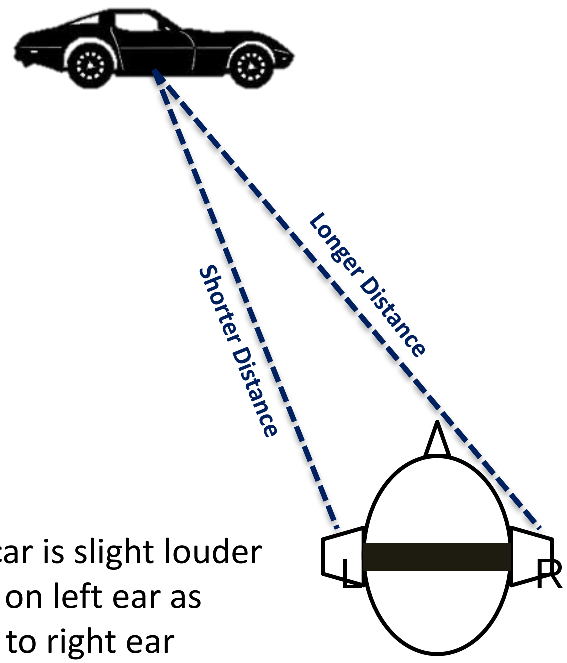
Main Binaural Cues due to Head Related Transfer Function (HRTF)

- Sound wave scatters of torso (~45 cm), head (~20cm) and ears (~4 cm)
- Also scatter off surrounding (>2m)
- Model scattering effect independently
- Ears act as directional acoustic probes.



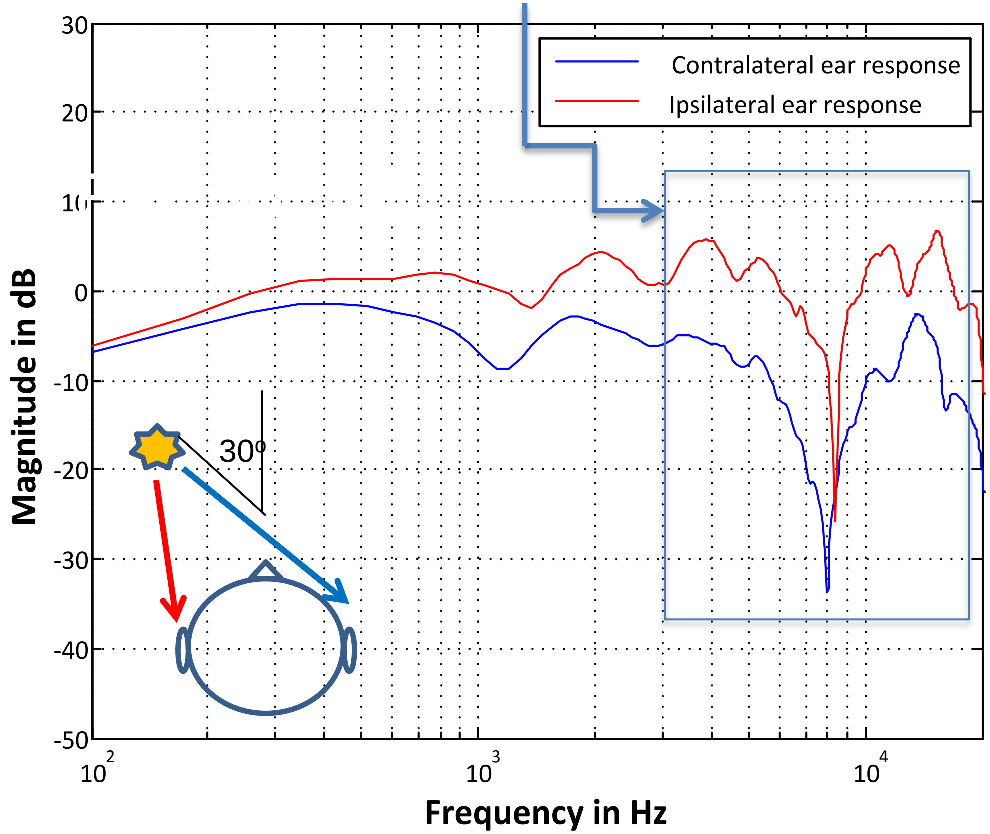


Head Related Transfer Functions (HRTF)



Sound of car is slight louder and faster on left ear as compared to right ear

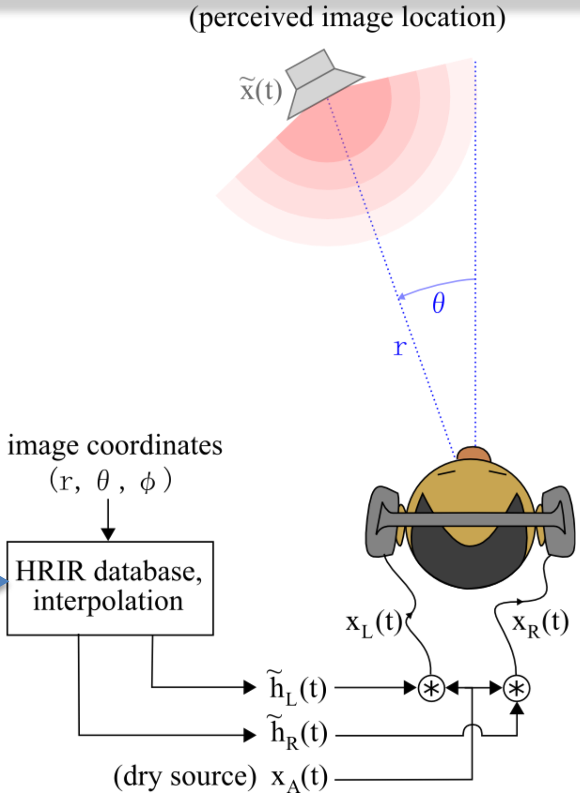
Many high-frequency details due to pinna scattering



HRTF of KEMAR dummy head for an angle of 30 degree azimuth*

*W. G. Gardner and K. D. Martin, "HRTF Measurements of a Kemar," Journal of the Acoustical Society of America, vol. 97, pp. 3907-3908, June 1995.

Implementation of 3D sound in headphones



Images from Wikipedia

- Head related impulse response (HRIR)
- Encodes the acoustic propagation between the sound source and the listener's ears
- Characteristic to the position of the sound source with respect to the listener
- Highly dependent on the morphology of listener.

Binaural Recording

Binaural Recording can be done either at the eardrum (or at blocked ear canal) of a listener or a dummy head

Binaural Recording on human subjects and dummy heads

- Played back using a stereo set of loudspeakers or a headphone
- Encapsulates all the directional information generated by the interaction of the sound with the listener's morphology

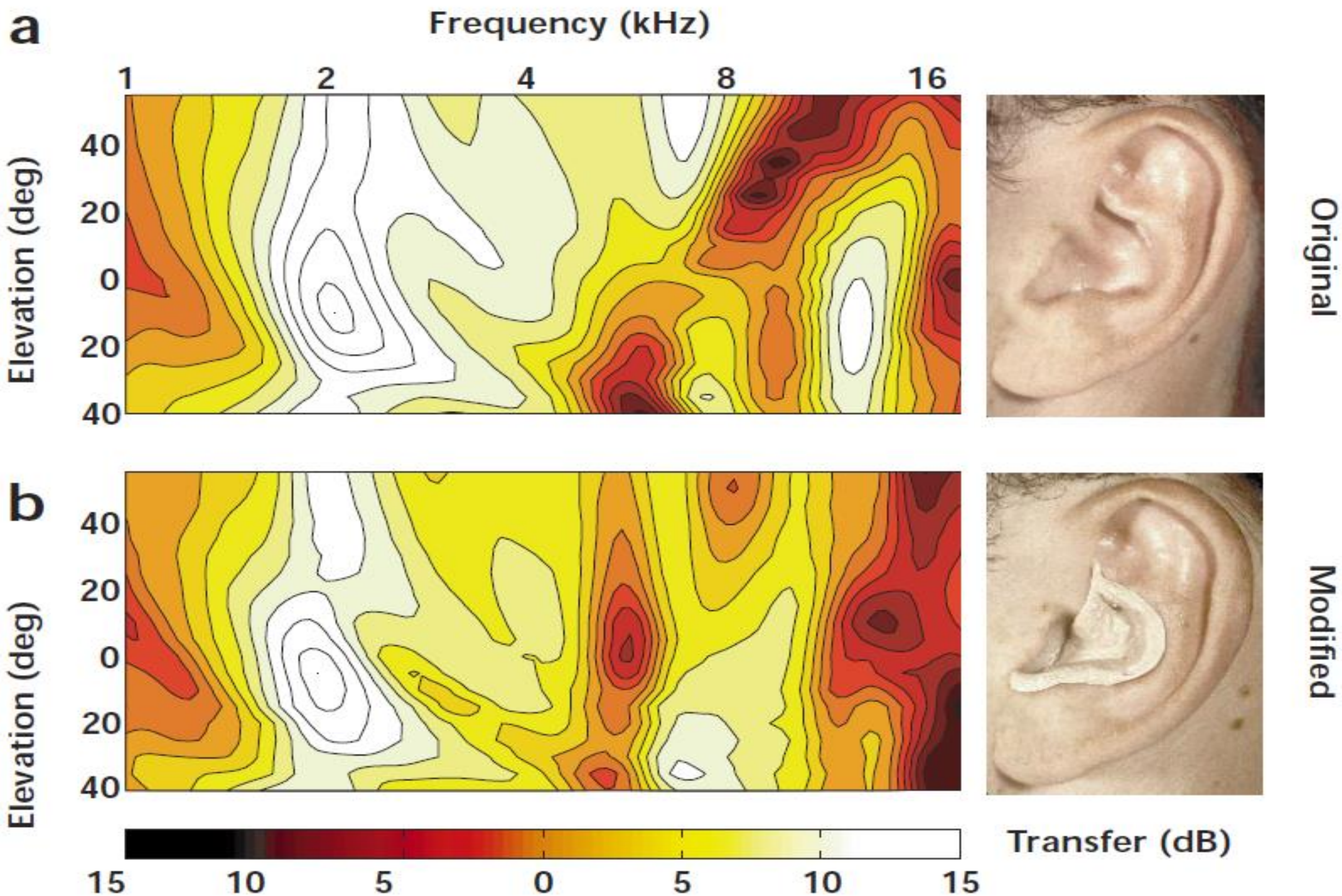
Individual Sound Filtering (Earprint)

Variation in Pinna morphology

Pinna of human subjects taken from the CIPIC database

- Human pinna is found to be as **idiosyncratic as the fingerprint**
- HRTFs are highly individual and differs substantially from one subject to the other
- For perfect 3D audio playback, **individualized recordings/HRTFs** and **individualized headphone equalization** are required

Highly Individualized Ear's Response



Picture extracted from
Paul M. Hofman, "Relearning sound localization with new ears," nature *neuroscience* • volume 1 no 5 • september 1998

CIPIC Anthropometry Measurements

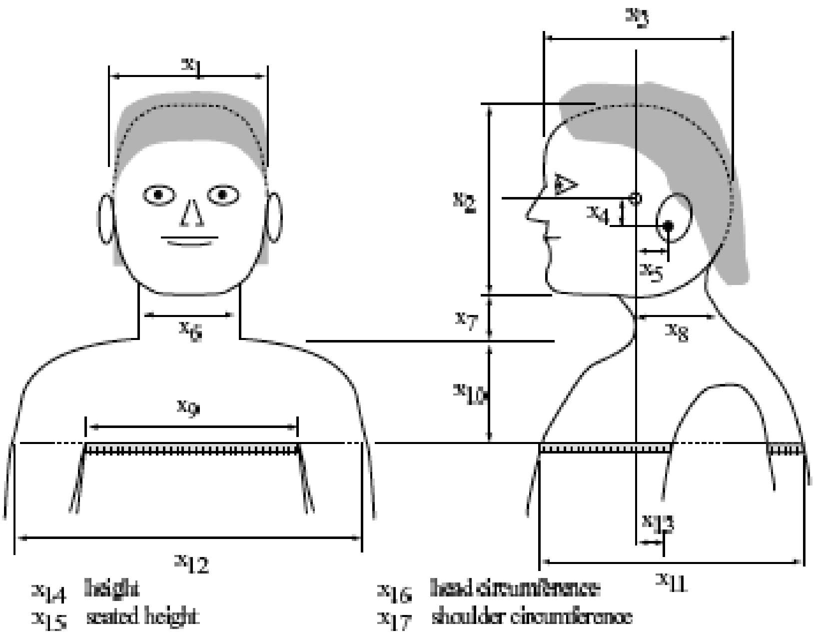


Figure 2: *Head and torso measurements*

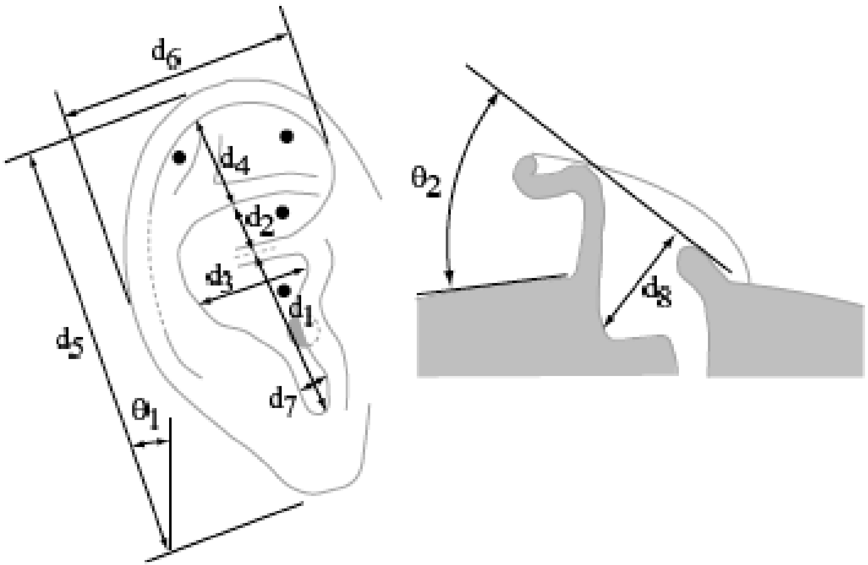


Figure 3: *Pinna measurements*

Var	Measurement	μ	σ	%
x_1	head width	14.49	0.95	13
x_2	head height	21.46	1.24	12
x_3	head depth	19.96	1.29	13
x_4	pinna offset down	3.03	0.66	43
x_5	pinna offset back	0.46	0.59	254
x_6	neck width	11.68	1.11	19
x_7	neck height	6.26	1.69	54
x_8	neck depth	10.52	1.22	23

3D Audio Using Stereo Headphones

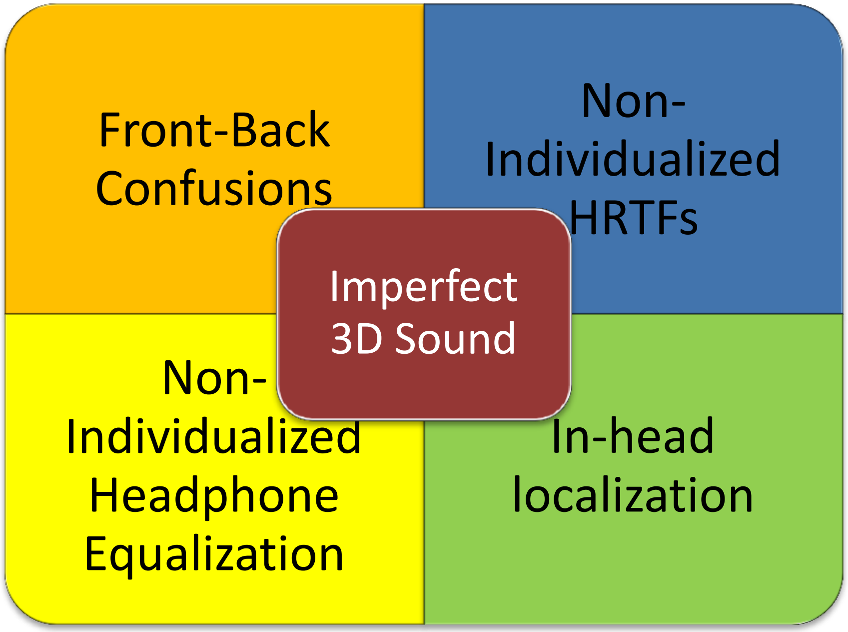
- Least expensive way to add realism to game and movie; instead of loudspeaker.
- Can produce intimately close sounds.

Image from: http://www.srstechnologies.com/content.aspx?id=426

3D audio reproduction using headphones is degraded due to front-back confusion and in-the-head localization. By overcoming these issues, we are able to faithfully recreate 3D audio using headphones (as in natural listening).

Binaural audio reproduction over headphones

Binaural audio is highly idiosyncratic



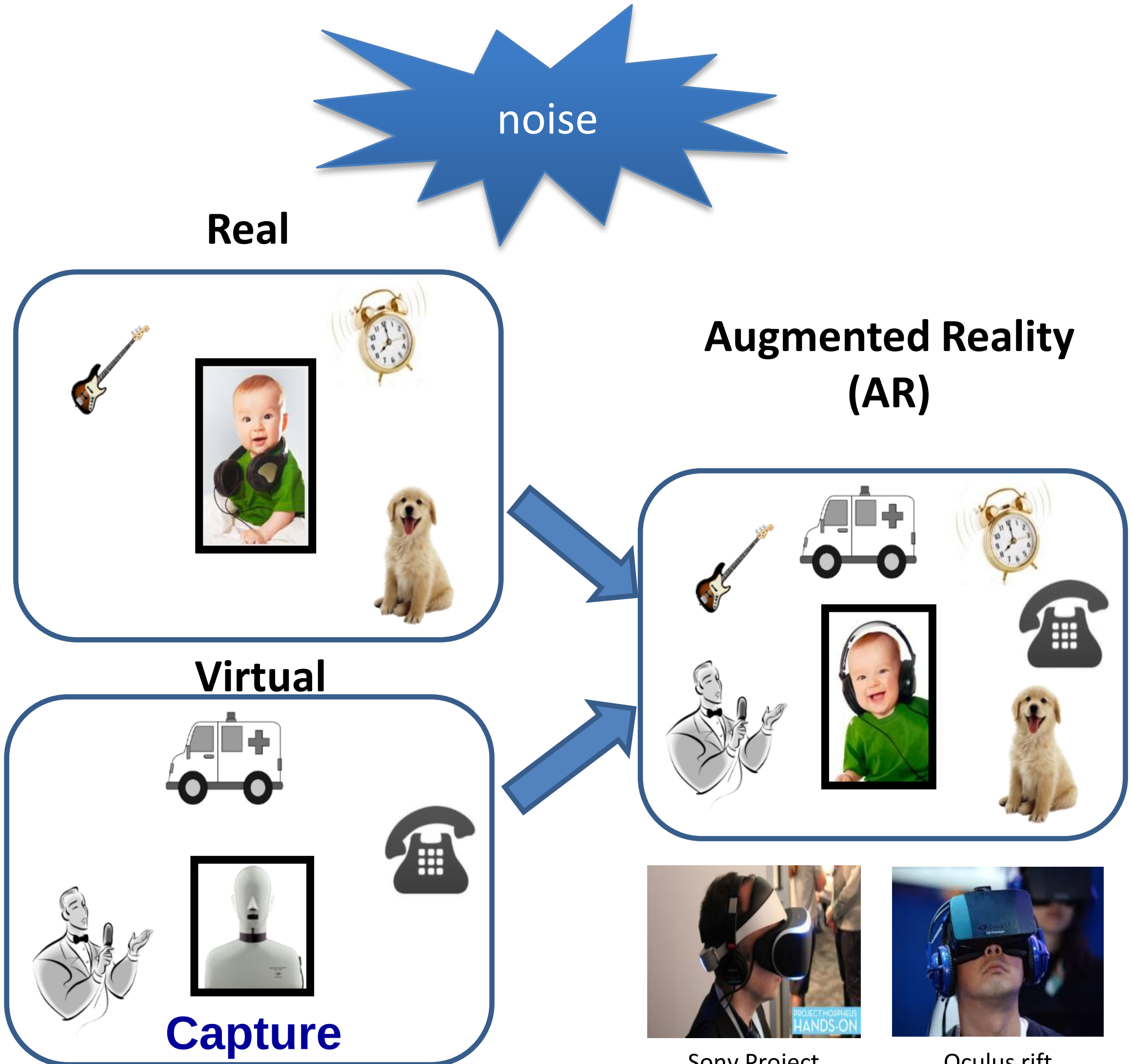
Use of non-individual HRTFs degrades the veracity of the perception of 3D sound

Solution:

Individualizing the non-individual HRTFs can improve the 3D audio playback using headphones

Other Challenges and New Applications

- Natural Sound Rendering
- Active Noise Reduction
- Extension to Augmented Reality Applications
- Assisted Listening
- **Binaural Signal Processing…**



Sony Project Morpheus

Oculus rift

Benefits of Spatial (binaural)Listening

Binaural Localization →	**Improve situation awareness and Provide augmented information to visual**
Cocktail party effect →	**Improve intelligibility for remote sensing and Better discriminate or selective attention**
Active Noise Control →	**Better conversation and ability to hear others in telecommunication.**
Natural Listening →	**Improve immersive listening and enhance "being there" experience**

Adapted from Begault

Key References on Fundamentals of 3D Sound

[1] D. R. Begault, *3-D sound for virtual reality and multimedia*: AP Professional, 2000.

[2] V. R. Algazi and R. O. Duda, "Headphone-based spatial sound," *Signal Processing Magazine, IEEE,* vol. 28, no. 1, pp. 33-42, Jan. 2011.

[3] R. Nicol, *Binaural Technology*: AES, 2010.

[4] H. Møller, M. F. Sørensen, D. Hammershøi, and C. B. Jensen, "Head-related transfer functions of human subjects," *J. Audio Eng. Soc.,* vol. 43, no. 5, pp. 300-321, May 1995.

[5] F. Rumsey, Spatial Audio. Oxford, UK: Focal Press, 2001.

[6] T. Holman, Surround sound up and running 2nd ed., MA: Focal Press, 2008.

[7] J. Blauert, Spatial hearing: The psychophysics of human sound localization. Cambridge, MA, USA: MIT Press, 1997.

[8] B. S. Xie, Head-related transfer function and virtual auditory display, 2nd edition. J. Ross Publishing, US, 2013.

[9] W. G. Gardner, and K. D. Martin, “HRTF Measurements of a KEMAR,” J. Acoust. Soc. Am., vol., vol. 97, pp. 3907-3908, 1995.

[10] W. Gardner, “3-D audio using loudspeakers,” PhD thesis, School of Architecture and planning, MIT, USA, 1997.

[11] V. R. Algazi, R. O. Duda, D. M. Thompson, and C. Avendano, “The CIPIC HRTF database,” in Proc. IWASPAA, New Paltz, NY, USA, Oct. 2001.

[12] J. Blauert (Ed.), The Technology of Binaural Listening, Springer-Verlag Berlin Heidelberg, 2013.

[13] J. Breebaart, and C. Faller, Spatial audio processing: MPEG Surround and other applications, John Wiley & Sons, 2007

[14] H. Møller, “Fundamentals of binaural technology,” Appl. Acoust., vol. 36, pp. 171–218, 1992.

Module II
Natural Sound Rendering for Headphones

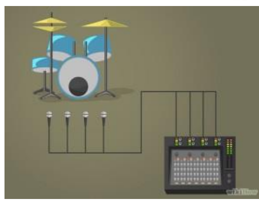
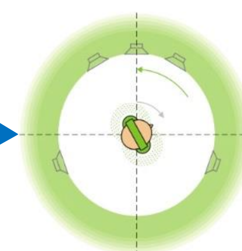
K. Sunder, J. He, E. L. Tan, and W. S. Gan, "Natural sound rendering for headphones," IEEE Signal Processing Magazine, vol. 32, no. 2, pp. 100-113, Mar. 2015.

To achieve natural sound rendering in headphones

Natural sound rendering essentially refers to rendering of the spatial sound using headphones to create an immersive listening experience and the sensation of "being there" at the venue of the acoustic event.

- **Differences** between natural listening and headphone listening;
- **Challenges** for rendering sound in headphone to mimic natural listening;
- How can **signal processing** techniques help?
- How to **integrate** these techniques?
- And **subjective evaluation**

Challenges and solutions

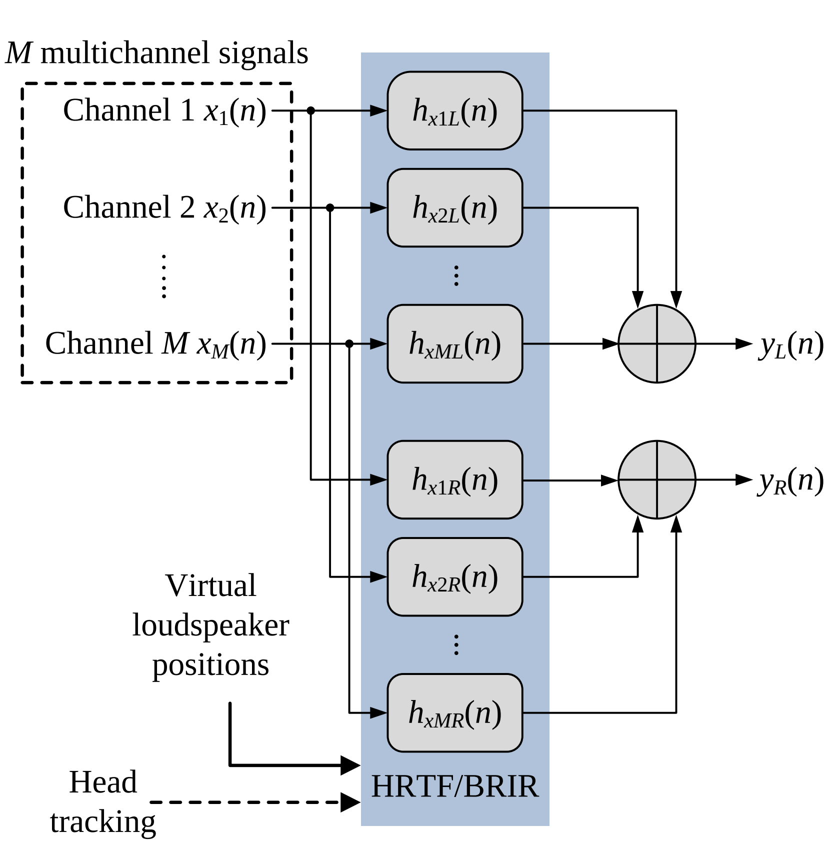
Listening	Headphone listening	Signal processing	Natural listening
Source		Virtualization	
		Sound scene decomposition	
Medium		Equalization	
Receiver		Individualization	
		Head tracking	

D. R. Begault, *3-D sound for virtual reality and multimedia*: AP Professional, 2000.

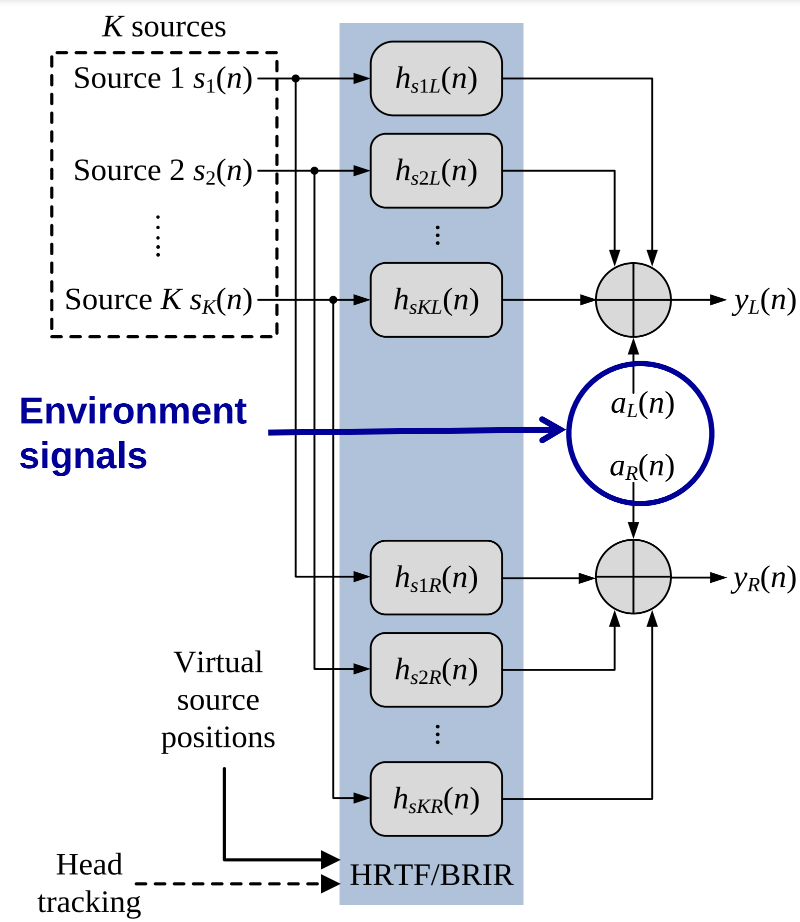
Signal processing techniques

1. **Virtualization**:
   - to match the desired playback for the digital media content;
2. **Sound scene decomposition**:
   - to optimally facilitate the separate rendering of sound sources and sound environment;
3. **Individualization**:
   - to compensate for the lost or altered individual filtering of sound in headphone listening;
4. **Equalization**:
   - to preserve the original timbral quality of the source and alleviate the adverse effect of the inherent headphone response;
5. **Head tracking**:
   - to adapt to the dynamic head movements of the listener.

Virtualization



(a) Virtualization of multichannel loudspeaker signals



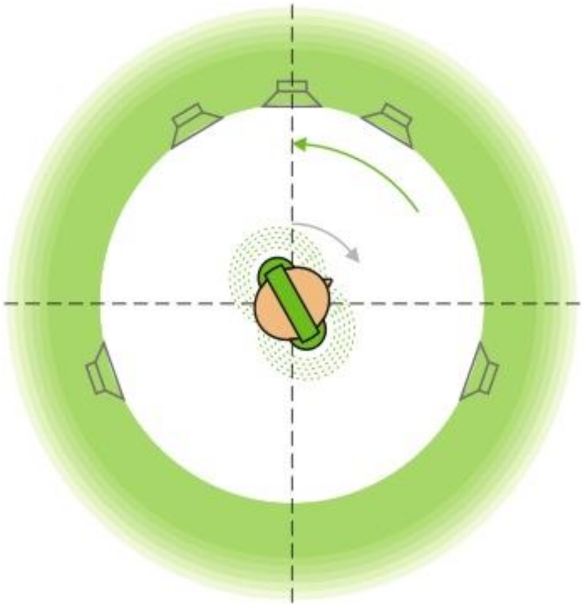
(b) Virtualization of source and environment signals

J. Breebaart and E. Schuijers, "Phantom materialization: a novel method to enhance stereo audio reproduction on headphones," *IEEE Trans. Audio, Speech, and Language Processing,* vol. 16, no.8, pp. 1503-1511, Nov. 2008.

Virtualization with head tracking

- Head movement information is tracked by a sensor (e.g., accelerometer, gyroscope, magnetometer, camera);
- Adapt to the changes of sound scene with respect to head movements;
- Reduce front-back confusions, azimuth localization errors;
- Concern of head tracking latency (80ms).

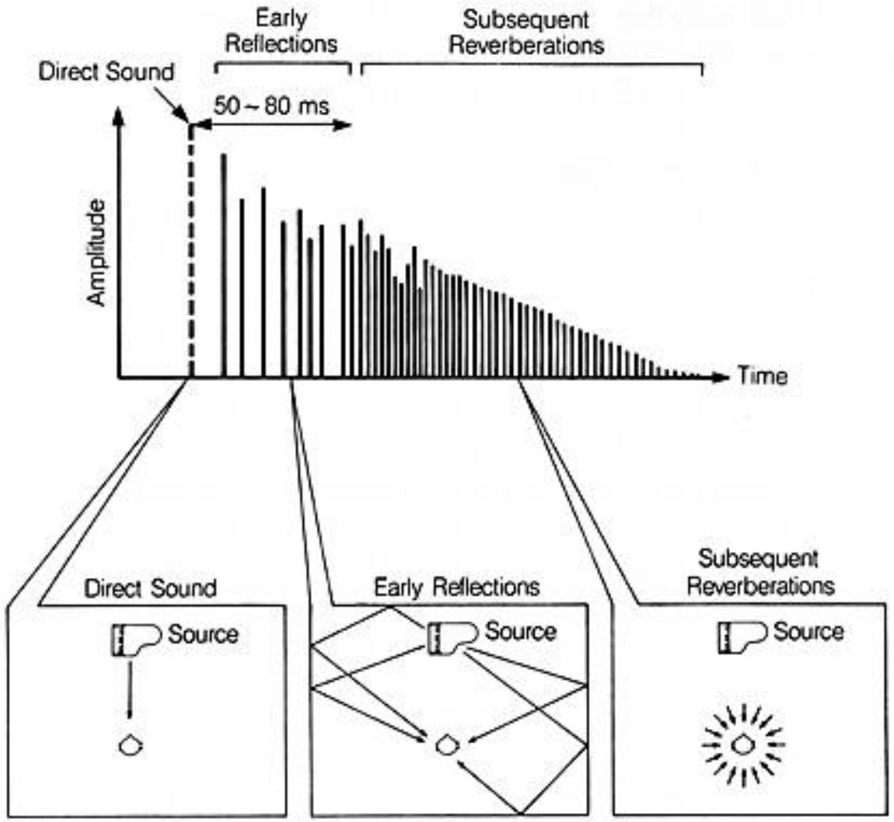
Source from http://3dsoundlabs.com/en/



Source from north-america.beyerdynamic.com

Virtualization: further considerations

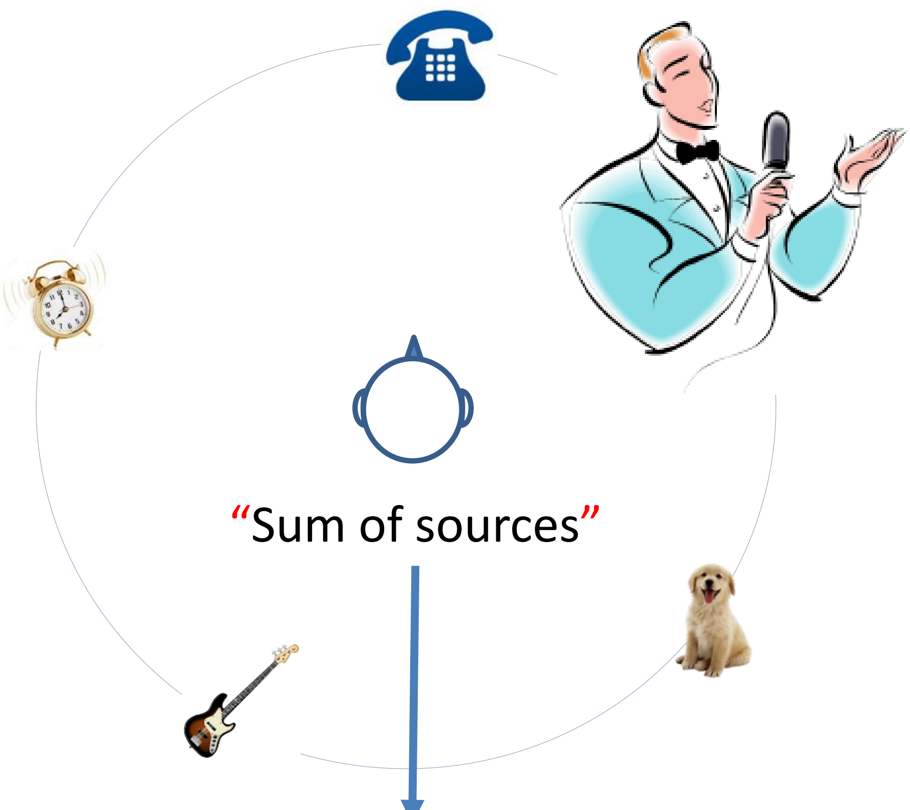
- Add reverberation
  - Externalization of the sound sources, and enhance depth perception;
  - Rendering of the sound environment;
  - How to select correct amount of reverberation.



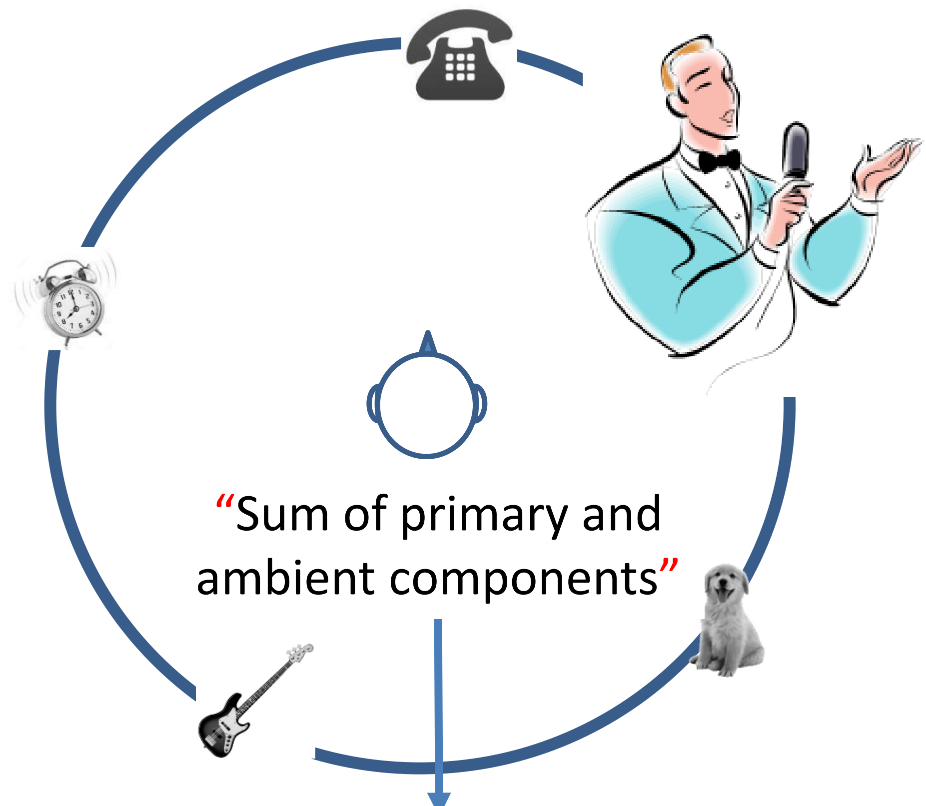
Source from
http://www.torgny.biz/Recording%20sound_2.htm

Sound scene decomposition: overview

Aim: to obtain useful information about the original sound scene from given mixtures, and facilitate natural sound rendering.



Blind source separation



Primary ambient extraction

K. Sunder, J. He, E. L. Tan, and W. S. Gan, "Natural sound rendering for headphones," IEEE Signal Processing Magazine, Mar. 2015.

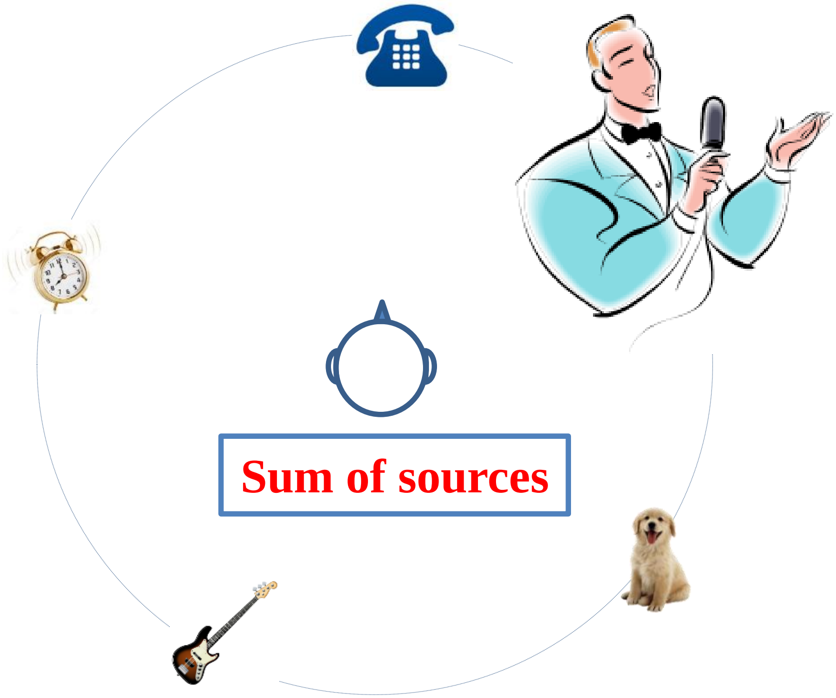
Sound scene decomposition: comparison

Techniques	Blind Source Separation	Primary Ambient Extraction
Objective	To obtain useful information about the original sound scene from given mixtures, and facilitate natural sound rendering	
Basic model	1. Multiple sources sum together 2. Sources are independent	1. Dominant sources + Environmental signal 2. Primary components are highly correlated; 3. Ambient components are uncorrelated
Common characteristics	1. Usually no prior information, only mixture signals 2. Perform extraction/separation based on various signal models 3. Require objective as well as subjective evaluation	
Typical applications	Speech, music	Movie, gaming
Limitations	1. Small number of sources 2. Sparseness/disjoint 3. No/simple environment	1. Small number of sources 2. Sparseness/disjoint 3. Low ambient power 4. Primary ambient uncorrelated

K. Sunder, J. He, E. L. Tan, and W. S. Gan, “Natural sound rendering for headphones,” IEEE Signal Processing Magazine, Mar. 2015.

Sound scene decomposition: BSS

Objective:
to extract the K sources from M mixtures

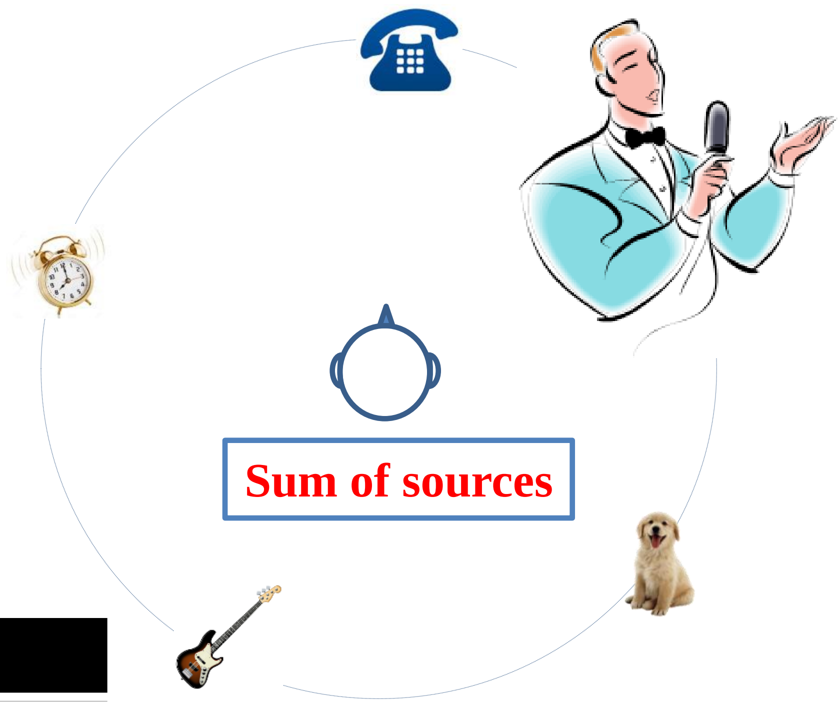


Mixtures = function (gain, source, time difference, model error)

$$x_m(n) = \sum_{k=1}^{K} g_{mk} s_k(n - \tau_{mk}) + e_m(n), \quad \forall m \in \{1, 2, \ldots, M\}$$

Sound scene decomposition: BSS

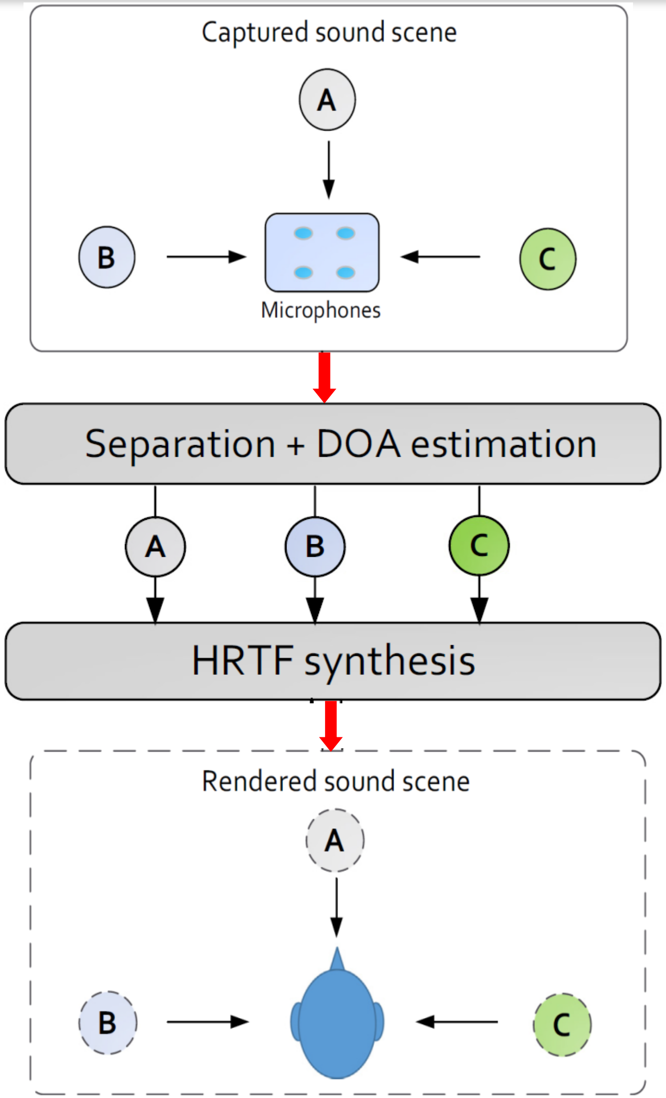
Objective:
to extract the K sources from M mixtures



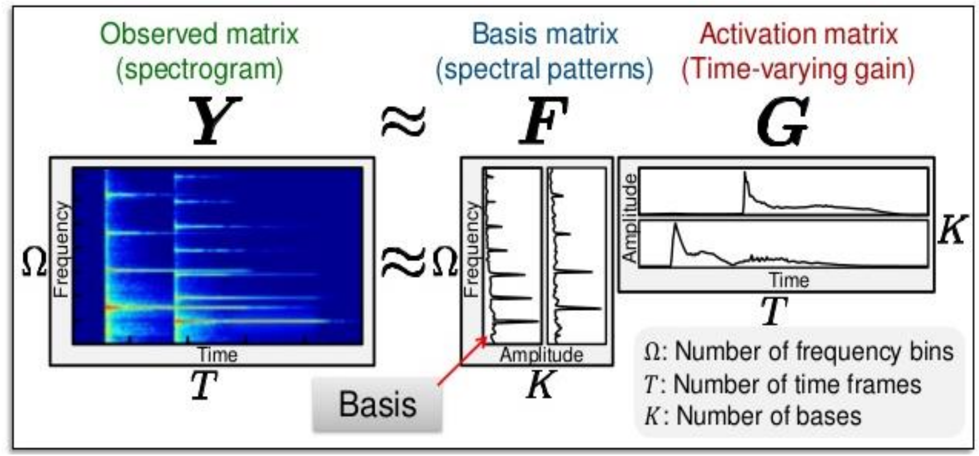
Case		Typical techniques
M = K		ICA
M > K		ICA with PCA, Least-squares
M < K	M > 2	ICA with sparse solutions
	M = 2	Time-frequency masking
	M = 1	NMF, CASA

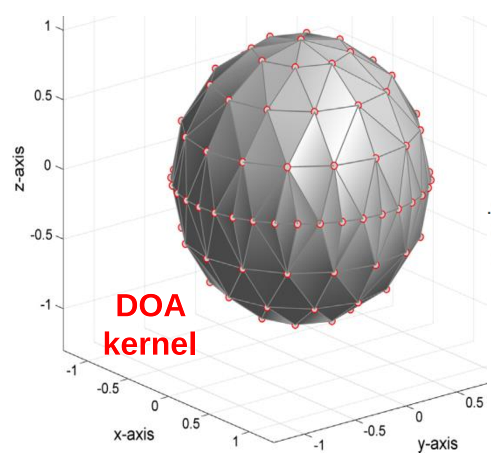
ICA : Independent component analysis
PCA : Principal component analysis
NMF : Non-negative matrix factorization;
CASA: Computational auditory scene analysis

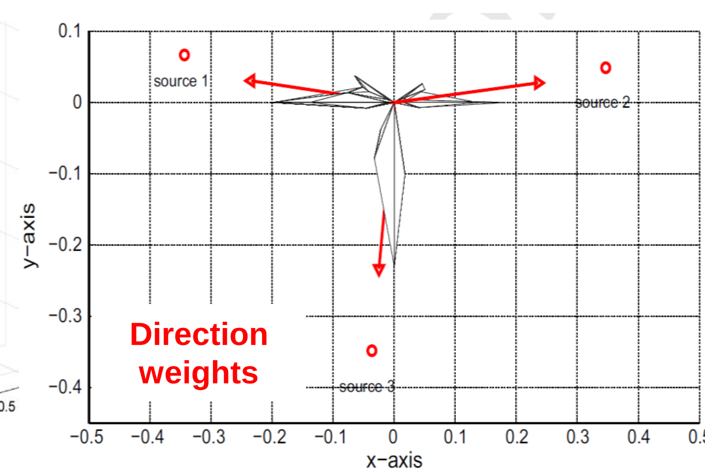
One example using NMF



NMF



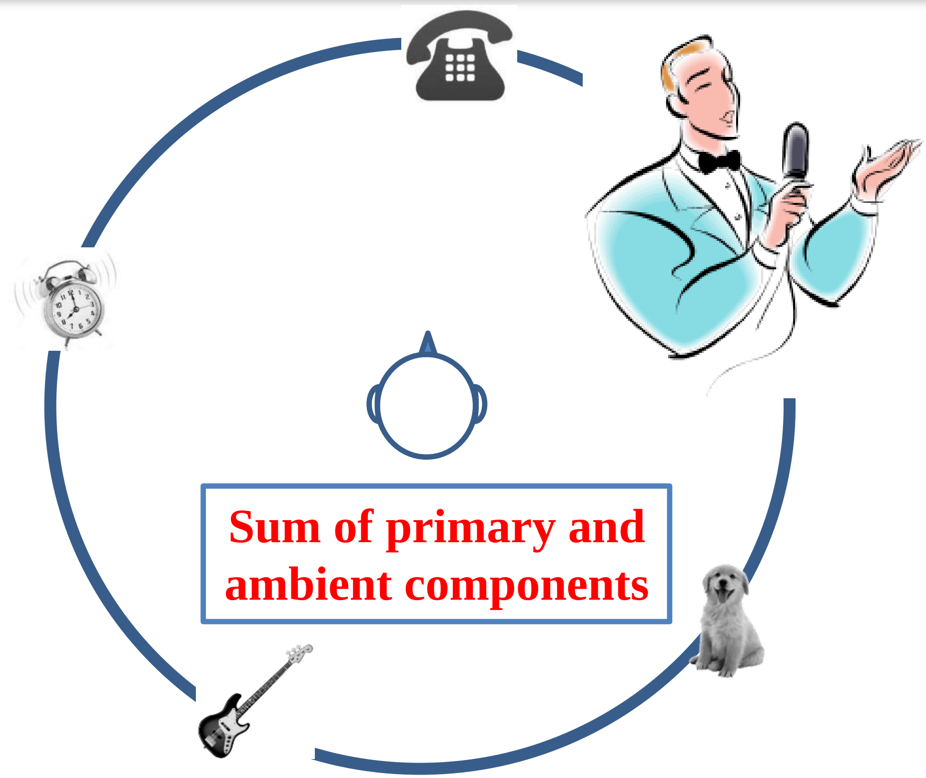




J. NiKunen et al. “Binaural rendering of microphone array captures based on source separation,” Speech Communication, 2015.

Sound scene decomposition: PAE

Objective:
to extract the primary and ambient components from M (M = 2, stereo) mixtures

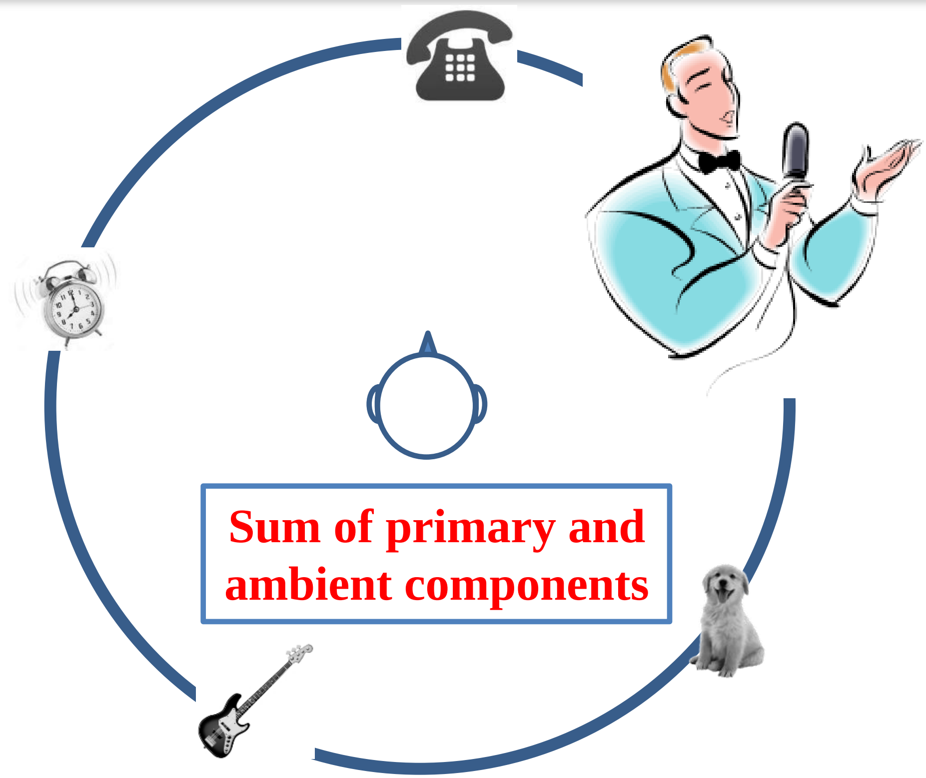


Mixtures = primary component + ambient component

$$x_m(n) = p_m(n) + a_m(n)$$

Sound scene decomposition: PAE

Objective:
to extract the primary and ambient components from M (M = 2, stereo) mixtures



Case		Typical techniques
Basic model	Channel-wise	Time frequency masking
	Combine channels	Linear estimation, Ambient spectrum estimation
Complex model		Time/phase shifting, Classification, Subband, Pairing up two channels, etc.

Definitions with Stereo Signal Model

Signal = Primary + Ambient

$$\mathbf{x}_0 = \mathbf{p}_0 + \mathbf{a}_0$$

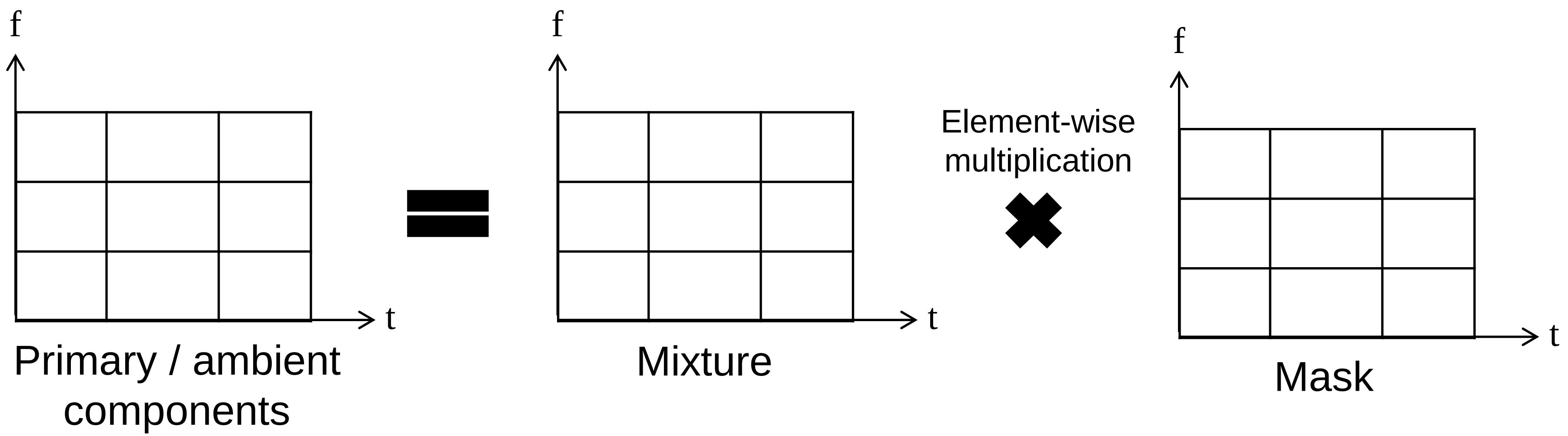
$$\mathbf{x}_1 = \mathbf{p}_1 + \mathbf{a}_1$$

Assumptions

Primary components highly correlated	$\mathbf{p}_1 = k\mathbf{p}_0$
Ambient components uncorrelated	$\mathbf{a}_0 \perp \mathbf{a}_1$
Primary ambient components uncorrelated	$\mathbf{p}_i \perp \mathbf{a}_j$
Ambient power balanced	$P_{\mathbf{a}_0} = P_{\mathbf{a}_1}$

J. He, E. L. Tan and W. S. Gan, "Linear estimation based primary-ambient extraction for stereo audio signals," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 22, no. 2, pp. 505-517, Feb. 2014.

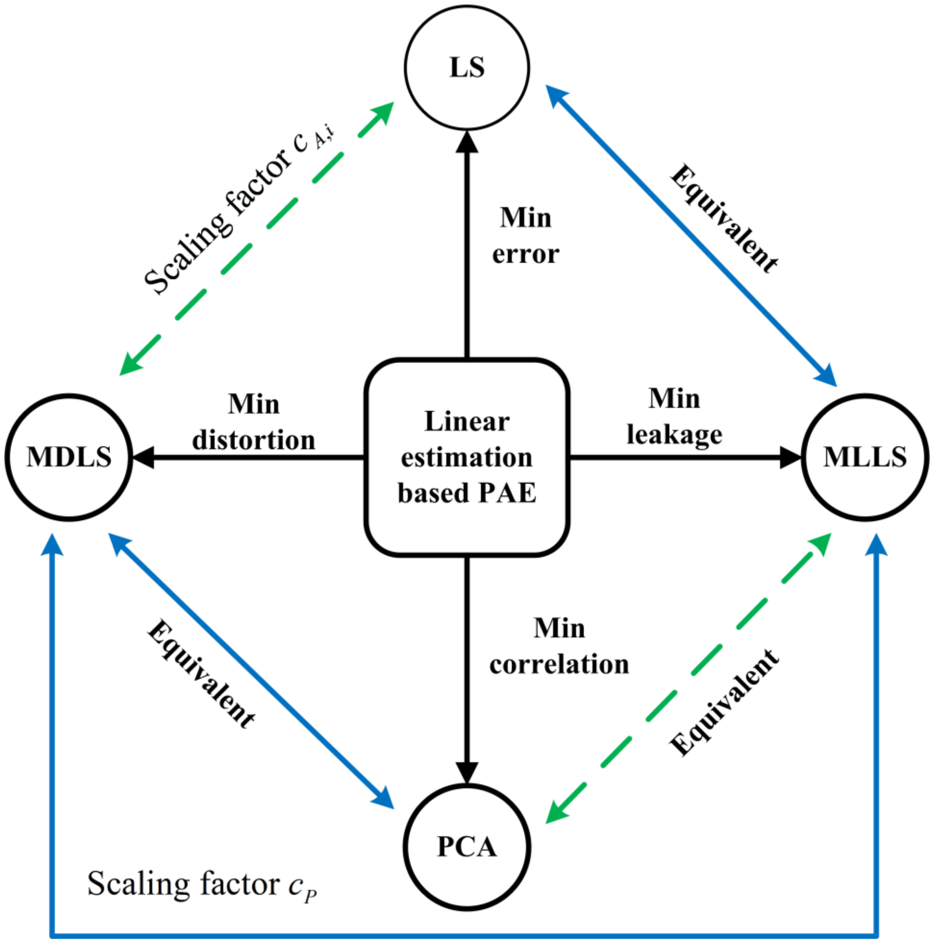
PAE: time frequency masking



Mask can be constructed using

- **Inter-channel coherence [Avendano and Jot, 2004]**
- **Pairwise correlation [Thompson et al., 2012]**
- **Equal level of ambience [Merimaa et al., 2007]**
- **Diffuseness [Pulkki, 2007]**

PAE: linear estimation



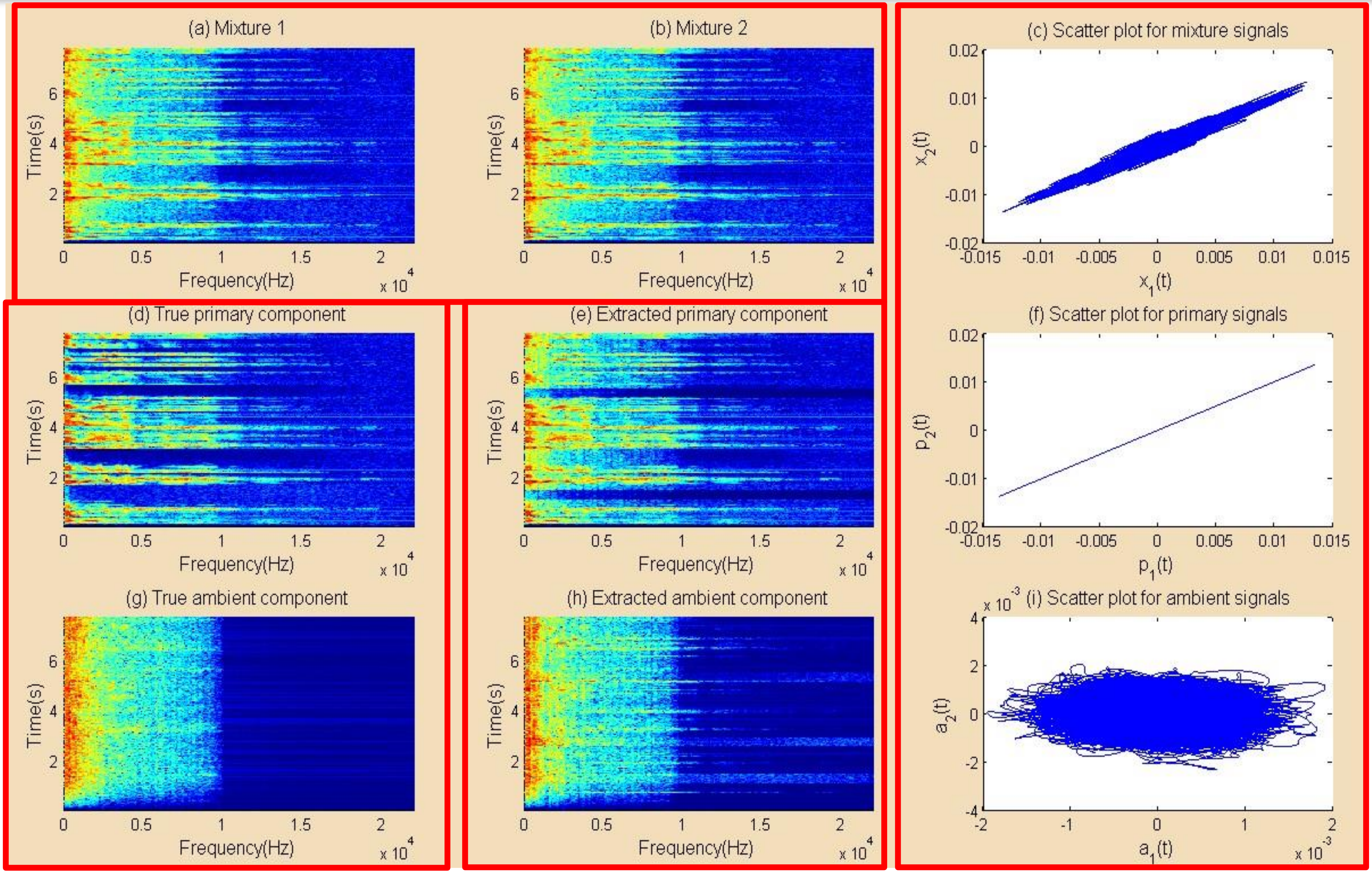
$$\begin{bmatrix} \hat{p}_0(n) \\ \hat{p}_1(n) \\ \hat{a}_0(n) \\ \hat{a}_1(n) \end{bmatrix} = \begin{bmatrix} w_{\mathrm{P}0,0} & w_{\mathrm{P}0,1} \\ w_{\mathrm{P}1,0} & w_{\mathrm{P}1,1} \\ w_{\mathrm{A}0,0} & w_{\mathrm{A}0,1} \\ w_{\mathrm{A}1,0} & w_{\mathrm{A}1,1} \end{bmatrix} \begin{bmatrix} x_0(n) \\ x_1(n) \end{bmatrix}$$

Objectives and relationships of four linear estimation based PAE approaches.

- **Blue** solid lines represent the relationships in the **primary** component;
- **Green** dotted lines represent the relationships in the **ambient** component.
- **MLLS**: minimum leakage LS
- **MDLS**: minimum distortion LS

J. He, E. L. Tan, and W. S. Gan, "Linear estimation based primary-ambient extraction for stereo audio signals," *IEEE/ACM Trans. Audio, Speech, and Language Processing,* vol. 22, no.2, pp. 505-517, 2014.

PAE: an example from least-squares



PAE: ambient spectrum estimation

Signal model $\mathbf{X}_0 = \mathbf{P}_0 + \mathbf{A}_0,\ \mathbf{X}_1 = \mathbf{P}_1 + \mathbf{A}_1 \qquad \mathbf{A}_c = |\mathbf{A}| \square e^{j\boldsymbol{\theta}_c},\ \forall c \in \{0,1\},$

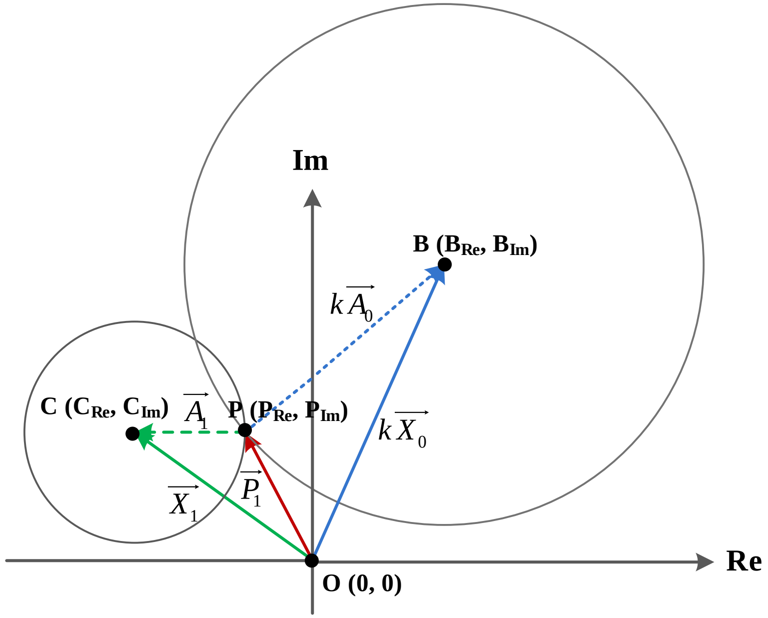
Ambient Phase Estimation (APE)

$$\mathbf{A}_c = (\mathbf{X}_1 - k\mathbf{X}_0)./(e^{j\boldsymbol{\theta}_1} - ke^{j\boldsymbol{\theta}_0}) \square e^{j\boldsymbol{\theta}_c},$$

$$\mathbf{P}_c = \mathbf{X}_c - \mathbf{A}_c, \forall c \in \{0,1\}.$$

⇨ Find $\boldsymbol{\theta}_0, \boldsymbol{\theta}_1$

Ambient Magnitude Estimation (AME)



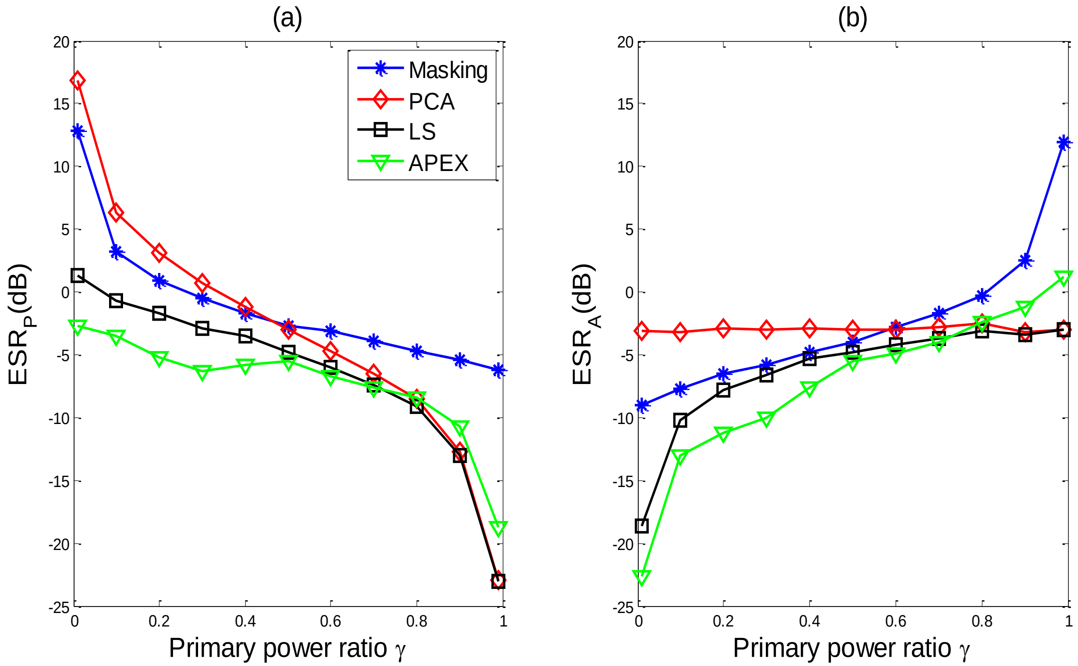
⇨ Find $\mathbf{r} = |\mathbf{A}|$

Sparsity constraint $\text{APES}: \hat{\theta}_1^* = \arg\min_{\hat{\theta}_1} \|\hat{\mathbf{P}}_1\|_1, \quad \text{or} \quad \text{AMES}: \hat{\mathbf{r}}^* = \arg\min_{\hat{\mathbf{r}}} \|\hat{\mathbf{P}}_1\|_1$

J. He, E. L. Tan, and W. S. Gan, "Primary-ambient extraction using ambient spectrum estimation for immersive spatial audio reproduction," IEEE/ACM Trans. Audio, Speech, Lang. Process., vol. 23, no. 9, pp. 1431-1444, Sept. 2015.

PAE: ambient spectrum estimation using sparsity

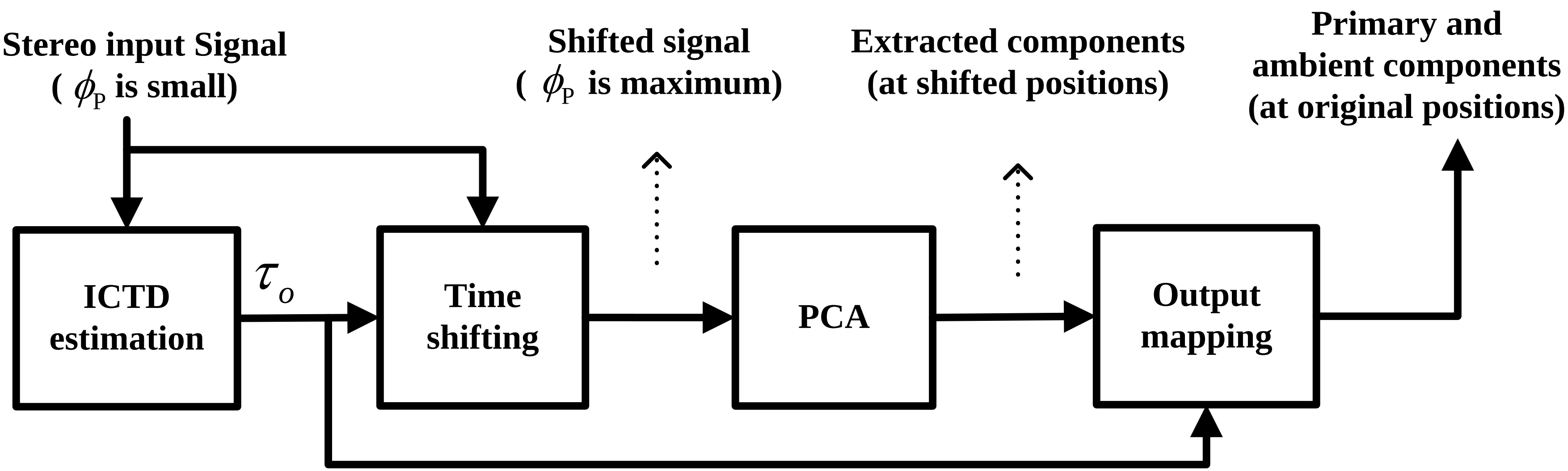
$$\text{Approximate efficient solution APEX}: \hat{\boldsymbol{\theta}}_1^* = \begin{cases} \angle \mathbf{X}_1 & ,\forall k > 1 \\ \angle(\mathbf{X}_1 - \mathbf{X}_0), & \forall k = 1 \end{cases}$$



Performance of Ambient Spectrum Estimation approaches:

- **Lower estimation error (ESR reduction: 3-6 dB average);**
- **Robust to variation on ambient magnitude difference (up to 10 dB);**
- **Validated in subjective listening tests.**

PAE: time shifting



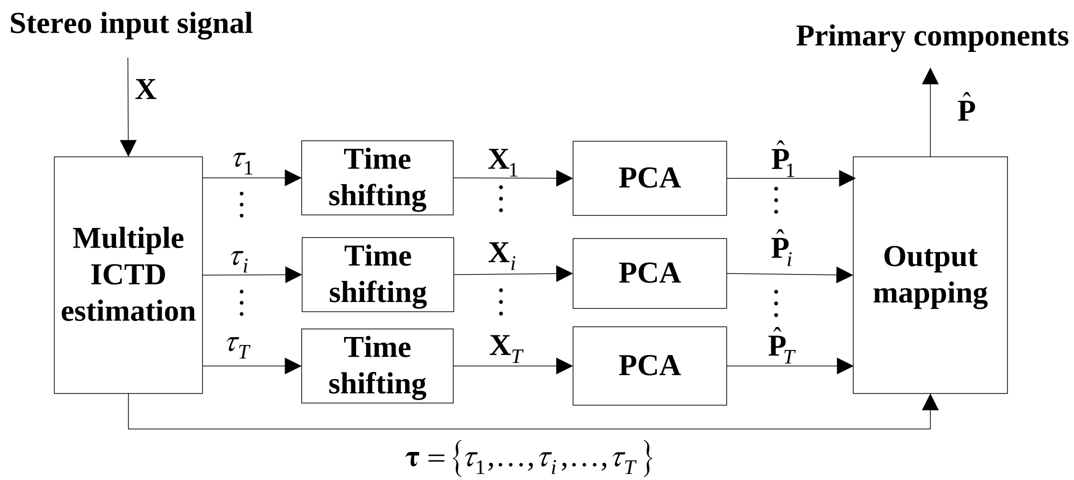
For mixture signals with partially correlated primary components

- **More accurate estimation of model parameter;**
- **Lower extraction error;**
- **Closer estimation of the spatial attributes;**
- **Increase of computational load.**

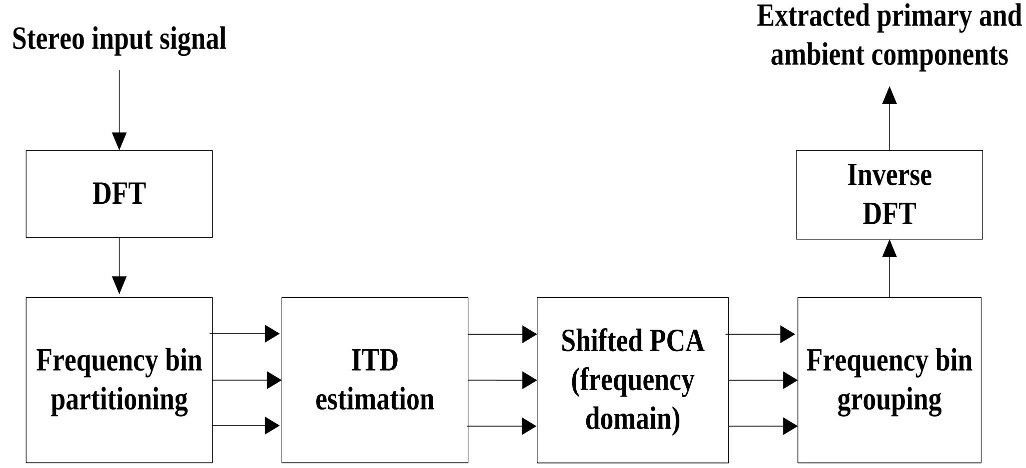
J. He, W. S. Gan, and E. L. Tan, "Time-shifting based primary-ambient extraction for spatial audio reproduction," IEEE/ACM Trans. Audio, Speech, Lang. Process., vol. 23, no. 10, pp. 1576-1588, Oct. 2015.

PAE: multiple sources

Multi-shifting PAE with ICC based output weighting



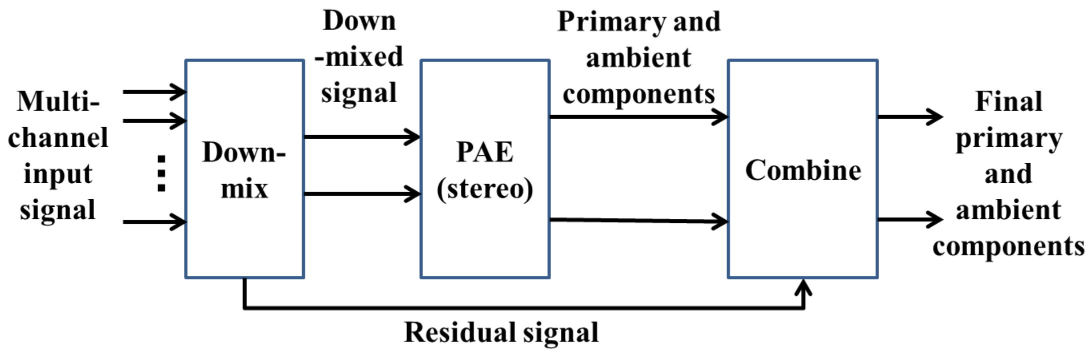
Subband PAE with frequency bin partitioning



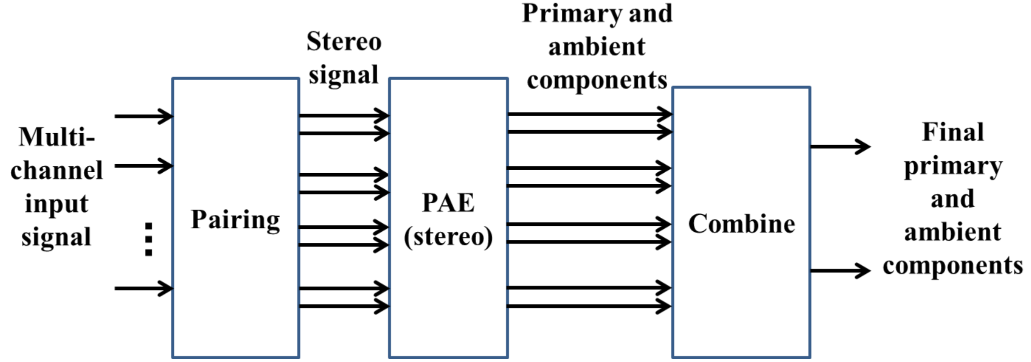
J. He, and W. S. Gan, “Multi-shift principal component analysis based primary component extraction for spatial audio reproduction,” in Proc. ICASSP, Brisbane, Australia, Apr. 2015, pp. 350-354.
J. He, E. L. Tan, and W. S. Gan, "A study on the frequency-domain primary-ambient extraction for stereo audio signals," in *Proc. ICASSP*, Florence, Italy, 2014, pp. 2892-2896.

PAE: from stereo to multiple channels

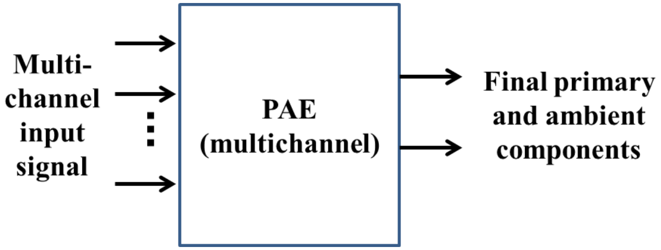
1. Using down-mix



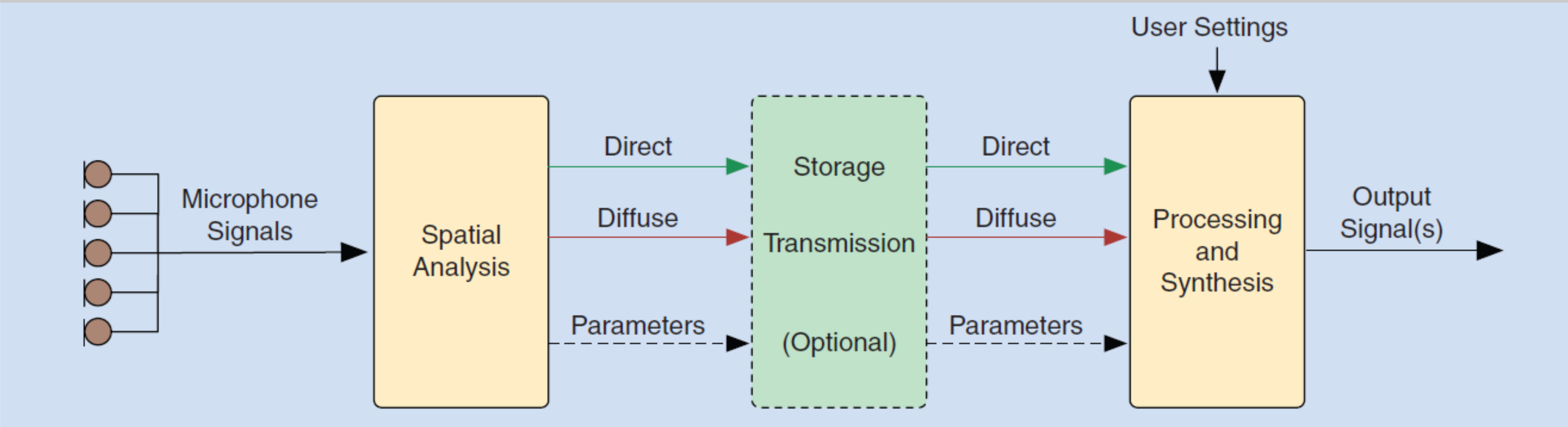
2. Using pairing

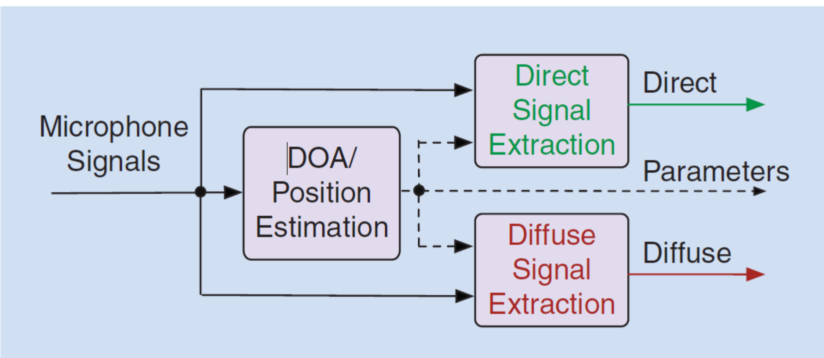


3. Direct



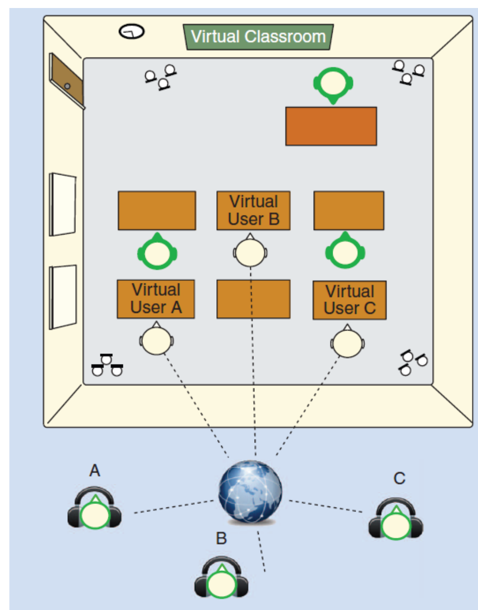
Related: Parametric spatial sound processing



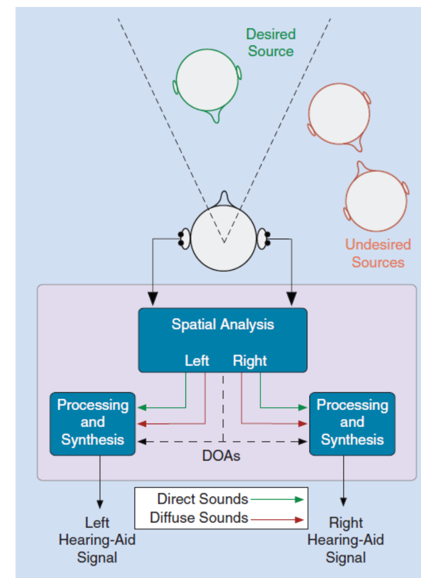


Signal extraction:

1. Single-channel wiener filters;
2. Multiple-channel filters: LCMV;



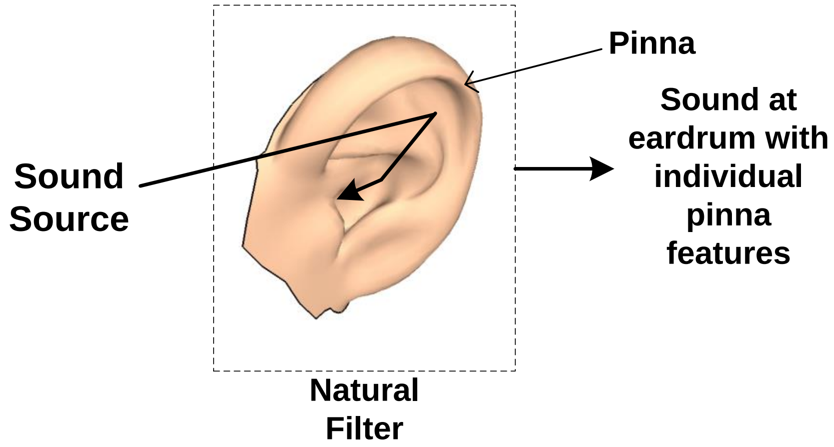
[FIG7] A virtual classroom scenario.

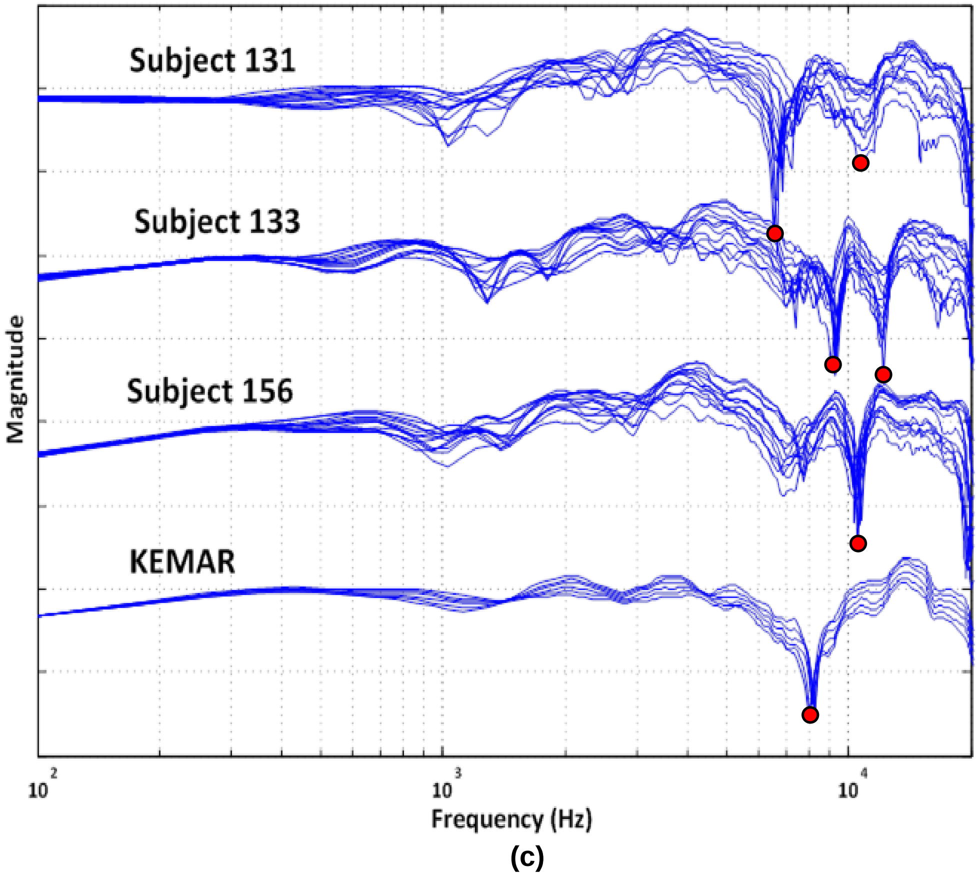


[FIG8] A general parametric spatial sound processing scheme for binaural hearing aids.

Konrad Kowalczyk, Oliver Thiergart, Maja Taseska, Giovanni Del Galdo, Ville Pulkki, and Emanuël A.P. Habets, "Parametric Spatial Sound Processing," IEEE Signal Processing Magazine, vol. 32, no. 2, Mar 2015.

Individualization: the need





Variation of HRTFs (Idiosyncratic)

S. Xu, Z. Li, and G. Salvendy, "Individualization of head-related transfer function for three-dimensional virtual auditory display: a review," in *Virtual Reality*, ed: Springer, 2007, pp. 397-407.

Overview of HRTF individualization techniques

To obtain individualized HRTF/perception

- ❑ Acoustical measurements
- ❑ Anthropometry
- ❑ Training/tuning
- ❑ Frontal projection

Individualization: acoustical measurements

Air Force Research Laboratory, US

Nagaoka University of Technology, Japan

ISVR, University of Southampton, UK

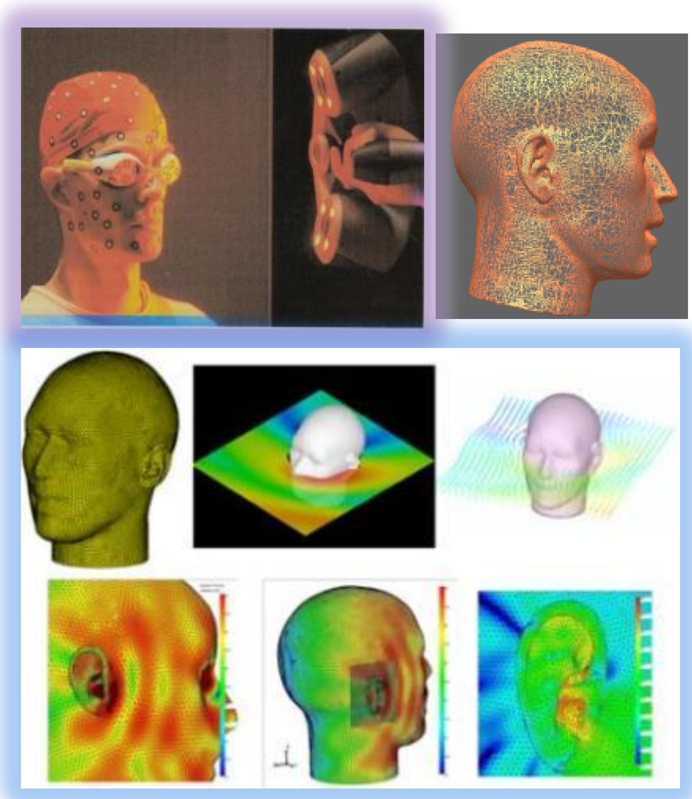
South China University of Technology, China

Tohoku University, Japan

Summary of popular HRTF databases

Databases	(Subjects, Directions)
IRCAM France http://recherche.ircam.fr/equipes/salles/listen	(51, 187)
CIPIC , UC Davis http://interface.cipic.ucdavis.edu/sound/hrtf.html	(45,1250)
University of Maryland http://www.isr.umd.edu/Labs/NSL/Software.htm	(7,2093)
Tohoku University, Japan http://www.ais.riec.tohoku.ac.jp/lab/db-hrtf	(3,454)
Nagoya University, Japan http://www.sp.m.is.nagoya-u.ac.jp/HRTF/database.html	(96,72)
Austrian Academy of Sciences http://www.kfs.oeaw.ac.at/index.php?option=com_content&view=article&id=608:ari-hrtf-database&catid=158:resources-items&Itemid=606&lang=en	(70,1550)
TU Berlin (3m,2m,1m,0.5m) http://audio.qu.tu-berlin.de/?p=641	(KEMAR,360)
MIT Lab http://sound.media.mit.edu/resources/KEMAR.html	(KEMAR,710)
Oldenburg University (0.8m,3m) http://medi.uni-oldenburg.de/hrir/html/documentation.html	(HATS,365)
SDAC, KAIST (0.2,0.6,1m) http://sdac.kaist.ac.kr/research/index.php?mode=area&act=DownHRTFDatabase	(HATS, 100)
Nagoaka University (1.5 m) http://www.nagaoka-ct.ac.jp/ee/lab_syano/index_e.html	(SAMRAI dummy head + 3 subjects, 72 azim, 8 elev)
DSP Lab @ NTU (0.35,0.45,0.50,0.60,0.75,0.8,1,1.4m) http://eeeweba.ntu.edu.sg/DSPLab/dsplabhrtf/	(HATS + 3 subjects, 72)

Individualization: anthropometry



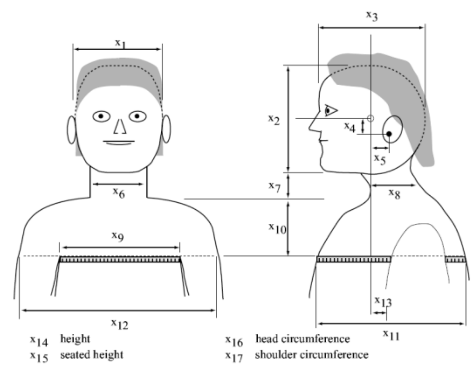


Figure 2: *Head and torso measurements*

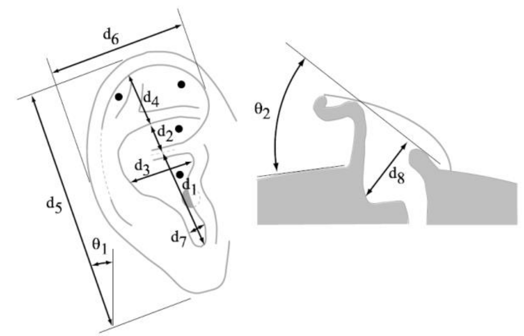
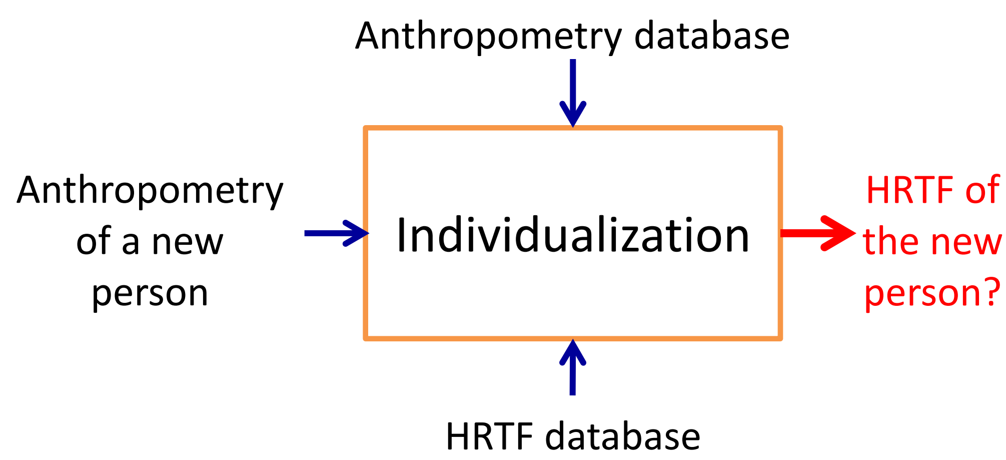


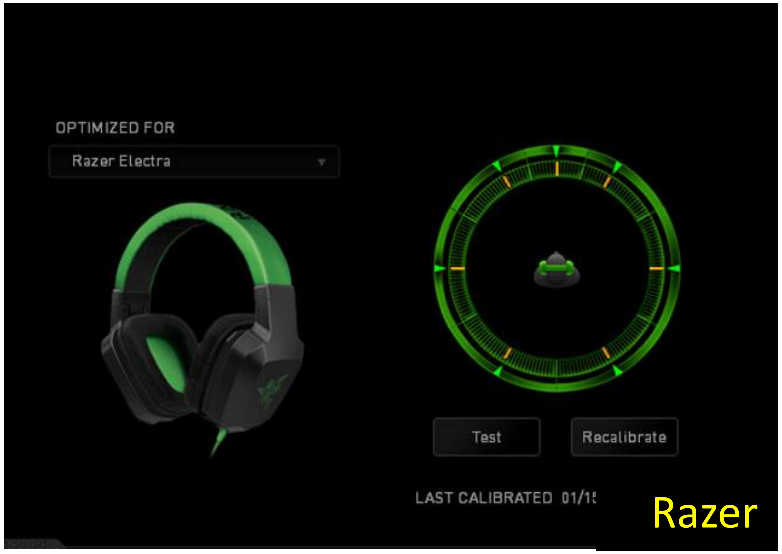
Figure 3: *Pinna measurements*

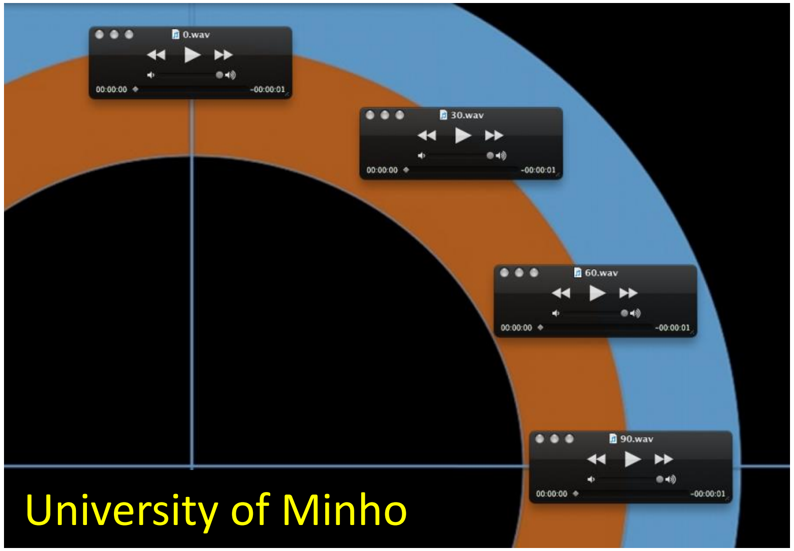


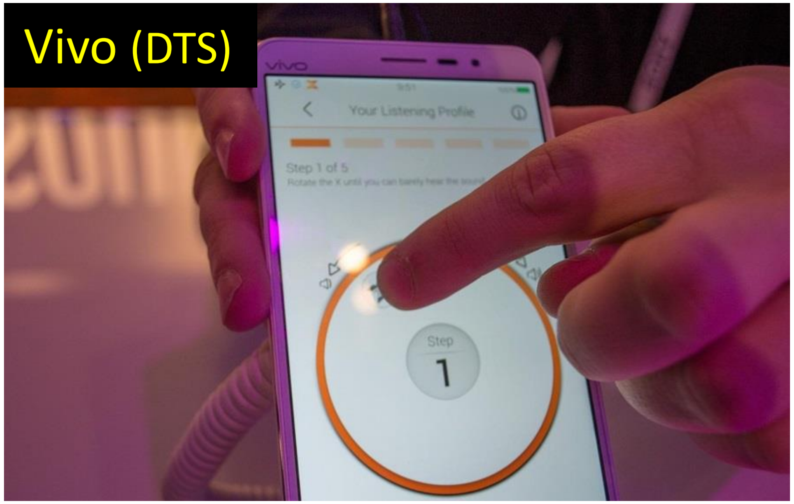
Obtain HRTFs numerically by:
1. Solving of acoustic equation
2. Numerical methods: Finite element method (FEM), Boundary element method (BEM)

Individualization: anthropometry

Individualization: training/tuning



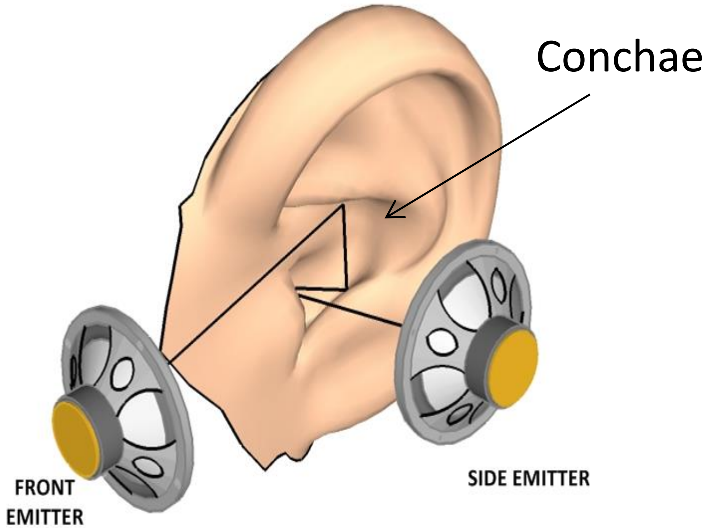


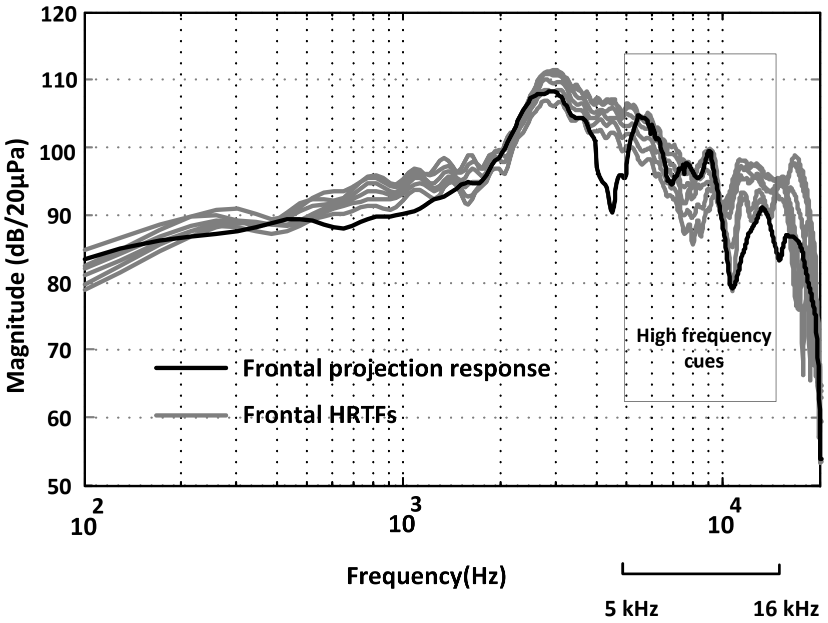


Individualization: training/tuning

Source from https://www.youtube.com/watch?v=pOtN-KMWTeM

Individualization: frontal projection





- No additional measurements and listening experiments required
- Reduce front-back confusion by > 50%;
- Zero user effort, plug and play (automatic during playback)

K. Sunder, E. L. Tan, and W. S. Gan, "Individualization of binaural synthesis using frontal projection headphones," *J. Audio Eng. Soc.*, vol. 61, no. 12, pp. 989-1000, Dec. 2013.

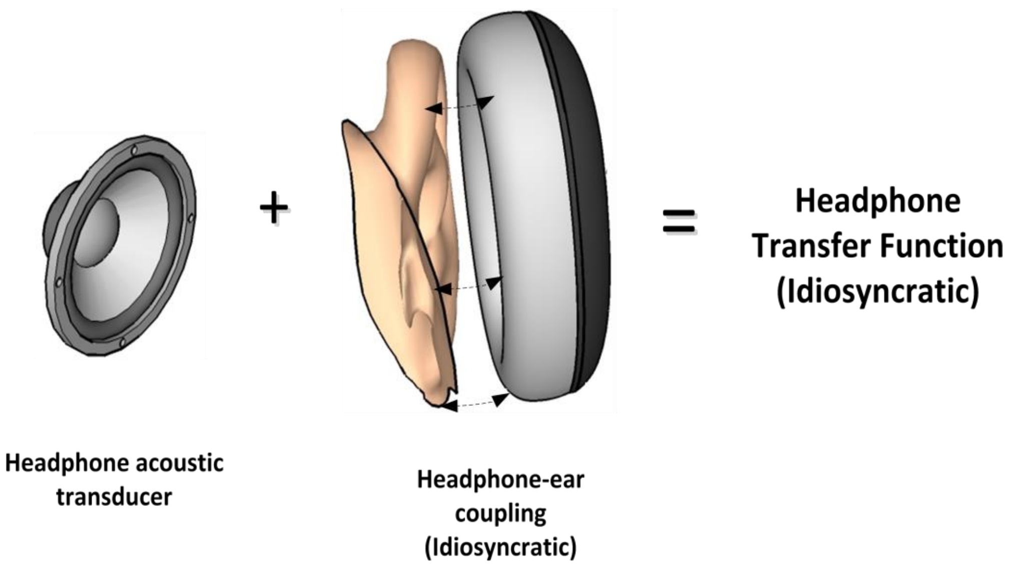
Individualization: summary

Obtain Individual Features	Techniques	Pros and Cons	Performance
Acoustical measurements	Individual measurements, IRCAM France, CIPIC, Tohuko Univ., etc.	Ideal, accurate Tedious, requires high precision	Reference for individualization techniques
Anthropometry	3D mesh, pictures; Numerical Solutions: PCA, FEM,BEM, ANN	Need a large database; Requires high resolution imaging; Expensive	Uses the correlation between individual HRTF and anthropometric data
Training/tuning	PCA weight tuning, Tune magnitude Spectrum, Selection from HRTF database	Directly relates to perception; requires regular training;	Obtains the best HRTFs perceptually
Frontal projection	Frontal Projection Headphones	No additional measurement, Type-2 EQ	Automatic customization, reduced front-back confusions
Non-individualized HRTF	Generalized HRTF	Easy to implement, Poor localization	Not an individualization technique

Equalization

Headphone is not acoustically transparent:

- Headphone colors the input sound spectrum;
- Affects the free-field characteristics of the sound pressure at the ear

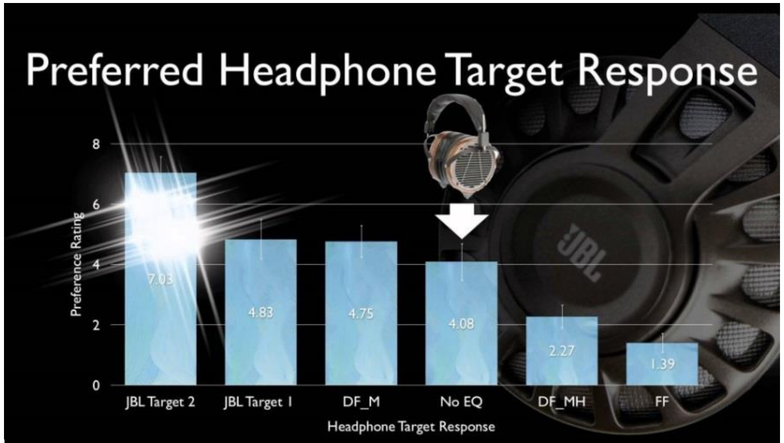


Breakdown of headphone transfer function (HPTF)

Decoupled equalization for binaural/stereo

Aim: Emulate the reproduction in a reference field

- Free-field:
  - Target: free-field front loudspeaker response
- Diffuse-field and other reference curves:
  - Target: response of diffuse-field, or a reference room
  - Lesser inter-individual variability



Source from http://seanolive.blogspot.sg/2014/01/the-perception-and-measurement-of.html

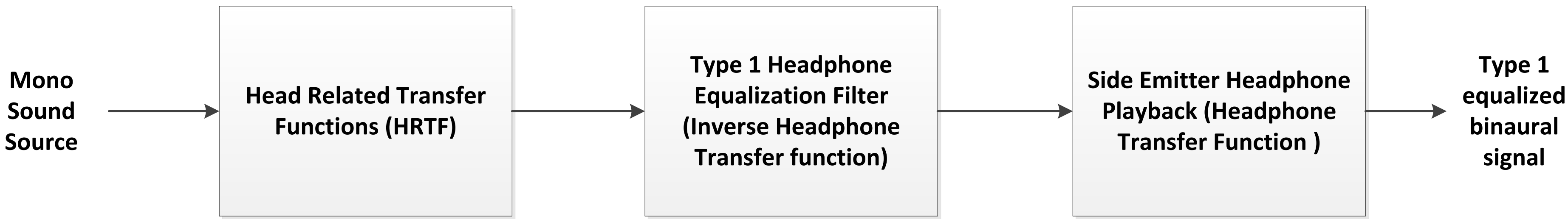
Non-decoupled equalization for binaural

Aim: Spectrum at eardrum is the individual HRTF features

- For conventional headphone
- For front projection headphone

Conventional equalization (Type 1 EQ)

Headphone is not acoustically transparent, therefore the effect of the headphone must be removed.



Equalization process : Removing the headphone transfer function

$$\boldsymbol{Y(\omega) = S(\omega) \cdot HRTF(\omega) \cdot \frac{1}{HPTF(\omega)} \cdot HPTF(\omega)}$$

Where, $Y(\omega)$ = Equalized Binaural Signal
$S(\omega)$ = Source Signal Spectrum
$HRTF(\omega)$ = Head Related Transfer Function (Left/Right)
$HPTF(\omega)$ = Headphone Transfer Function (Left/Right)

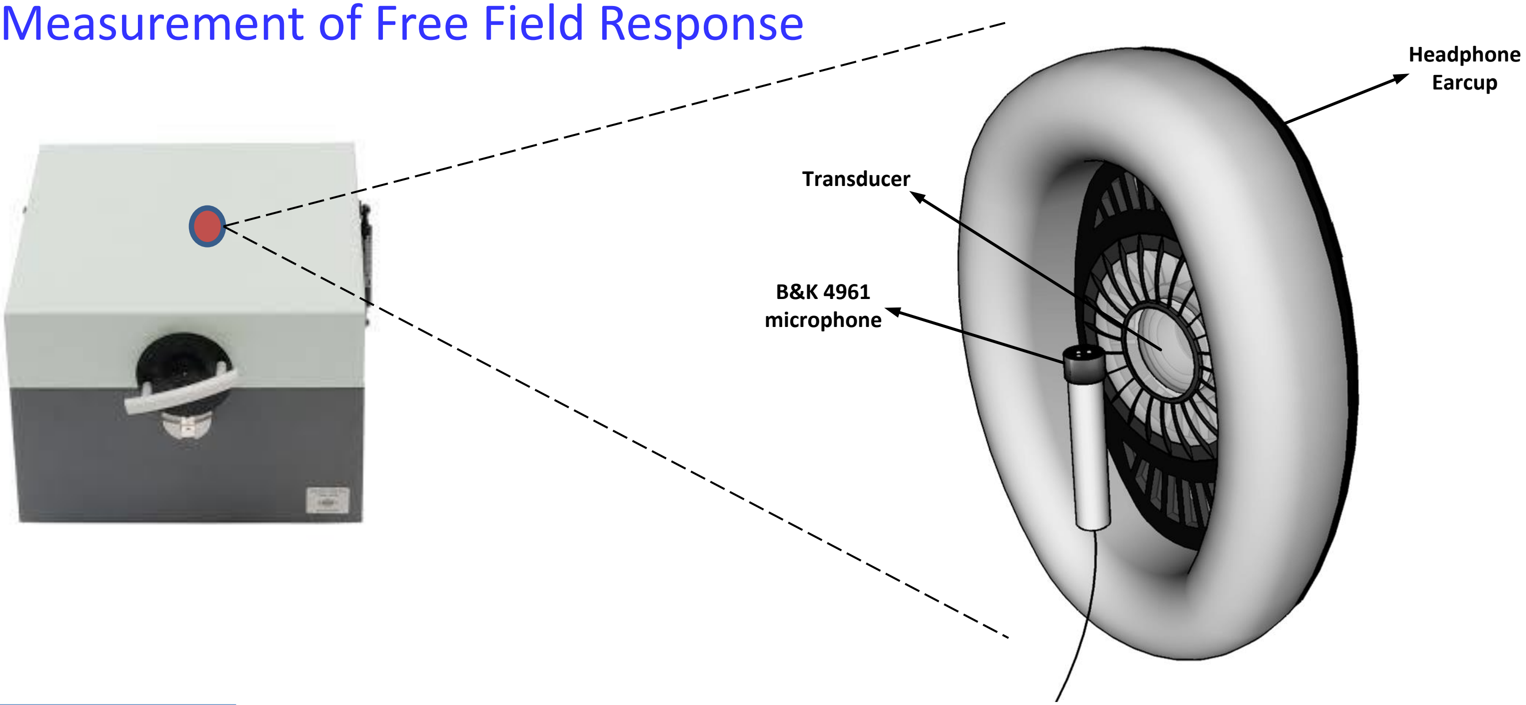
And, $\frac{1}{HPTF(\omega)}$ = Equalization Filter

Dependent on individual pinna feature

Type 2 EQ (for frontal emitters)

- Reflections/diffractions created by the interactions with the pinna due to the frontal projection are important and should be retained.
- Does not include headphone-ear coupling.
- Equalizing to the free field response of the headphone with the ear-cup.

Measurement of Free Field Response



Sunder, Kaushik, Ee-Leng Tan, and Woon-Seng Gan. "Individualization of Binaural Synthesis Using Frontal Projection Headphones." *Journal of the Audio Engineering Society* 61.12 (2013): 989-1000.

Type 2 EQ (for frontal emitters)

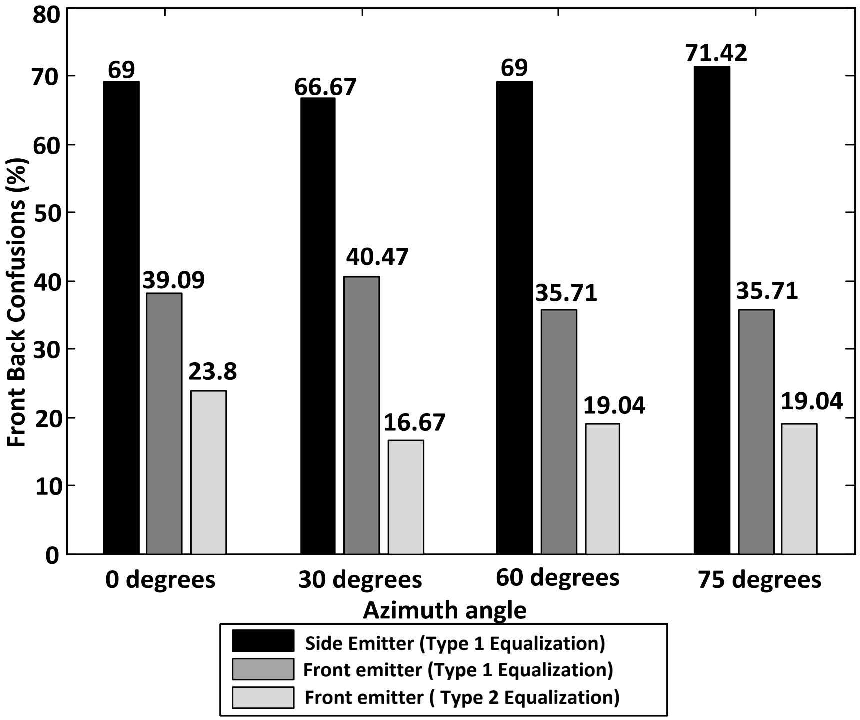
- Reflections/diffractions created by the interactions with the pinna due to the frontal projection are important and should be retained.
- Does not include headphone-ear coupling.
- Equalizing to the free field response of the headphone with the ear-cup.

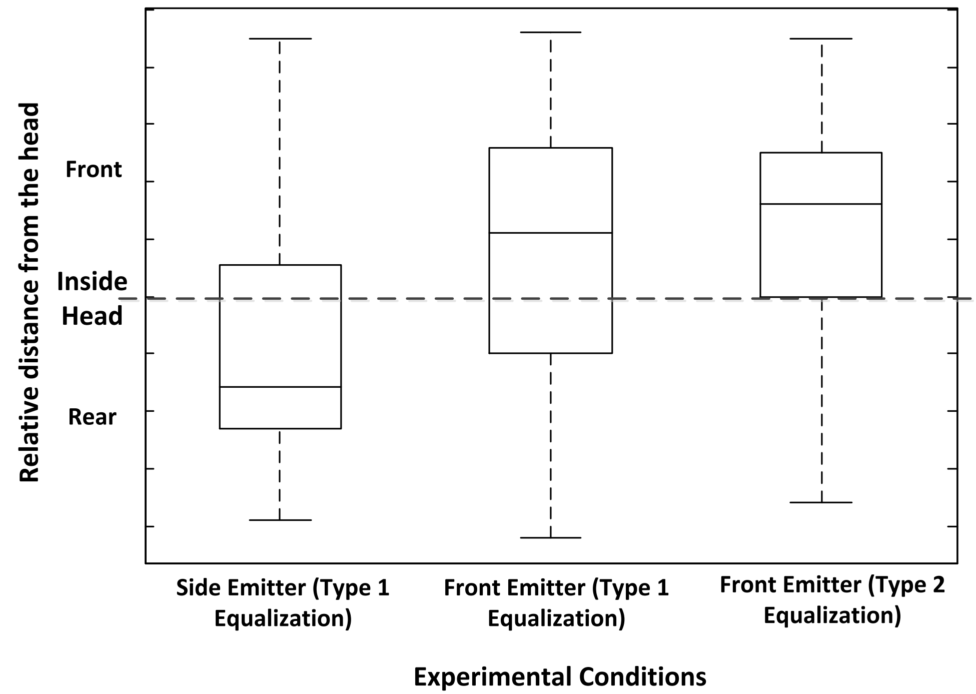
$$Y(\omega) = S(\omega)\cdot HRTF(\omega)\cdot \frac{1}{FFR(\omega)}\cdot HPTF(\omega)$$

$$HPTF(\omega) = FFR(\omega)\cdot PC(\omega)$$

Where, $Y(\omega)$ = Equalized Binaural Signal
$S(\omega)$ = Source Signal Spectrum
$HRTF(\omega)$ = Head Related Transfer Function (Left/Right)
$HPTF(\omega)$ = Headphone Transfer Function (Left/Right)
$FFR(\omega)$ = Free Field Response of the Frontal Emitter
$PC(\omega)$ = Personalized Pinna Cues generated by frontal projection

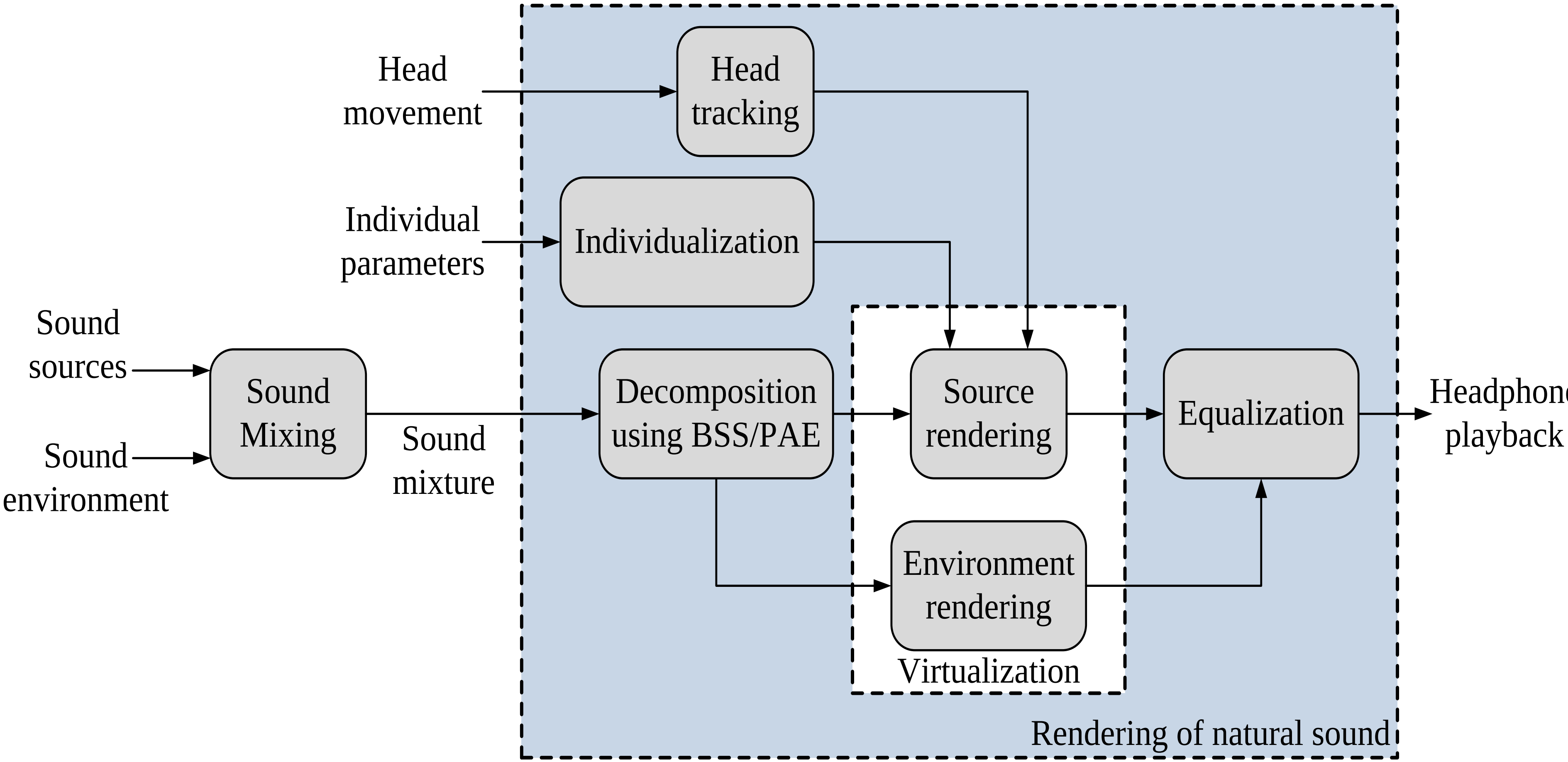
Subjective validation of equalization





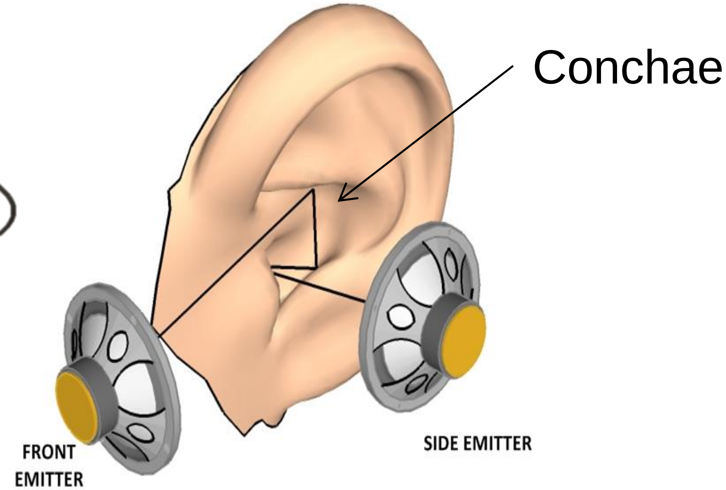
- Type-2 EQ works best for the frontal emitter playback
- ANOVA results show
  - Type of emitter has a significant effect on rate of reversal
  - Type-2 EQ has a significant effect in reducing F/B reversal

Integration

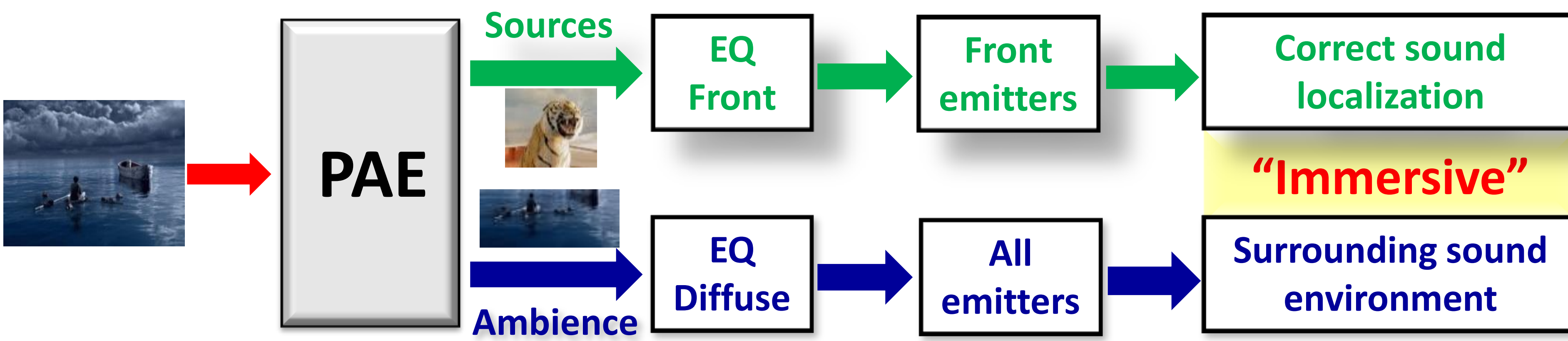


K. Sunder, J. He, E. L. Tan, and W. S. Gan, "Natural sound rendering for headphones," IEEE Signal Processing Magazine, Mar. 2015.

3D Audio Headphone: an example



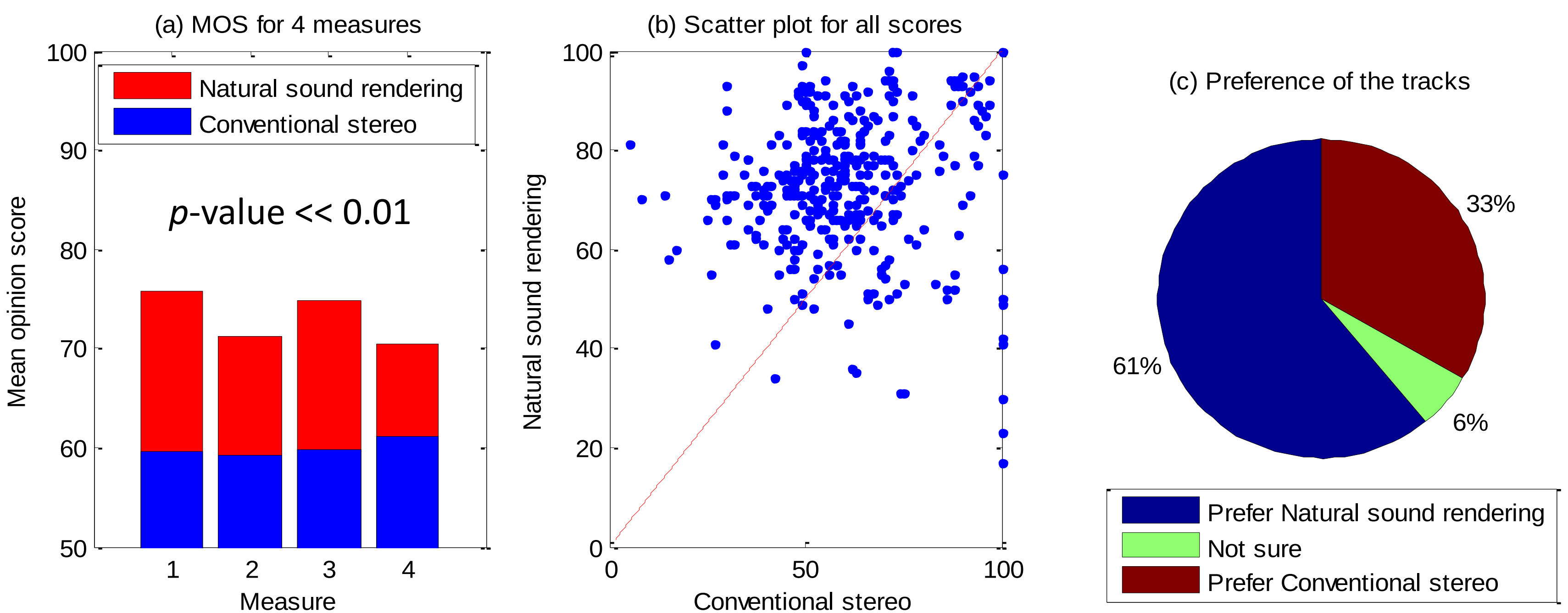
Key features
- ✓ Patented structure with strategic-positioned emitters;
- ✓ Individualization via frontal projection; no measurements or training required;
- ✓ Recreate an immersive perception of sound objects with surrounding ambience;
- ✓ Compatible with all existing sound formats.



W. S. Gan and E. L. Tan, "Listening device and accompanying signal processing method," US Patent 2014/0153765 A1, 2014.

Subjective evaluation

- Conventional stereo system: stereo headphone
- Natural sound rendering system: 3D headphone
- **Stimuli**: binaural, movie and gaming tracks;
- **4 measures**: Sense of direction, externalization, ambience, and timbral quality;
- 18 subjects, score of 0-100, and overall preference.



Conclusions

- Advent of high speed, low power, and low cost **embedded processors** fueling a strong growth in portable and wearable applications.
- Many opportunities for **new innovations** in spatial audio rendering for assistive technologies; being-there communications; immersive AR/VR gaming; and interactive entertainments.
- **Seamless integration** of real and virtual sound objects to achieve natural listening.

> " … future headphones are becoming more **content-aware**, **location-aware**, **listener-aware**, and hence become more intelligent and assistive."
>
> IEEE Signal Processing Magazine, March '15

Key References on Natural Sound Rendering

[1] D. R. Begault, *3-D sound for virtual reality and multimedia*: AP Professional, 2000.

[2] S. Spors, H. Wierstorf, A. Raake, F. Melchior, M. Frank, and F. Zotter, "Spatial sound with loudspeakers and its perception: A review of the current state," *Proc. IEEE,* vol. 101, no. 9, pp. 1920-1938, Sep. 2013.

[3] V. Pulkki, "Spatial sound reproduction with directional audio coding," *J. Audio Eng. Soc.,* vol. 55, no.6, pp. 503-516, Jun. 2007.

[4] S. Olive, T. Welti, and E. McMullin, "Listener Preferences for Different Headphone Target Response Curves," in *Proc. 134th Audio Engineering Society Convention*, Rome, Italy, May 2013.

[5] M. M. Goodwin and J. M. Jot, "Binaural 3-D audio rendering based on spatial audio scene coding," in *Proc. 123rd Audio Engineering Society Convention*, New York, Oct. 2007.

[6] J. Breebaart and E. Schuijers, "Phantom materialization: a novel method to enhance stereo audio reproduction on headphones," *IEEE Trans. Audio, Speech, and Language Processing,* vol. 16, no.8, pp. 1503-1511, Nov. 2008.

[7] C. Avendano and J. M. Jot, "A frequency-domain approach to multichannel upmix," *J. Audio Eng. Soc.,* vol. 52, no.7/8, pp. 740-749, Jul. 2004.

[8] C. Faller, "Multiple-loudspeaker playback of stereo signals," *J. Audio Eng. Soc.,* vol. 54, no.11, pp. 1051-1064, Nov. 2006.

[9] F. Menzer and C. Faller, "Stereo-to-binaural conversion using interaural coherence matching," in *Proc. 128th Audio Engineering Society Convention,* London, UK, May 2010.

[10] D. R. Begault, E. M. Wenzel, and M. R. Anderson, "Direct comparison of the impact of head tracking, reverberation, and individualized head-related transfer functions on the spatial perception of a virtual speech source," *J. Audio Eng. Soc.,* vol. 49, no. 10, pp. 904-916, Oct. 2001.

[11] C. Faller and F. Baumgarte, “Binaural cue coding—part II: schemes and applications,” *IEEE Trans. Speech Audio Process.,* vol. 11, no. 6, pp. 520–531, Nov. 2003.

[12] V. R. Algazi and R. O. Duda, "Headphone-based spatial sound," *Signal Processing Magazine, IEEE,* vol. 28, no. 1, pp. 33-42, Jan. 2011.

[13] R. Nicol, *Binaural Technology*: AES, 2010.

[14] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent component analysis*. New York: John Wiley & Sons, 2004.

[15] M. D. Plumbley, T. Blumensath, L. Daudet, R. Gribonval, and M. E. Davies, "Sparse representations in audio and music: from coding to source separation," *Proc. IEEE,* vol. 98, no. 6, pp. 995-1005, Jun. 2010.

[16] O. Yilmaz and S. Rickard, "Blind separation of speech mixtures via time-frequency masking," *IEEE Trans. Signal Processing,* vol. 52, no.7, pp. 1830-1847, Jul. 2004.

Key References on Natural Sound Rendering

[17] T. Virtanen, "Sound source separation in monaural music signals," PhD Thesis, Tampere University of Technology, 2006.

[18] E. Vincent, N. Bertin, R. Gribonval, and F. Bimbot, "From blind to guided audio source separation," *IEEE Signal Processing Magazine,* vol. 31, no. 3, pp. 107-115, 2014.

[19] D. Wang and G. J. Brown, *Computational auditory scene analysis: Principles, algorithms, and applications.* NJ: Wiley-IEEE Press, 2006.

[20] J. Merimaa, M. M. Goodwin, and J. M. Jot, "Correlation-based ambience extraction from stereo recordings," in *Proc. 123rd Audio Engineering Society Convention,* New York, Oct. 2007.

[21] J. He, E. L. Tan, and W. S. Gan, "Linear estimation based primary-ambient extraction for stereo audio signals," *IEEE/ACM Trans. Audio, Speech, and Language Processing,* vol. 22, no.2, pp. 505-517, 2014.

[22] J. He, E. L. Tan, and W. S. Gan, "Time-shifted principal component analysis based cue extraction for stereo audio signals," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP'13)*, Canada, May 2013, pp. 266-270.

[23] J. He, E. L. Tan, and W. S. Gan, "A study on the frequency-domain primary-ambient extraction for stereo audio signals," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP'14)*, Florence, Italy, 2014, pp. 2892-2896.

[24] J. Thompson, B. Smith, A. Warner, and J. M. Jot, “Direct-diffuse decomposition of multichannel signals using a system of pair-wise correlations,” in *Proc. 133rd Audio Eng. Soc. Conv.*, San Francisco, 2012.

[25] H. Møller, M. F. Sørensen, D. Hammershøi, and C. B. Jensen, "Head-related transfer functions of human subjects," *J. Audio Eng. Soc.,* vol. 43, no. 5, pp. 300-321, May 1995.

[26] S. Xu, Z. Li, and G. Salvendy, "Individualization of head-related transfer function for three-dimensional virtual auditory display: a review," in *Virtual Reality*, ed: Springer, 2007, pp. 397-407.

[27] G. Enzner, “3D-continuous-azimuth acquisition of head-related impulse responses using multi-channel adaptive filtering,” *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA),* Oct. 2009, pp. 325-328.

[28] M. Rothbucher, M. Durkovic, H. Shen, and K. Diepold, "HRTF customization using multiway array analysis," in *Proc. 18th European Signal Processing Conference (EUSIPCO'10)*, Aalborg, August 2010, pp. 229-233.

[29] R. O. Duda, V. R. Algazi, and D. M. Thompson, "The use of head-and-torso models for improved spatial sound synthesis," in *Proc. 113th Audio Engineering Society Convention*, Los Angeles, Oct. 2002.

[30] D. N. Zotkin, J. Hwang, R. Duraiswaini, and L. S. Davis, "HRTF personalization using anthropometric measurements," in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA'03)*, New York, Oct. 2003, pp. 157-160.

Key References on Natural Sound Rendering

[31] J. C. Middlebrooks, "Individual differences in external-ear transfer functions reduced by scaling in frequency," *J. Acoust. Soc. Amer.,* vol. 106, no. 3, pp. 1480-1492, Sep. 1999.

[32] K. J. Fink and L. Ray, "Tuning principal component weights to individualize HRTFs," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP'12)*, Kyoto, Mar. 2012, pp. 389-392.

[33] K. Sunder, E. L. Tan, and W. S. Gan, "Individualization of binaural synthesis using frontal projection headphones," *J. Audio Eng. Soc.,* vol. 61, no. 12, pp. 989-1000, Dec. 2013.

[34] A. Bondu, S. Busson, V. Lemaire, and R. Nicol, "Looking for a relevant similarity criterion for HRTF clustering: a comparative study," in *Proc. 120th Audio Engineering Society Convention*, Paris, France, May 2006.

[35] H. Møller, D. Hammershoi, C. B. Jensen, and M. F. Sorensen, "Transfer characteristics of headphones measured on human ears," *J. Audio Eng. Soc.,* vol. 43, no. 4, pp. 203-217, Apr. 1995.

[36] V. Larcher, J. M. Jot, and G. Vandernoot, "Equalization methods in binaural technology," in *Proc. 105th Audio Engineering Society Convention,* SanFrancisco, Sep. 1998.

[37] A. Kulkarni and H. S. Colburn, "Variability in the characterization of the headphone transfer-function," *J. Acoust. Soc. Amer.,* vol. 107, no. 2, pp. 1071-1074, Feb. 2000.

[38] H. Møller, C. B. Jensen, D. Hammershøi, and M. F. Sørensen, "Design criteria for headphones," *J. Audio Eng. Soc.,* vol. 43, no. 4, pp. 218-232, Apr. 1995.

[39] W. S. Gan and E. L. Tan, "Listening device and accompanying signal processing method," US Patent 2014/0153765 A1, 2014.

[40] K. Sunder, J. He, E. L. Tan, and W. S. Gan, “Natural sound rendering for headphones,” IEEE Signal Processing Magazine, vol. 32, no. 2, pp. 100-113, Mar. 2015.

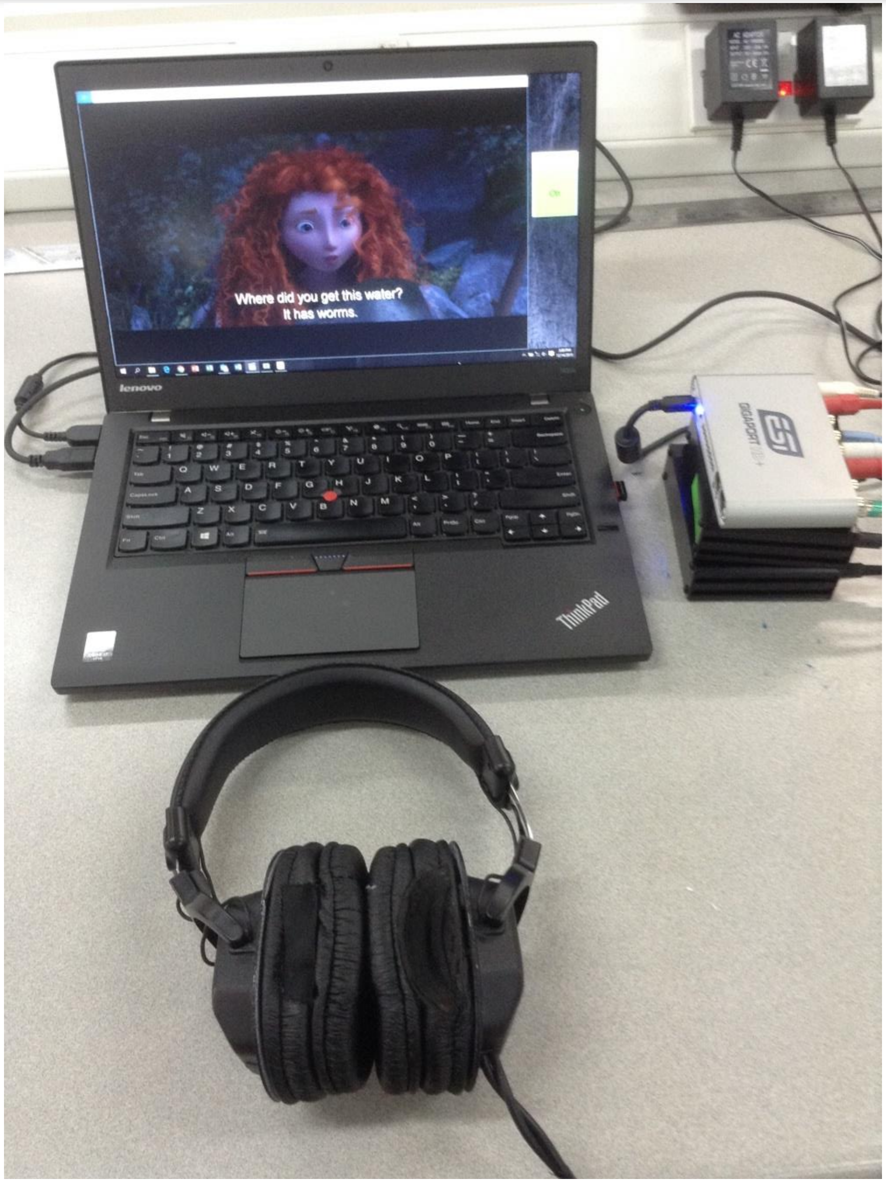
[41] J. He, E. L. Tan, and W. S. Gan, “Primary-ambient extraction using ambient spectrum estimation for immersive spatial audio reproduction,” IEEE/ACM Trans. Audio, Speech, Lang. Process., vol. 23, no. 9, pp. 1431-1444, Sept. 2015.

[42] J. He, W. S. Gan, and E. L. Tan, “Time-shifting based primary-ambient extraction for spatial audio reproduction,” IEEE/ACM Trans. Audio, Speech, Lang. Process., vol. 23, no. 10, pp. 1576-1588, Oct. 2015.

[43] J. He, W. S. Gan, and E. L. Tan, “Primary-ambient extraction using ambient phase estimation with a sparsity constraint,” IEEE Signal Process. Letters, vol. 22, no. 8, pp. 1127-1131, Aug. 2015.

[44] J. He, and W. S. Gan, “Multi-shift principal component analysis based primary component extraction for spatial audio reproduction,” in Proc. ICASSP, Brisbane, Australia, Apr. 2015, pp. 350-354.

Demo of 3D audio headphones



Module III
ANC Headphones

Active Noise Control

- Active noise control (ANC) uses additional secondary sources to produce anti-noise that cancel the undesired noise.
- Principle:
  1. Mathematics: $x + (-x) = 0$
  2. Physics (Superposition): Anti-noise of equal amplitude and opposite phase is combined with the primary noise, resulting in the cancellation of both noises

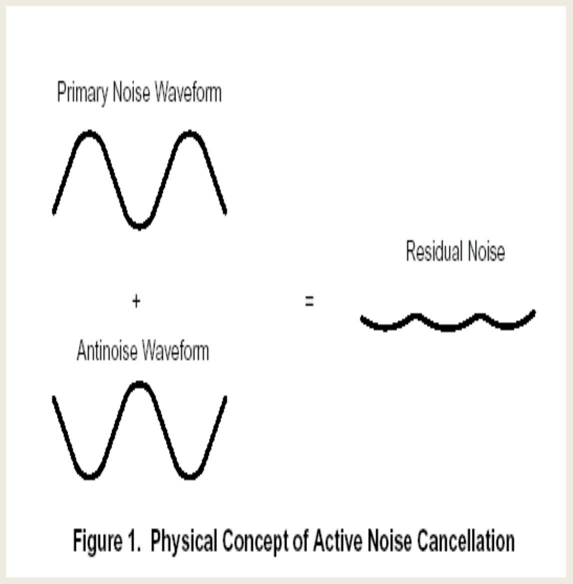
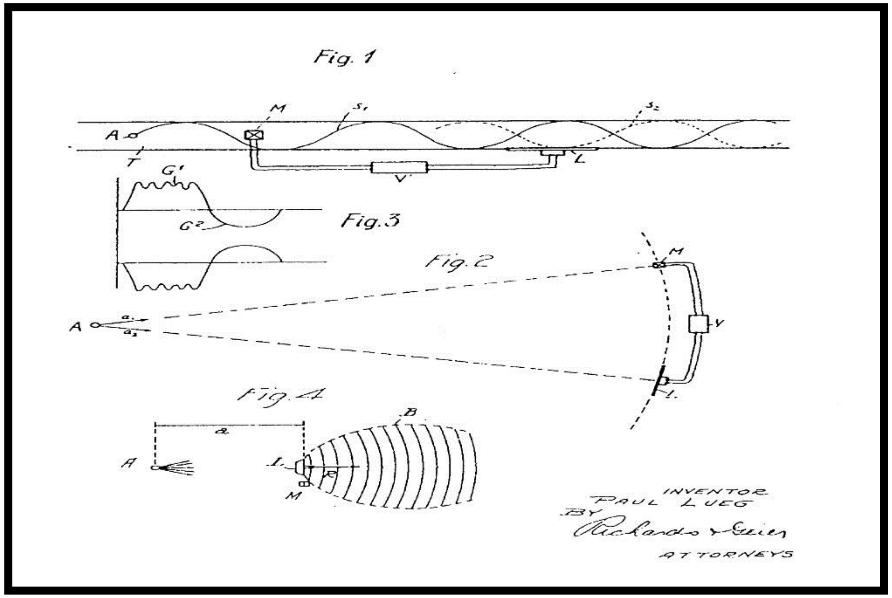


Figure 1. Physical Concept of Active Noise Cancellation



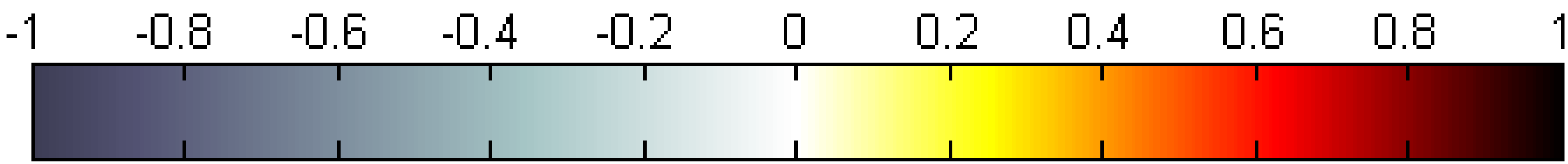
Leug, 1933

Sound Field Interaction of 2 Point Sources

Destructive

Constructive

Normalised sound pressure

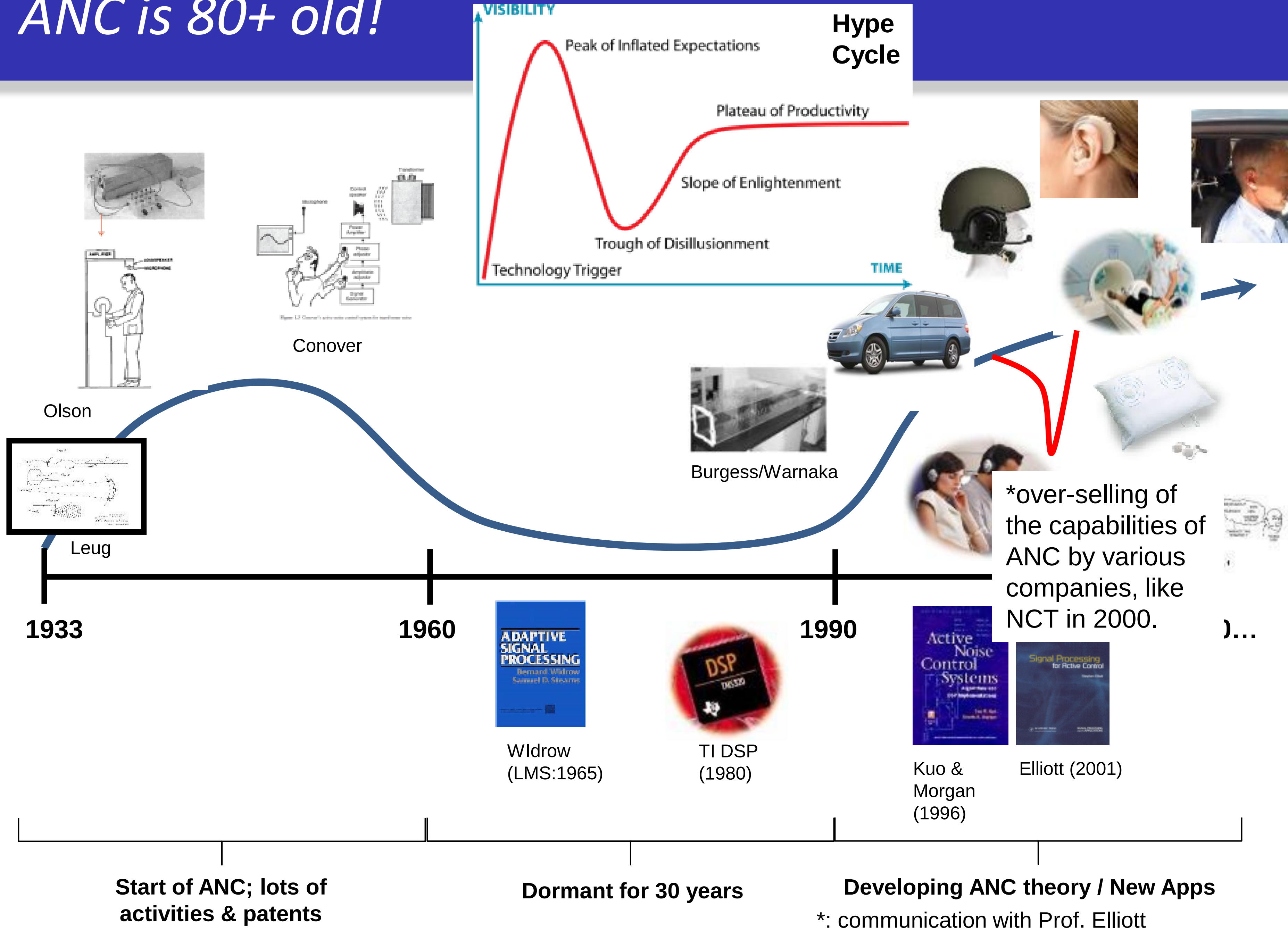


ANC dated back to Chinese Martial Art !

金庸小说“倚天屠龙记”

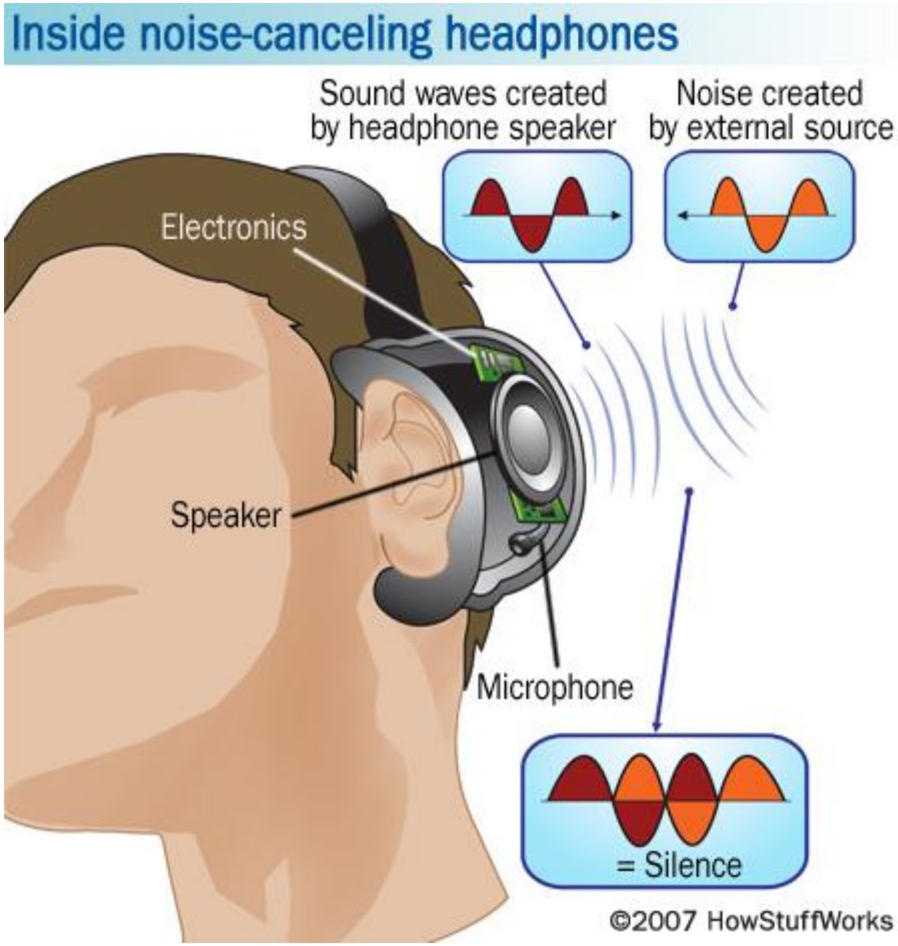
- 张无忌一惊，不及趋避，足尖使劲，拔身急起，斜飞而上，只听得飕飕两声轻响，跟着“啊”的一下长声呼叫。他在半空中转过头来，只见何太冲和班淑娴的两柄长剑并排插在鲜于通胸口。原来何氏夫妇纵横半生，却当众败在一个后辈手底，无论如何咽不下这口气去，两人拾起长剑，眼见张无忌正俯身在点鲜于通的穴道，对望一眼，心意相通，点了点头，突然使出一招“无声无色”，同时疾向他背后刺去。这招“无声无色”是昆仑派剑学中的绝招，**必须两人同使，两人功力相若，内劲相同，当剑招之出，劲力恰恰相反，于是两柄长剑上所生的荡激之力、破空之声，一齐相互抵消。**这路剑招本是用于夜战，黑暗中令对方难以听声辨器，事先绝无半分朕兆，白刃已然加身，但若白日用之背后偷袭，也令人无法防备。

ANC is 80+ old!


VISIBILITY
Hype Cycle
Peak of Inflated Expectations
Plateau of Productivity
Slope of Enlightenment
Trough of Disillusionment
Technology Trigger
TIME
Olson
Conover
Leug
Burgess/Warnaka
*over-selling of the capabilities of ANC by various companies, like NCT in 2000.
1933
1960
1990
ADAPTIVE SIGNAL PROCESSING
Bernard Widrow
Samuel D. Stearns
WIdrow (LMS:1965)
DSP
TI DSP (1980)
Active Noise Control Systems
Kuo & Morgan (1996)
Signal Processing for Active Control
Elliott (2001)
Start of ANC; lots of activities & patents
Dormant for 30 years
Developing ANC theory / New Apps
*: communication with Prof. Elliott

Applications

Active Noise Control Headset



Commercial ANC Headphones

- **Bose** Quiet Comfort 15 Acoustic Noise Canceling Headphones (US$365)
- **Sony** MDR-NC100D Digital Noise Canceling Earbuds (US$131)
- **Creative** HN-900 Noise Canceling Headphones (US$159)
- **Sennheiser** MM 550 Bluetooth Wireless Headphones (US$475)
- **Digital Silence** Ambient Noise Canceling Earbud (US$62)
- **Blackbox** i10 Active Noise Rejection Earbud-powered by iPod battery (US$127)



ANC Headphones

- With many promising applications, we are witnessing the Golden Age of ANC
  - Low cost and accurate sensors and actuators
  - High speed, low cost and low power consumption embedded processors
- For successful consumer applications:
  - ANC needs to be **combined or integrate** with other functions
  - Allow **sharing hardware** resources (e.g. amplifier, loudspeakers, microphones etc.)
  - **Reuse software** code or tap on existing audio functions in digital implementation.
  - Many opportunities to **innovate**
- **Outline some of the work carried out by the research community to integrate ANC with other functions**

1st Paper to look into Integration of ANC with Audio System (from Kuo et al.)

- ***Appears in the 1993 IEEE Transaction on Consumer Electronics***
- ***Won the 1st place in the IEEE Consumer Electronics Society Chester Sall Awards in 1993.***

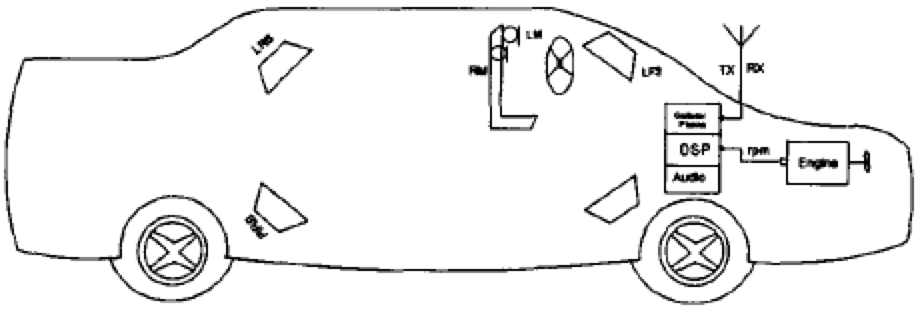


Figure 1. Integrated Hands-Free Cellular Phone, Active Noise Control and Audio System

Using multiple adaptive filters to integrate:

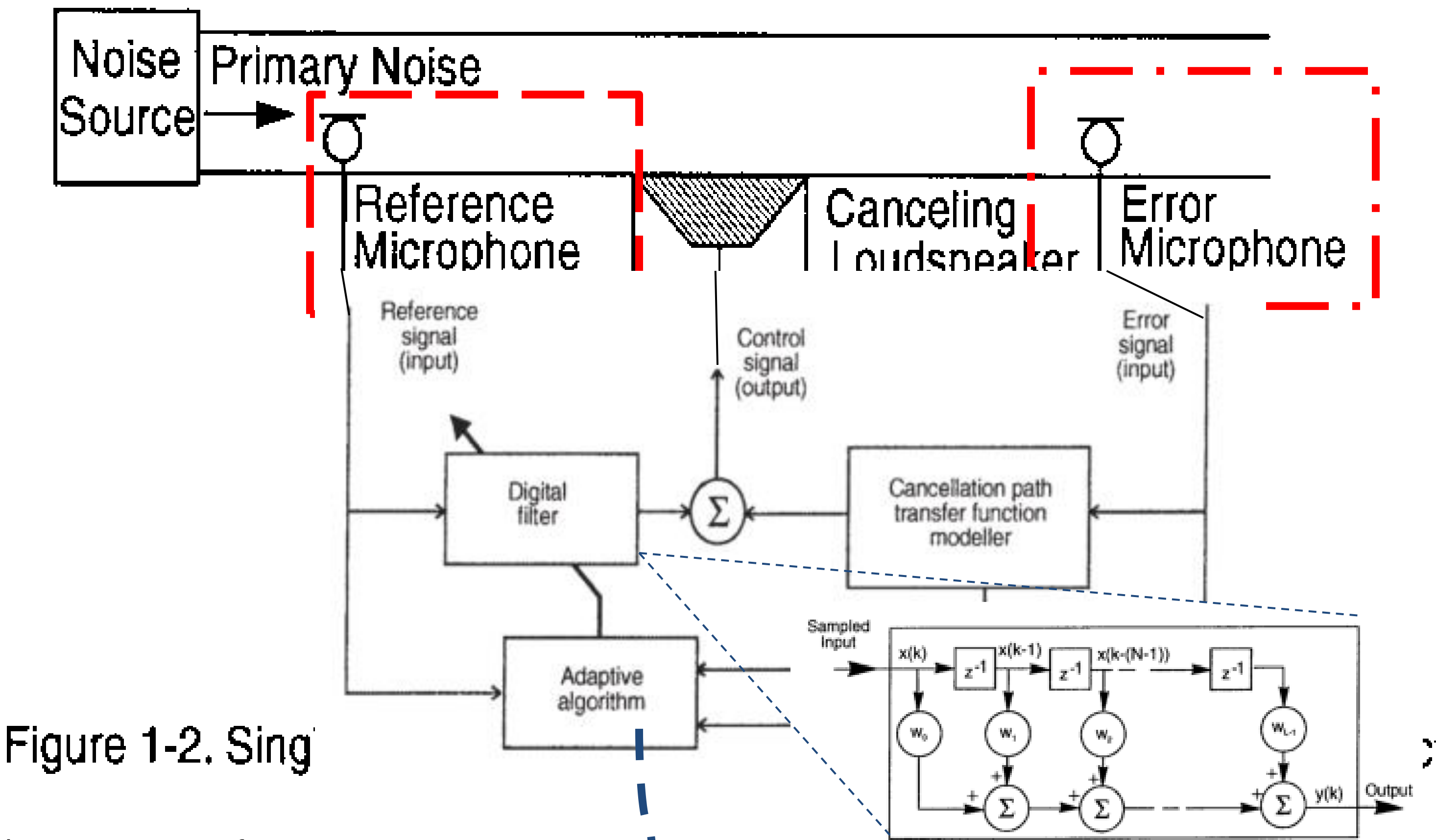
(i) Active noise control
(ii) Acoustic echo cancellation
(iii) Adaptive noise cancellation
(iv) Adaptive musical interference suppression

This paper lays the foundation of many integrated ANC + Audio systems/papers

(picture from)

Kuo, S.M.; Chuang, H.; Mallela, P.P.; "Integrated automotive signal processing and audio system", IEEE Trans Consumer Electronics, Volume 39, Issue 3, August 1993, pp. 522-532

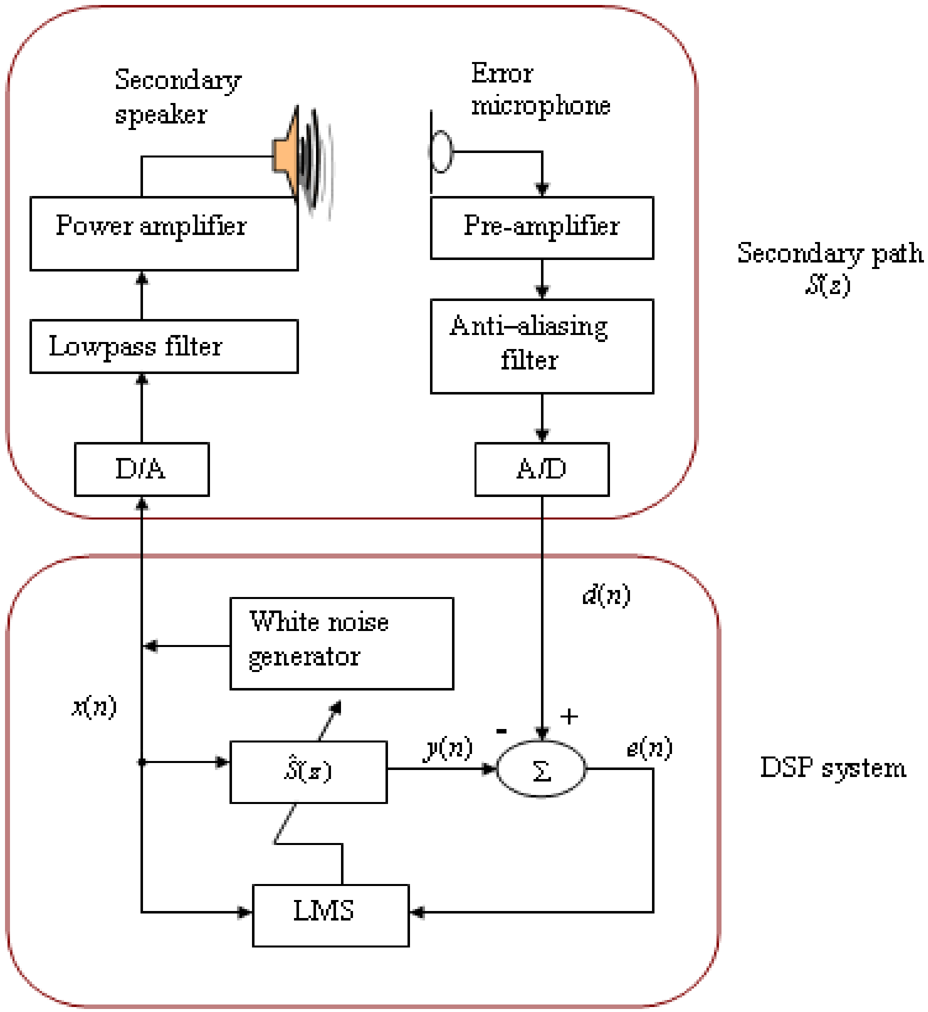
Basic Definitions in ANC: A Simple Duct Application



Picture courtesy of Sen M Kuo

FXLMS Algorithm: $\mathbf{w}(n+1) = \mathbf{w}(n) + \mu\,\mathbf{x}'(n)e(n)$

Offline Secondary Path Modeling



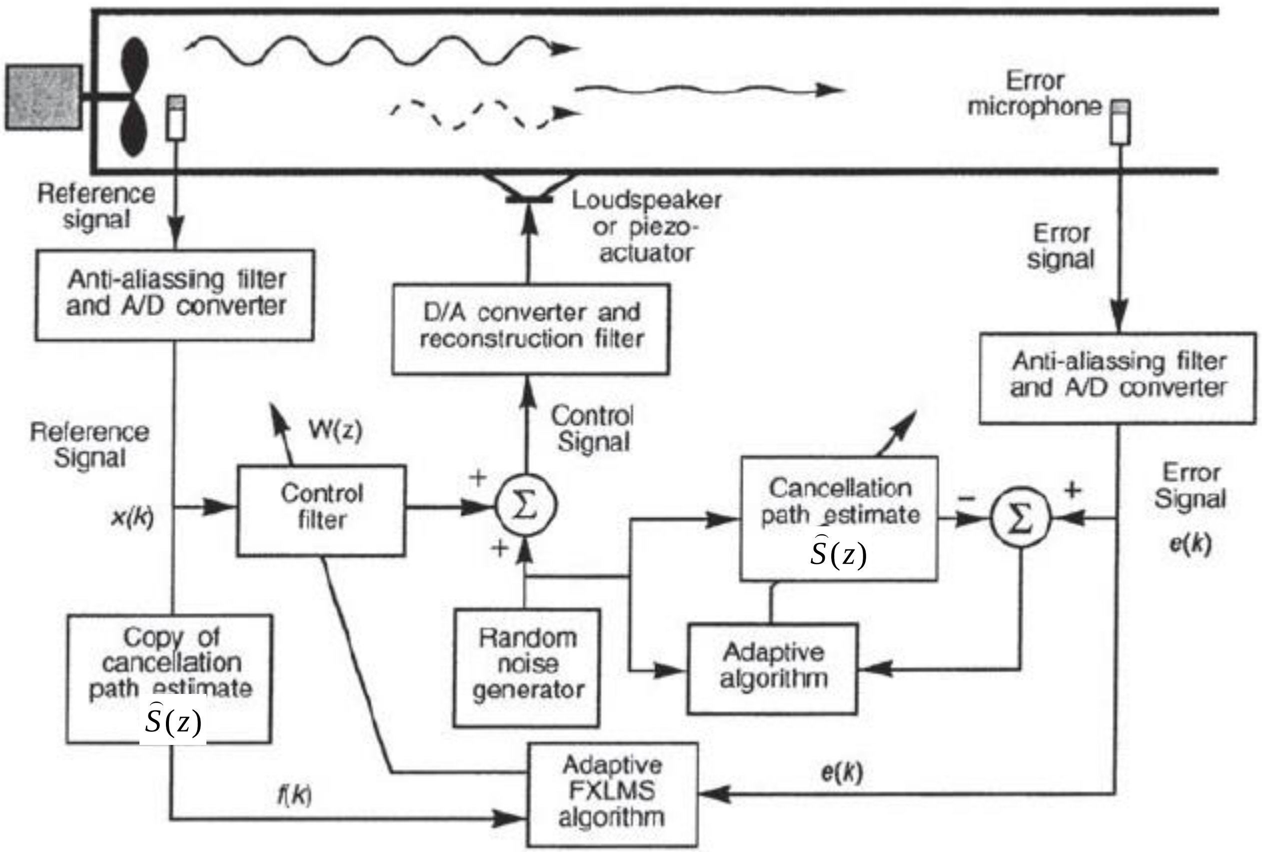
Experimental setup for off-line modeling of secondary-path *S*(*z*).

Picture courtesy of Sen M Kuo

Note:

- Offline secondary path modeling performed before the actual operation of ANC.
- The estimated model is used in actual operation.
- $S_{est}(z)$ can be fixed or adaptively update.
- Within the limit of slow adaptation, the algorithm will converge with nearly of **90° phase error** between $S_{est}(z)$ and $S(z)$.

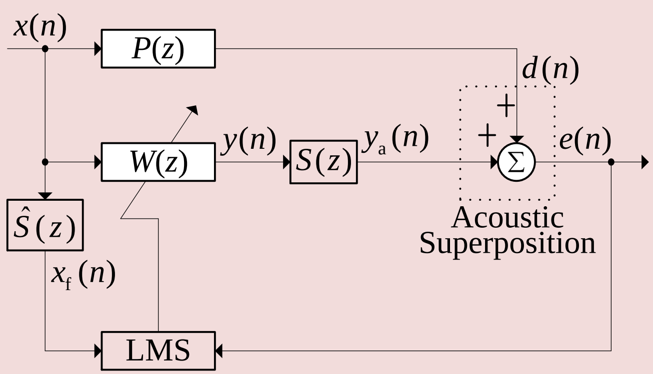
ANC System Layout with On-line Modeling



Picture courtesy of Colin Hansen

Configurations of Adaptive ANC System

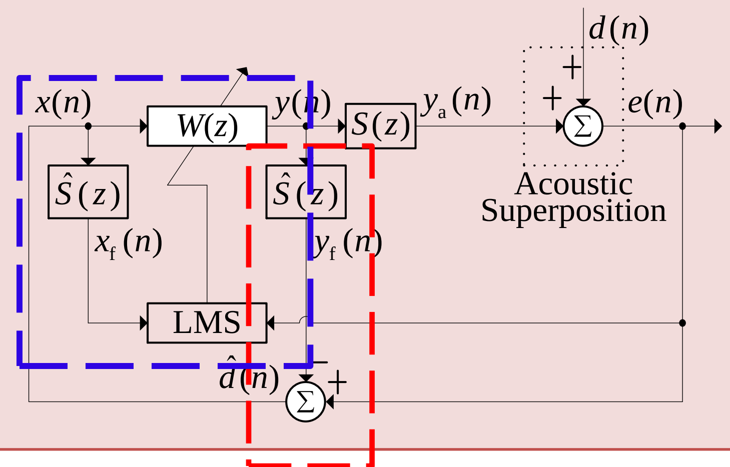
Feedforward ANC



- Broadband and narrowband FFANC
- 2 sensors and 1 actuator (1-dim case)

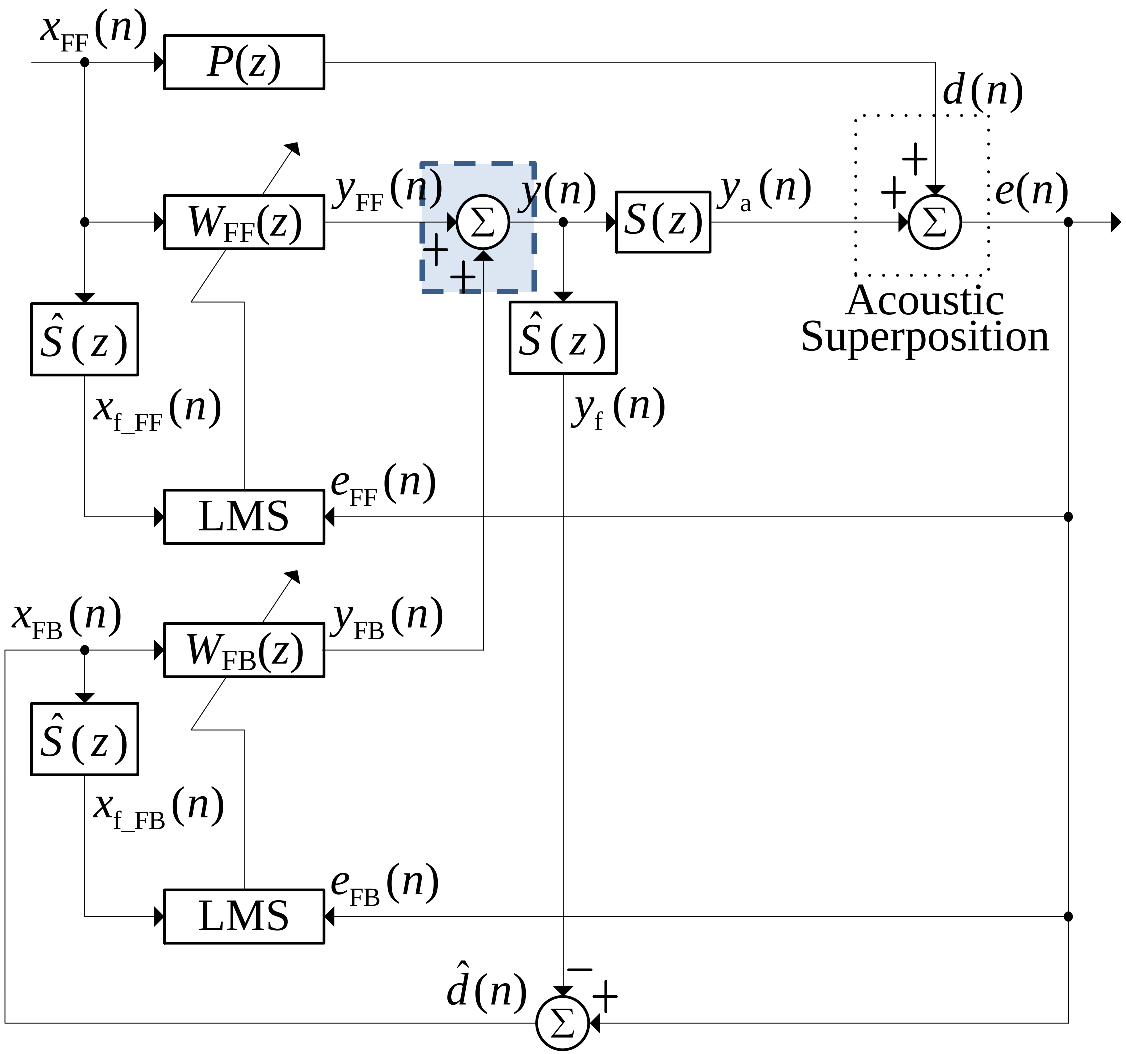
$$W^o(z) = \frac{P(z)}{S(z)}.$$

Feedback ANC



- Also known as internal model control (**IMC**) feedback ANC.
- 1 sensor and 1 actuator (1-dim case)
- Estimate the primary noise based on error signal and adaptive filter output.
- Adaptive filter loop & Synthesis loop
- Generally good in cancelling out predictable tonal noise.

Hybrid FF and FB ANC



- Dual role in canceling primary noise from reference sensor & residual noise pickup by error sensor.
- Good performance for both narrow and broadband noise and offer flexibility in ANC design.
- May come with extra computational cost.

More info in Prof. Kuo's classic textbook on Active Noise Control

Performance Hierarchy of ANC

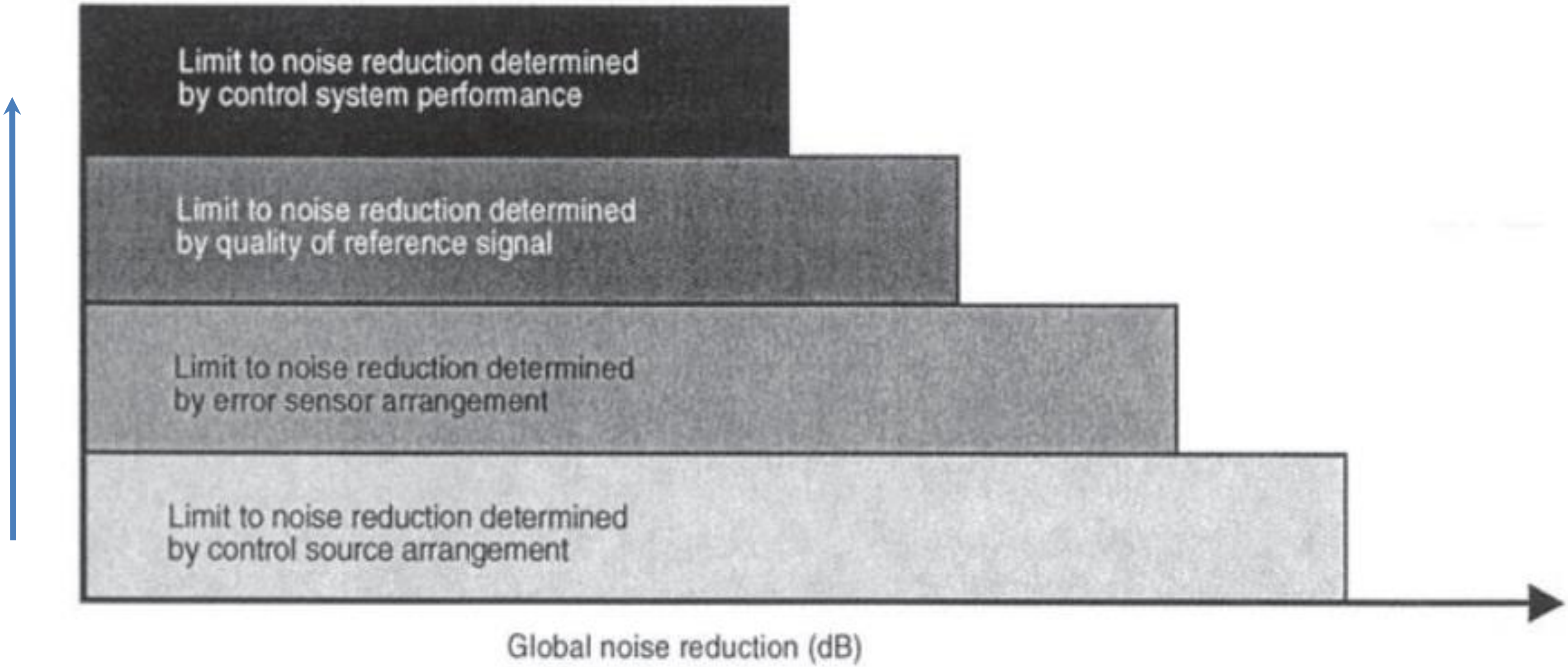


Figure 2.5 Performance hierarchy of an active noise control system.

Picture extracted from Hansen's book

Integrated ANC with other functions

A. Wearable Devices

- **Headsets**
- **Motorcycle Helmet**
- **Hearing Aids**

B. Other Applications:

- Automotive
- Snoring cancellation
- Incubator
- MRI

Use carefully fitted, and if possible, noise-cancelling earphones/headphones. If suited to the individual user, earphones and headphones allow music to be heard clearly at lower levels of volume. Noise-cancelling earphones and headphones cut down background noise, so that you can hear sounds at lower volumes than otherwise needed.

a noisy
e from
ers.
venues

io devices.
olume
audio
y would
e.

ure noise
bout the

Use carefully fitted, and if possible, noise-cancelling earphones/headphones. If suited to the individual user, earphones and headphones allow music to be heard clearly at lower levels of volume. Noise-cancelling earphones and headphones cut down background noise, so that you can hear sounds at lower volumes than otherwise needed.

Limit time spent engaged in noisy activities

The duration of the exposure to noise is one of the key factors contributing to overall sound energy levels. There are ways to minimize the duration. It is advisable to:

Have short listening breaks. When going to nightclubs, discotheques, bars, pubs, sporting events and other noisy places, take short listening breaks to help reduce the overall duration of noise exposure.

risk for noise-induced hearing loss from your personal audio device. Applications or "apps" accessible through the smartphones can help by displaying noise intensity levels in decibels and indicating whether exposure to a particular level of sound is risky. Know your product, its safety features and its safe listening level.

Heed the warning signs of hearing loss

Seek help from a hearing health care professional in case of tinnitus or difficulty in hearing high-pitched sounds such as doorbells, telephones or alarm clocks; understanding speech, especially over the telephone; or following conversations in noisy environments, such as in restaurants or venues for other social gatherings.

Get regular hearing check-ups

Take advantage of the services offered by schools, workplaces and communities for periodic hearing check-ups, as such screening can help to identify the onset of hearing loss at an early stage.

International Ear Car Day on 3rd March '15

Make Listening Safe

Make Listening Safe. Once you lose your hearing, it won't come back!

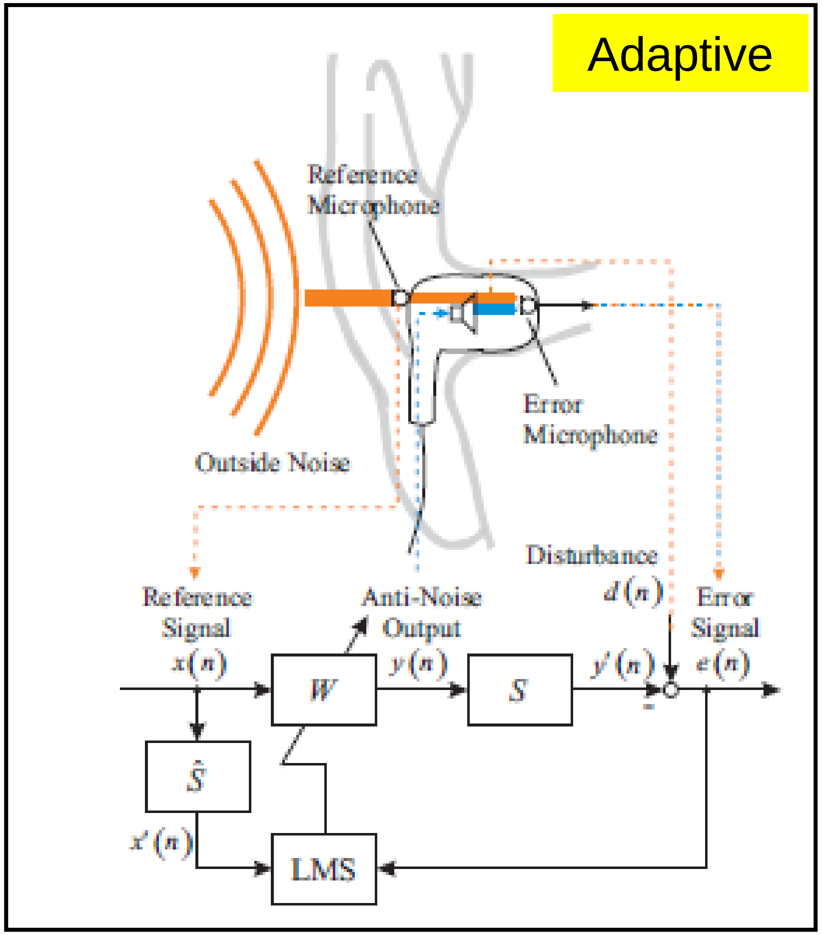
Department for Management of NCDs, Disability, Violence and Injury Prevention (NVI)
World Health Organization
20 Avenue Appia
CH-1211 Geneva 27
Switzerland
Tel +41-22-791-2064
www.who.int/pbd/deafness/activities/MLS

World Health Organization

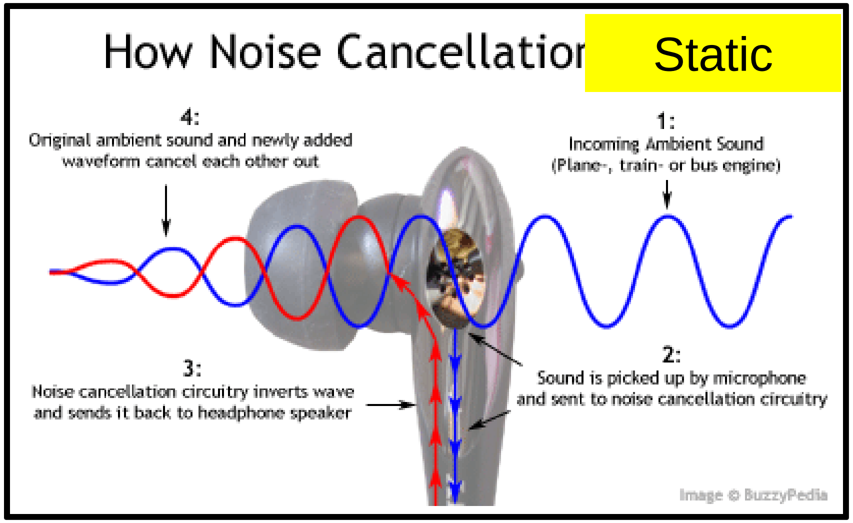
Applying ANC to Headphones/Earphones

- Music playback or received speech signal
- Communication
- Perceptual Compensation
- Noise Reduction
- Augmented Reality Headsets

Earphones: Adaptive vs Static ANC



Picture from Priese



Picture from internet

In general, the adaptive filter always achieves about 20 dB attenuation in the frequency range between 100 Hz and 1000 Hz; but static ANC results in only 10 dB reduction.

S Priese et al, "Adaptive feedforward control for active noise cancellation in-ear headphones," 164th Meeting of the Acoustical Society of America, Kansas City, Missouri, 22 - 26 October 2012

Commercial ANC Headphones with digital adaptation

- Some of the commercial ANC headphones are still based on analog controller
- Recent years have seen more digital adaptive ANC in commercial products and patents:

Brand	Model Number/Name	Weblink
Sennheiser	Sennheiser S1 Digital Aviation Headset	http://en-us.sennheiser.com/aviation-headset-pilot-headset-active-noise-cancelling-s1-digital
Sony	MDR-10R Noise Cancelling Headphones	http://www.sony.com.sg/product/mdr-10rnc
Sony	MDR-ZX750BN/B Bluetooth & Digital Noise Cancelling Headphones (Black)	http://www.sony.com.sg/product/mdr-zx750bn
Samsung	Samsung Level Over	http://www.amazon.com/Samsung-Cancelling-Wireless-Headphones-Smartphones/dp/B00KGGK738
Beats	Beats Noise Cancelling Headphones	http://sg.beatsbydre.com/headphones/

ANC Headset with audio playback

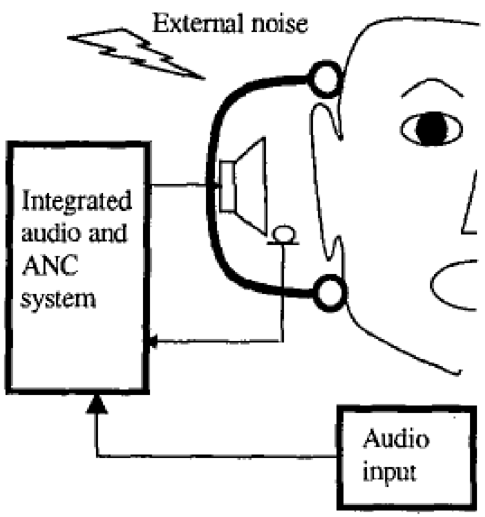
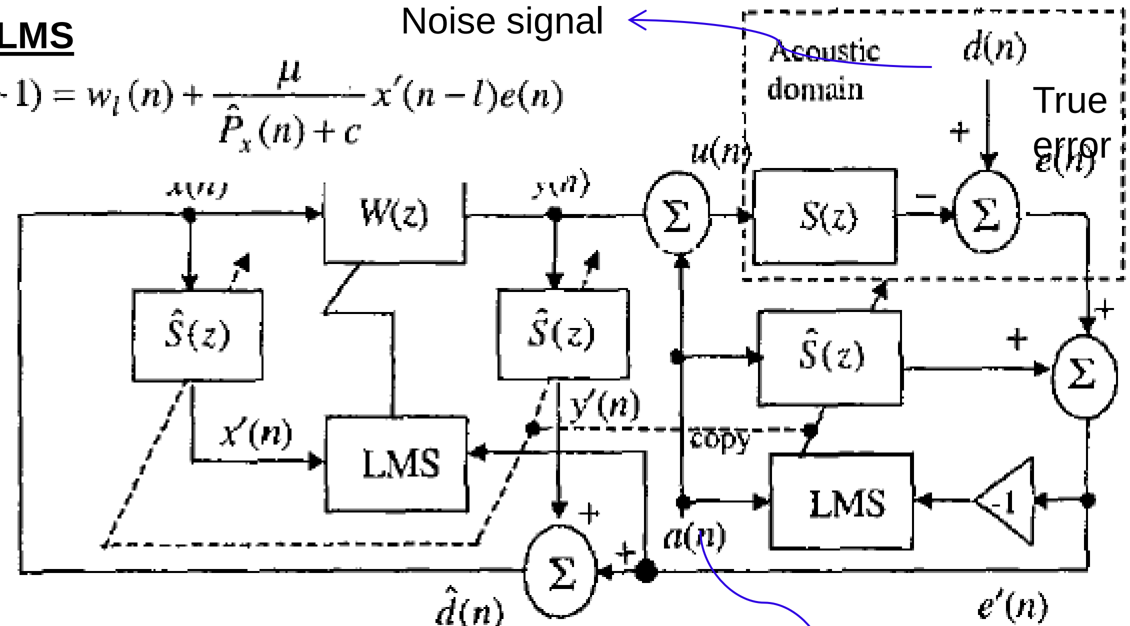


Fig. 1 Integrated audio and ANC headsets

FXNLMS

$$w_l(n+1) = w_l(n) + \frac{\mu}{\hat{P}_x(n)+c} x'(n-l)e(n)$$



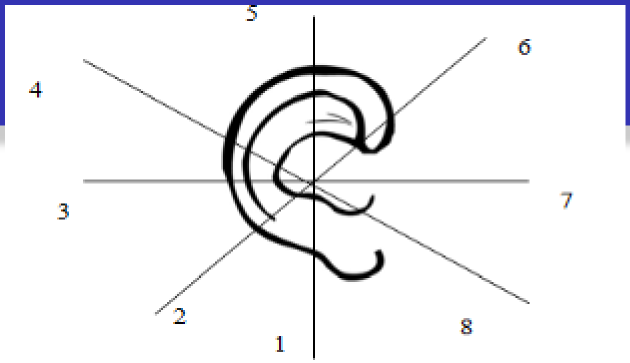
$$E'(z) = D(z) - A(z)S(z) - Y(z)S(z) + \hat{S}(z)A(z).$$

Assuming $\hat{S}(z) = S(z)$

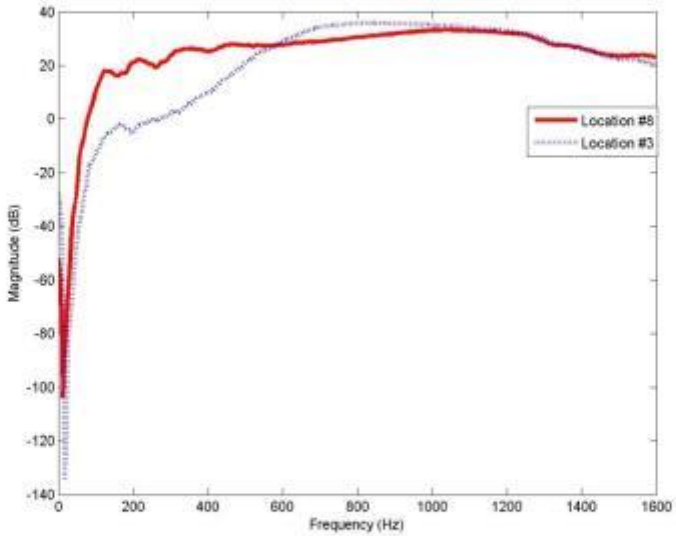
$$E'(z) = D(z) - Y(z)S(z)$$

Pseudo error does not contain any audio signal

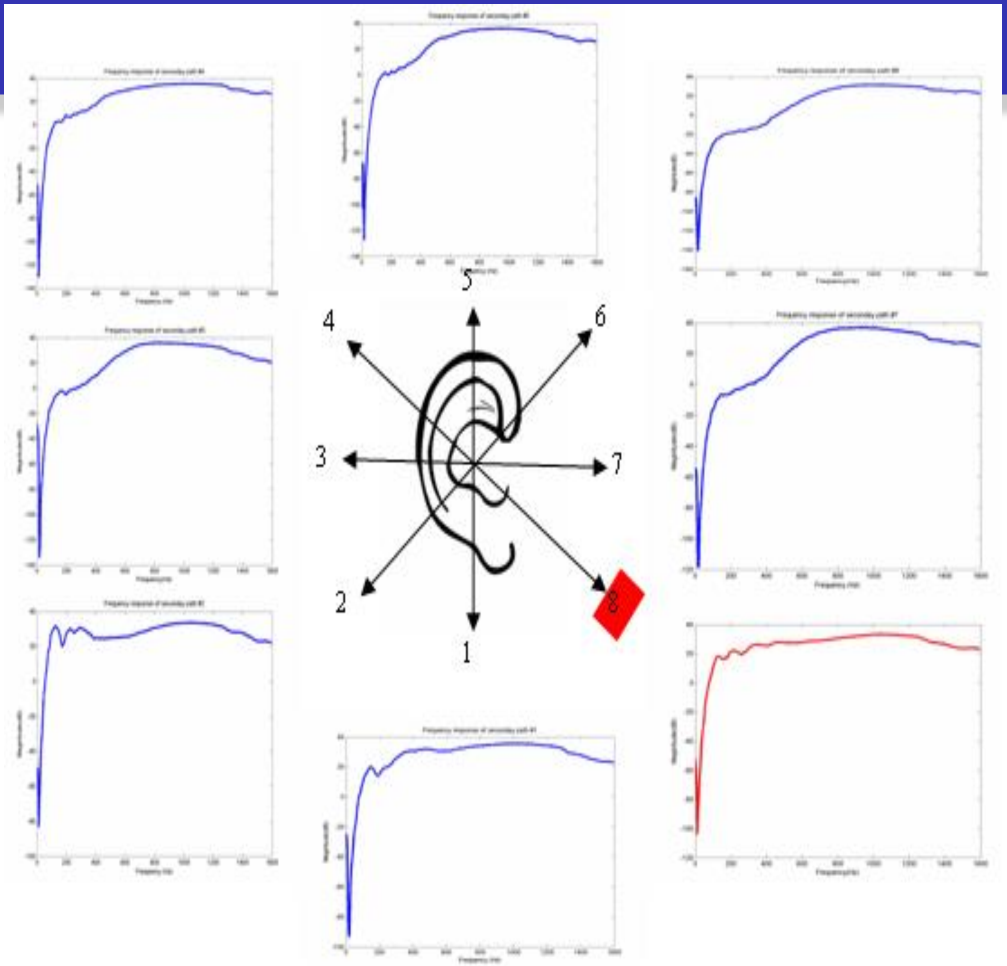
Woon S. Gan and Sen M. Kuo, "AN INTEGRATED AUDIO AND ACTIVE NOISE CONTROL HEADSETS", IEEE Transactions on Consumer Electronics, Vol. 48, No. 2, MAY 2002



Possible error microphone locations corresponding to ear surface



Magnitude responses of secondary path *S*(*z*) at microphone locations #3 and #8



Frequency responses of secondary-path *S*(*z*) at microphone locations #1 to #8

Optimum location #8 for the error microphone

Results

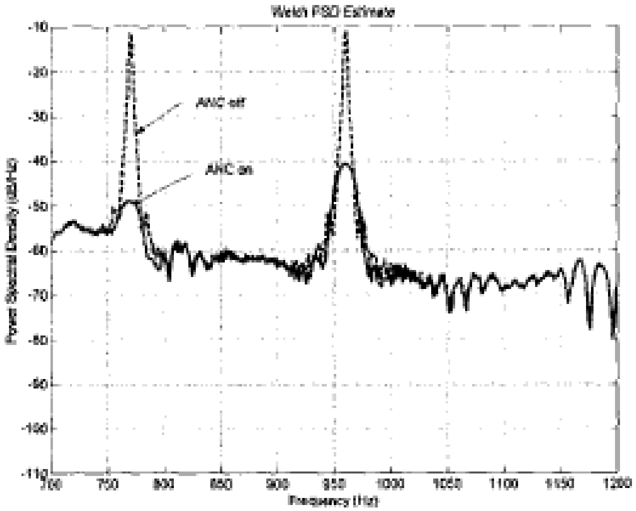


Fig.4. Noise spectrum for the error signal with (dotted line) and without (solid line) using IFBANC under repeating siren disturbance

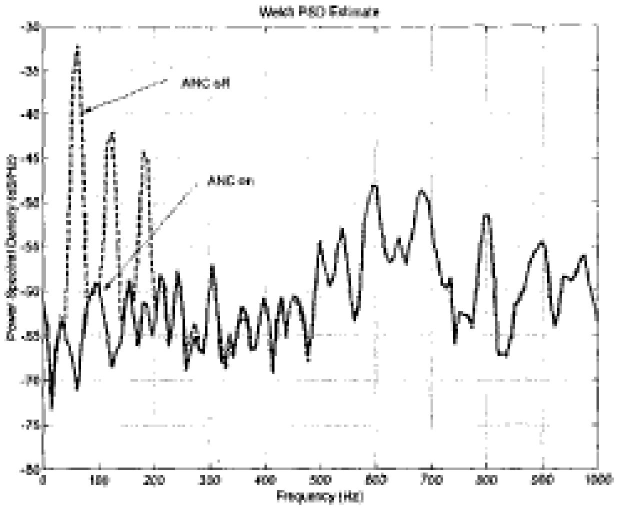


Fig. 5. Noise spectral for the error signals with (dotted line) and without (solid line) using IFBANC under an engine disturbance.

Advantage of the integrated approach

- Good estimation of the true residual noise $e'(n)$ without interfering with the audio signal $a(n)$;
- Large step size can be used in adapting the cancellation filter $W(z)$;
- The adaptive feedback ANC technique provides a more accurate noise cancellation since the microphone is placed inside the ear-cup of the headset;
- The system uses single microphone per ear cup, thus produces a compact, lower power consumption and a cheaper solution;
- The audio signal can be neatly used to drive both on-line and offline modelling of the secondary path transfer function.

Extension to the Integrated ANC+Audio Playback

- Include communication features in integrated ANC + Audio playback.
- Our paper also compare the noise cancellation performance with commercial-off-the-shelve ANC headphones
- Our adaptive ANC integrated headphones outperformed the commercial one by 15-20 dB at low frequency below 183 Hz.
- More objective and subjective evaluation can be conducted to test out other parameters of the ANC headphones.

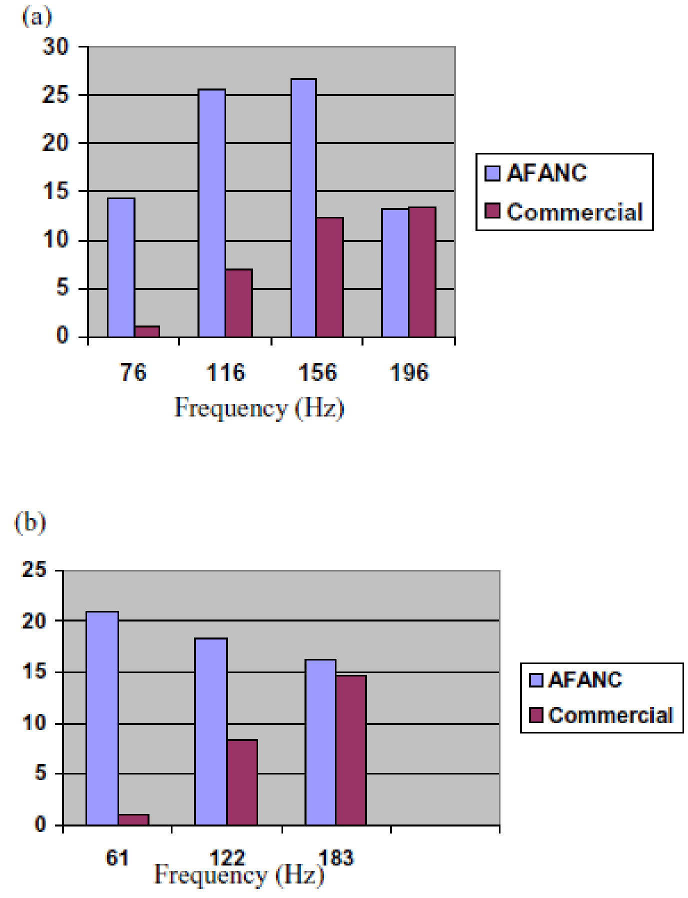
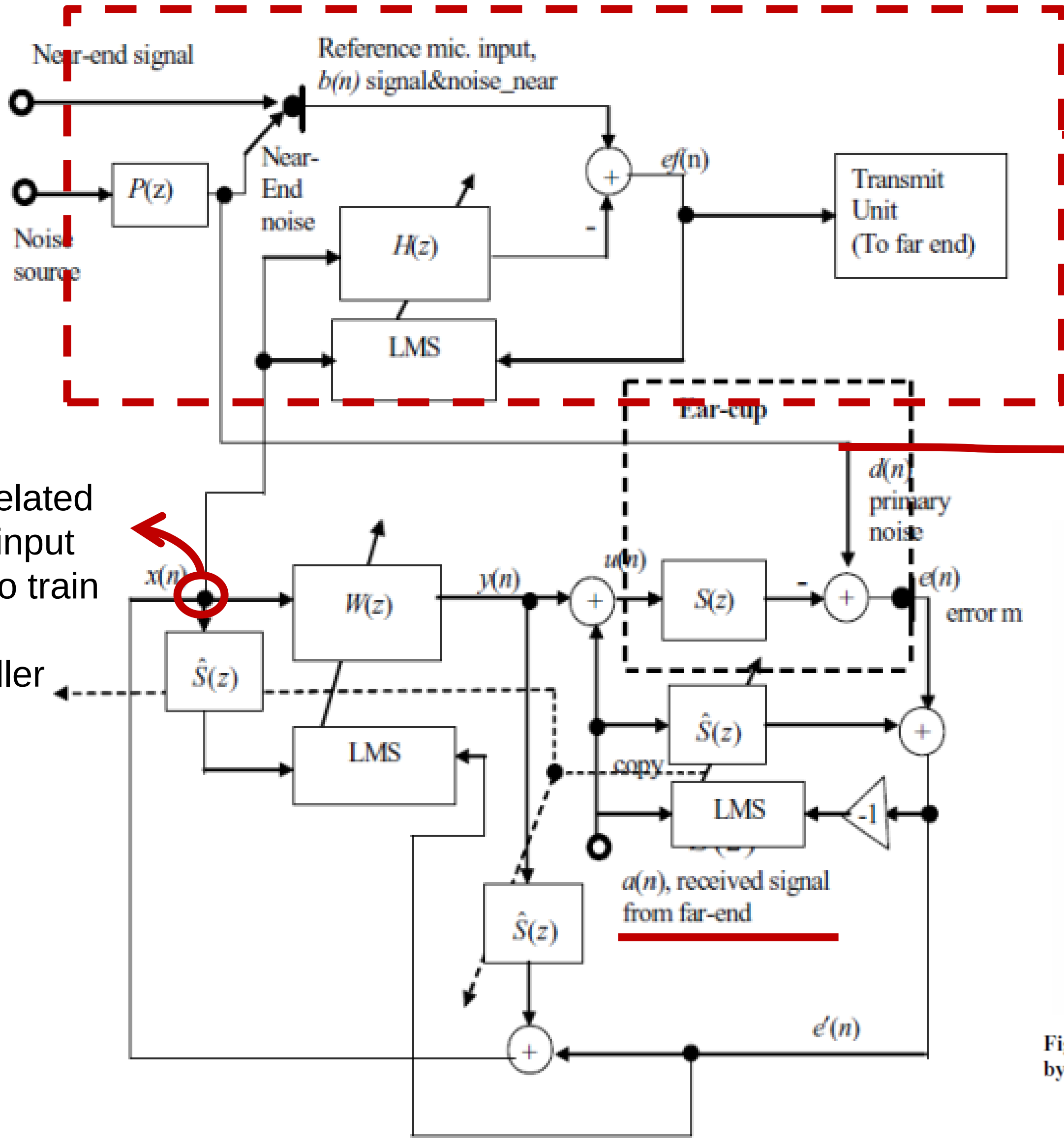


Figure 8. Net noise cancellation comparison for the proposed AFANC and a commercial noise canceling headset for engine noise at (a) 2,200 rpm and (b) 3,700 rpm

Woon S. Gan, Sohini Mitra and Sen M. Kuo, "Adaptive Feedback Active Noise Control Headset: Implementation, Evaluation and Its Extensions", IEEE Transactions on Consumer 976 Electronics, Vol. 51, No. 3, AUGUST 2005

ANC Headset with audio playback and speech communication



To remove near-end noise before transmission. Use an adaptive noise canceller. **Note:** ***a purely electrical cancellation; not acoustic.***

Acoustic path

A correlated noise input used to train Noise canceller

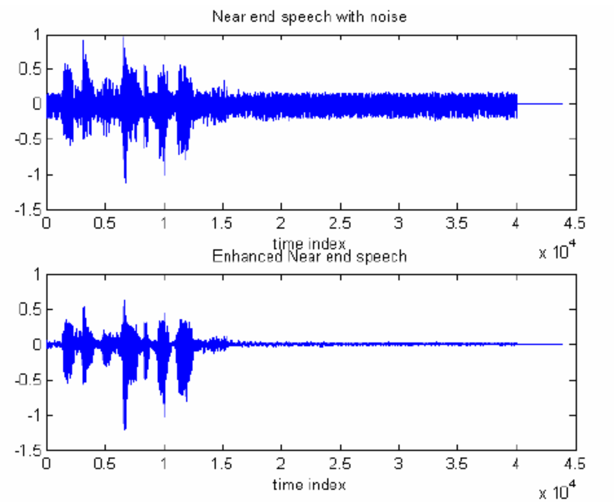
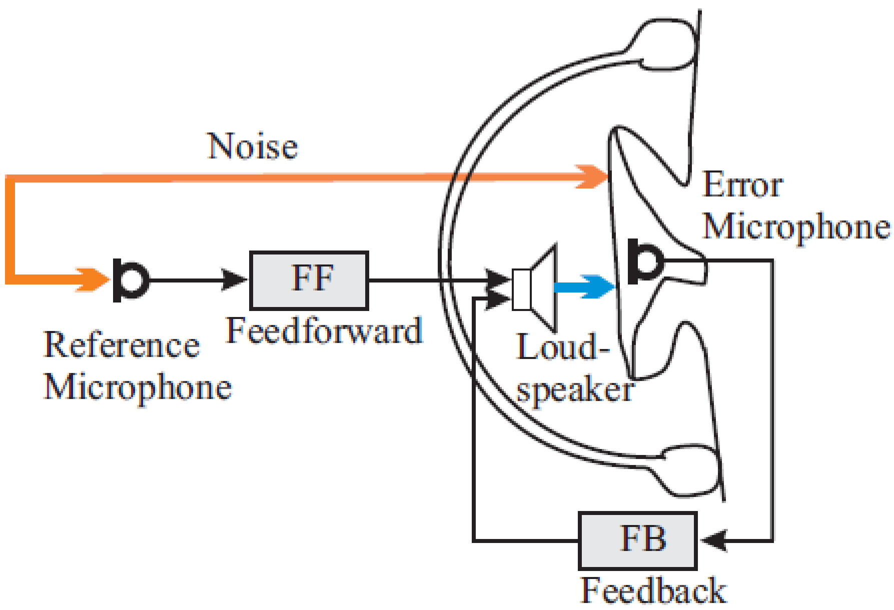


Figure 13. The top and bottom plots show the original signal corrupted by engine noise and enhanced near-end speech respectively

Extension to a Hybrid ANC system



- Combined both feedforward and feedback ANC techniques to further improved on its broadband noise cancellation performance.
- Also perform better than the feedback ANC for narrow frequency separation engine noise. (~ 6 to 10 dB better)

Yong-Kim Chong, Liang Wang, See-Chiat Ting and Woon-Seng Gan," INTEGRATED HEADSETS USING THE ADAPTIVE HYBRID ACTIVE NOISE CONTROL SYSTEM", ICICS 2005

Broadband Feedback ANC

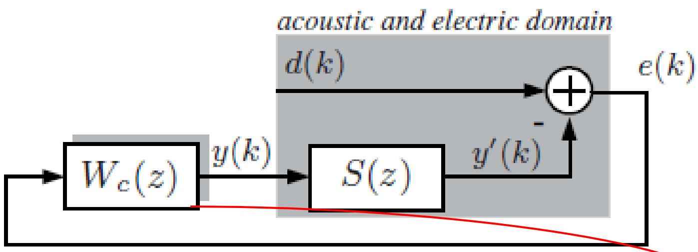


Fig. 2. Model for a classical feedback ANC system.

Analog

Combined Analog-Digital

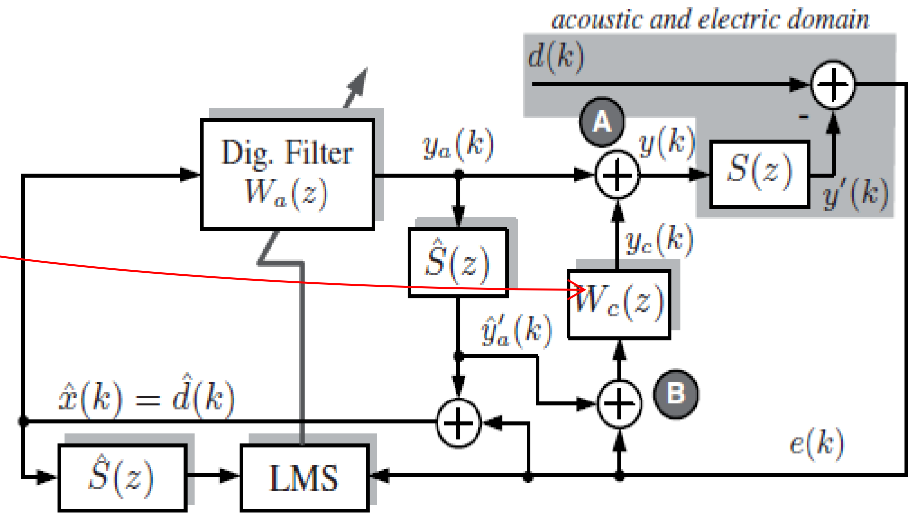


Fig. 4. The novel hybrid feedback ANC system based on the FxLMS approach.

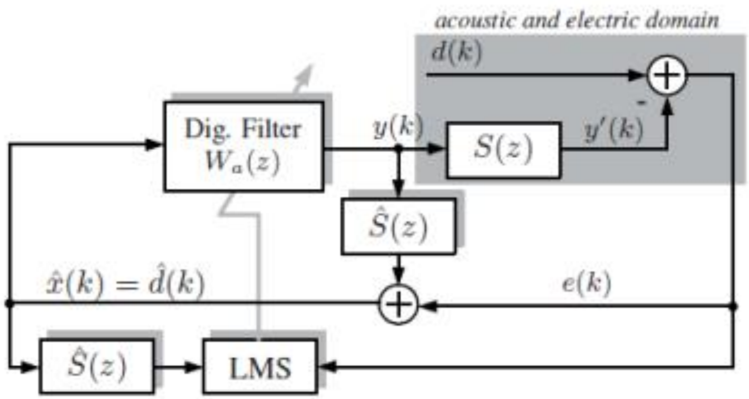


Fig. 3. Model for an adaptive feedback ANC system involving the FxLMS approach.

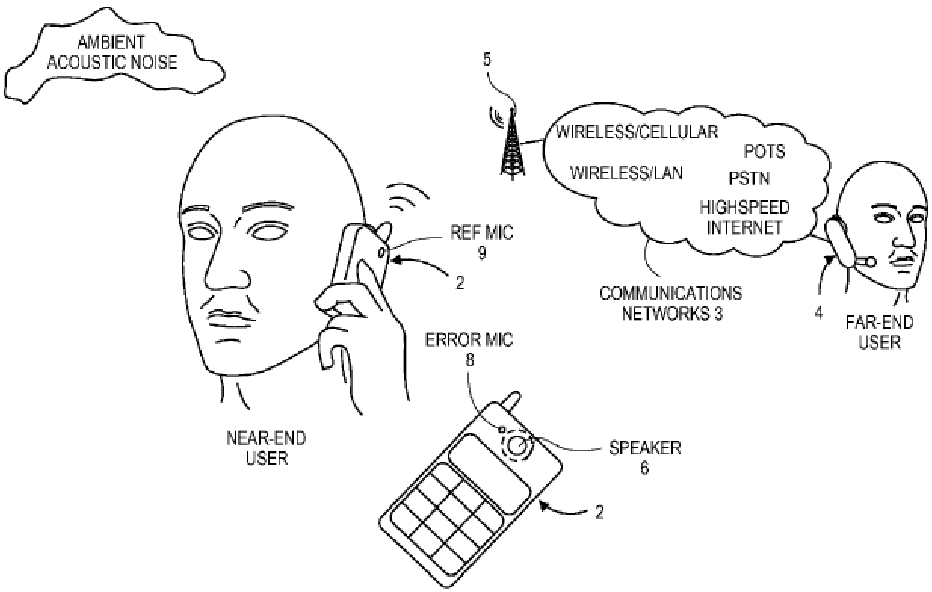
Digital

In general, attenuation of noise in lower frequency areas as well as periodic noise components in all frequency areas

Picture from

Thomas Schumacher, et al, " ACTIVE NOISE CONTROL IN HEADSETS: A NEW APPROACH FOR BROADBAND FEEDBACK ANC," ICASSP 2011

Apple Patent Application on ANC for iPhone (US20130259250 A1)



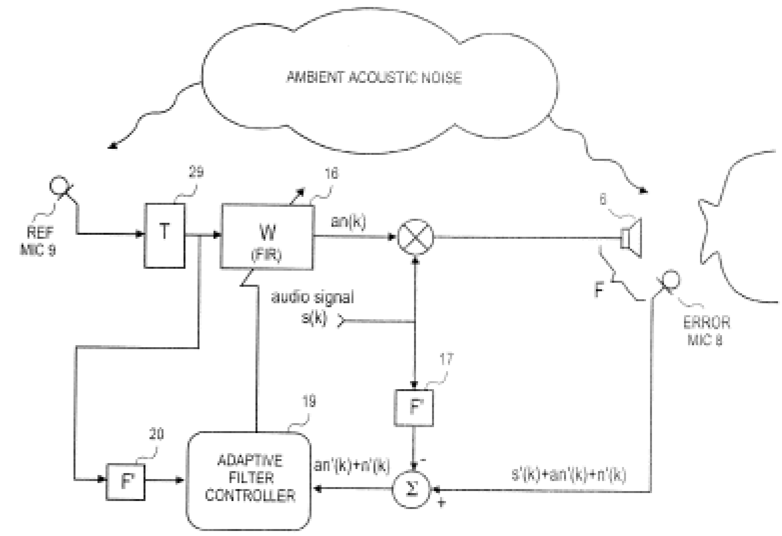


FIG. 2

- Patent application on a Feedforward ANC system for use in a portable audio device using an adaptive digital filter and a reference/error microphones.
- Novelty is adding a **non-adaptive pre-shaping filter (minimum phase)** and has > 2dB more gain over low frequency.
- Integrating speech/audio signal into the FF ANC system.

Picture from patent

Another Apple Patent on ANC Decision in a Portable Audio Device (US8515089)

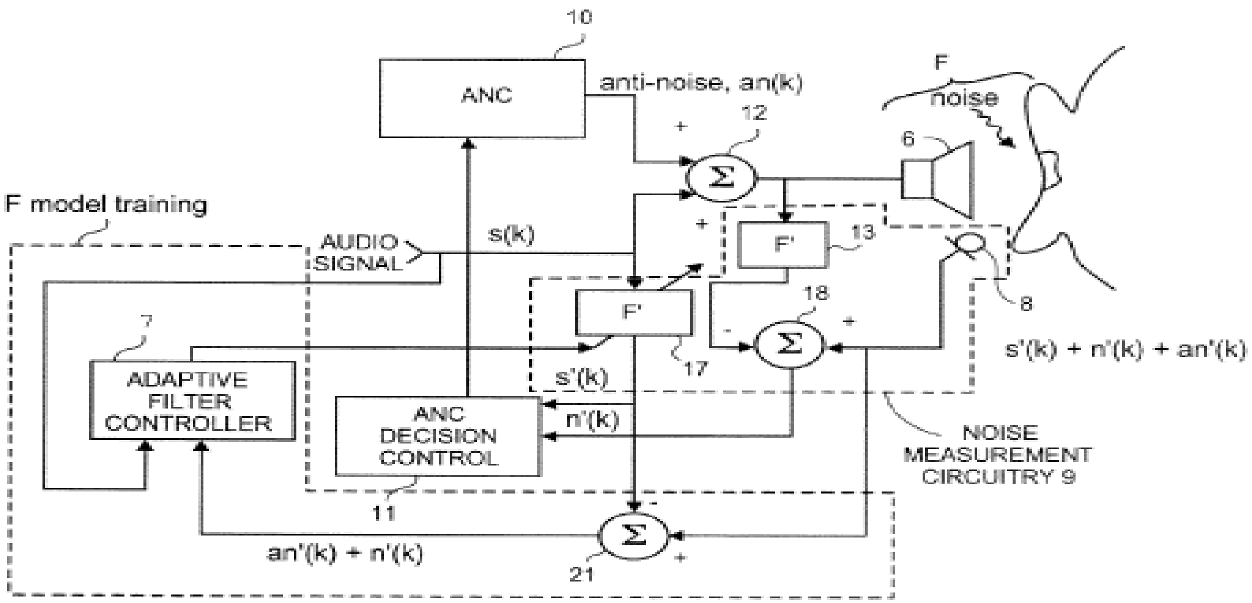


FIG. 2

- To control the ambient acoustic noise outside the device that may be heard by a user of the device.
- Depending on the signal strength of the sound emitted from the earpiece speaker and the ambient acoustic noise to **activate/deactivate ANC.**

Picture from patent

ANC Patent Application from Fraunhofer (US2015/0003625 A1)

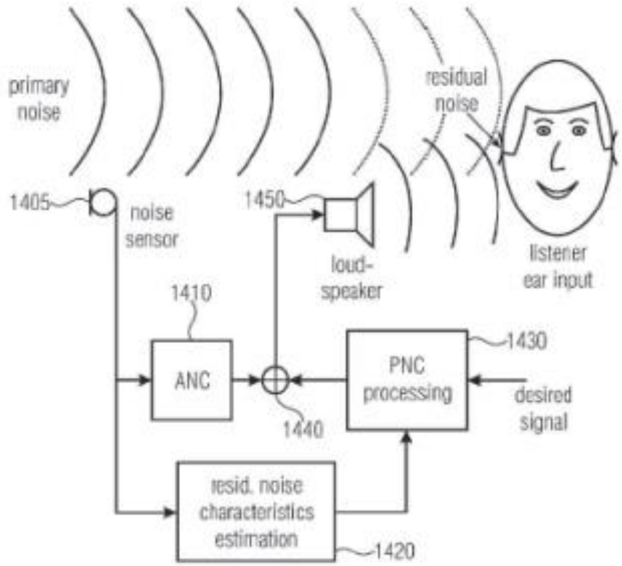


FIGURE 14

- **<u>Objective</u>**: Restore the original audio loudness and timbre
- Improve perceived quality of sound reproduction of an audio
- Include a perceptual noise compensator to generate noise-compensated signal based on **audio signal** and **residual noise characteristic.**

Picture from patent

Conclusions on ANC Headphones

- Fundamental theories, algorithms, and experiments of ANC have been well established over the last few decades. ***Focus now on how we can deploy and integrate ANC in existing and other new applications***.
- Witnessing an increased activities from the industry applying ANC techniques into their products (especially in **ANC headsets**, **hearing aids** and **automotive** applications)
- Issues with maintainability and costs become lesser concern with ***better, reliable, small form factor, and low cost*** sensors → suited for wearable consumer devices.

Key References on ANC Headphones

[1] P. Leug, “Process of silencing sound oscillations,” U.S. Patent 2043416, 1936.

[2] E. D. Simshauser and M. E. Hawley, “The noise canceling headset -an active ear defender,” Journal of the Acoustical Society of America, vol. 27, pp. 207–217, 1955.

[3] S. M. Kuo and D. R. Morgan, Active noise control systems: algorithms and DSP implementations. New York, NY, USA: Wiley Interscience, 1996.

[4] S. J. Elliott, Signal Processing for Active Control. San Diego, CA.: Academic Press, 2001.

[5] Y. Kajikawa, W. S. Gan and S. M. Kuo, “Recent advances on active noise control: open issues and innovative applications,” APSIPA Trans. Sign. Inf. Process., vol. 1, pp. 1–21, Aug. 2012.

[6] Kuo, S.M.; Chuang, H.; Mallela, P.P.; "Integrated automotive signal processing and audio system", IEEE Trans Consumer Electronics, Volume 39, Issue 3, August 1993, pp. 522-532.

[7] S Priese et al, “Adaptive feedforward control for active noise cancellation in-ear headphones,” 164th Meeting of the Acoustical Society of America, Kansas City, Missouri, 22 - 26 October 2012

[8] Woon S. Gan and Sen M. Kuo, “AN INTEGRATED AUDIO AND ACTIVE NOISE CONTROL HEADSETS”, IEEE Transactions on Consumer Electronics, Vol. 48, No. 2, MAY 2002

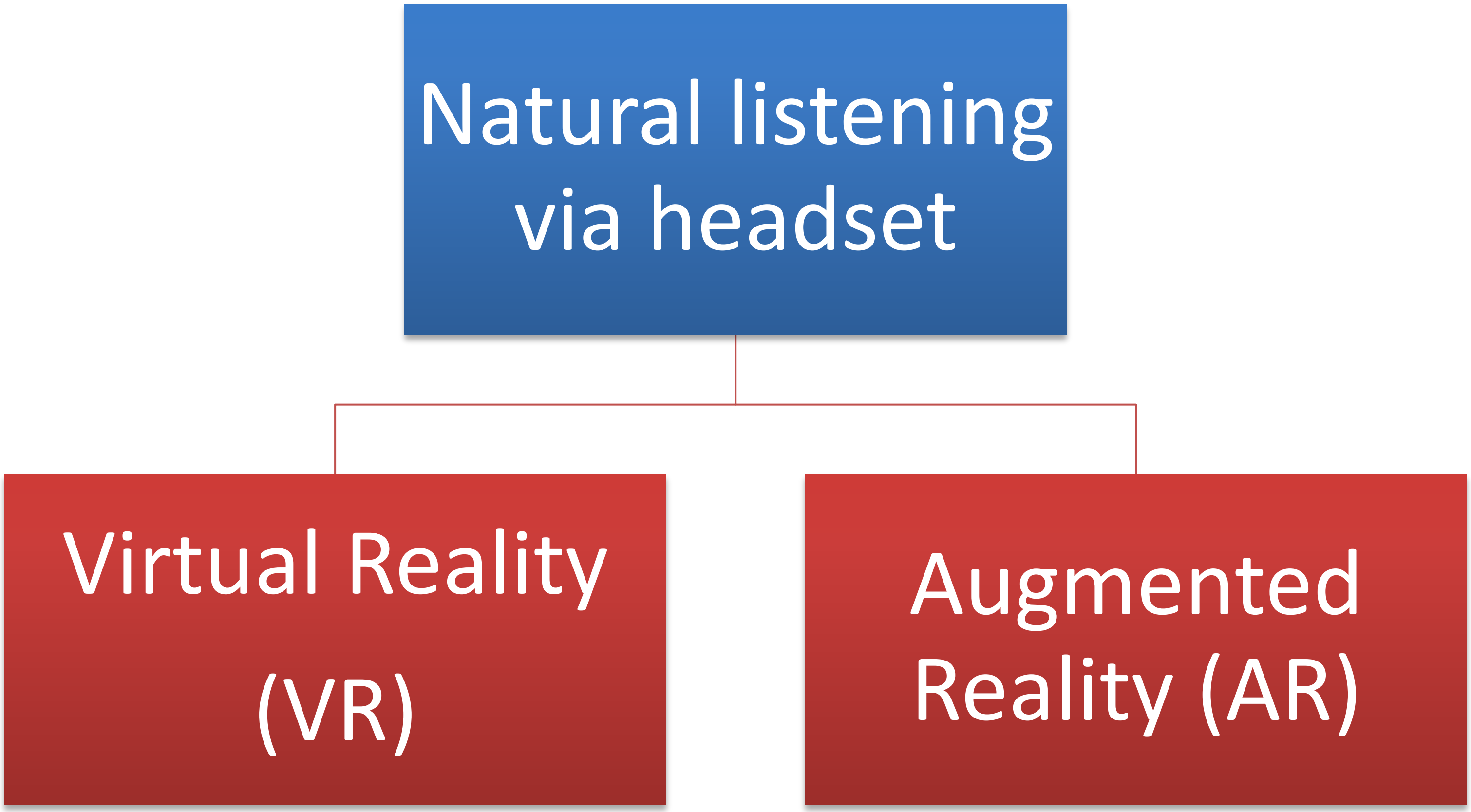
[9] Thomas Schumacher, et al, “ ACTIVE NOISE CONTROL IN HEADSETS: A NEW APPROACH FOR BROADBAND FEEDBACK ANC,” ICASSP 2011

[10] Yong-Kim Chong, Liang Wang, See-Chiat Ting and Woon-Seng Gan,” INTEGRATED HEADSETS USING THE ADAPTIVE HYBRID ACTIVE NOISE CONTROL SYSTEM”, ICICS 2005

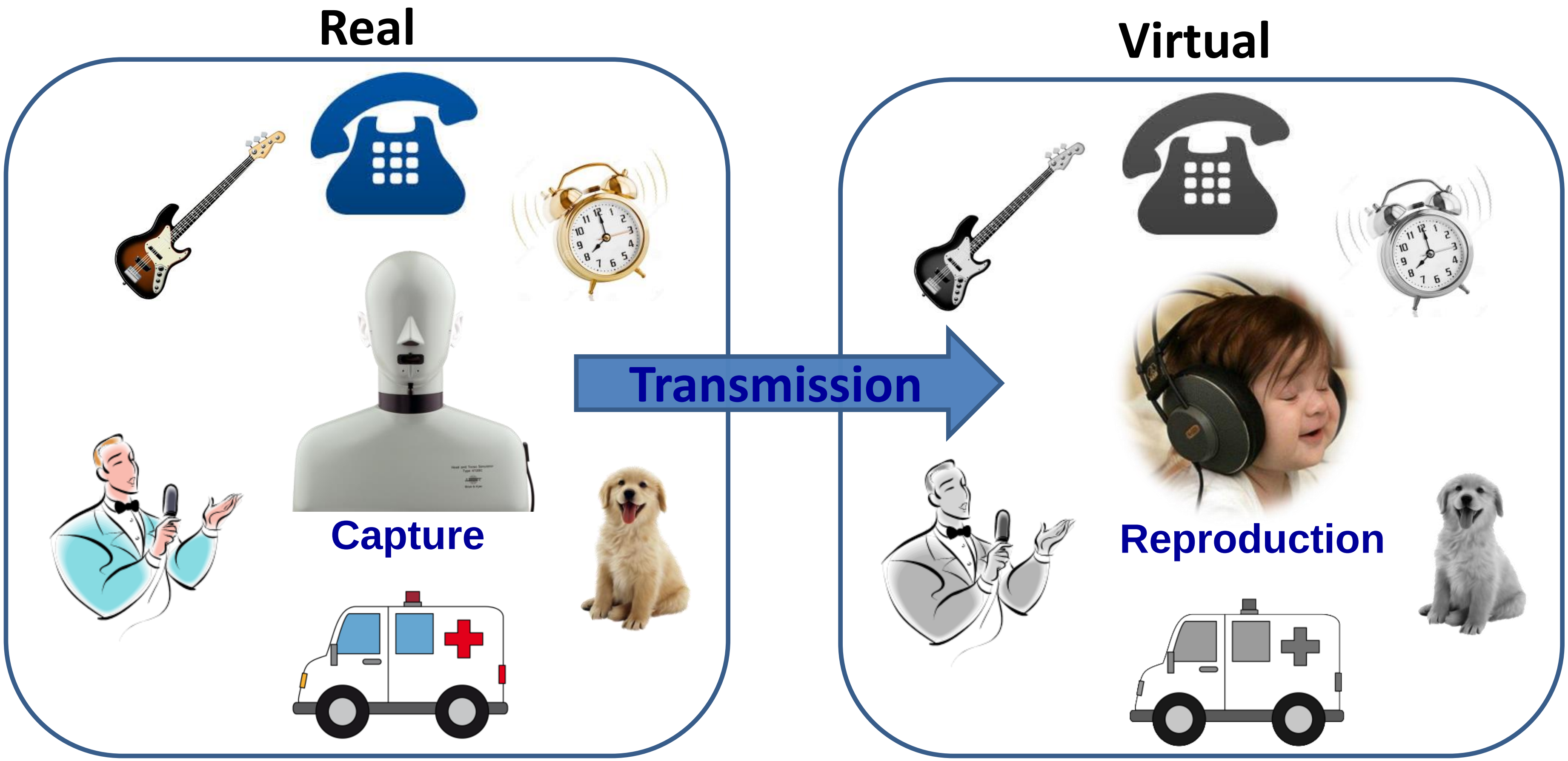
[11] Woon S. Gan, Sohini Mitra and Sen M. Kuo, “Adaptive Feedback Active Noise Control Headset: Implementation, Evaluation and Its Extensions”, IEEE Transactions on Consumer 976 Electronics, Vol. 51, No. 3, AUGUST 2005

Module IV
Natural Augmented Reality Headsets

Natural Listening for VR and AR



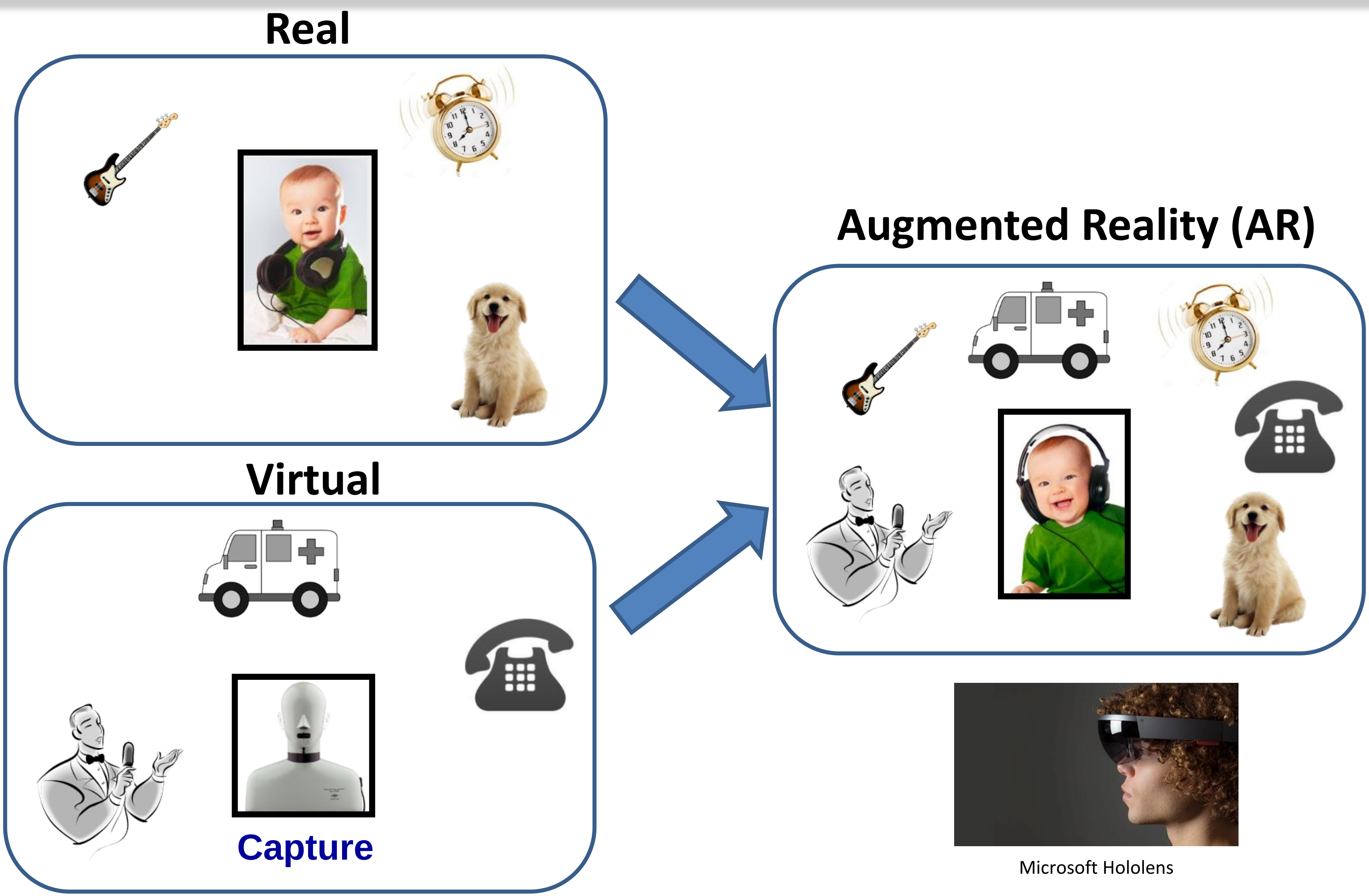
Natural Listening in Virtual Reality



Sony Project Morpheus


Oculus rift

Natural Listening in Augmented reality



Microsoft Hololens

Augmented Reality

Augmented reality is changing the way we live in real world

Wearable VR/AR devices:

Google-glass

Oculus rift

Microsoft Hololens

VR/AR applications:

Navigation

VR and AR world

Gaming

Related Works

1. *Augmented reality audio (ARA) headset* using in-ear headphones and external binaural microphones to assist the listener with pseudo-acoustic scenes[1].

Problem addressed: Blockage of natural sounds coming from outside. Using binaural microphones to capture, process and playback so as to make ARA headset acoustically transparent.

2. Surround sound reproduction over headphones with binaural microphones positioned inside ear cup near ear opening [2].

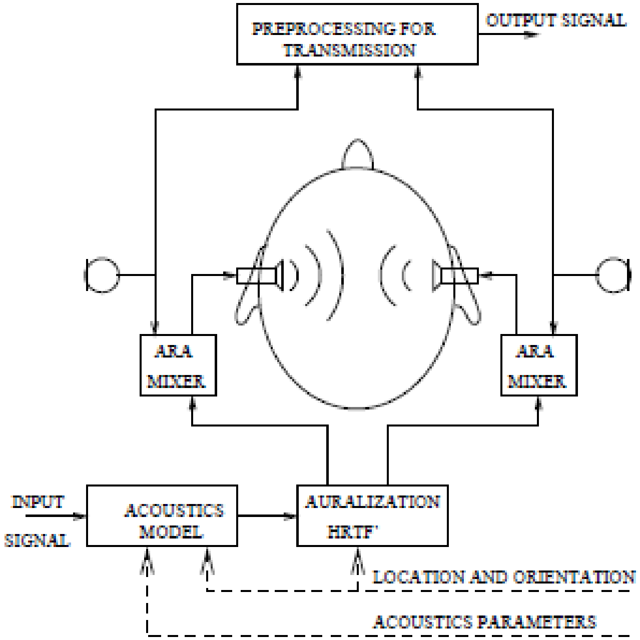
Problem addressed: Large localization errors due to non-individualized equalization of headphones. Using ANC technique to calibrate the system for every individual to achieve sound reproduction same as a multichannel setup.

[1] A. Härmä, J. Jakka, M. Tikander, M. Karjalainen, T. Lokki, J. Hiipakka, *et al.*, "Augmented reality audio for mobile and wearable appliances," *Journal of the Audio Engineering Society,* vol. 52, pp. 618-639, 2004
[2] D. W. Schobben and R. M. Aarts, "Personalized multi-channel headphone sound reproduction based on active noise cancellation," *Acta acustica united with acustica,* vol. 91, pp. 440-450, 2005.

ARA Headset - Overview

- Use of closed in-earphones to capture the external sound, mix with the virtual sounds and playback
- Basic idea is to relay the external sounds unaltered with minimum latency (<1ms)

Left and middle: ARA headset (Philips SHN2500)
Right: Prototype ARA mixer [2]

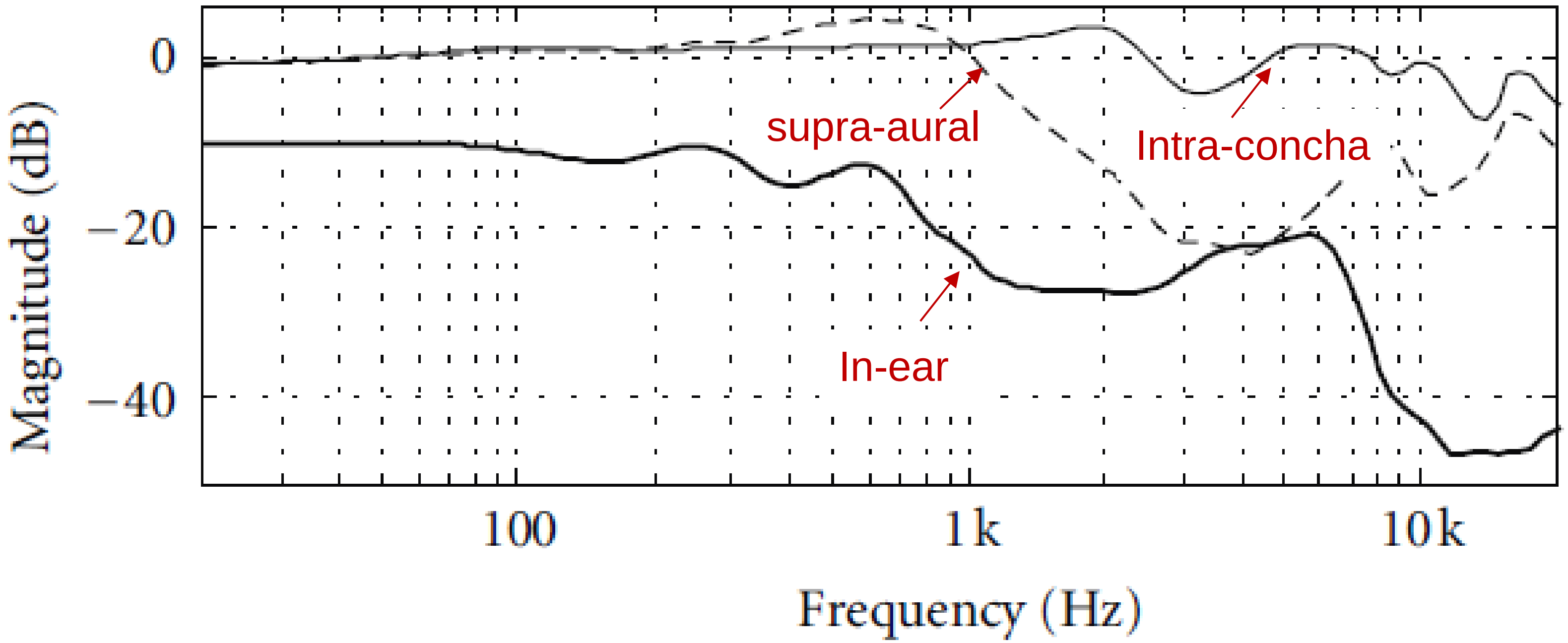


ARA headset system diagram [1]

[1] A. Härmä, J. Jakka, M. Tikander, M. Karjalainen, T. Lokki, J. Hiipakka, *et al.*, "Augmented reality audio for mobile and wearable appliances," *Journal of the Audio Engineering Society,* vol. 52, pp. 618-639, 2004
[2] M. Tikander, M. Karjalainen, & V. Riikonen , "An augmented reality audio headset". In *Proc. of the 11th Int. Conf. on Digital Audio Effects (DAFx-08), Espoo, Finland*, (2008, September).

ARA Headset (challenges)

- Blockage of external sounds by ARA headset

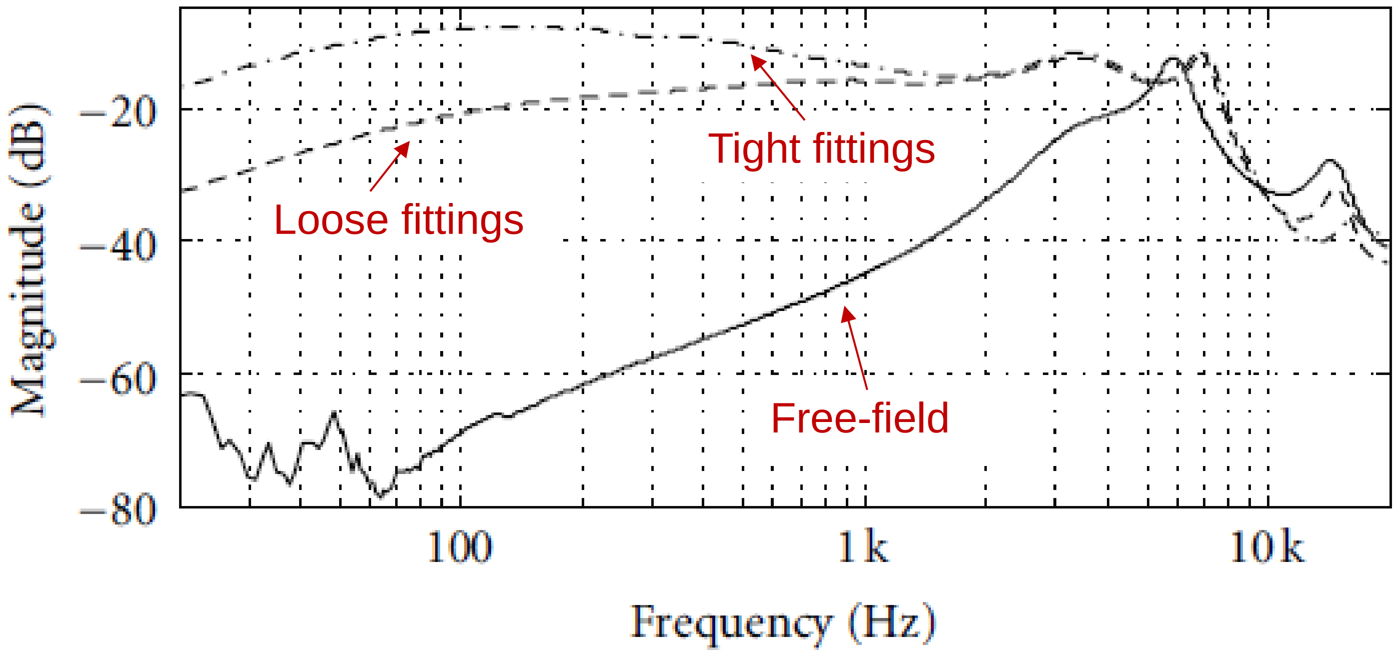


- Closed in-ear headphones with good fittings is necessary for good performance of ARA headset
- Loose fittings can dramatically change the attenuation

[1] J. Rämö, & V. Välimäki, (2012). Digital Augmented Reality Audio Headset.*Journal of Electrical and Computer Engineering*, *2012*.

ARA Headset (challanges)

- Closed in-ear phones modify the ear canal resonance (pressure chamber principle)

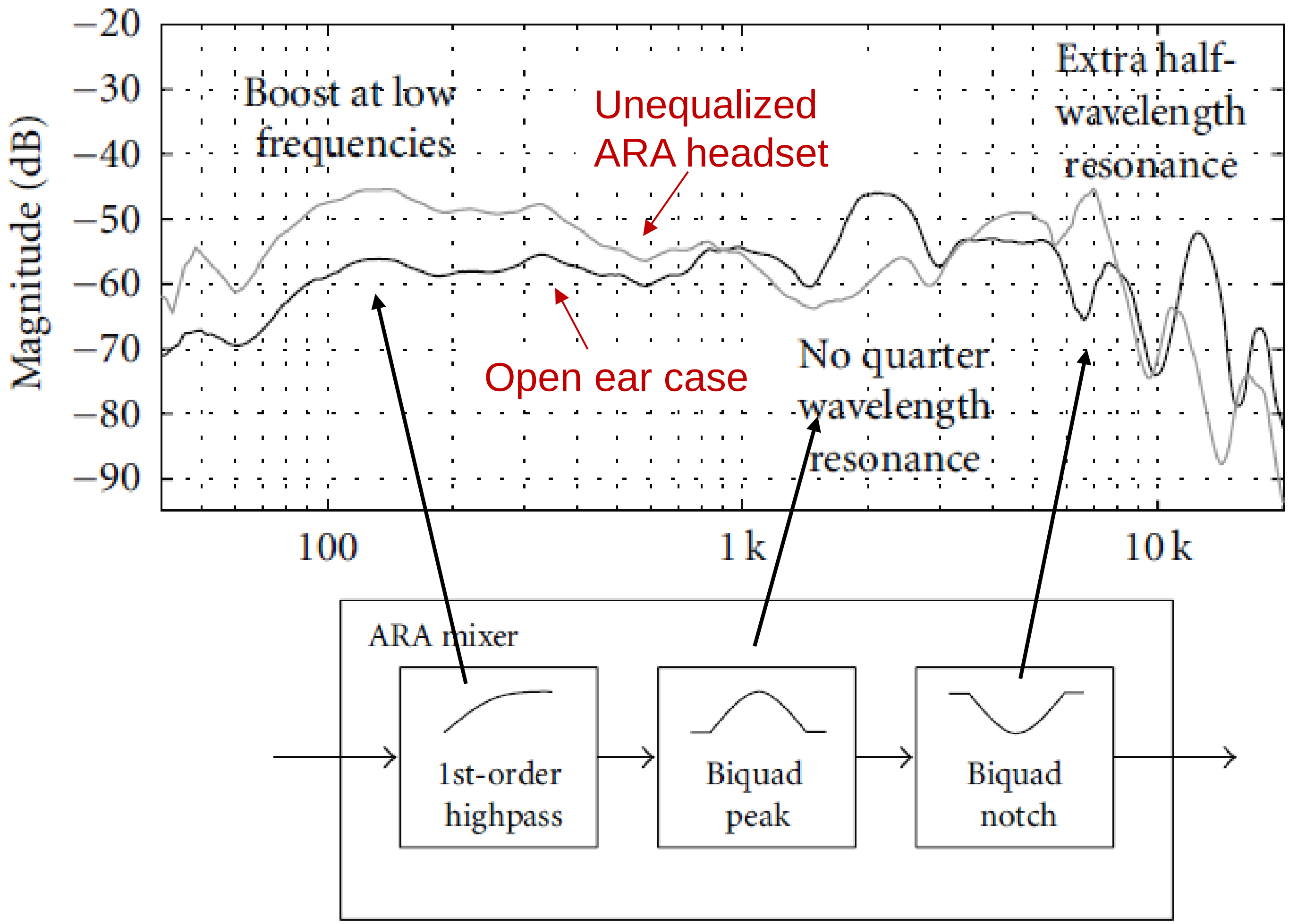


- Difficult to predict the headphone response if loosely fitted

[1] J. Rämö, & V. Välimäki, (2012). Digital Augmented Reality Audio Headset.*Journal of Electrical and Computer Engineering*, *2012*.

ARA Headset (Design of ARA mixer)

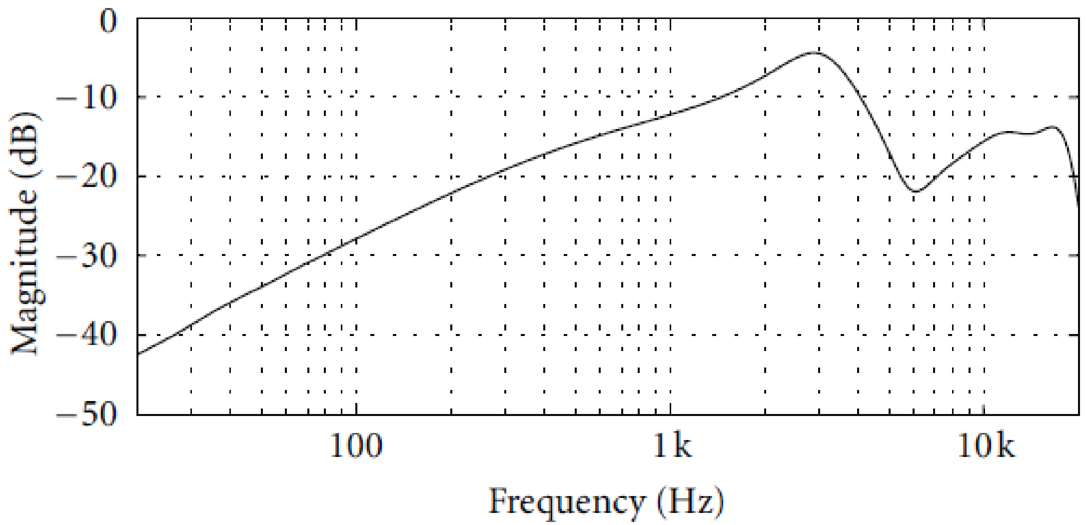
- ARA headset equalization

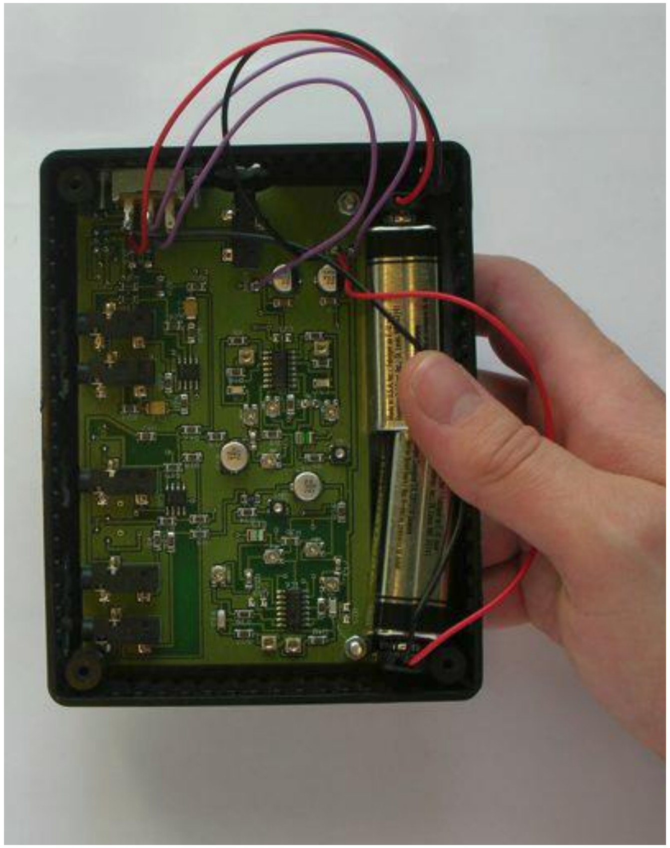


[1] J. Rämö, & V. Välimäki, (2012). Digital Augmented Reality Audio Headset.*Journal of Electrical and Computer Engineering*, *2012*.

ARA Headset

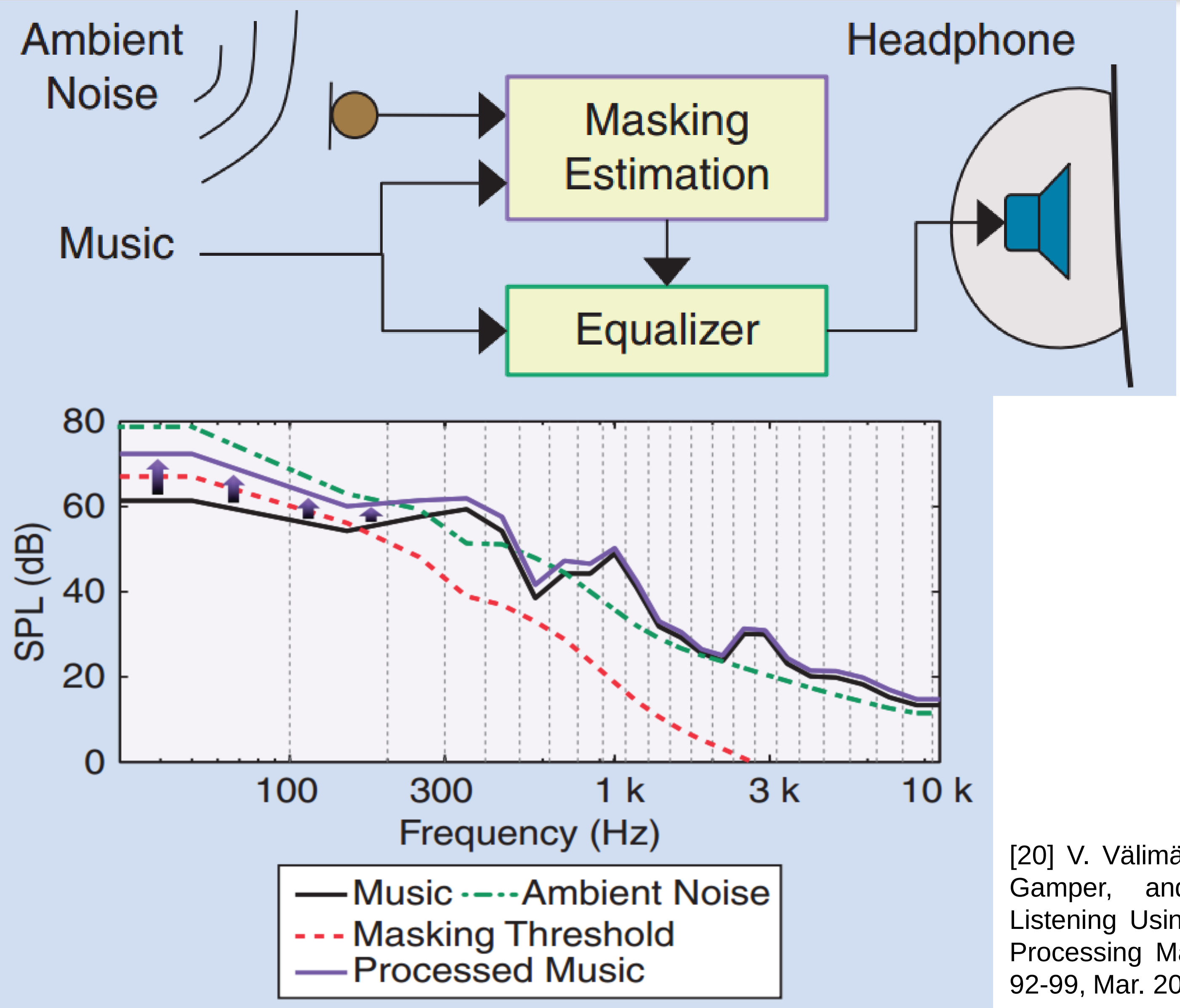
- Generic ARA equalization based on 4 individual measurements





[1] J. Rämö, & V. Välimäki, (2012). Digital Augmented Reality Audio Headset.*Journal of Electrical and Computer Engineering*, *2012*.

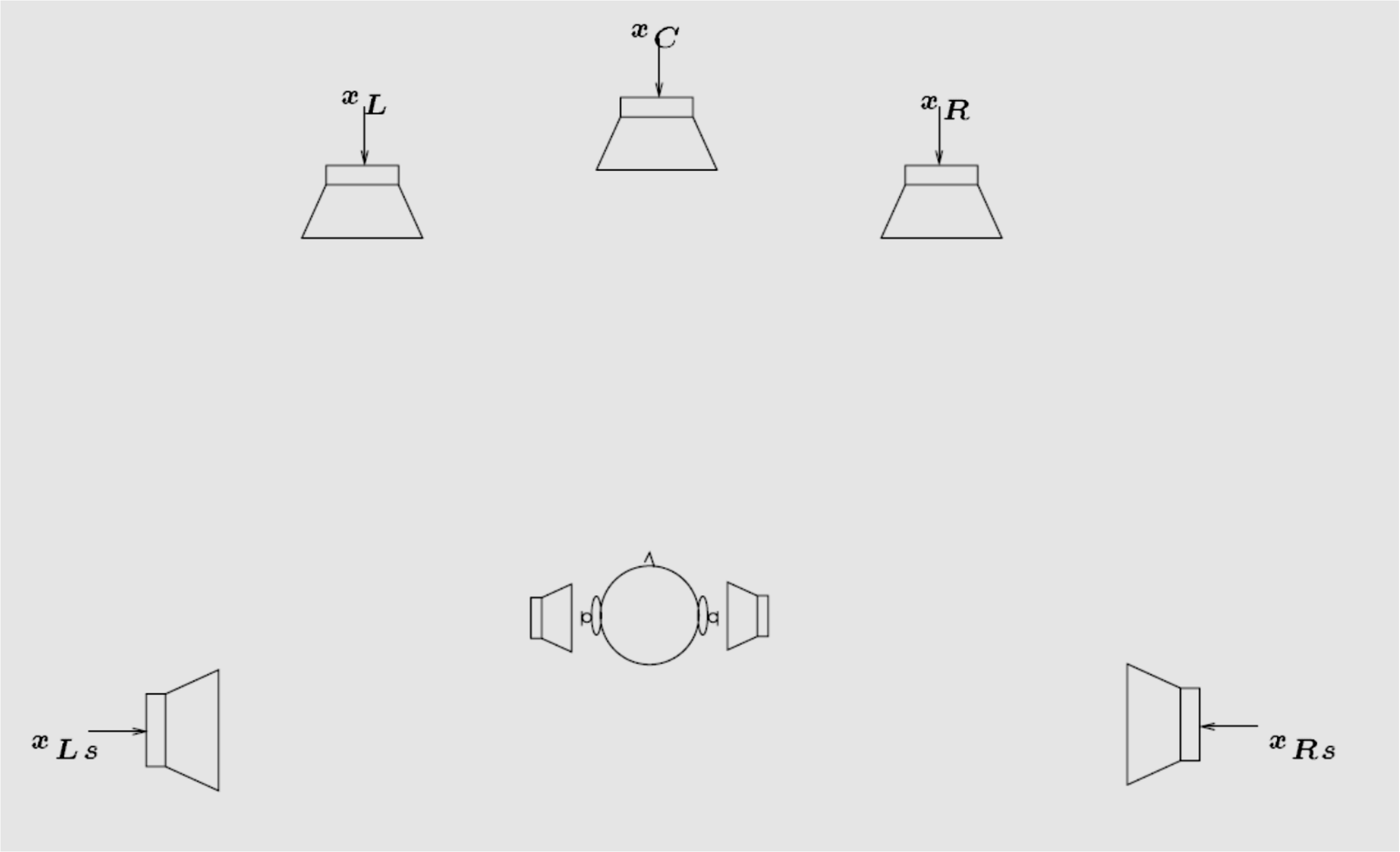
Masking in ARA Headset



[20] V. Välimäki, A. Franck, J. Rämö, H. Gamper, and L. Savioja, "Assisted Listening Using a Headset," IEEE Signal Processing Magazine, vo. 32, no. 2, pp. 92-99, Mar. 2015.

3D headphone sound reproduction using ANC

- Emulating a 5 channel sound reproduction setup through headset with two microphones

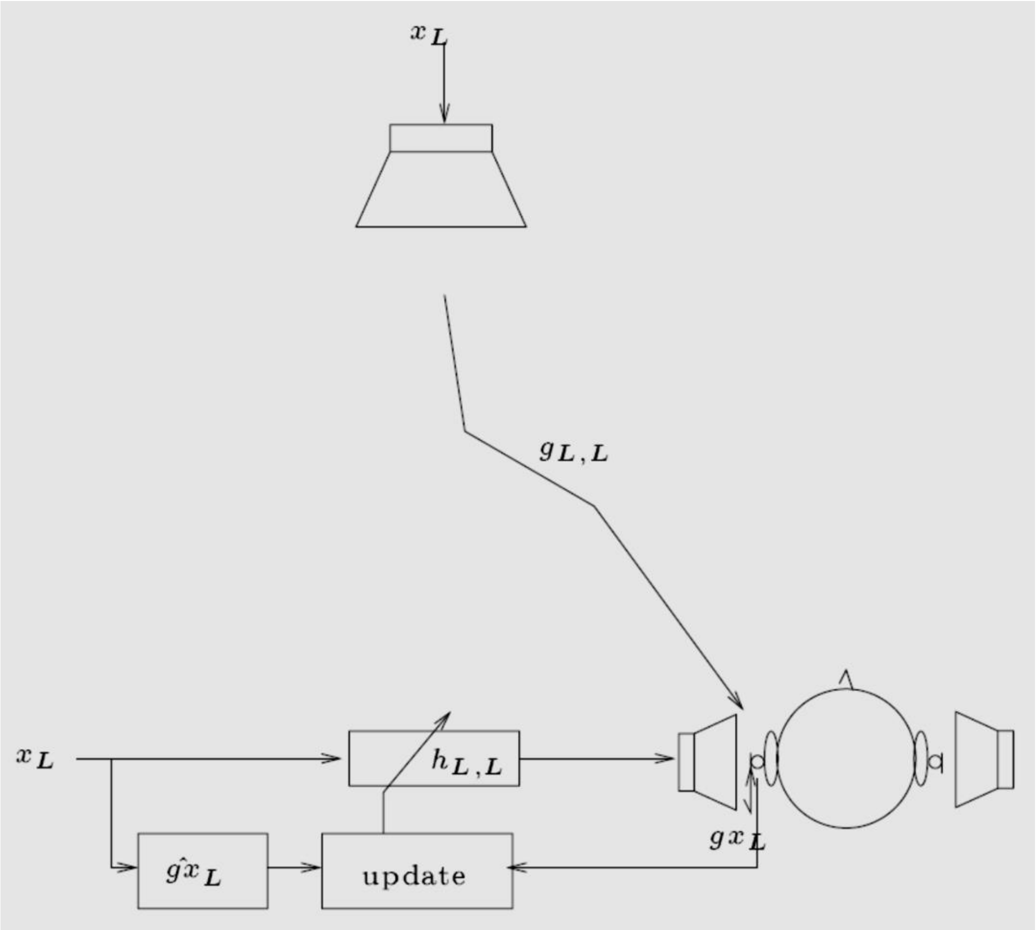


[2] D. W. Schobben and R. M. Aarts, "Personalized multi-channel headphone sound reproduction based on active noise cancellation," *Acta acustica united with acustica,* vol. 91, pp. 440-450, 2005.

3D headphone sound reproduction using ANC

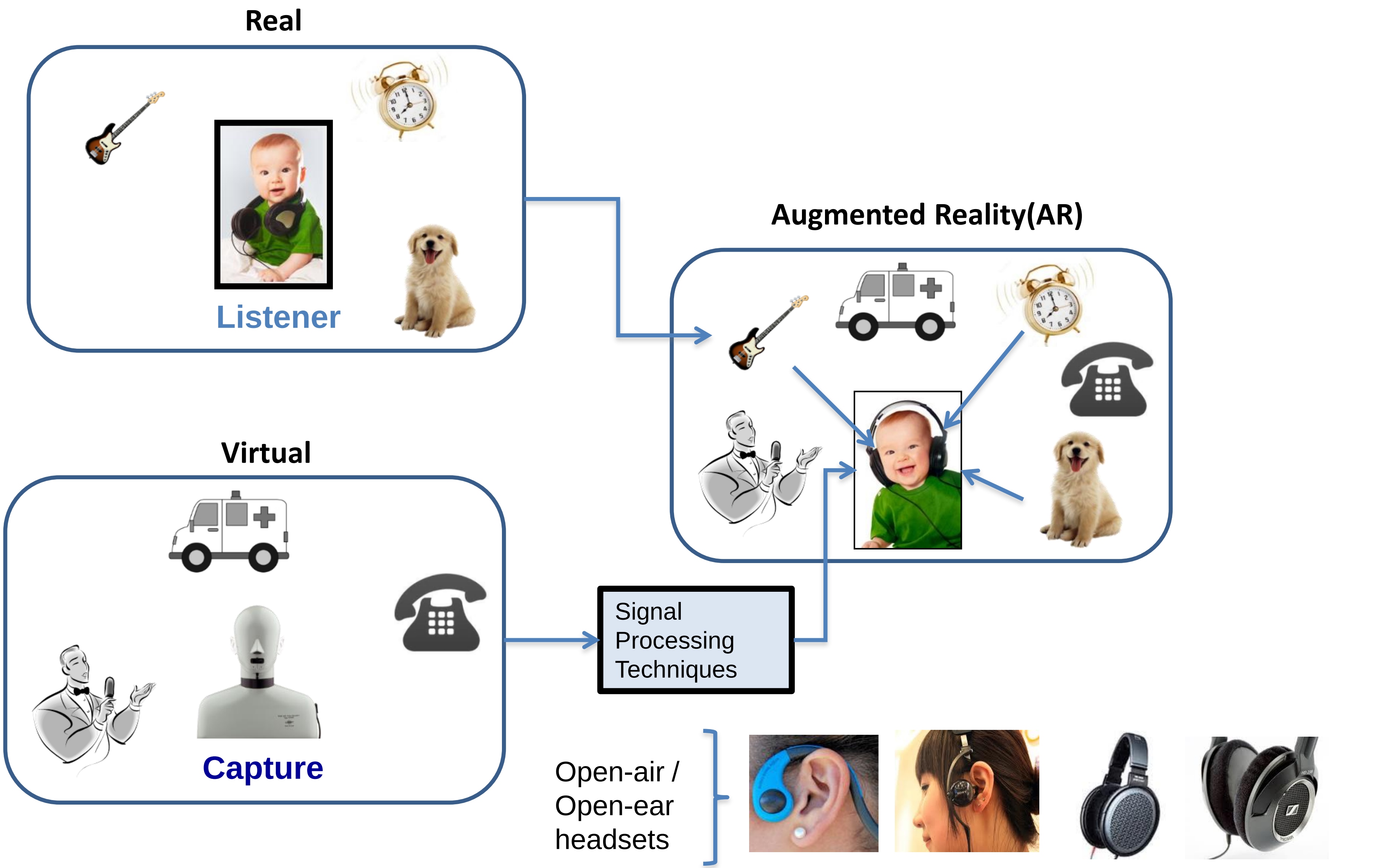
- Offline calibration process using adaptive process FxLMS

Headphone with integrated microphones



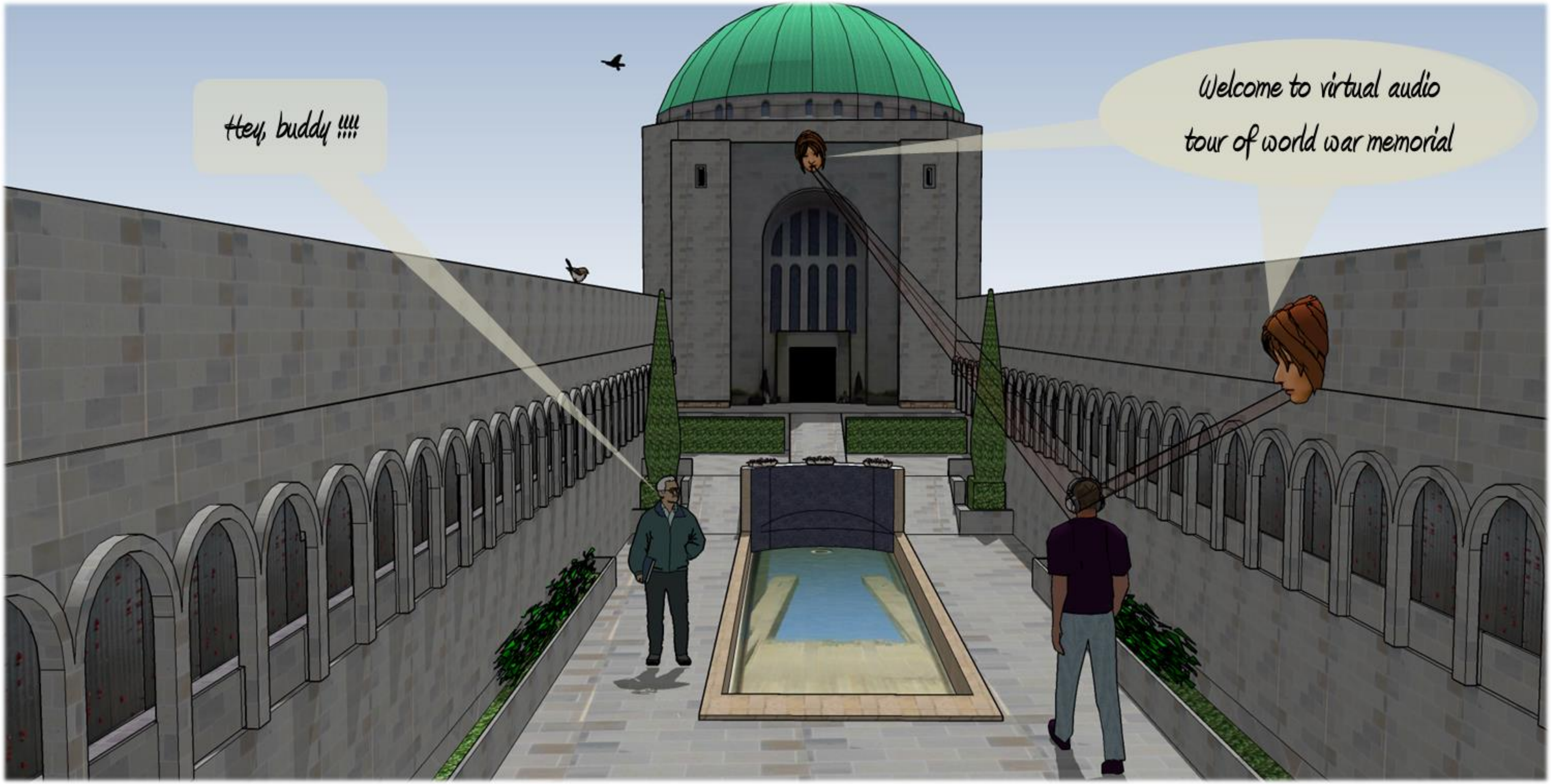
[2] D. W. Schobben and R. M. Aarts, "Personalized multi-channel headphone sound reproduction based on active noise cancellation," *Acta acustica united with acustica,* vol. 91, pp. 440-450, 2005.

Acoustic-Hear-Through AR (open-air headphones)



NAR headset potential applications:

Augmented Audio tour



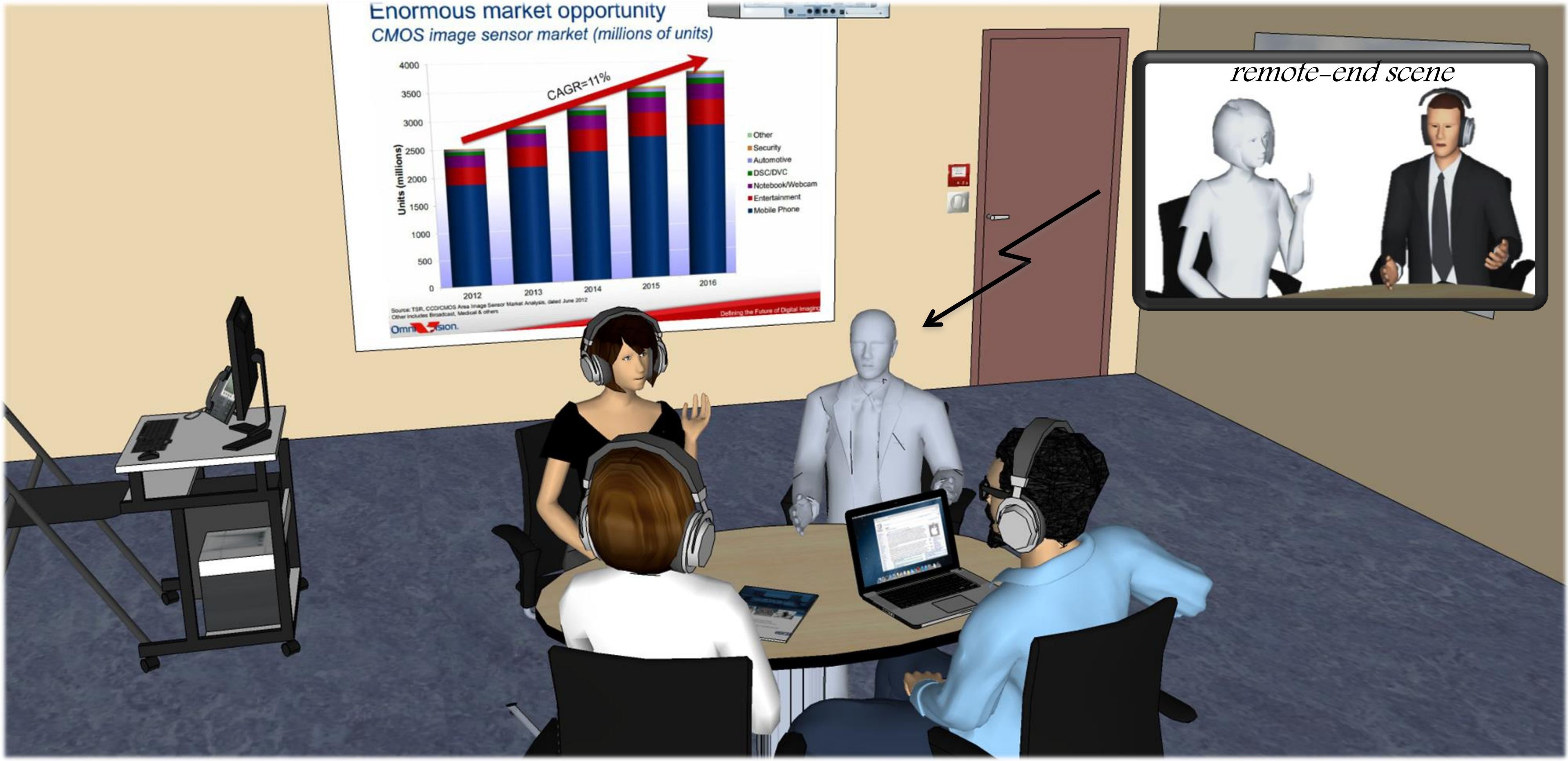
NAR headset potential applications:

Augmented Assistive listening for Visually Impaired

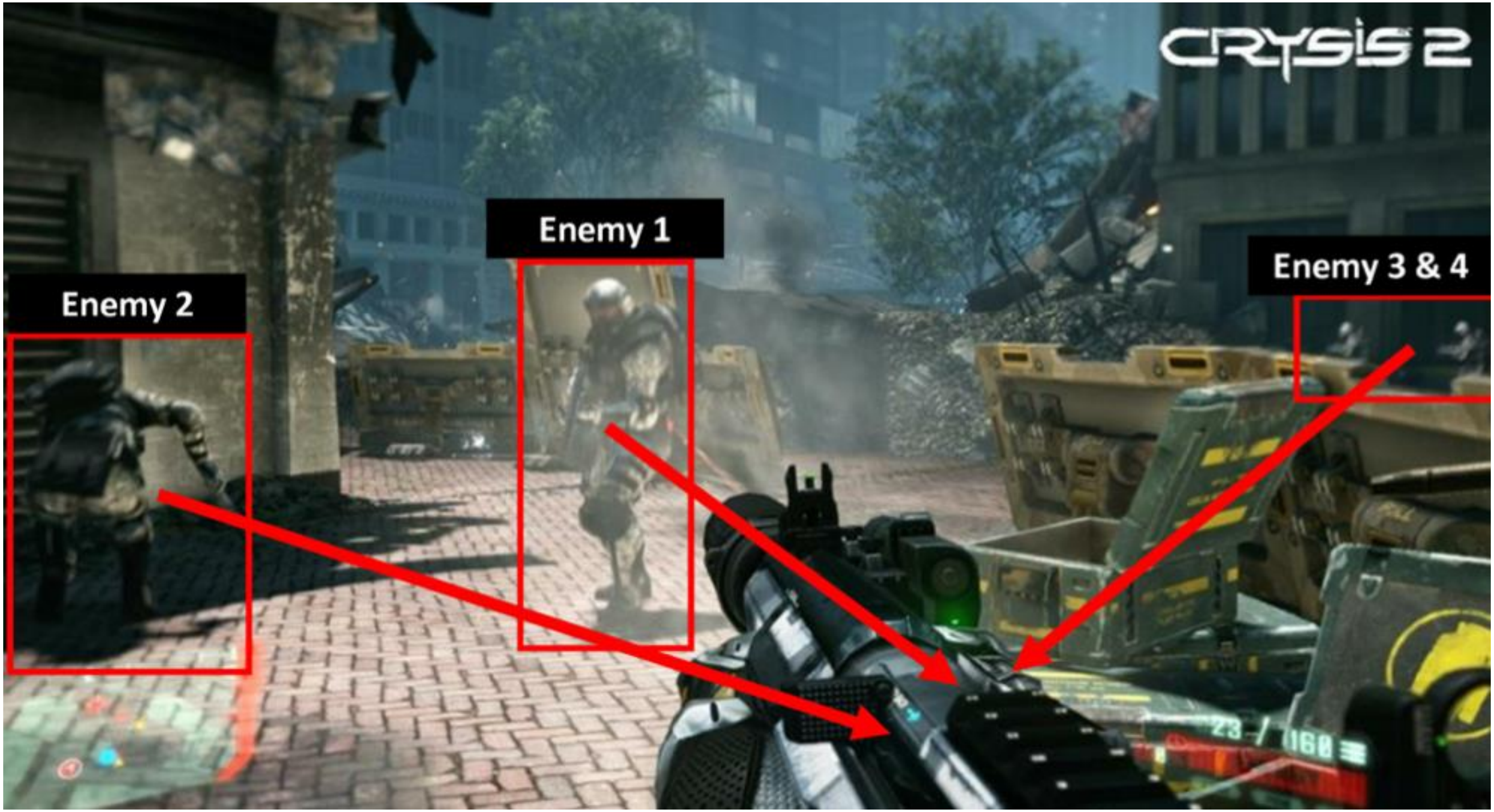
Assistive Listening is also needed for the normal person!

NAR headset potential applications:

Augmented Multi-party Conferencing



Using HRTF in Gaming to Create Audio Depth



Enemy 1 & 2 sound closer compared to enemy 3 & 4

Image from: CRYSIS 2 game

Proposed Natural Augmented Reality (NAR) Headset

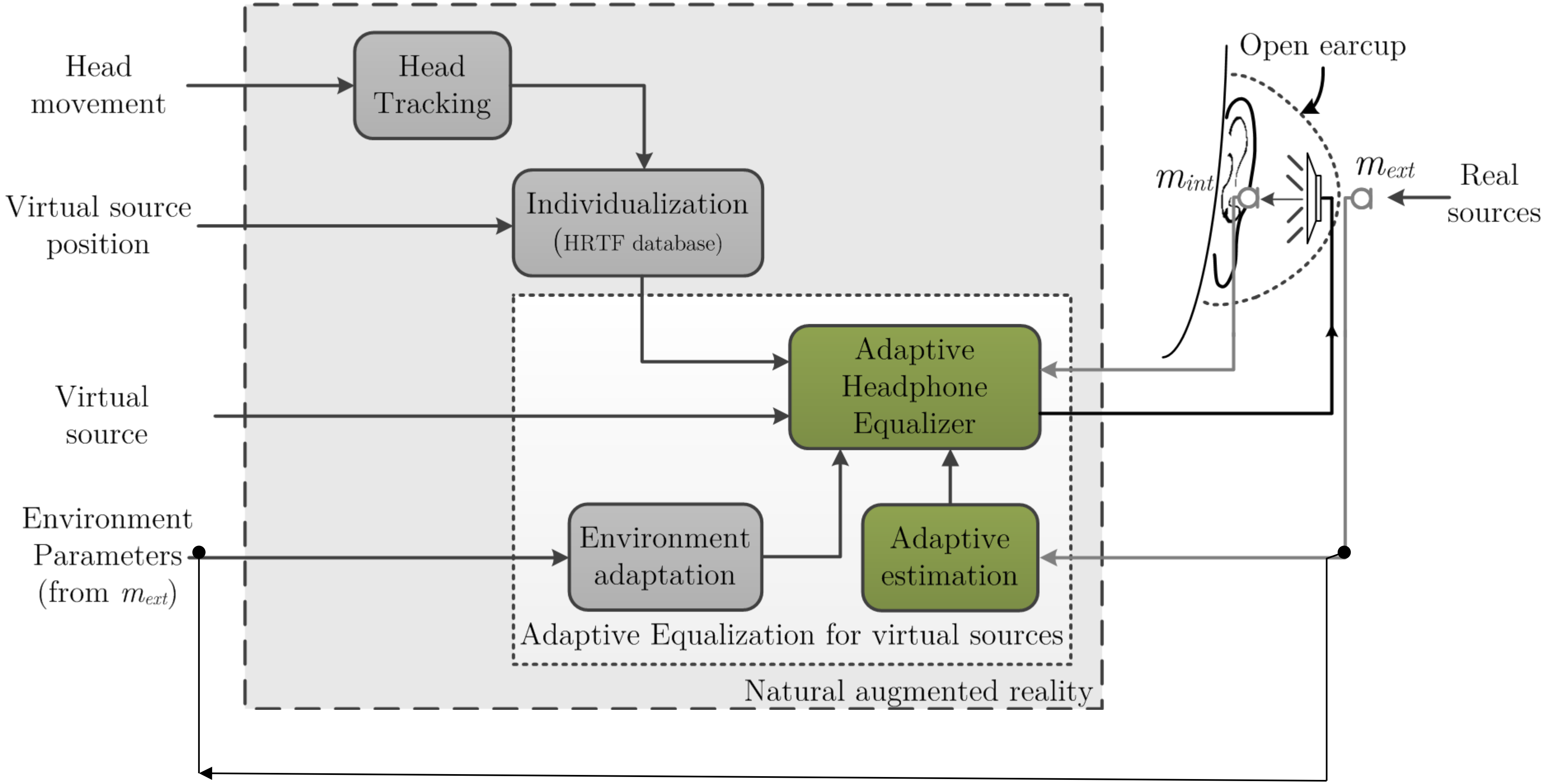
Proposed a headset structure with two pairs (*int*/*ext*) binaural microphones attached to the earcups.

- NAR headset with **2 pairs** of binaural microphones
- Adaptive equalization methods to compensate for sonic difference between real and headphone playback virtual signals
- Natural mixing of real and virtual sources

NAR headset prototype

- Internal microphone used as **error microphone** to adapt the virtual sound at ear canal to natural sound
- External microphone used as **reference microphone** to capture the sound event from the surroundings (real sounds)

NAR Headset Overview Block Diagram



Headphones Isolation Characteristics

- NAR headsets should not block the direct sounds coming from physical sound sources
- We analyze the headphone isolation characteristics of different type of headphones

The performance of the measured HRTFs can be evaluated by a spectral distortion (SD) score in dB given by

$$SD = \sqrt{\frac{1}{K}\sum_{k=1}^{K}\left(20log\frac{|H(f_k)|}{|\hat{H}(f_k)|}\right)^2}$$

K: Number of frequency bins

where

$|H(f_k)|$ is the magnitude response of reference HRTF i.e., measured without headphones from loudspeakers

$|\hat{H}(f_k)|$ is the magnitude response of HRTFs measured with the headphones from loudspeakers

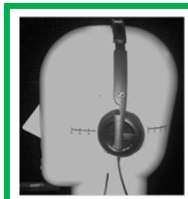
Smaller Sd score indicates a closer magnitude response to the reference

Headphones Isolation Characteristics

- Spectral distortion scores for 4 different type of headphones

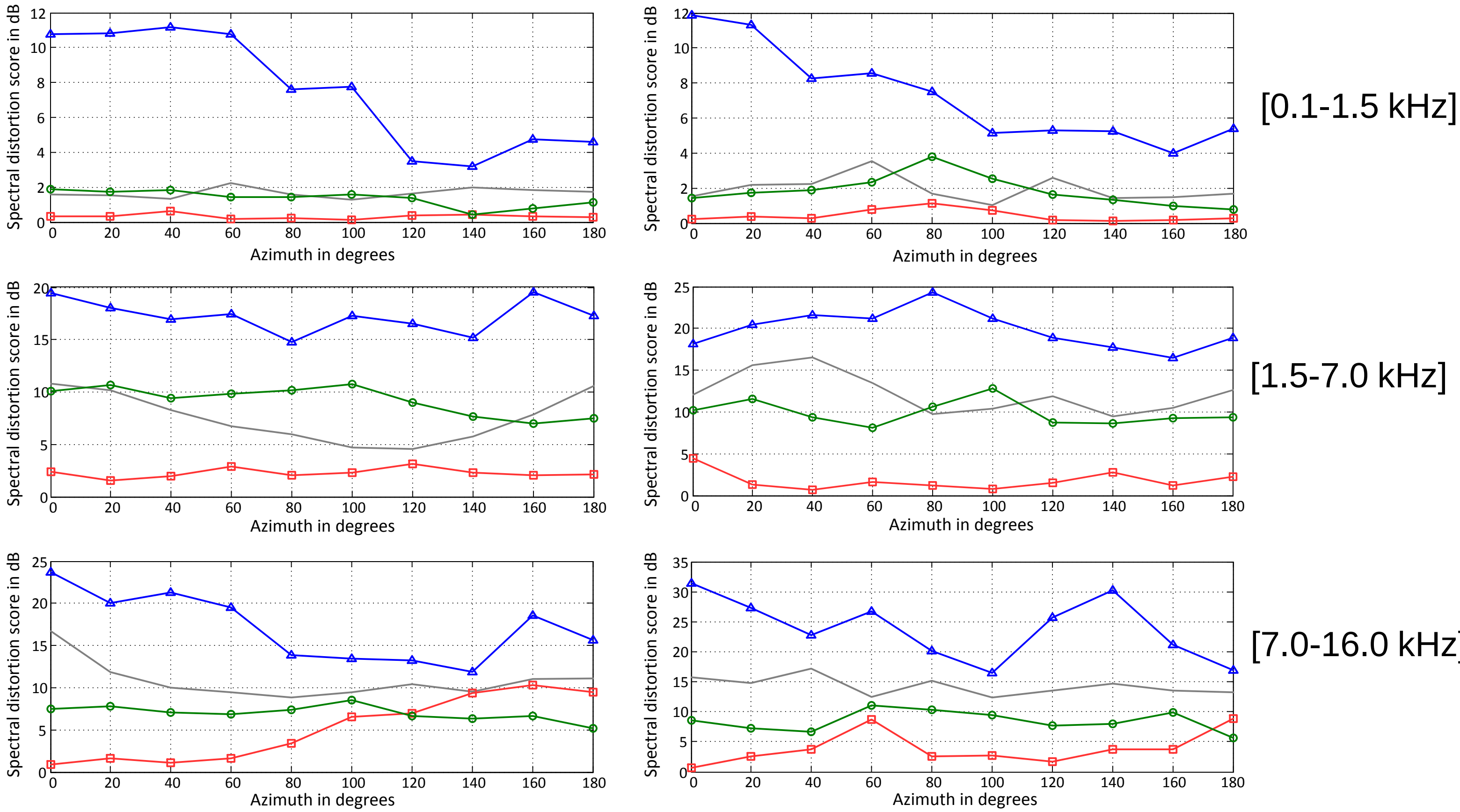
(b) AKG K702
Open Circumaural

(c) Sony PFR-V1
Open personal field speaker



(d) Sennheiser px 100-IIi
Open Supra-aural

(e) ATH M30
Closed circumaural

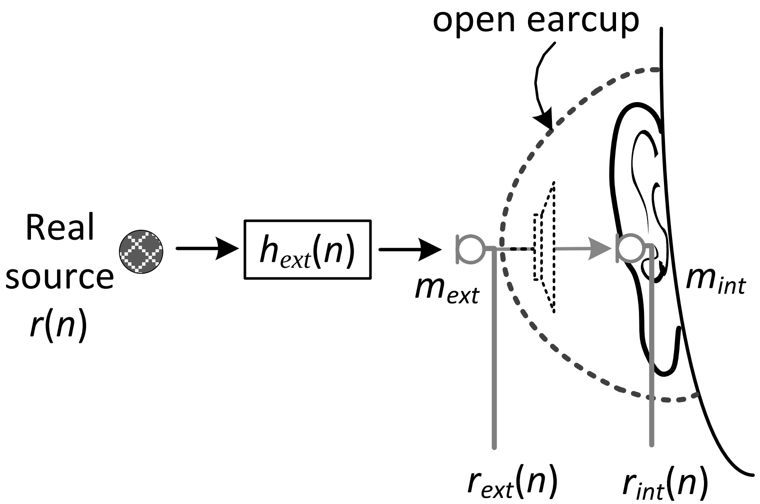


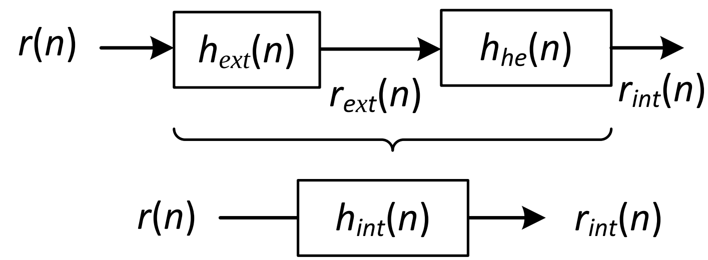
Contralateral (left) and ipsilateral (right) ears

NAR headset: (A) Normal Mode

Normal mode – Only external sound source present (No additional processing in NAR headset)

We choose **open headphones** so as to allow direct sound without much attenuation

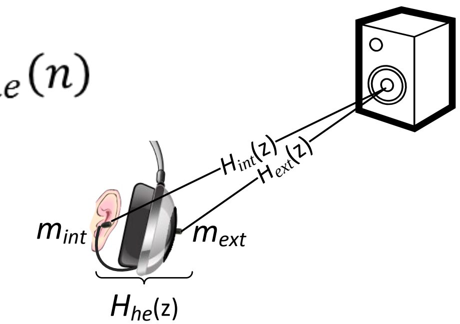




$$h_{int}(n) = h_{ext}(n) * h_{he}(n)$$

$\Downarrow z$

$$H_{he}(z) = \frac{H_{int}(z)}{H_{ext}(z)}$$

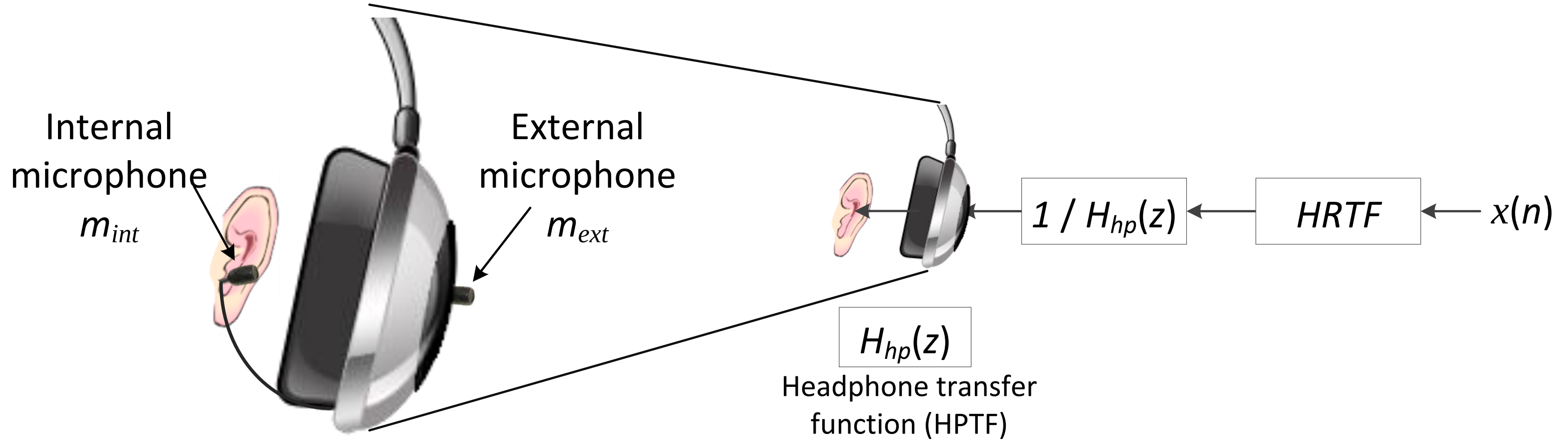


Account for the ear cup plus pinnae

Natural sounds pass through the open ear cup of the AR headset and reach the ear opening.

NAR headset : (B) Binaural Synthesis Mode

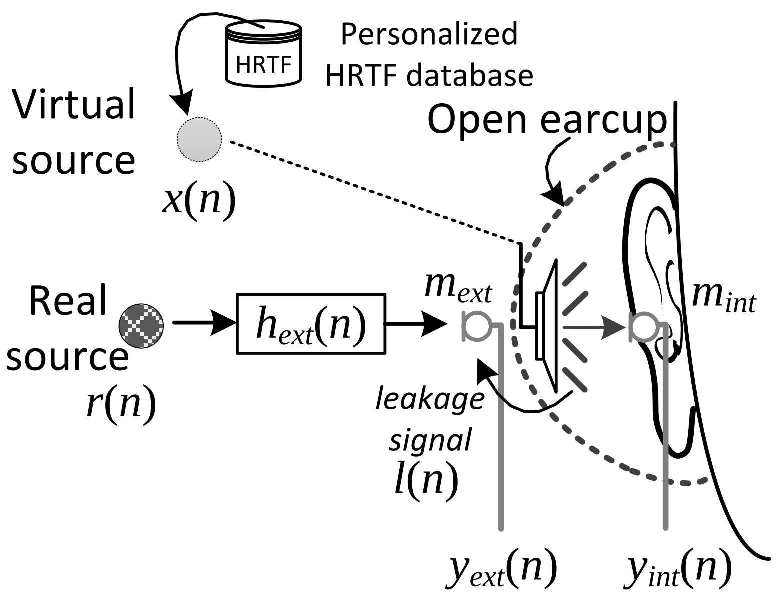
Binaural Synthesis using headphones playback

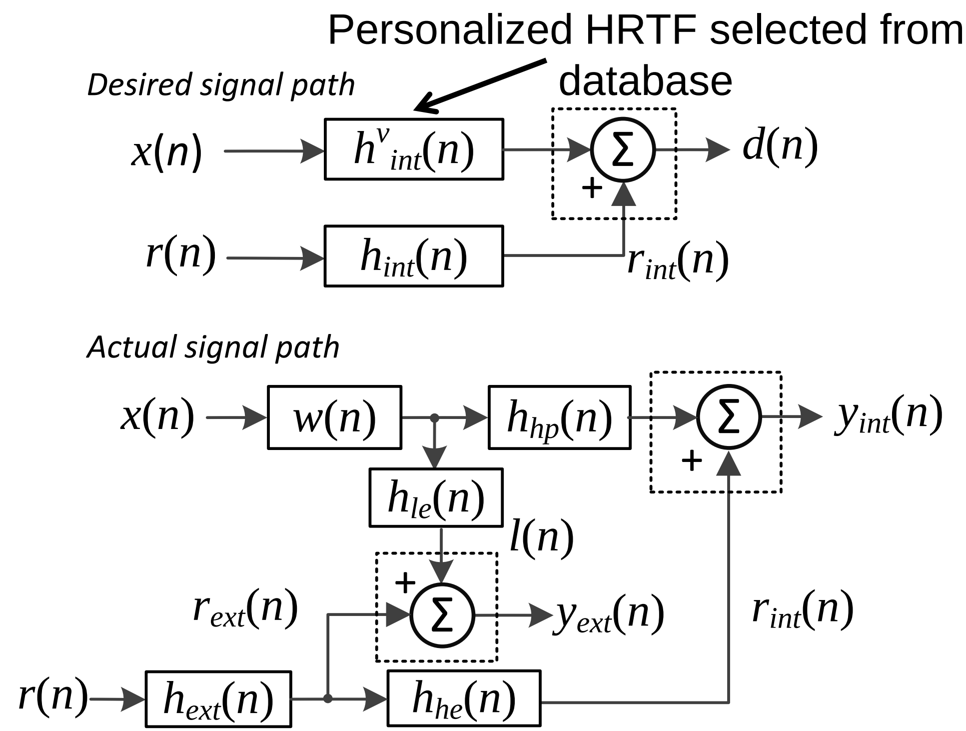


- Virtual monaural signal convolves with HRTF of the virtual object.
- Individual HPTF effect must be removed using compensation filter
- Compensation can be done by:
  - Direct inversion of HPTF, which may not be available[1]
  - Using an adaptive algorithm like FxLMS, which is more effective [2]

[1] M. Bouchard, S. G. Norcross, and G. A. Soulodre, "Inverse filtering design using a minimal-phase target function from regularization," in *AES Convention 121*, 2006.
[2] S. M. Kuo and D. Morgan, *Active noise control systems: algorithms and DSP implementations*: John Wiley & Sons, Inc., 1995.

Augmented reality mode – **virtual sound reproduction** in the **presence of external signals**





$l(n)$: Leakage from headphone to external microphone, m_{ext}

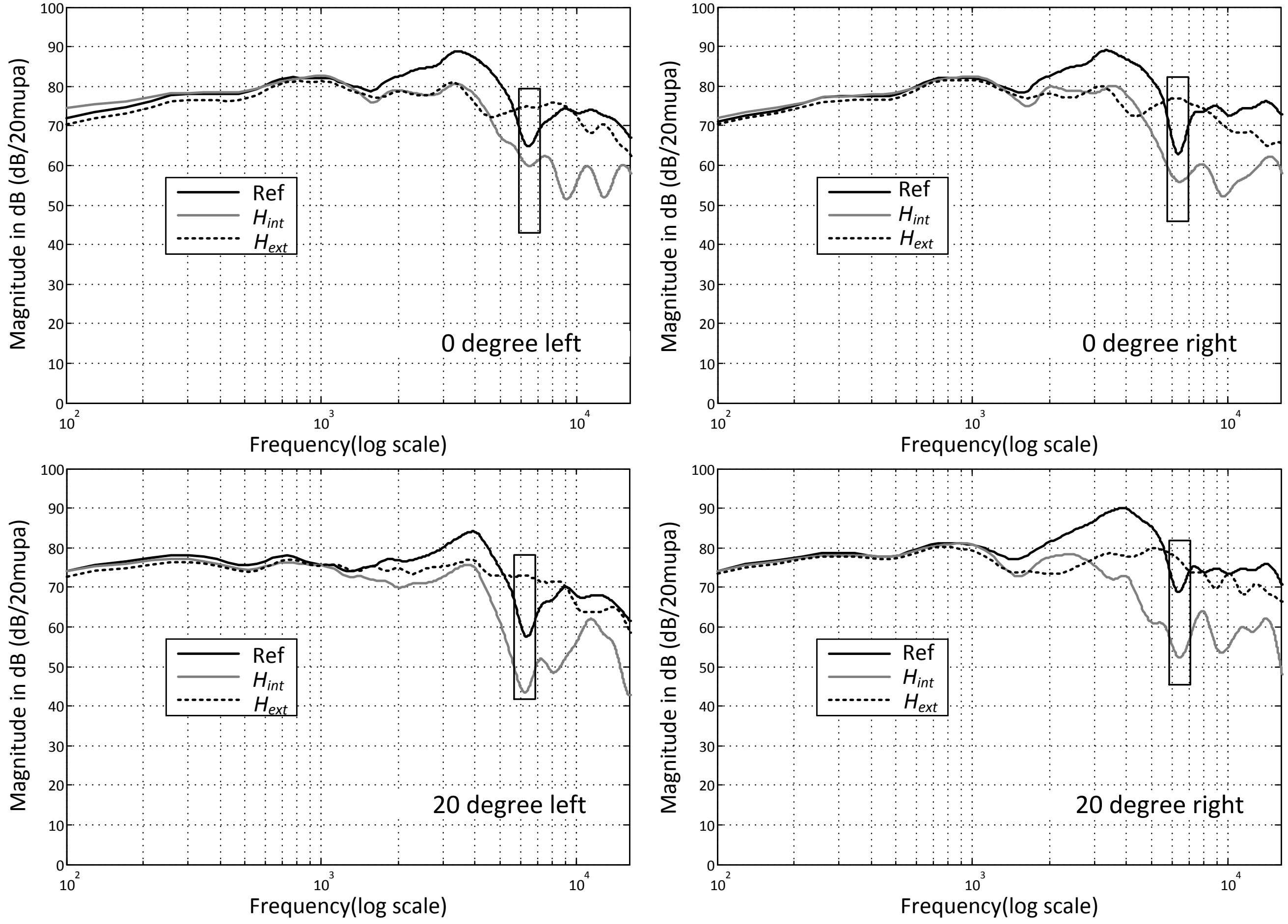
$$x_{int}(n) = w(n) * h_{hp}(n) * x(n)$$

$$W(z) = \frac{H^{v}_{int}(z)}{H_{hp}(z)}$$

Aim: To reproduce virtual sources as if they sound similar to physical sources, without being affected by external sounds

NAR headset : *int* and *ext* Transfer Functions

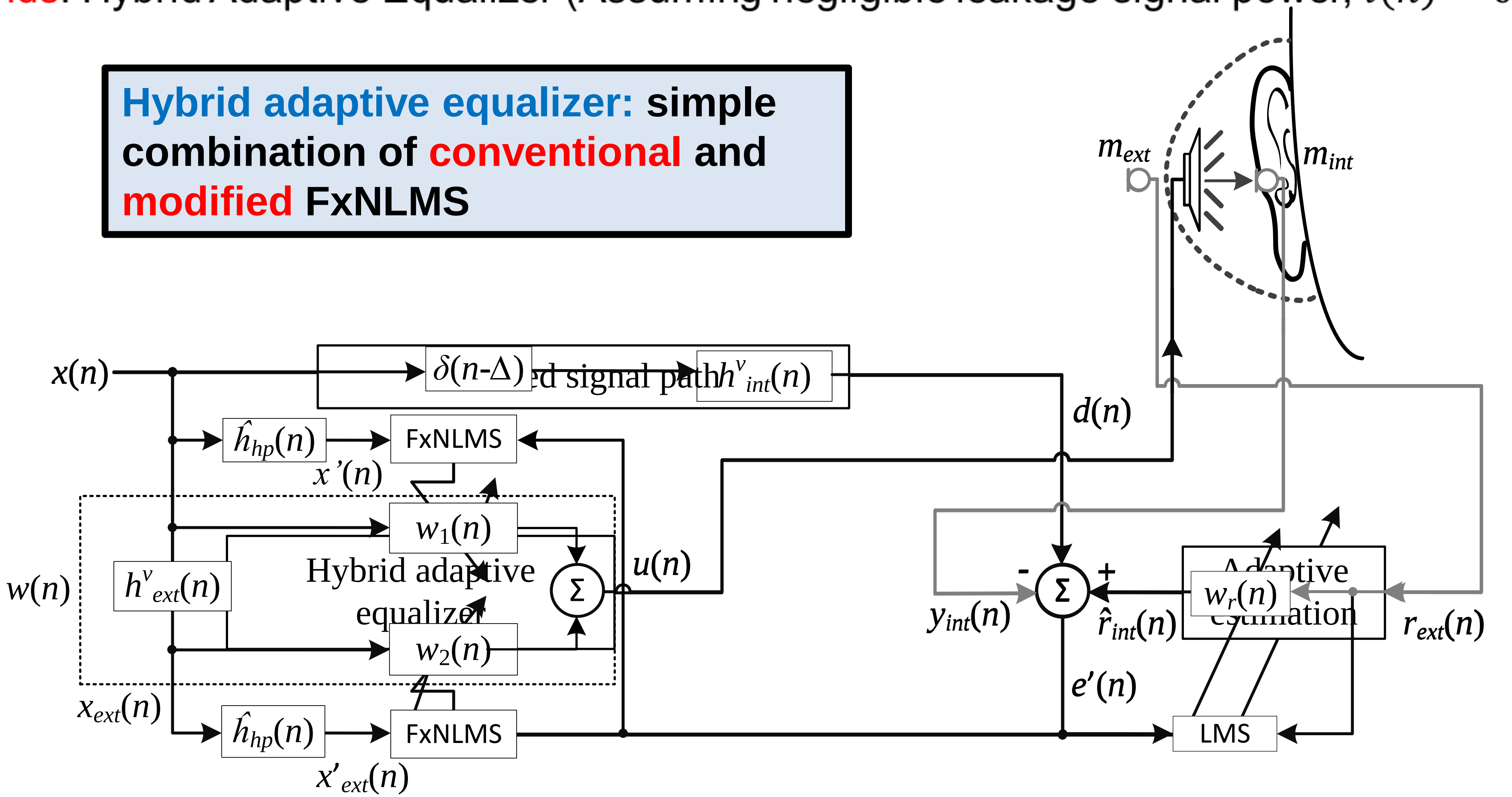
- $H_{int}(z)$ is an approximate HRTF with additional headphone effects
- $H_{ext}(z)$ contains all individual related characteristics and environment minus the pinnae specific notch and headphone shell reflections



NAR headset: Mixing virtual augmented signal with real ext source

Augmented reality mode – virtual sound reproduction in the presence of external sounds: Hybrid Adaptive Equalizer (Assuming negligible leakage signal power, $l(n) = 0$)

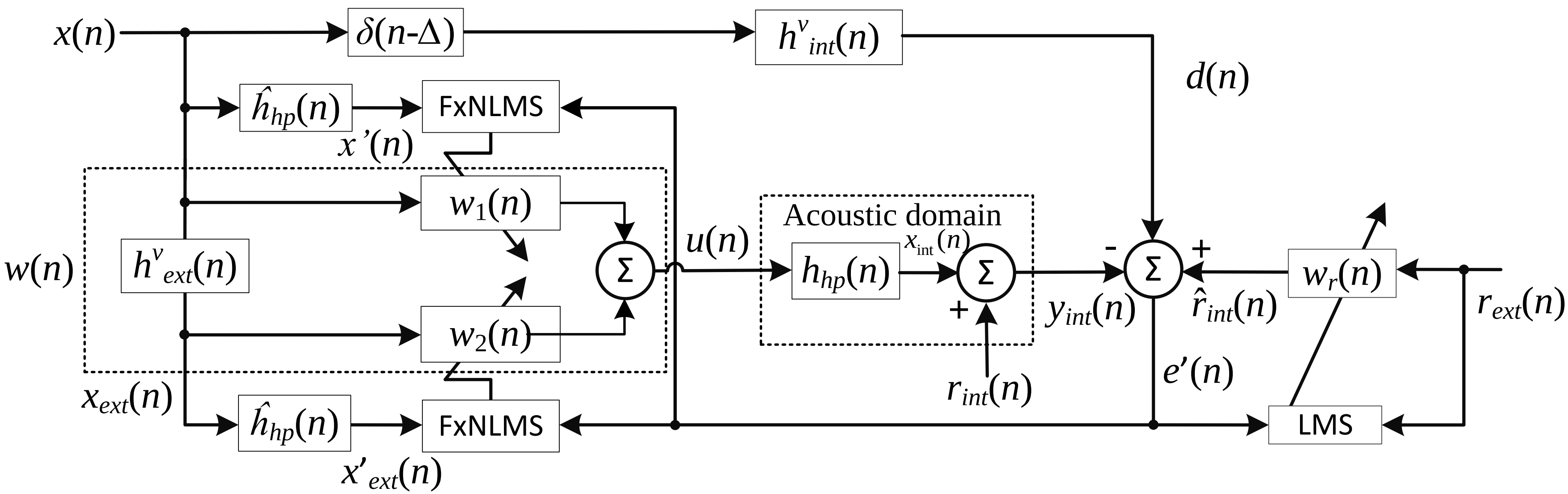
Hybrid adaptive equalizer: simple combination of conventional and modified FxNLMS



R Ranjan, Woon-Seng Gan, "Natural Listening over Headphones in Augmented Reality using Adaptive Filtering Techniques," *IEEE/ACM Transactions of Audio, speech and Language Processing*, Vol. 23 no. 11, 1998-2002 (2015).

NAR headset: Mixing virtual augmented signal with real ext source

Augmented reality mode – virtual sound reproduction in the presence of external sounds: Hybrid Adaptive Equalizer (Assuming negligible leakage signal power, $l(n) = 0$)



$w_1(n)$: Adaptive filter corresponding to *conventional FXNLMS*

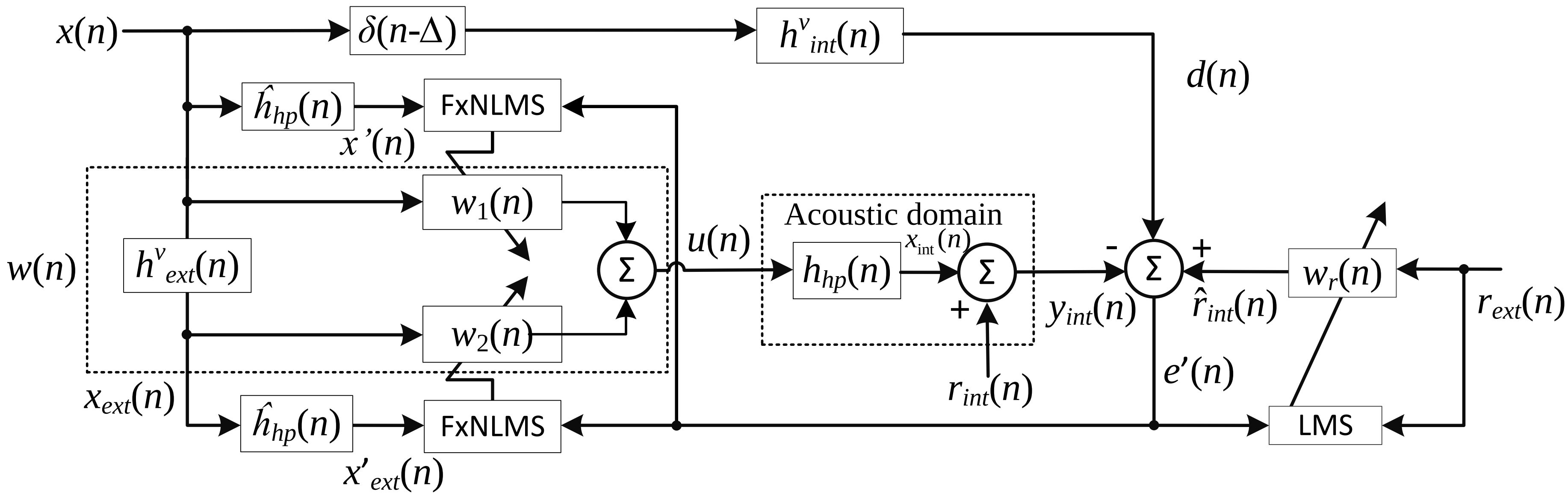
- Slower convergence rate due to presence of secondary path transfer function

$w_2(n)$: Adaptive filter corresponding to *Modified FxNLMS*

- Faster convergence rate by introducing spatial filter, $h^v{}_{ext}(n)$ in the secondary path but slightly higher steady state error (shorter filter taps)

NAR headset: Mixing virtual augmented signal with real ext source

Augmented reality mode – virtual sound reproduction in the presence of external sounds: Hybrid Adaptive Equalizer (Assuming negligible leakage signal power, $l(n) = 0$)



Hybrid adaptive filters :

$$w(n) = w_1(n) + h^v{}_{ext}(n) * w_2(n)$$

$$\Downarrow z$$

$$W(z) = W_1(z) + H^v{}_{ext}(z)\, W_2(z):$$

- Spatial information retained in $h^v{}_{ext}(n)$ results in faster convergences and smaller MSE using hybrid adaptive filters.

NAR headset: Mixing virtual augmented signal with real ext source

Augmented reality mode – virtual sound reproduction in the presence of external sounds

Weight update equations:

$$\mathbf{w_1}(n+1) = \mathbf{w_1}(n) + \mu \frac{\mathbf{x}'(n)}{\|\mathbf{x}'(n)\|^2} e'(n)$$

$$\mathbf{w_2}(n+1) = \mathbf{w_2}(n) + \mu \frac{\mathbf{x}'_{ext}(n)}{\|\mathbf{x}'_{ext}(n)\|^2} e'(n)$$

$$\mathbf{w_r}(n+1) = \mathbf{w_r}(n) - \mu\, \mathbf{r}_{ext}(n)\, e'(n)$$

Optimal solution for hybrid adaptive filter:

$$W^o(z) = \alpha\, W_1^o(z) + \beta\, H^v_{ext}(z) W_2^o(z),$$

where,

$$W^o(z) = W_1^o(z) = \frac{H_{int}(z) z^{-\Delta}}{H_{hp}(z)}$$

and,

$$W_2^o(z) = \frac{H^v_{int}(z) z^{-\Delta}}{H^v_{ext}(z) H_{hp}(z)}$$

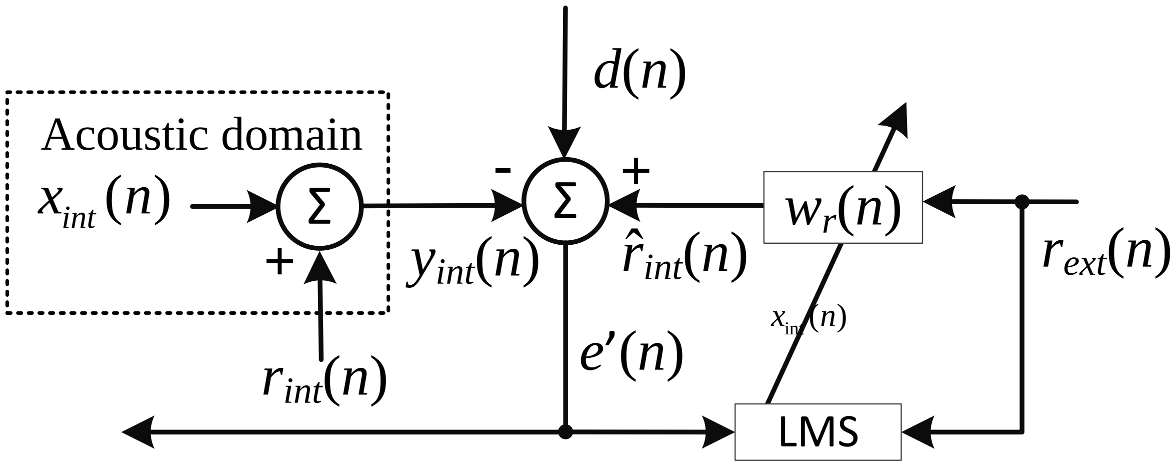
$$\forall\ \alpha + \beta = 1\,; \quad 0 \le \alpha, \beta \le 1$$

Optimal solution for adaptive estimation filter:

$$W_r^o(z) = \frac{R_{int}(z)}{R_{ext}(z)} = \frac{H_{int}(z)}{H_{ext}(z)} = H_{he}(z)$$

The optimal solution for adaptive estimation filter is the headphone effect transfer function between the two microphone positions

Optimal solution for hybrid adaptive filter can be expressed as linear combination of optimal solutions for two FxNLMS adaptive filters



Augmented superimposed signal:

$$y_{int}(n) = x_{int}(n) + r_{int}(n),$$

where,

$$x_{int}(n) = h_{hp}(n) * u(n)$$

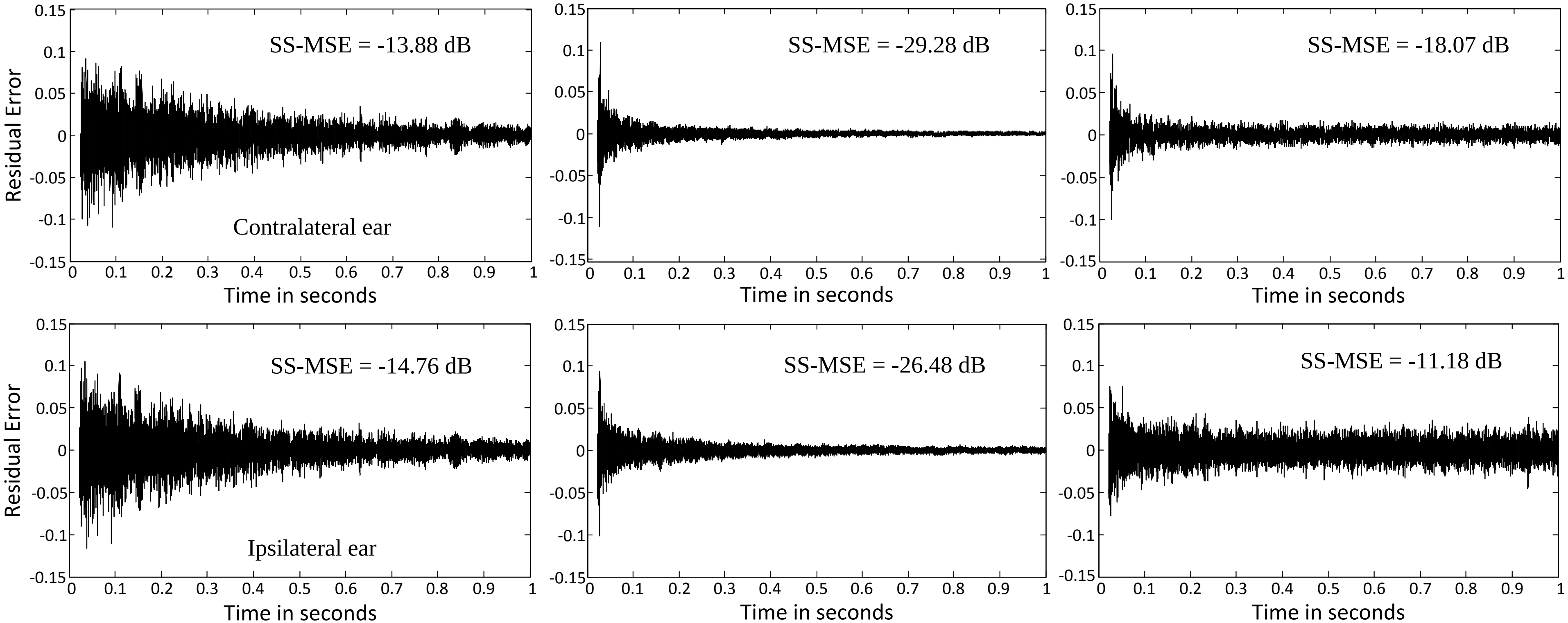
Error signal:

0

$$\begin{aligned} e'(n) &= \{d(n) + \hat{r}_{int}(n)\} - y_{int}(n) \\ &= \{d(n) - x_{int}(n)\} + \{-(r_{int}(n) - \hat{r}_{int}(n))\} \\ &= e_v(n) + e_r(n) \end{aligned}$$

Results: Augmented Reality mode

Hybrid FxNLMS performance comparison with or without external sound source present but no adaptive estimation process i.e., $w_r(n) = 0$



(a) Conventional FxNLMS with no real source present

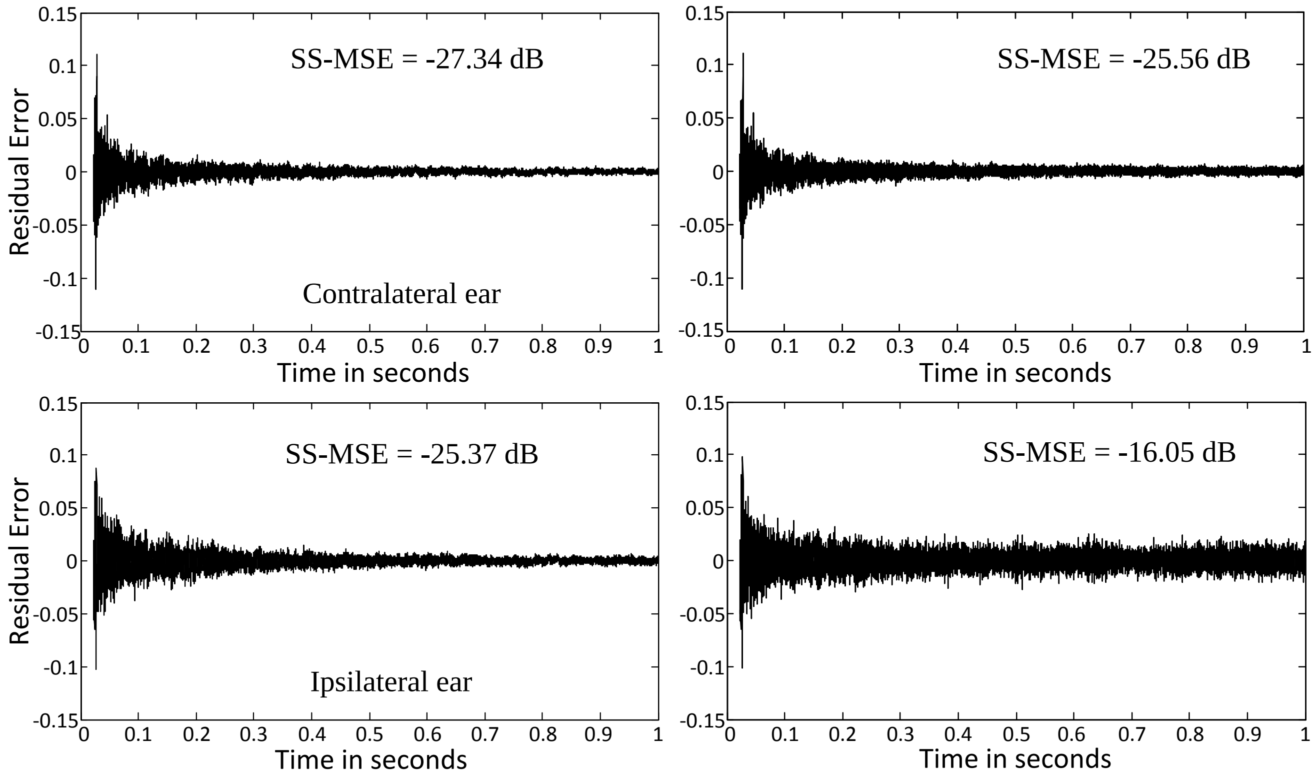
(b) Hybrid adaptive equalizer with no real source present

(c) Hybrid adaptive equalizer with real source present

Clearly, presence of external sound sources affect the adaptive process and results in larger steady state error

Results: Augmented Reality mode

Hybrid FxNLMS performance with adaptive estimation i.e., adaptive $w_r(n)$ Vs off-line estimation i.e, fixed $w_r(n)$ as average $h_{he}(n)$ filter to estimate the real signals



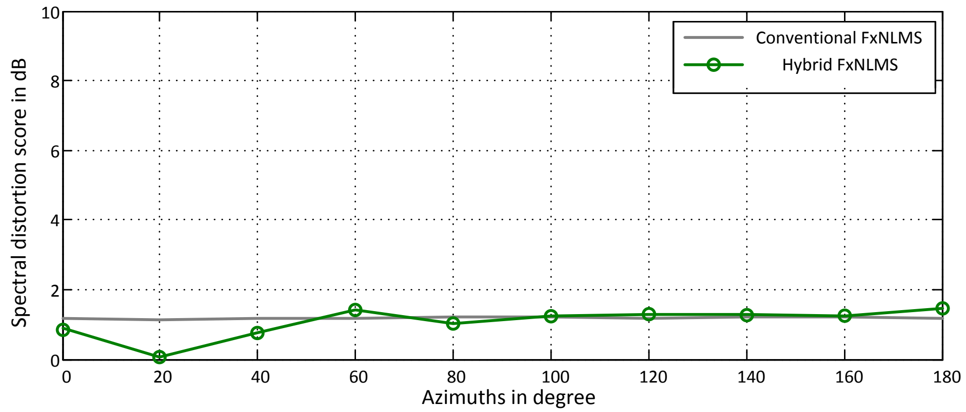
(a) Hybrid adaptive equalizer with adaptive estimation

(b) Hybrid adaptive equalizer with off-line estimation

Hybrid FxNLMS with adaptive estimation works equally well even in the presence of real sounds reproducing the virtual sources as close as possible to real sources.

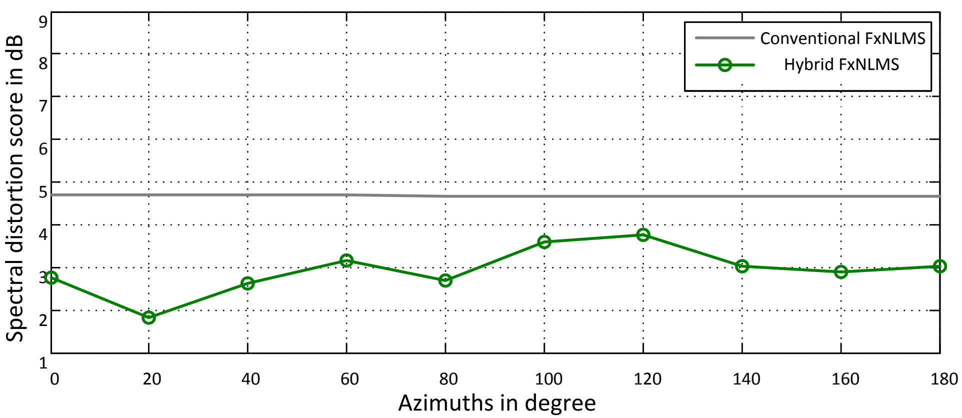
Results: Augmented Reality mode

Spectral distortion score comparison for Hybrid Adaptive Equalizer (100-16000 Hz)



(Contralateral ear)

(Ipsilateral ear)



Clearly, Hybrid FxNLMS is optimal for all the azimuths as compared to conventional and modified FxNLMS

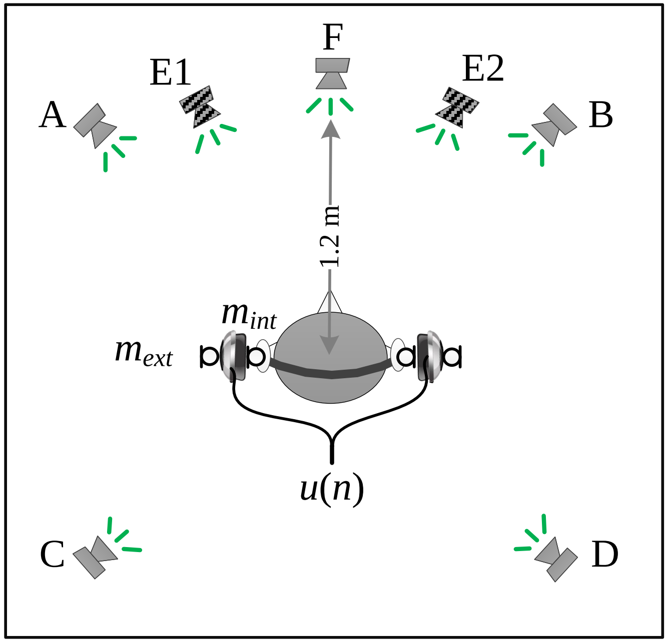
Listening Test

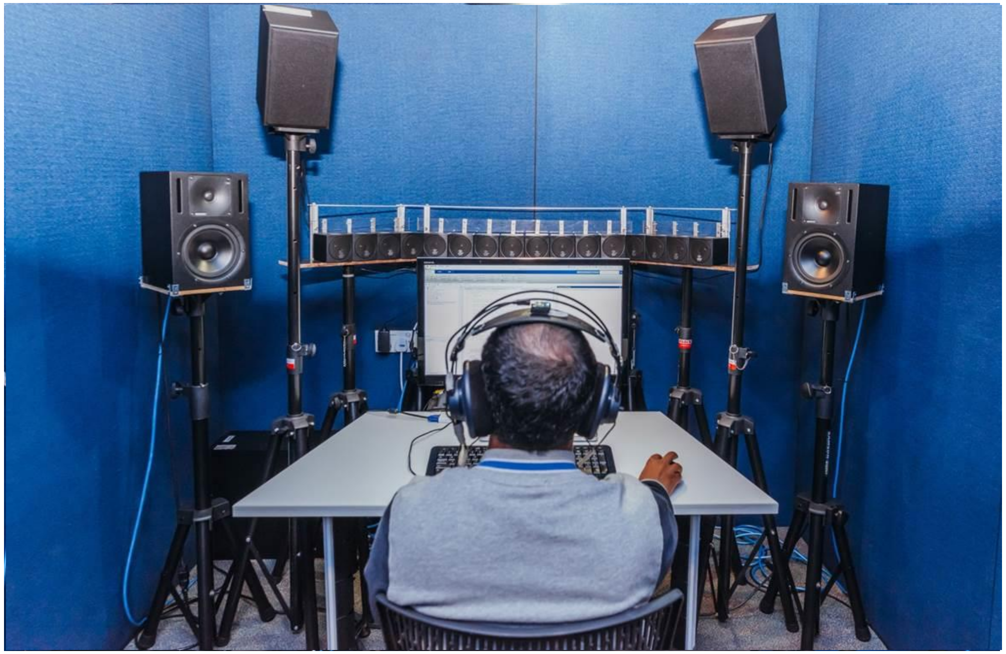
Three main objectives:

- Naturalness: Does virtual playback sound source feels natural to you?
- Sound similarity: Does virtual sounds similar to real speaker sound?
- Source position similarity: Does virtual sound source coming from position in 3D space as real source?

Experiment setup

- 7 loudspeakers: 5 in horizontal plane and 2 in median plane





Listing Test Set up (: Elevated speaker; : Azimuth speaker)

Listening test overview

Conducted in two phases:

1. Listeners' BRIRs (automated) measurements
   - Head Tracker to ensure "look ahead" and still head position
2. Listening test based on individual's BRIRs measurements
   - Fixed listening position
   - No head movements allowed (future: headtracker)

Listening test sets

Three different listening sets are carried out as follows:

- **SET *1***: Perceptual similarity test of a male speech signal (5 sec)
- **SET *2***: Perceptual similarity test for playback of two male speech signals from two different directions (3.5 sec each)
- **SET *3***: Perceptual similarity test for superposition of a speech signal with ambient sound (6 sec)

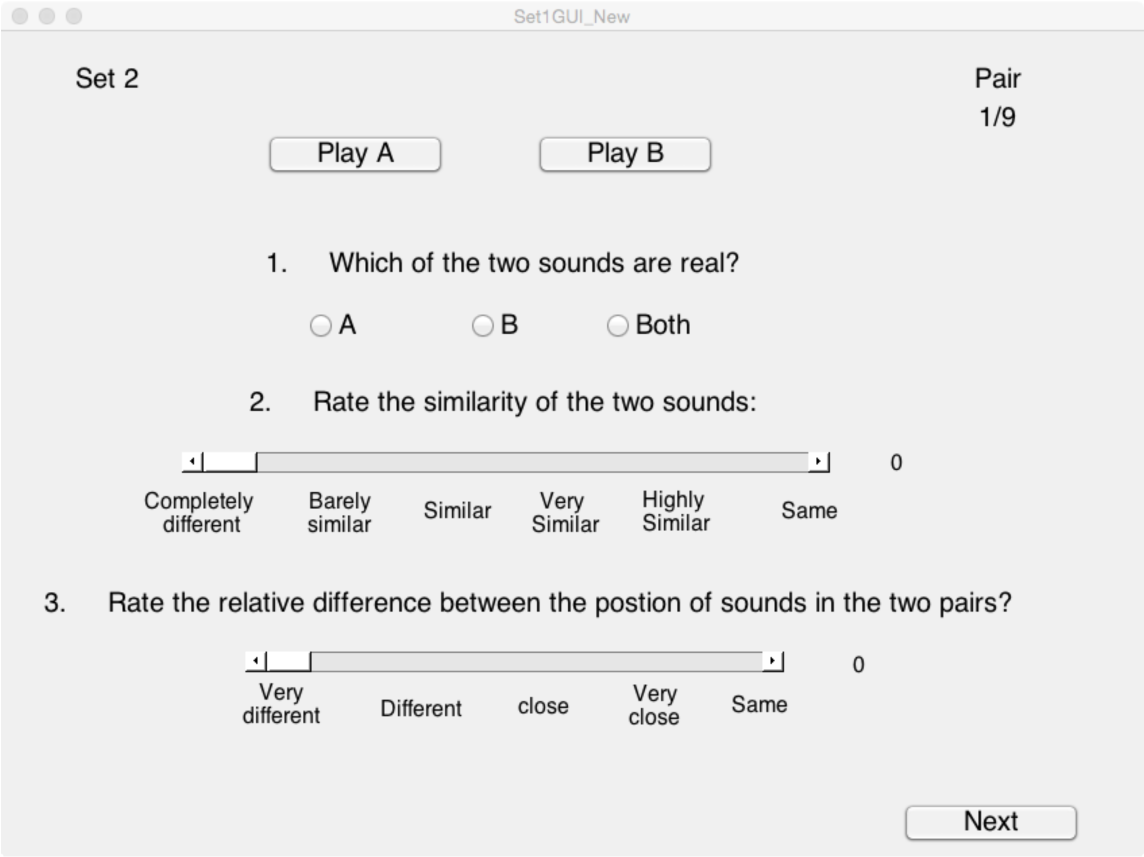
Listening SETs

SET 1:

- Total seven speaker positions
- Two hidden anchors (coming from the same physical speakers)

SET 2:

- Total 6 speaker pairs
- Two hidden anchors

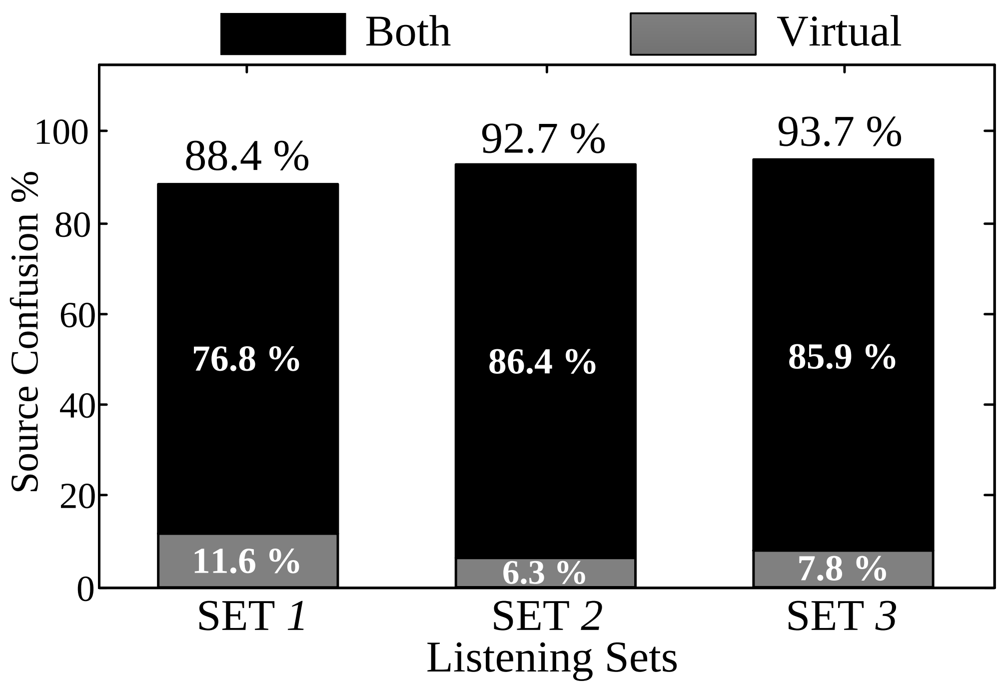


SET 3:

- Total 4 pairs
- speech from F and ambience from surroundings (2ch and 4ch)
- Two methods: Adaptive eq. with and without adaptive estimation
- One hidden anchor

Listening Test Results : Source Confusion

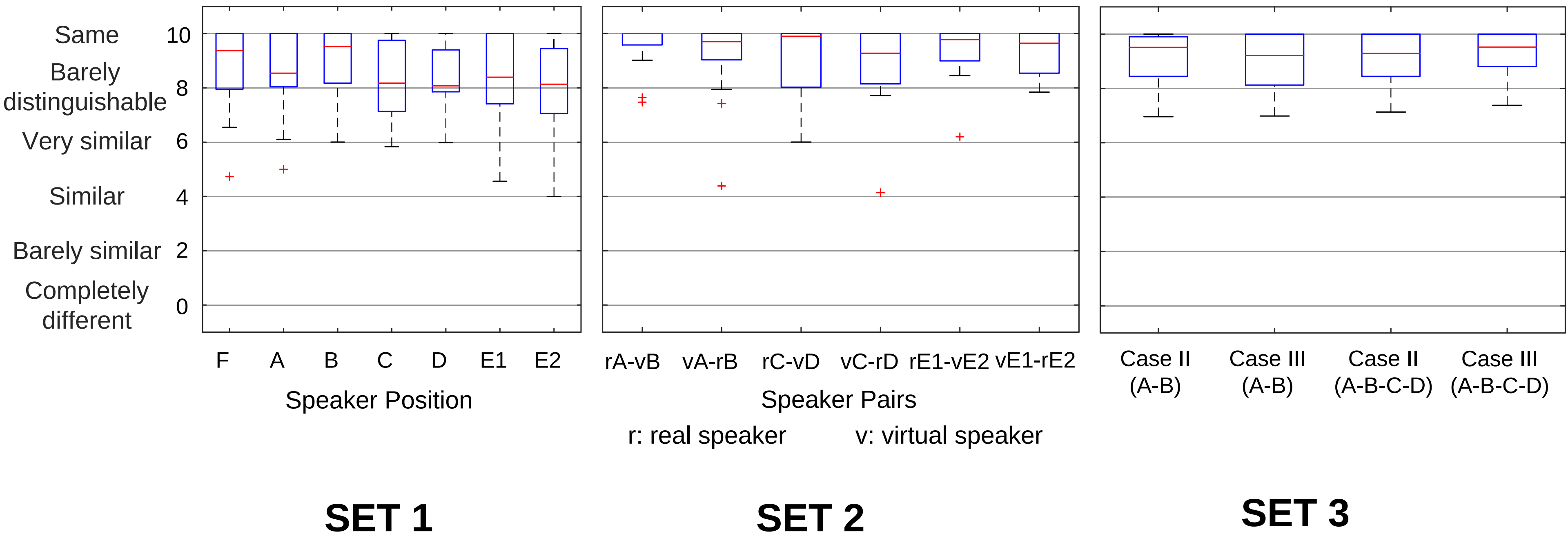
Source confusion measure in percentage as responses for subjects who marked either virtual sound as real or both sounds as real



- Clearly, more than 75% source confusion for the case where subjects marked both sounds as real implying virtual source perceived natural by most of the listener
- Source confusion further increases when more source present.

Listening Test Results : Sound Similarity

Sound similarity: Subjective score of 0 to 10

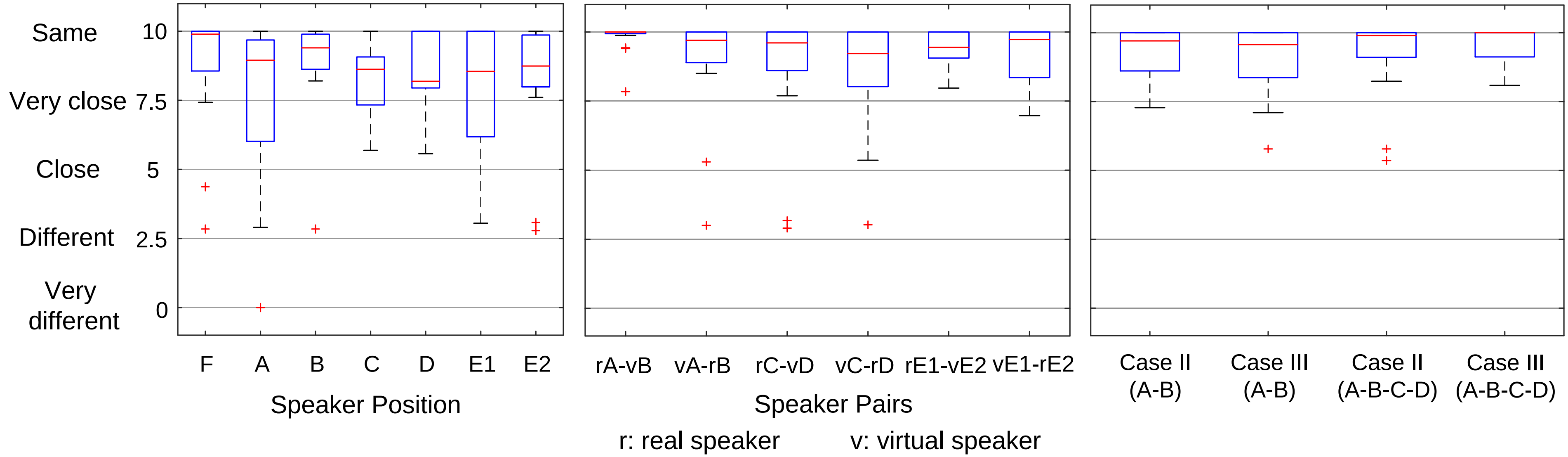


Mean subjective score with 99% confidence interval for sound similarity

SET 1	SET 2	SET 3
8.44 (7.89 - 9.02)	9.19 (8.59 - 9.79)	9.13 (8.61 - 9.65)

Listening Test Results : Source position Similarity

Source position similarity: Subjective score of 0 to 10



SET 1 **SET 2** **SET 3**

Mean subjective score with 99% confidence interval for Source position similarity

SET 1	SET 2	SET 3
8.26 (7.33 - 9.19)	9.09 (8.32 - 9.87)	9.28 (8.77 - 9.78)

NAR Headset Extension

S. No.	Extensions based on current limitations of the NAR headset
E1	NAR headset should adapt to the change in external environment (H_{int} and H_{ext})
E2	Individualized HRIR acquisition using NAR headset
E3	Adaptive equalization for any type of virtual signals (speech, music, etc.)
E4	Detection and Fast Estimation of Headphone Transfer Function in NAR Headset
E5	Adaptive Equalization for Non-stationary Virtual signals with Adaptive Estimation of External signals
E6	ANC mode for the NAR headset in presence of unwanted/unpleasant ambient sound
E7	NAR headset with head tracking to include dynamic motion cues

Conclusions on NAR Headsets

- Adaptive filtering techniques presented to enable natural listening in augmented reality environments using NAR headset with two pairs of binaural microphones.
- Hybrid adaptive equalizer based on FxNLMS for virtual sound reproduction to equalize the NAR headset for an individual
- Faster convergence and smaller steady state residual error.
- Proposed hybrid adaptive equalizer able to attenuate the error signal by more than 25 dB, and implying virtual signals are perceptually reproduced very close to the real signals.
- Listening test result shows very high source confusion % for virtual sources
- Very high sound similarity and source position similarity were observed between real and augmented sounds

Key References on NAR Headsets

[1] R. T. Azuma, "A survey of augmented reality,"Presence, vol. 6, pp. 355–385, 1997.

[2] T. Nilsen, S. Linton, and J. Looser, "Motivations for augmented reality gaming," in Proc. FUSE, 2004, vol. 4, pp. 86–93.

[3] T. Sielhorst, M. Feuerstein, and N. Navab, "Advanced medical displays: A literature review of augmented reality," J. Display Technol., vol. 4, pp. 451–467, 2008.

[4] T. Lokki, H. Nironen, S. Vesa, L. Savioja, A. Härmä, and M. Karjalainen, "Application scenarios of wearable and mobile augmented reality audio," in Proc. Audio Eng. Soc. Conv. 116, 2004.

[5] M. Billinghurst and H. Kato, "Collaborative augmented reality," Commun. ACM, vol. 45, pp. 64–70, 2002.

[6] T. Miyashita, et al., "An augmented reality museum guide," in Proc. 7th IEEE/ACM Int. Symp. Mixed Augment. Real., 2008, pp. 103–106.

[7] H. Ishii and B. Ullmer, "Tangible bits: Towards seamless interfaces between people, bits and atoms," inProc. ACM SIGCHI Conf. Human Factors in Comput. Syst., 1997, pp. 234–241.

[8] R. W. Lindeman, H. Noma, and P. G. de Barros, "Hear-through and mic-through augmented reality: Using bone conduction to display spatialized audio," inProc. 6th IEEE and ACM Int. Symp. Mixed Augment. Real., 2007, pp. 1–4.

[9] A. Härmä, J. Jakka, M. Tikander, M. Karjalainen, T. Lokki,and J. Hiipakkaet al., "Augmented reality audio for mobile and wearable appliances,"J. Audio Eng. Soc., vol. 52, pp. 618–639, 2004.

[10] M. Tikander, M. Karjalainen, and V. Riikonen, "An augmented reality audio headset," in Proc. 11th Int. Conf. Digital Audio Effects (DAFx-08), Espoo, Finland, 2008.

[11] J. Rämö and V. Välimäki, "Digital augmented reality audio headset," J. Elect. Comput. Eng., vol. 2012, p. 13, 2012, Article ID 457374.

Key References on NAR Headsets

[12] D. W. Schobben and R. M. Aarts, "Personalized multi-channel headphone sound reproduction based on active noise cancellation,"Acta Acusti. United Acust., vol. 91, pp. 440–450, 2005.

[13] J. Rämö, "Evaluation of an augmented reality audio headset and mixer," Ph.D. dissertation, Helsinki Univ. of Technol., Espoo, Finalnd, 2009.

[14] F. Brinkmann and A. Lindau, "On the effect of individual headphone compensation in binaural synthesis,"Fortschritte der Akustik: Tagungsband d. 36. DAGA, pp. 1055–1056, 2010.

[15] M. Bouchard, S. G. Norcross, and G. A. Soulodre, "Inverse filtering design using a minimal-phase target function from regularization," in Proc. Audio Eng. Soc. Conv. 121, 2006.

[16] S. T. Neely and J. B. Allen, "Invertibility of a room impulse response," J. Acoust. Soc. Amer., vol. 66, pp. 165–169, 1979.

[17] T. Nishino, N. Inoue, K. Takeda,and F. Itakura, "Estimation of HRTFs on the horizontal plane using physical features,"Appl. Acoust., vol. 68, pp. 897–908, 2007.

[18] R. Ranjan and W.-S. Gan, "Applying active noise control technique for augmented reality headphones," in Proc. Internoise, Melbourne, Australia, 2014.

[19] R. Ranjan and W.-S. Gan, "Natural Listening over Headphones in Augmented Reality Using Adaptive Filtering Techniques," IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, VOL. 23, NO. 11, NOVEMBER 2015

[20] V. Välimäki, A. Franck, J. Rämö, H. Gamper, and L. Savioja, "Assisted Listening Using a Headset," IEEE Signal Processing Magazine, vo. 32, no. 2, pp. 92-99, Mar. 2015.

Module V
Assisted Listening in Hearing Aids

Outline for Assisted Listening in Hearing Aids

1. Hearing loss and hearing aids
2. Noise reduction and speech enhancement
   - Acquisition
   - Localization
   - Enhancement
   - Presentation
3. Integration of ANC in hearing aids

Global Prevalence of Hearing Loss

- "Over 5% of the world's population – 360 million people – has disabling hearing loss."
- "Approximately 1/3 of persons over 65 years are affected by disabling hearing loss."

- WHO, Mar 2015

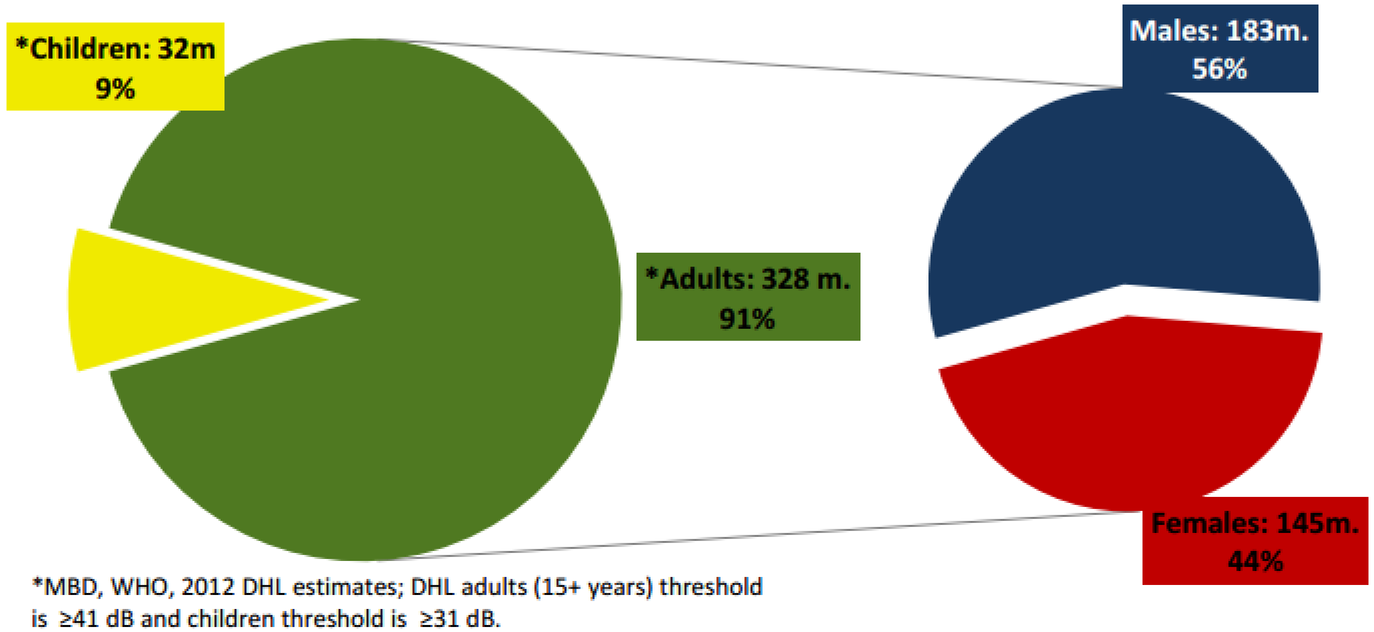
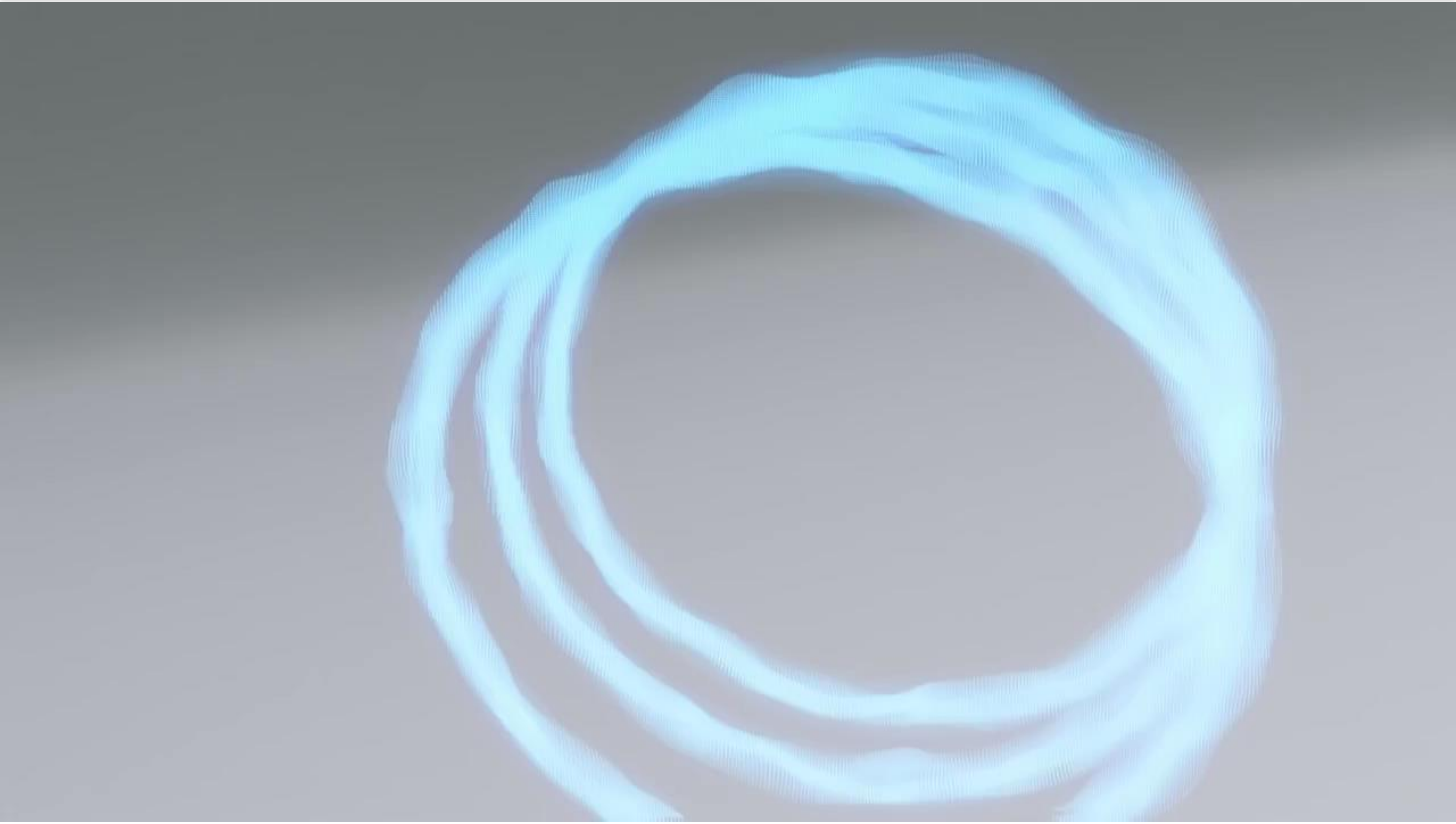


Figure from http://www.who.int/pbd/deafness/WHO_GE_HL.pdf?ua=1

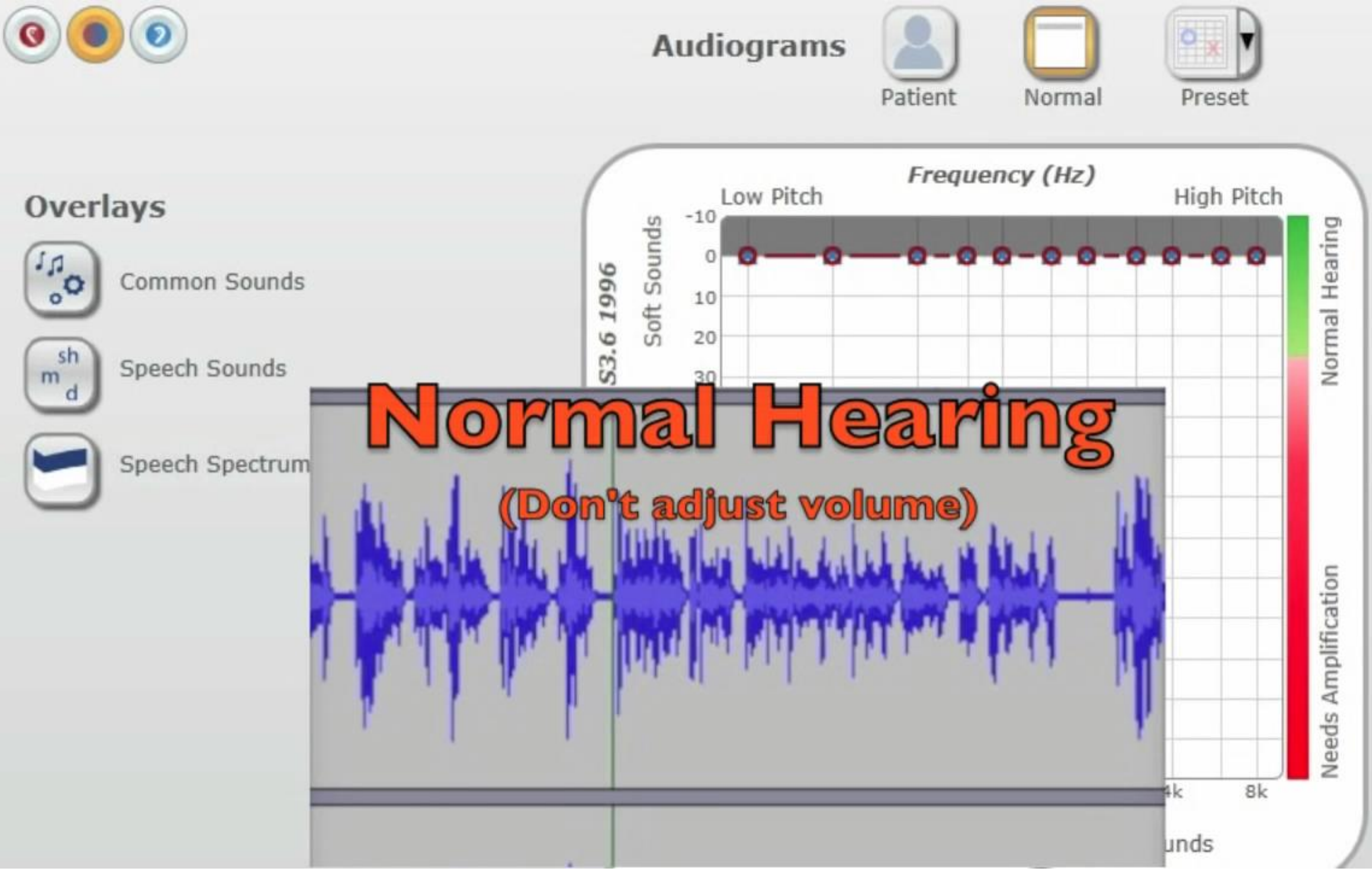
Hearing aids: for everyone to hear well

Hearing loss



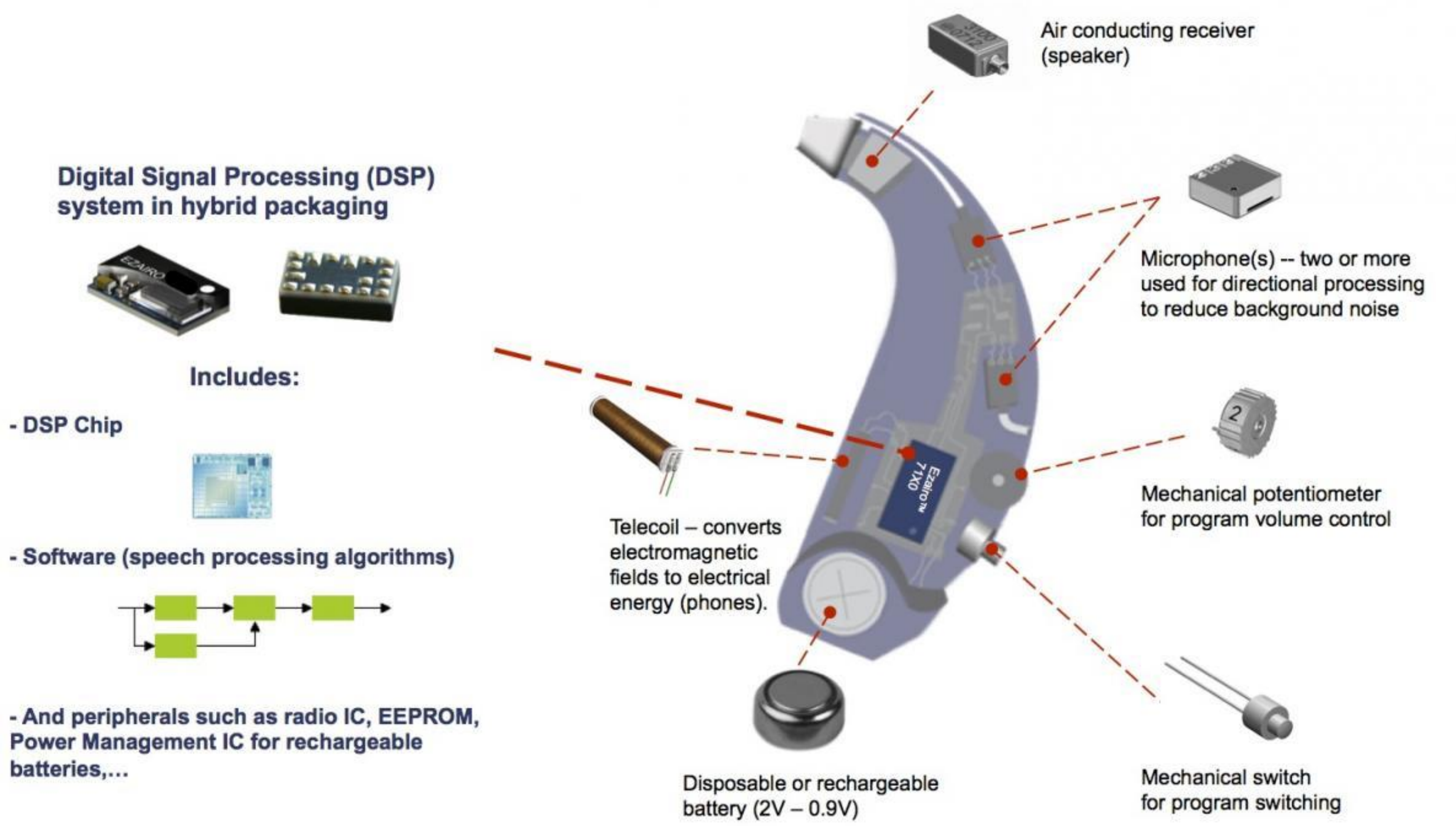
https://www.youtube.com/watch?v=2rhRo73F324

What it sounds like with hearing loss



https://www.youtube.com/watch?v=Bcz7AeBMLSc

Inside a hearing aids



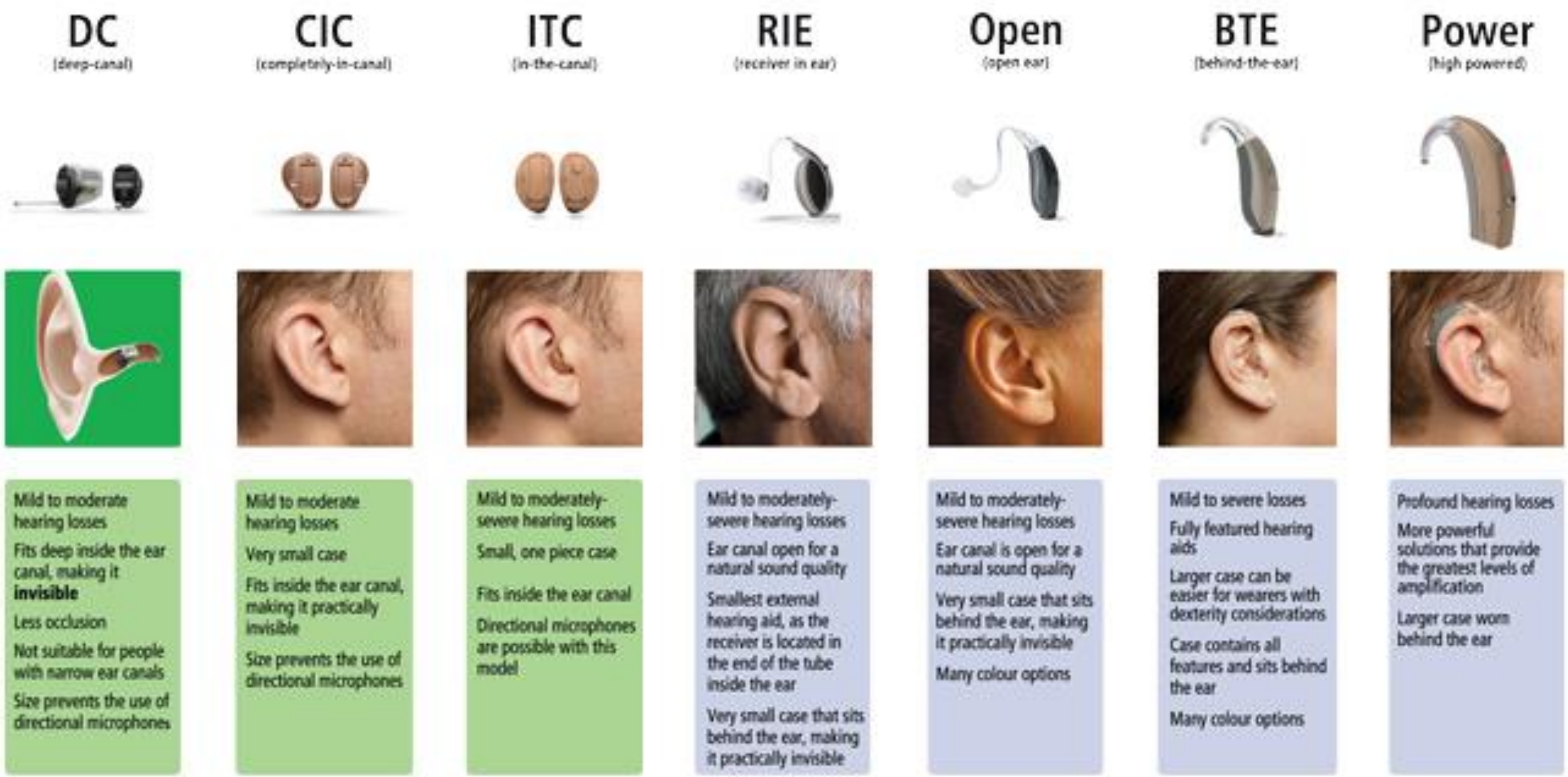
Source from
http://www.bdti.com/InsideDSP/2014/10/16/ONSemi

How hearing aids work

1. Sound goes in the Microphone.
2. Sound gets amplified.
3. Sound comes out the Speaker into your Ear.

Introduction to Hearing Aid Features by Steve Barber, HLAA-Wake Chapter

Style


DC
(deep-canal)
CIC
(completely-in-canal)
ITC
(in-the-canal)
RIE
(receiver in ear)
Open
(open ear)
BTE
(behind-the-ear)
Power
(high powered)
Mild to moderate hearing losses
Fits deep inside the ear canal, making it **invisible**
Less occlusion
Not suitable for people with narrow ear canals
Size prevents the use of directional microphones
Mild to moderate hearing losses
Very small case
Fits inside the ear canal, making it practically invisible
Size prevents the use of directional microphones
Mild to moderately-severe hearing losses
Small, one piece case
Fits inside the ear canal
Directional microphones are possible with this model
Mild to moderately-severe hearing losses
Ear canal open for a natural sound quality
Smallest external hearing aid, as the receiver is located in the end of the tube inside the ear
Very small case that sits behind the ear, making it practically invisible
Mild to moderately-severe hearing losses
Ear canal is open for a natural sound quality
Very small case that sits behind the ear, making it practically invisible
Many colour options
Mild to severe losses
Fully featured hearing aids
Larger case can be easier for wearers with dexterity considerations
Case contains all features and sits behind the ear
Many colour options
Profound hearing losses
More powerful solutions that provide the greatest levels of amplification
Larger case worn behind the ear

Technology

- Analog: Settings and Sound are both processed via analog technology.
- Digital Programmable: Settings are processed digitally, Sound is processed via analog technology.
- Full Digital: Both Settings and Sound are processed digitally.

Possible features

- Volume Control
- Telecoil
- Multiple Microphone Directionality
- Compression
- Clipping
- Direct Audio Input
- FM
- Bluetooth
- Programmability
- Speech Enhancement/Noise Reduction
- Frequency Shifting
- Earmold/Vent
- Remote Control

Hearing aid features

Feature	What does it do?	How does it help?
Channels	Separates the sound signal into discrete sections for processing.	Allows the hearing aid prescription to be customized across all pitches. The more channels in the hearing aid, the greater the ability to customize the frequency response.
Directional microphone systems	Gives preference to sounds coming from the front of the wearer and reduces sounds coming from other directions.	Improves speech understanding in background noise. Satisfaction is higher for hearing aids with directional microphone systems than for hearing aids without them.
Digital noise reduction	Determines if the signal contains unwanted noise and reduces the level of noise if present.	Makes the background noise less annoying and increases listening comfort. Digital noise reduction has been shown to be effective and preferred by hearing aid wearers.
Impulse noise reduction	Smooths quick impulse noises such as car keys rattling, typing on a keyboard and dishes rattling.	Improves listening comfort.
Feedback management	Reduces or eliminates whistling that can sometimes occur.	Improves listening comfort. Basic feedback management systems may reduce the overall amplification in order to remove the whistling. Advanced feedback management systems reduce or eliminate whistling without affecting overall amplification of the hearing aid.
Telecoil	Picks up signal from a compatible telephone or other electromagnetically looped system.	Improves signal to noise ratio and eliminates feedback because the signal bypasses the microphone and directly enters the processor. Commonly available in public places, like theatres and places of worship.
FM compatibility	Enables hearing aids to wirelessly connect with FM systems, sometimes via a special attachment called a boot.	Improves signal to noise ratio because the signal bypasses the microphone and directly enters the processor. Commonly used with children in educational settings.
Bluetooth compatibility	Enables hearing aids to wirelessly connect to mobile phones, MP3 players and other Bluetooth devices.	Improves signal to noise ratio and eliminates feedback or interference because the signal bypasses the microphone and directly enters the processor.
Wind noise reduction	Reduces the whooshing noise of wind blowing across the hearing aid microphone(s).	Improves listening comfort for people who spend time outdoors—such as golfers, boaters and walkers.
Data logging	Stores data about listening environments and user preferences.	Data can be viewed by hearing healthcare professional to improve fitting at follow-up.
Learning features	Logs settings that are set by the wearer for certain environments and then begins to make these changes automatically.	Gradually, the wearer will find that they need to adjust the volume or program less frequently, as the hearing aids become able to make these changes based on the sound environment.
Binaural processing	The two hearing aids communicate with each other.	This can be used to keep the hearing aids operating synchronously or to stream auditory signals from one aid to the other.

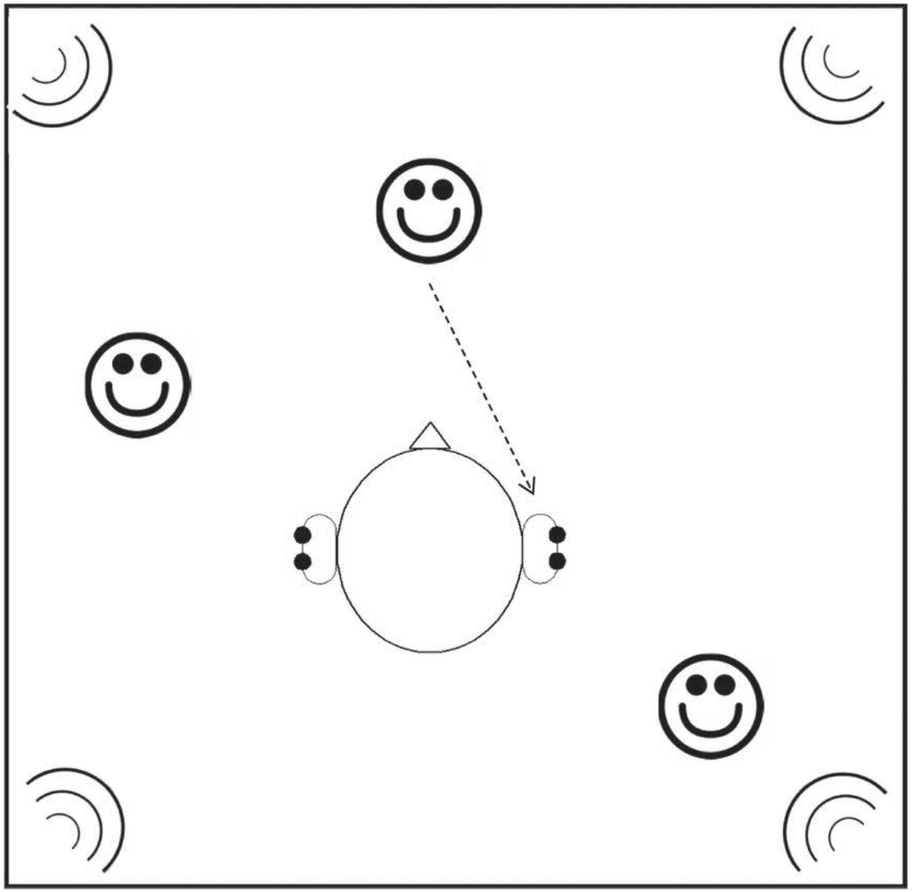
www.healthyhearing.com

http://www.healthyhearing.com/help/hearing-aids/technology

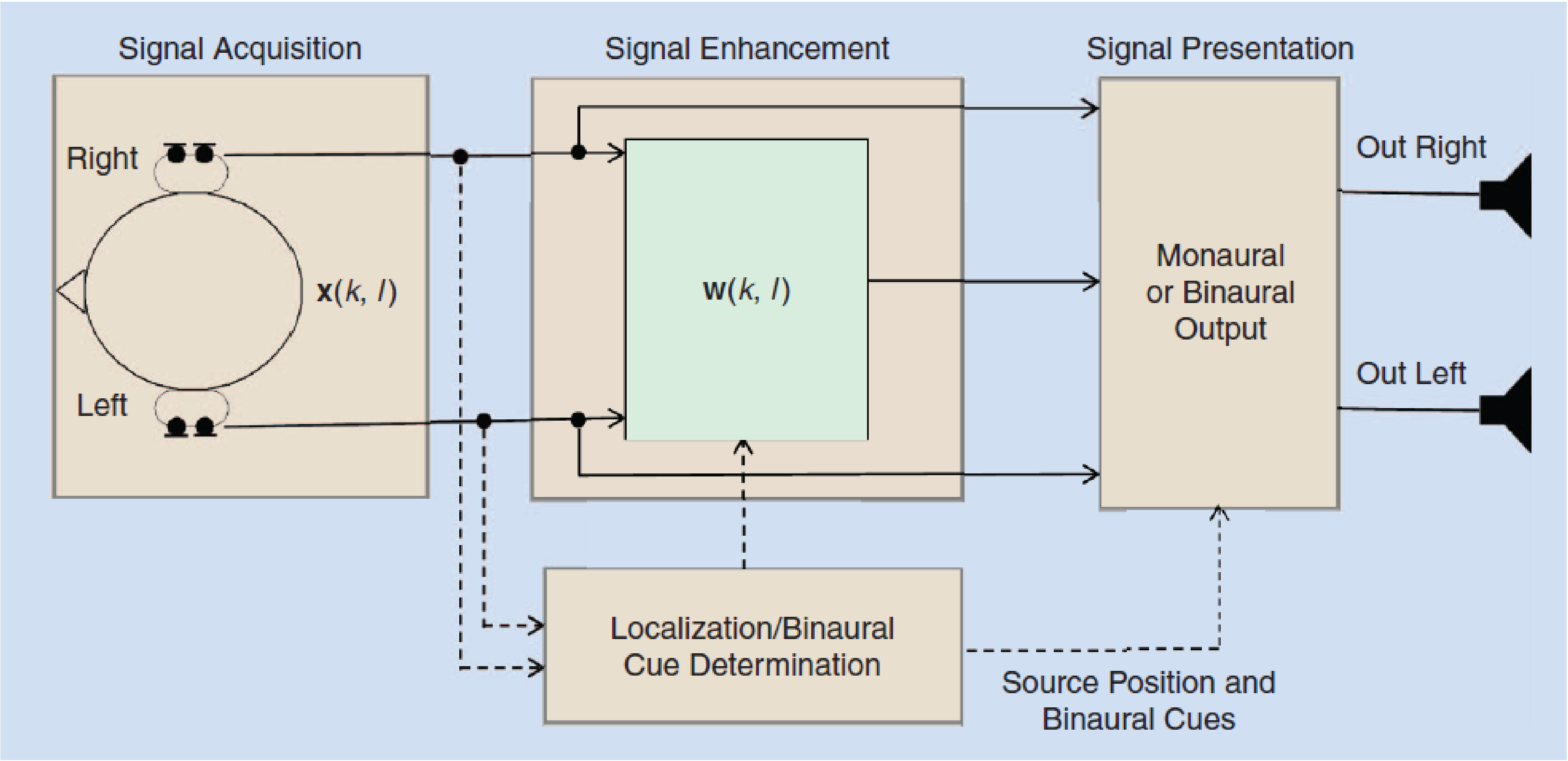
Noise reduction + binaural processing

- Single microphone
  - Adaptive analog filters
  - Spectral subtraction
  - Spectral enhancement

- Multi-microphone
  - Directional microphone elements
  - Two-microphone adaptive noise cancellation
  - Array with time-invariant weights
  - Two-microphone adaptive arrays



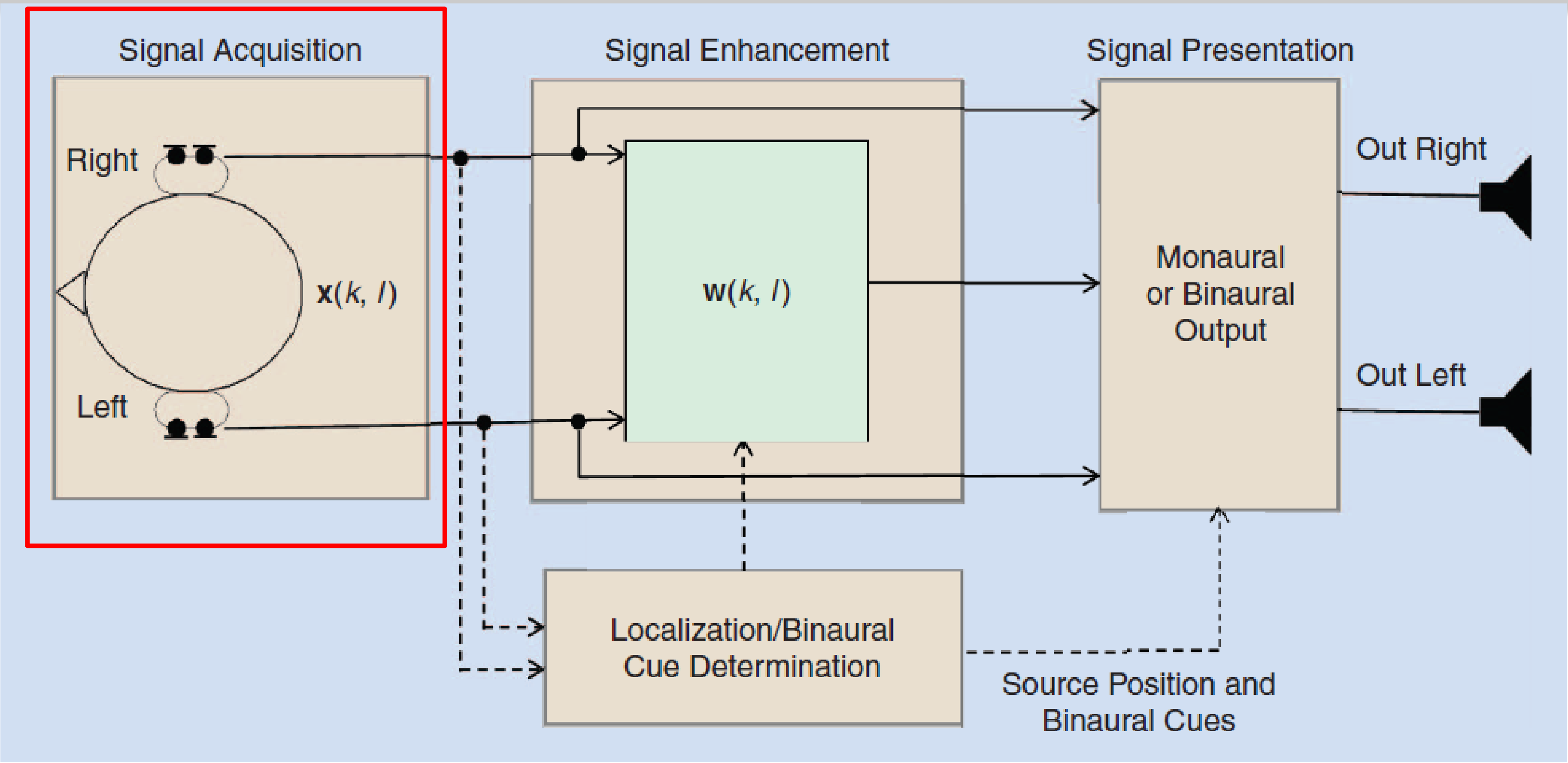
A general framework



Picture from

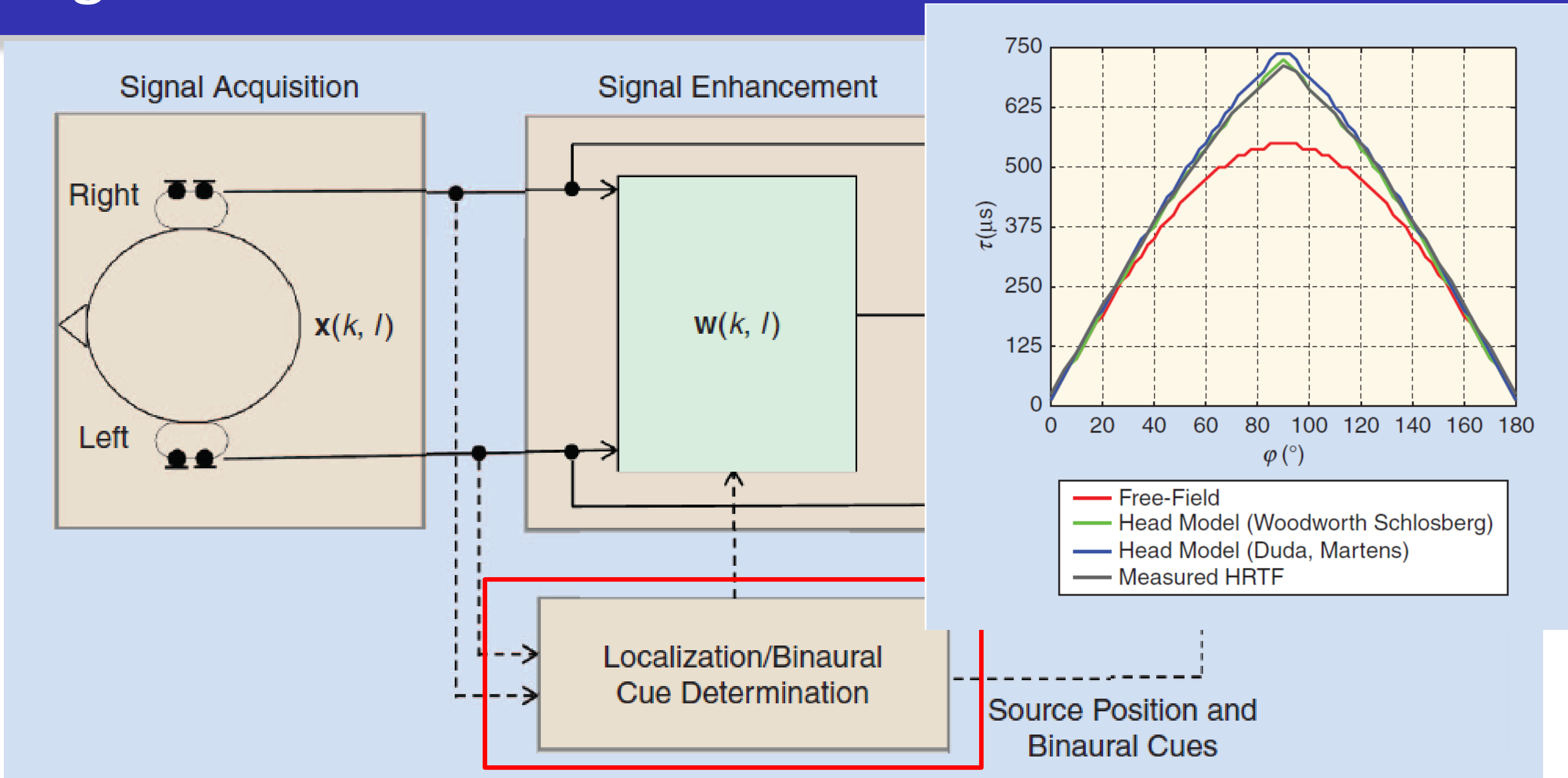
S. Doclo, W. Kellermann, S. Makino, and S. Nordholm, “Multichannel signal enchancement algorithms for assisted listening devices,” IEEE Signal Processing Magazine, vol. 32, no. 2, pp. 18-30, Mar. 2015.

A general framework: acquisition



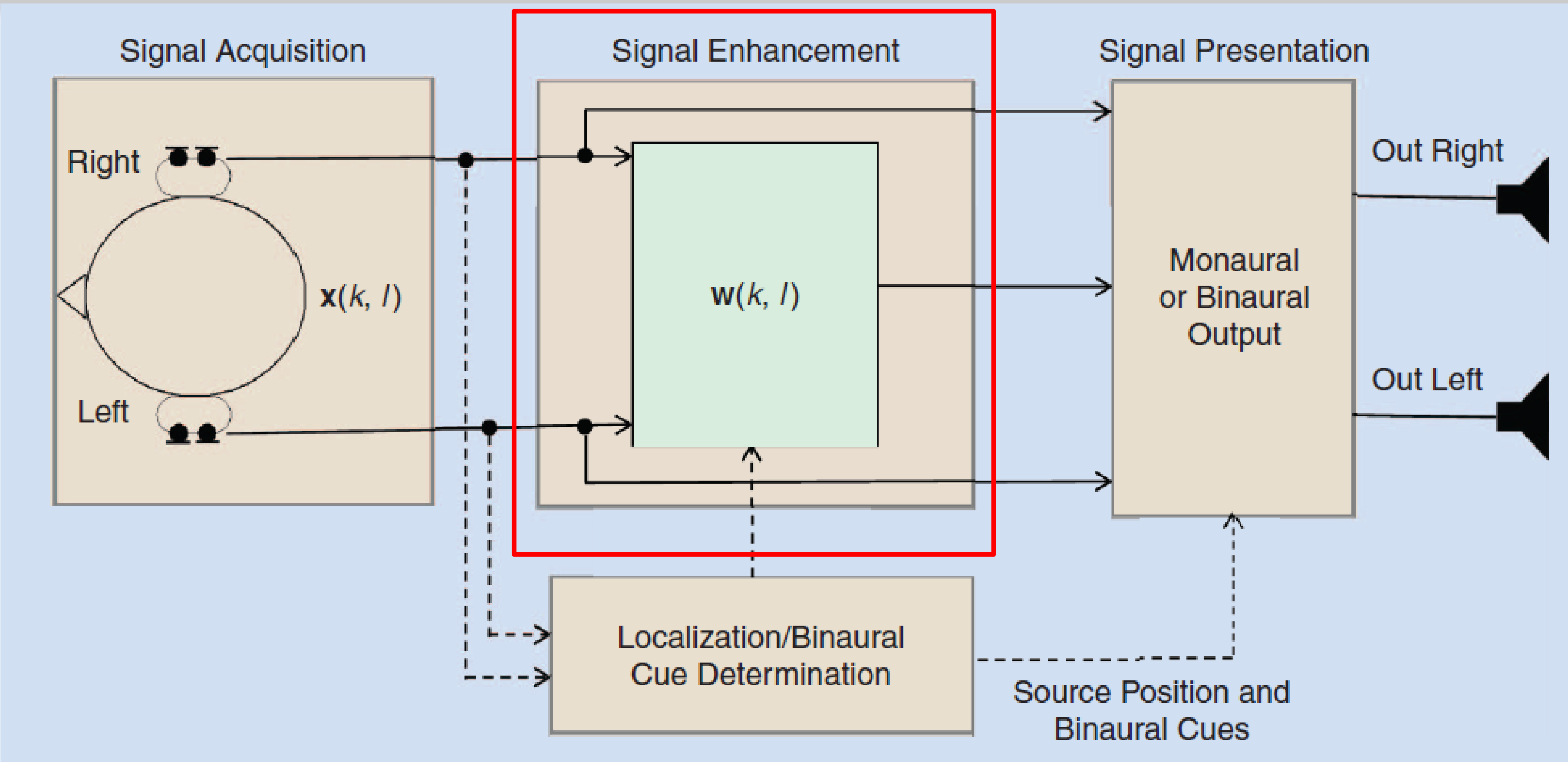
$$\mathbf{x} = \mathbf{h}_0 s_0 + \sum_{p=1}^{P-1} \mathbf{h}_p s_p + \mathbf{n} = \mathbf{h}_0 s_0 + \mathbf{v},$$

A general framework: localization



- Steered-response power (SRP);
- MUSIC;
- TDOA;
- GCC-PHAT; SRP-PHAT;
- BSS-based

A general framework: enhancement



Data-independent beamformer

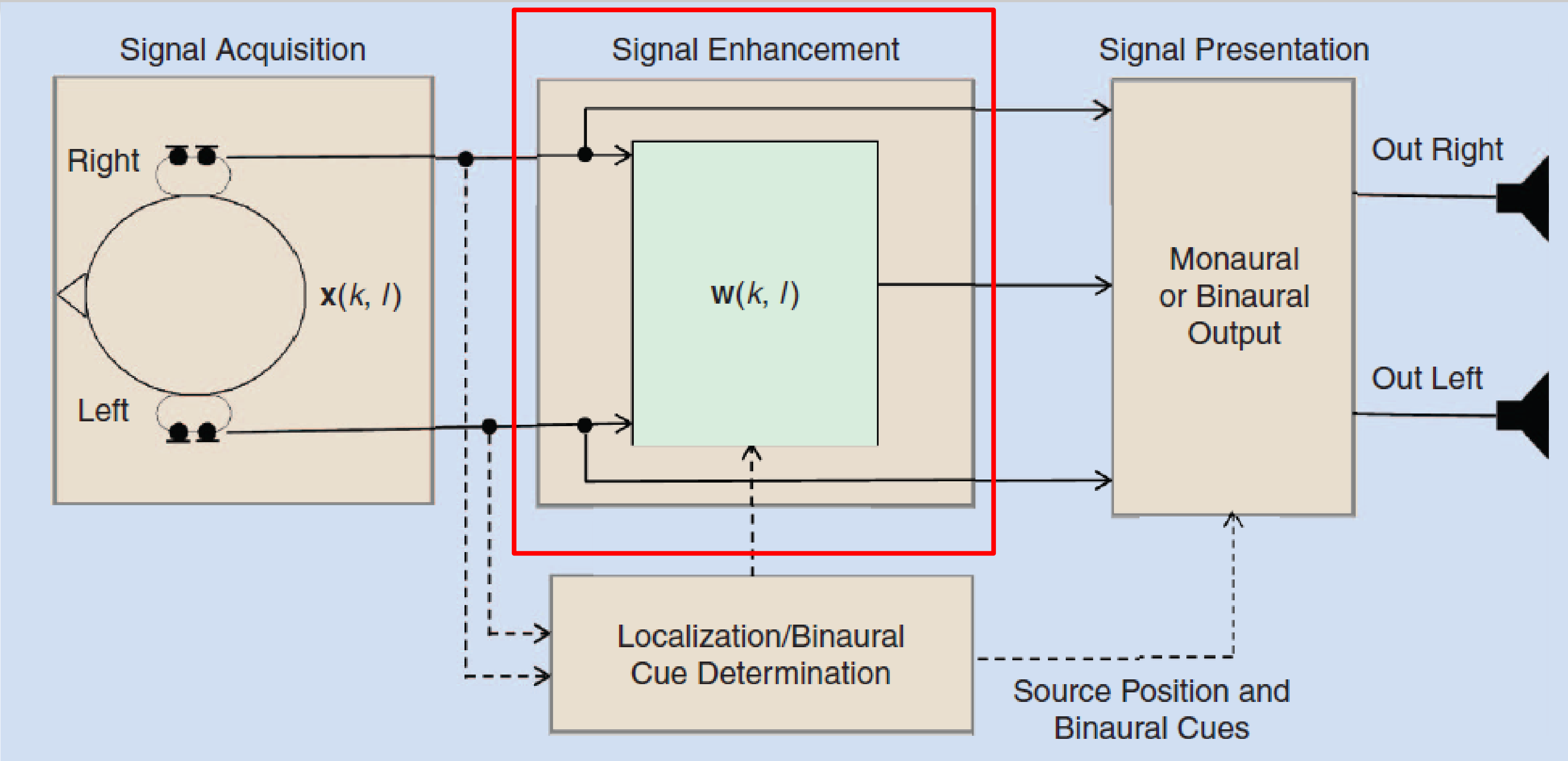
- Delay-and-sum
- Superdirective
- Differential

Need to know

- Target DOA
- Complete microphone topology

More suitable for monaural devices

A general framework: enhancement

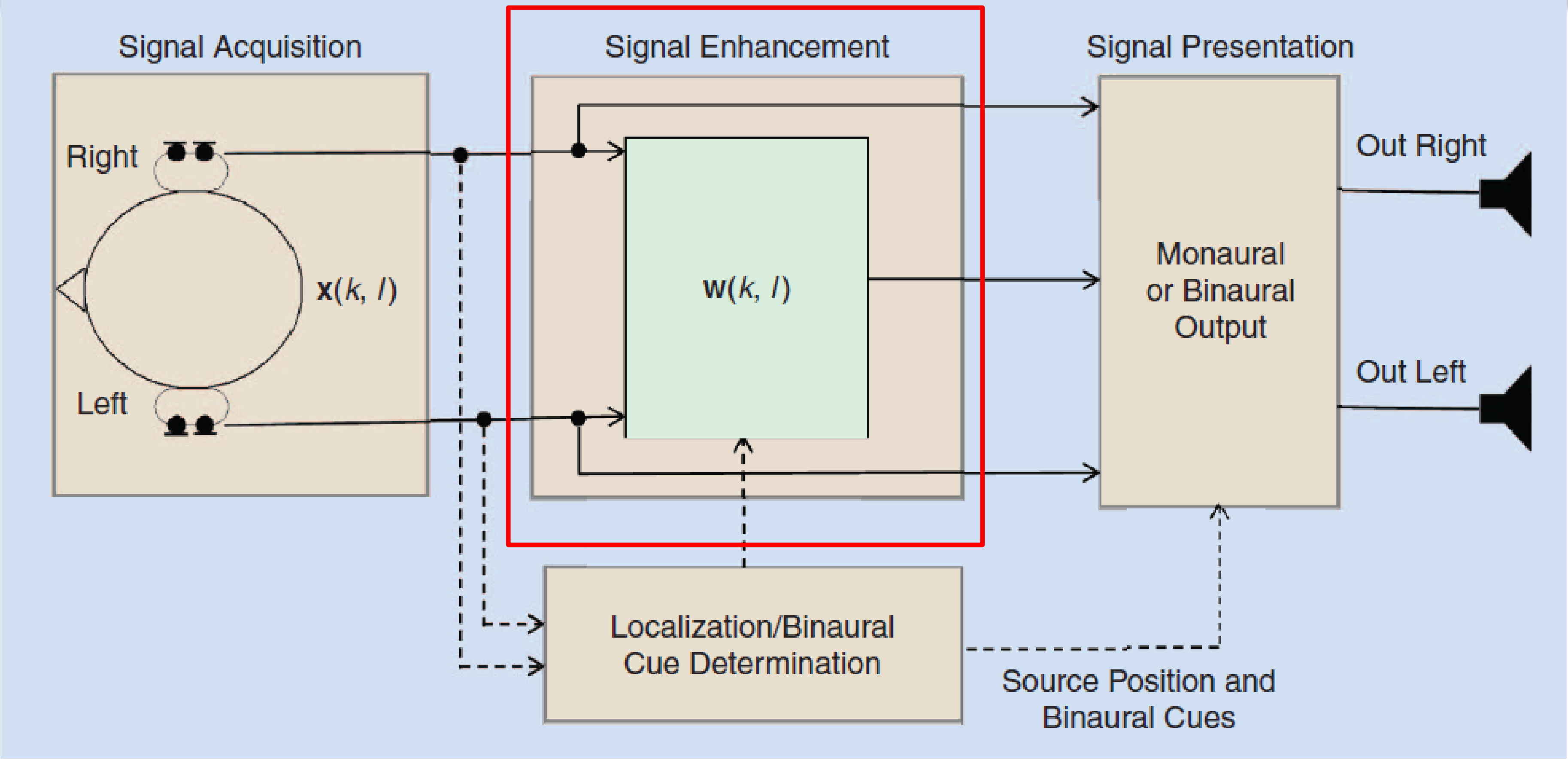


Data-dependent beamformer: statistically optimum

- Minimum Variance Distortionless Response (MVDR)
- Multichannel Wiener Filtering (MWF)
- Blind Source Separation (BSS)

Need to estimate the interference and noise statistics

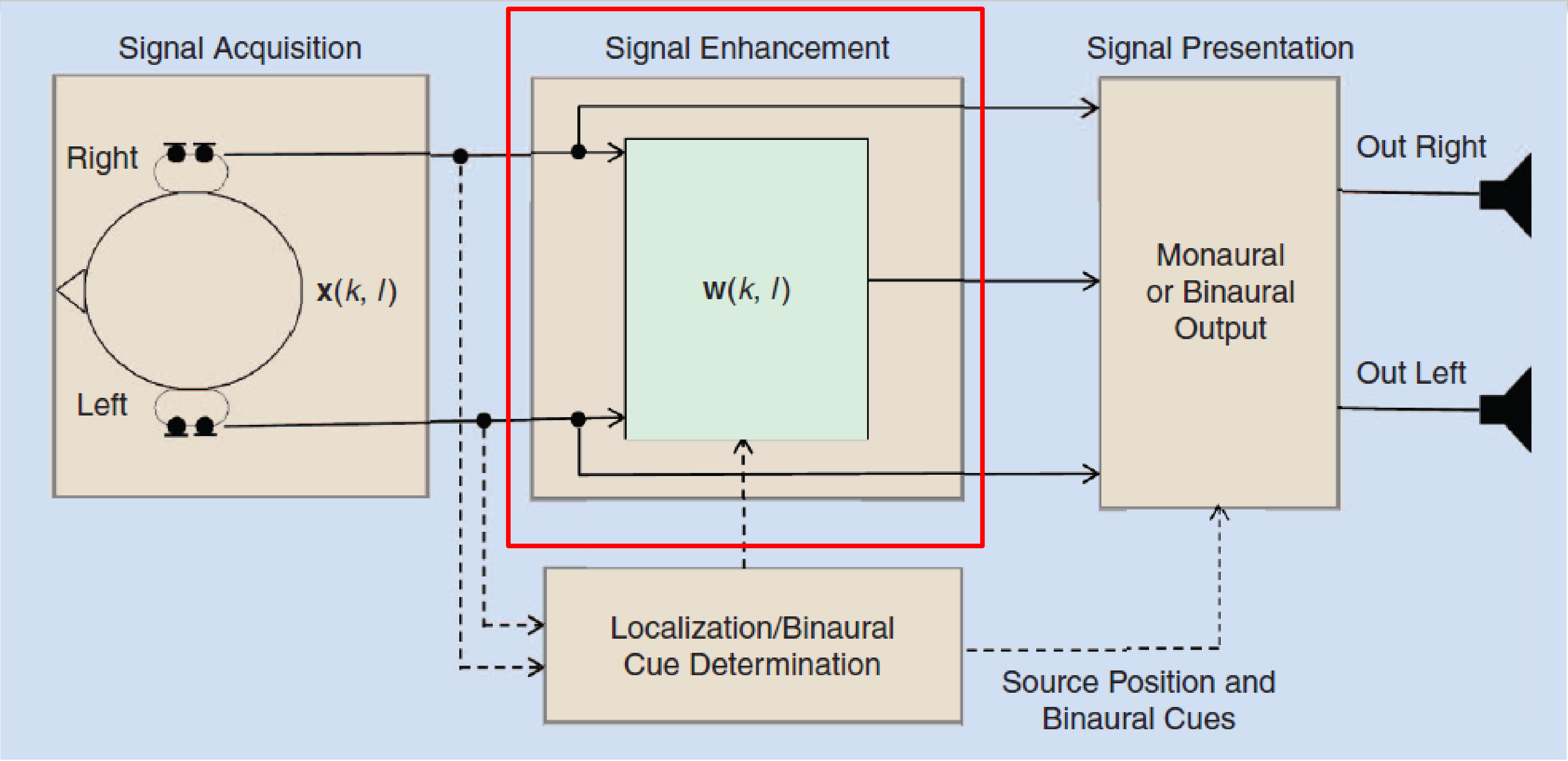
A general framework: enhancement



MVDR beamformer

$$\min_{\mathbf{w}} \mathbf{w}^H \mathbf{\Phi}_{xx} \mathbf{w}, \quad \text{s.t. } \mathbf{w}^H \mathbf{h}_0 = 1 \longrightarrow \mathbf{w}_{\text{MVDR}} = \frac{\mathbf{\Phi}_{vv}^{-1} \mathbf{h}_0}{\mathbf{h}_0^H \mathbf{\Phi}_{vv}^{-1} \mathbf{h}_0}$$

A general framework: enhancement

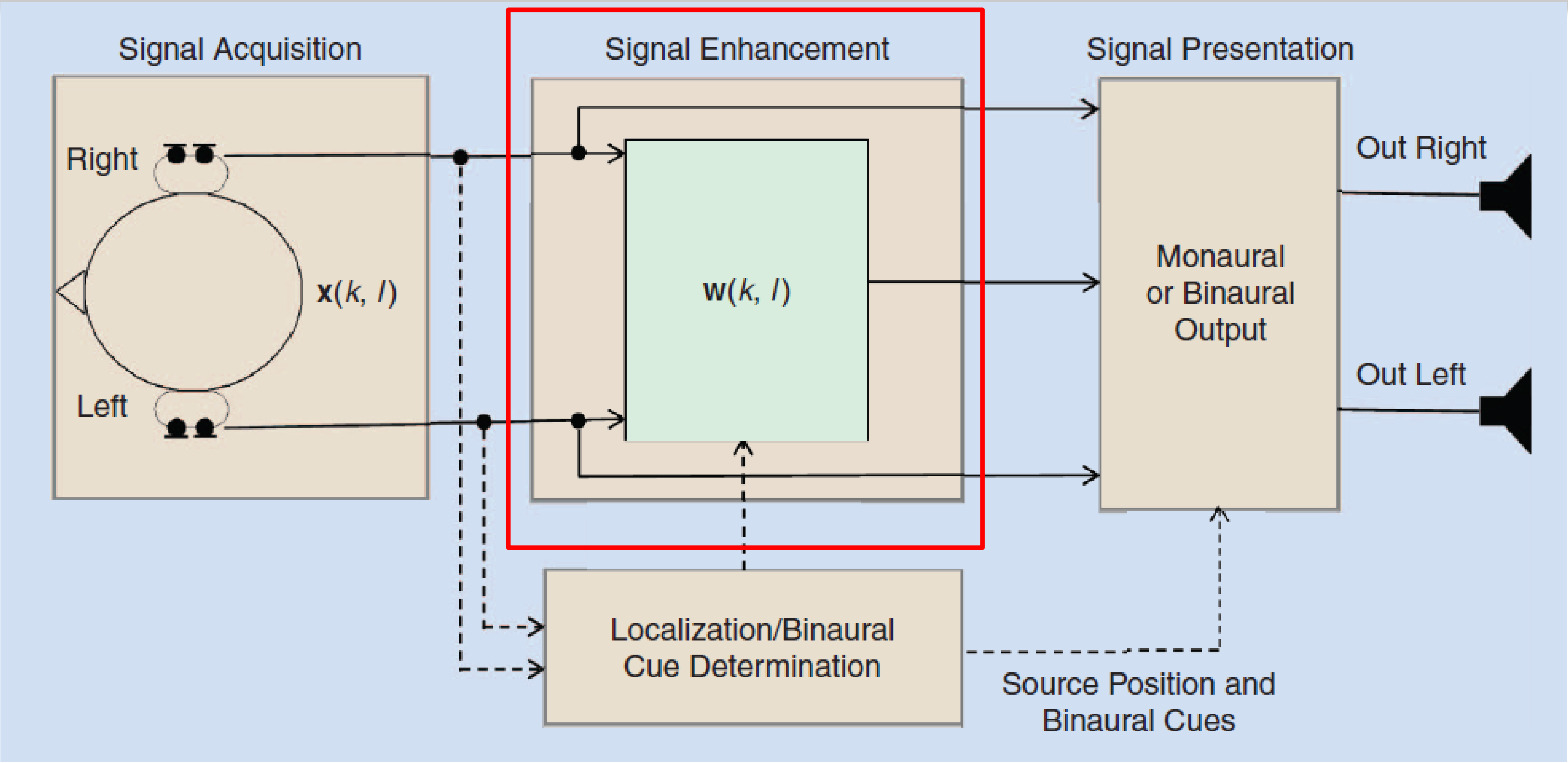


MVDR beamformer with relative transfer function

$$\tilde{\mathbf{h}}_0 = \frac{\mathbf{h}_0}{h_{0,r}}$$

$$\min_{\mathbf{w}} \mathbf{w}^H \mathbf{\Phi}_{xx}\mathbf{w}, \quad \text{s.t. } \mathbf{w}^H\mathbf{h}_0 = h_{0,r} \longrightarrow \quad \tilde{\mathbf{w}}_{\text{MVDR}} = \frac{\mathbf{\Phi}_{vv}^{-1}\tilde{\mathbf{h}}_0}{\tilde{\mathbf{h}}_0^H\mathbf{\Phi}_{vv}^{-1}\tilde{\mathbf{h}}_0}$$

A general framework: enhancement

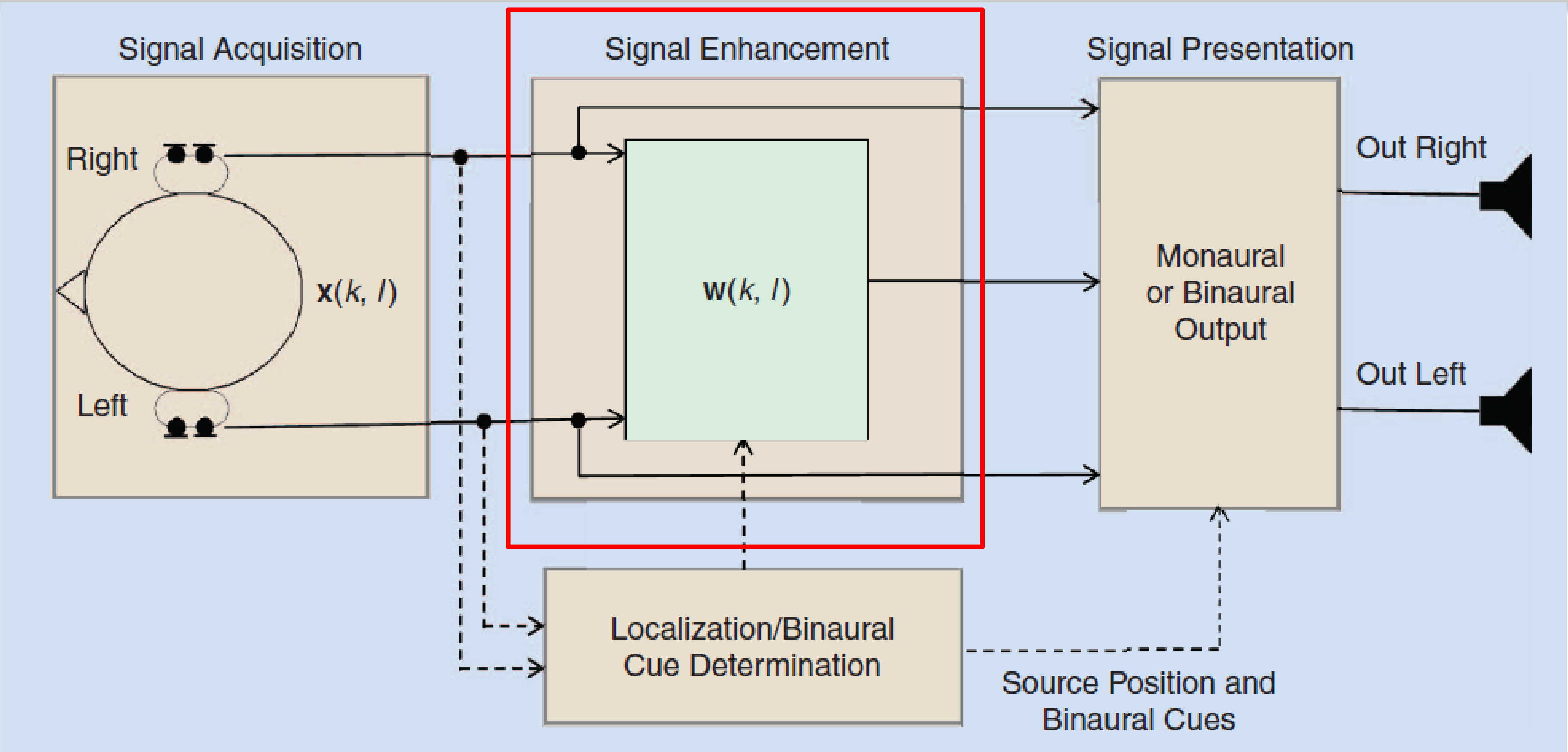


Multichannel Wiener Filter (MWF)

$$\min_{\mathbf{w}} E\left\{\left|s_0 - \mathbf{w}^H \mathbf{x}\right|^2\right\} \longrightarrow \mathbf{w}_{\mathrm{MWF}} = \mathbf{\Phi}_{\mathrm{xx}}^{-1}\mathbf{h}_0 \phi_{s_0 s_0}$$

PSD of s_0

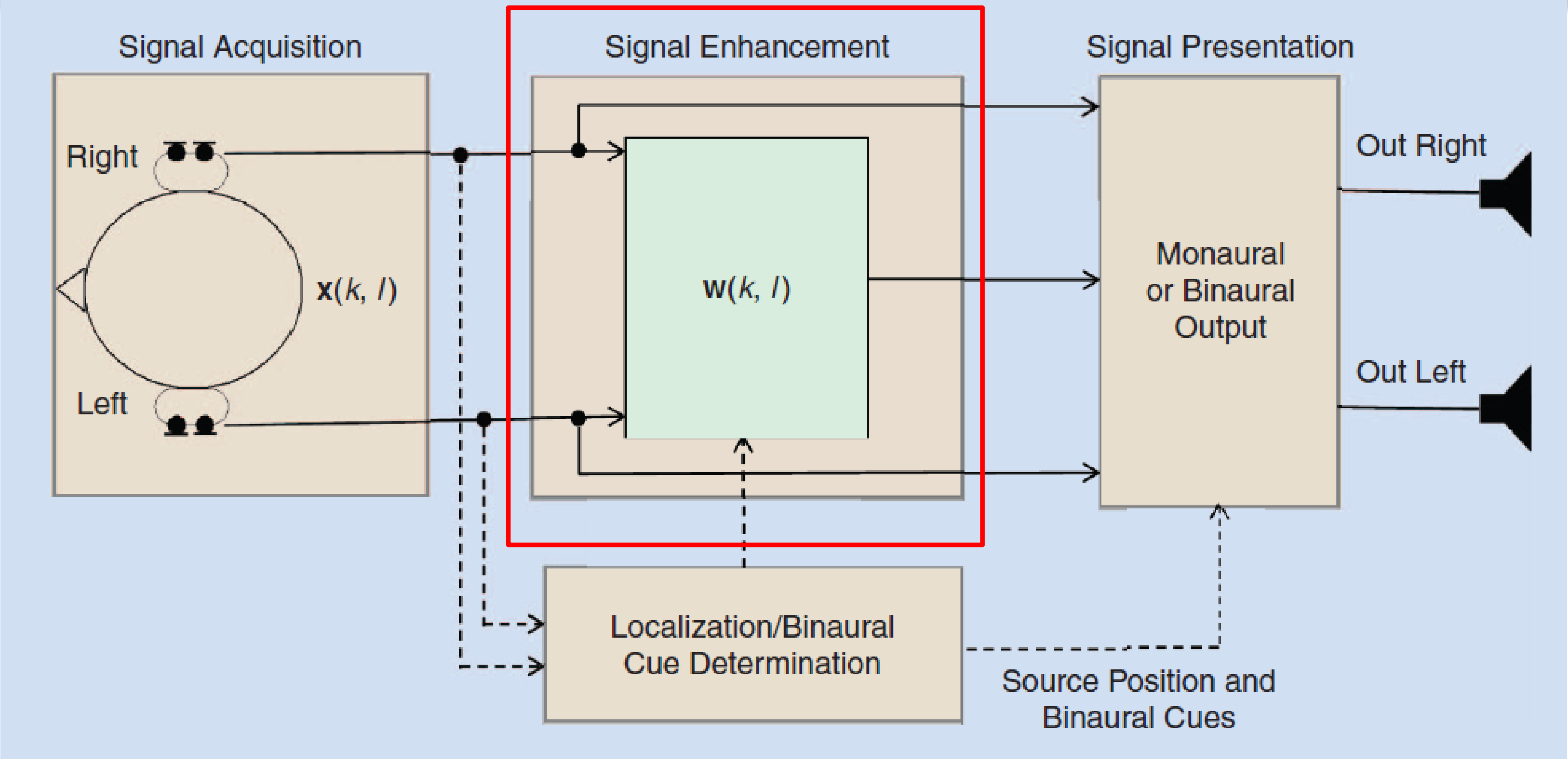
A general framework: enhancement



MWF with relative transfer function

$$\min_{\mathbf{w}} E\left\{\left|h_{0,r}s_0 - \mathbf{w}^H\mathbf{x}\right|^2\right\} \rightarrow \tilde{\mathbf{w}}_{\mathrm{MWF}} = \left(\phi_{s_0s_0}\mathbf{h}_0\mathbf{h}_0^H + \mathbf{\Phi}_{\mathrm{vv}}\right)^{-1}\phi_{s_0s_0}\mathbf{h}_0 h_{0,r}^*$$

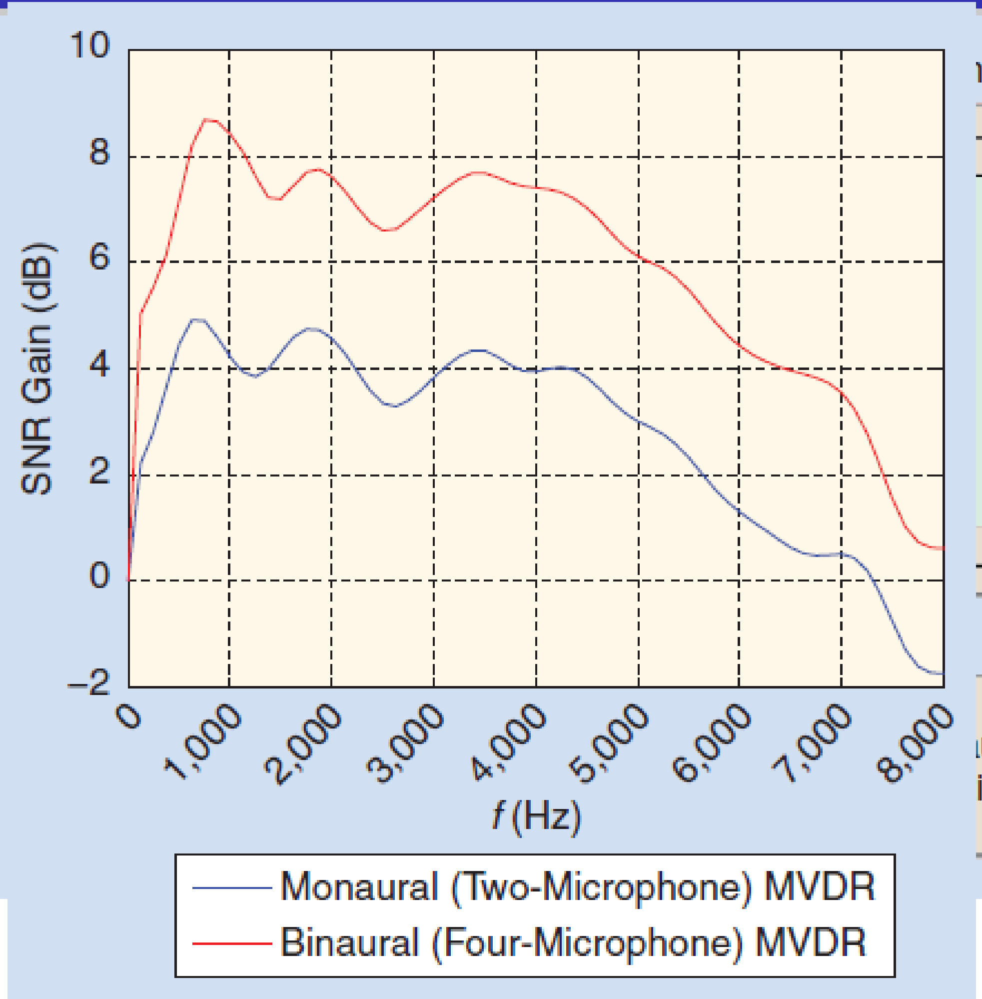
A general framework: enhancement

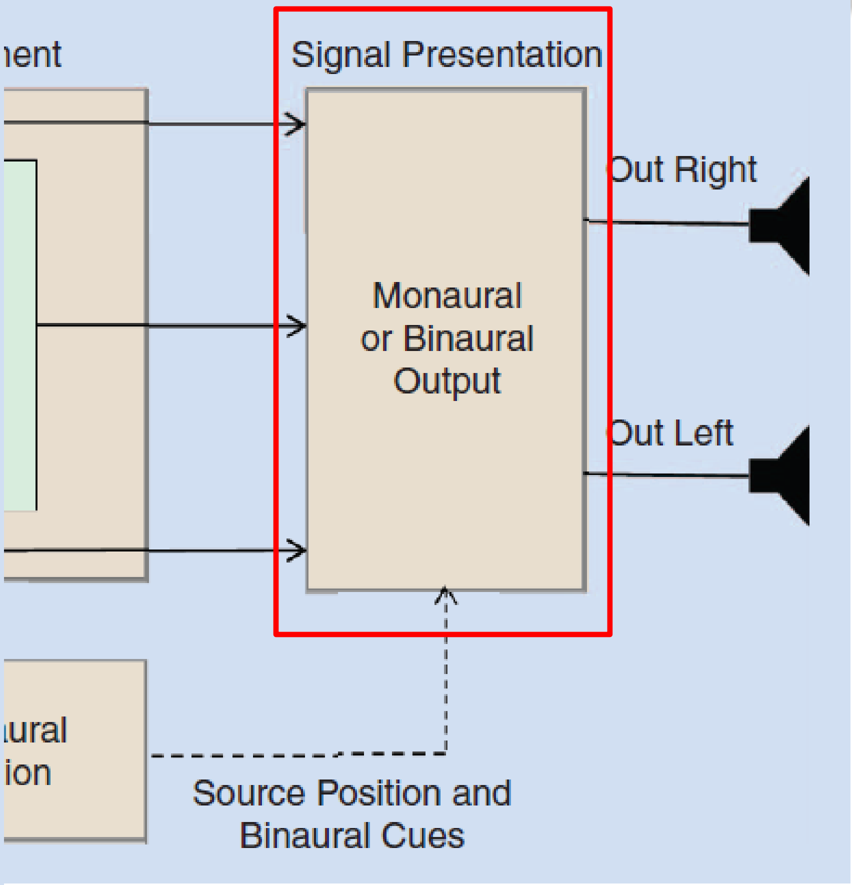


MWF with speech-distortion-weighted

$$\min_{\mathbf{w}} E\left\{\left|h_{0,r}s_0 - \mathbf{w}^H \mathbf{h}_0 \mathbf{s}_0\right|^2\right\} + \mu E\left\{\left|\mathbf{w}^H \mathbf{v}\right|^2\right\} \longrightarrow \tilde{\mathbf{w}}_{\mathrm{MWF}} = \left(\phi_{s_0 s_0}\mathbf{h}_0\mathbf{h}_0^H + \mu \mathbf{\Phi}_{\mathrm{vv}}^{-1}\right)\phi_{s_0 s_0}\mathbf{h}_0 h_{0,r}^*$$

A general framework: presentation

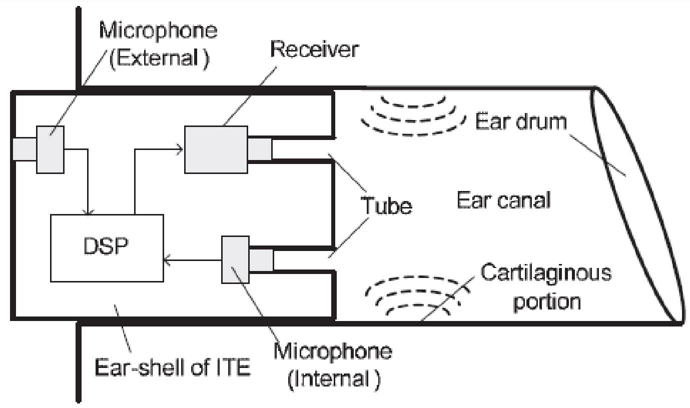




Binaural extraction:

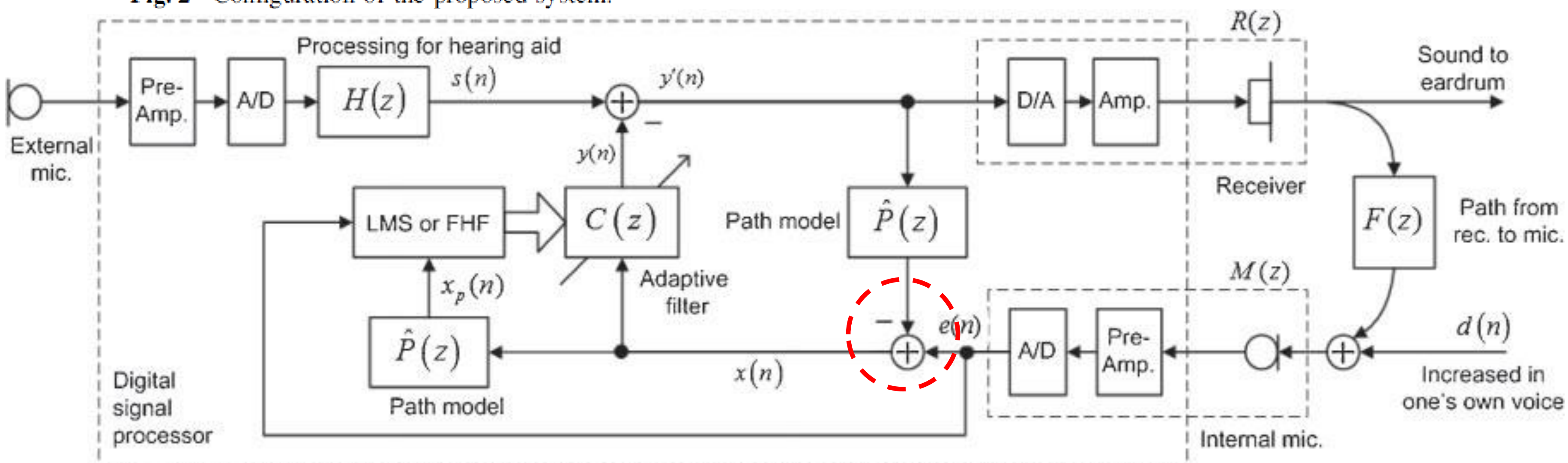
Also important to preserve the binaural cues of the residue noise to allow binaural unmasking

ANC in hearing aids with occlusion reduction (ITC hearing aids)



- Increased SPL at low frequencies when ear canal is blocked (~ 15 dB from 100-300 Hz)
- Occlusion reduction using FBANC.
- Need a fast adaptive algorithm to achieve good occlusion reduction.

Fig. 2 Configuration of the proposed system.

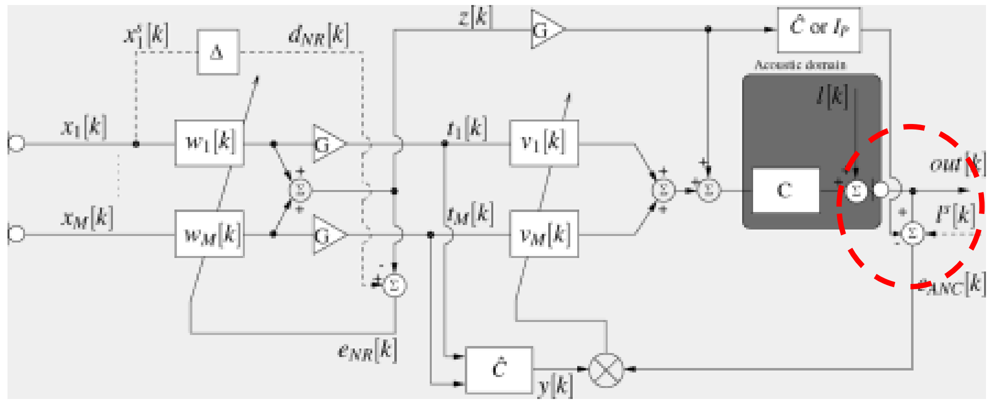


Picture from

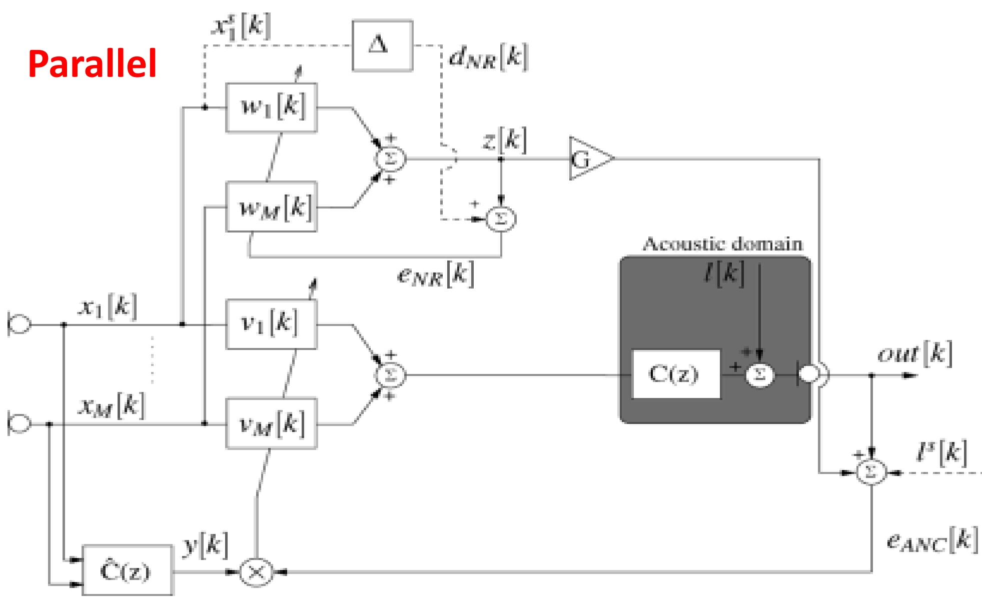
M. Sunohara, K Watanuki and M. Tateno, "Occlusion reduction system for hearing aids using active noise control technique", Acoust. Sci. & Tech. 35, 6 (2014)

Integrating ANC in Hearing Aids

Cascade



Parallel



- No earmold to prevent ambient sound from leakage into the ear canal → Lower SNR
- Attenuate leakage noise signal using ANC.
- Need an ear canal microphone

Picture from

R. Serizel, M. Moonen, J. Wouters, and S. H. Jensen, "Integrated Active Noise Control and Noise Reduction in Hearing Aids", IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, VOL. 18, NO. 6, AUG. 2010

Summary on hearing aids

- Various signal processing approaches have been applied in noise reduction for hearing aids.
- Incorporating multiple microphones and binaural processing have been shown to be beneficial and will be the key research topics in hearing aids.
- Personalization embedded with intuitive interface (in phones): hearing loss profile, or listening preference.

Read more

[1] H. Luts, K. Eneman, J. Wouters, M. Schulte, M. Vormann, M. Buechler, N. Dillier, R. Houben, W. A. Dreschler, M. Froehlich, H. Puder, G. Grimm, V. Hohmann, A. Leijon, A. Lombard, D. Mauler, and A. Spriet, "Multicenter evaluation of signal enhancement algorithms for hearing aids," J. Acoust. Soc. Amer., vol. 127, no. 3, pp. 2054–2063, Mar. 2010.

[2] M. Brandstein, and D. Ward, Microphone Arrays: Signal Processing Techniques and Applications. New York: Springer, 2001.

[3] J. Blauert, The Technology of Binaural Listening. New York: Springer, 2013.

[4] S. Makino, T.-W. Lee, and H. Sawada, Blind Speech Separation. New York: Springer, 2007.

[5] J. Blauert, Spatial Hearing: The Psychophysics of Human Sound Localisation. Cambridge, MA: MIT, 1997.

[6] R. Martin, U. Heute, and C. Antweiler, Advances in Digital Speech Transmission. Hoboken, NJ: Wiley, 2008.

[7] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," IEEE Trans. Antennas Propag., vol. 34, pp. 276–280, Mar. 1986.

[8] C. Knapp, and G. C. Carter, "The generalized correlation method for estimation of time delay," IEEE Trans. Acoust., Speech, Signal Processing, vol. 24, no. 4, pp. 320–327, 1976.

[9] J. Benesty, "Adaptive eigenvalue decomposition algorithm for passive acoustic source localization," J. Acoust. Soc. Amer., vol. 107, no. 1, pp. 384–391, Jan. 2000.

[10] A. Lombard, Y. Zheng, H. Buchner, and W. Kellermann, "TDOA estimation for multiple sound sources in noisy and reverberant environments using broadband independent component analysis," IEEE Trans. Audio, Speech, Lang. Processing, vol. 19, no. 6, pp. 1490–1503, Aug. 2011.

[11] R. O. Duda, and W. L. Martens, "Range dependence of the response of a spherical head model," J. Acoust. Soc. Amer., vol. 104, no. 5, pp. 3048–3058, 1998.

[12] D. L. Wang, and G. Brown, Computational Auditory Scene Analysis: Principles, Algorithms and Applications. New York: IEEE Press/Wiley-Interscience, 2007.

[13] T. May, S. van de Par, and A. Kohlrausch, "A probabilistic model for robust localization based on a binaural auditory front-end," IEEE Trans. Audio, Speech, Lang. Processing, vol. 19, no. 1, pp. 1–13, Jan. 2011.

[14] J. Scheuing, and B. Yang, "Disambiguation of TDOA estimation for multiple sources in reverberant environments," IEEE Trans. Audio, Speech, Lang. Processing, vol. 16, no.8, pp. 1479–1489, 2008.

Read more

[15] G. W. Elko, “Microphone array systems for hands-free telecommunication,” Speech Commun., vol. 20, no. 3–4, pp. 229–240, 1996.

[16] J. M. Kates, and M. R. Weiss, “A comparison of hearing-aid array-processing techniques,” J. Acoust. Soc. Amer., vol. 99, no.5, pp. 3138–3148, May 1996.

[17] S. Doclo, and M. Moonen, “Superdirective beamforming robust against microphone mismatch,” IEEE Trans. Audio, Speech, Lang. Processing, vol. 15, no. 2, pp. 617–631, Feb. 2007.

[18] E. Mabande, A. Schad, and W. Kellermann, “Design of robust superdirective beamformers as a convex optimization problem,” in Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing, Taipei, Taiwan, Apr. 2009, pp. 77–80.

[19] T. Lotter, and P. Vary, “Dual-channel speech enhancement by superdirective beamforming,” EURASIP J. Appl. Signal Processing, vol. 2006, no. 1, pp. 175–175, Jan. 2006.

[20] B. V. Veen, and K. Buckley, “Beamforming: A versatile approach to spatial filtering,” IEEE ASSP Mag., vol. 5, no. 2, pp. 4–24, 1988.

[21] J. Capon, “High resolution frequency-wavenumber spectrum analysis,” Proc. IEEE, vol.57, no. 7, pp. 1408–1418, Aug. 1969.

[22] L. J. Griffiths, and C. W. Jim, “An alternative approach to linearly constrained adaptive beamforming,” IEEE Trans. Antennas Propag., vol. 30, no. 1, pp. 27–34, Jan. 1982.

[23] O. Hoshuyama, A. Sugiyama, and A. Hirano, “A robust adaptive beamformer for microphone arrays with a blocking matrix using constrained adaptive filters,” IEEE Trans. Signal Processing, vol. 47, no. 10, pp. 2677–2684, Oct. 1999.

[24] S. Gannot, D. Burshtein, and E. Weinstein, “Signal enhancement using beamforming and non-stationarity with applications to speech,” IEEE Trans. Signal Processing, vol.49, no. 8, pp. 1614–1626, Aug. 2001.

[25] A. Krueger, E. Warsitz, and R. Haeb-Umbach, “Speech enhancement with a GSClike structure employing eigenvector-based transfer function ratios estimation,” IEEE Trans. Audio, Speech, Lang. Processing, vol. 19, no. 1, pp. 206–218, Jan. 2011.

[26] H. Cox, R. M. Zeskind, and M. M. Owen, “Robust adaptive beamforming,” IEEE Trans. Acoust., Speech, Signal Processing, vol. 35, no. 10, pp. 1365–1376, Oct. 1987.

[27] A. Spriet, M. Moonen, and J. Wouters, “Spatially pre-processed speech distortion weighted multi-channel Wiener filtering for noise reduction,” Signal Process., vol. 84, no.12, pp. 2367–2387, Dec. 2004.

[28] F. L . Luo, J. Y. Yang, C. Pavlovic, and A. Nehorai, “Adaptive null-forming scheme in digital hearing aids,” IEEE Trans. Signal Processing, vol. 50, no. 7, pp. 1583–1590, 2002.

[29] J. B. Maj, L. Royackers, M. Moonen, and J. Wouters, “Comparison of adaptive noise reduction algorithms in dual microphone hearing aids,” Speech Commun., vol. 48, no. 8, pp. 957–960, Aug. 2006.

Read more

[30] B. Cornelis, M. Moonen, and J. Wouters, "Speech intelligibility improvements with hearing aids using bilateral and binaural adaptive multichannel Wiener filtering based noise reduction," J. Acoust. Soc. Amer., vol. 131, no. 6, pp. 4743–4755, June 2012.

[31] S. Doclo, A. Spriet, J. Wouters, and M. Moonen, "Frequency-domain criterion for speech distortion weighted multichannel Wiener filter for robust noise reduction," Speech Commun., vol. 49, no. 7–8, pp. 636–656, July–Aug. 2007.

[32] L. W. Brooksand I. S. Reed, "Equivalence of the likelihood ratio processor, the maximum signal-to-noise ratio filter, and the Wiener filter," IEEE Trans. Aerosp. Electron. Syst., vol.8, no. 5, pp. 690–692, 1972.

[33] L. Parra, and C. Spence, "Convolutive blind separation of non-stationary sources," IEEE Trans. Speech Audio Processing, vol. 8, no. 3, pp. 320–327, May 2000.

[34] S. Araki, R. Mukai, S. Makino, T. Nishikawa, and H. Saruwatari, "The fundamental limitation of frequency domain blind source separation for convolutive mixtures of speech," IEEE Trans. Speech Audio Processing, vol. 11, no. 2, pp. 109–116, Mar. 2003.

[35] H. Buchner, R. Aichner, and W. Kellermann, "A generalization of blind source separation algorithms for convolutive mixtures based on second order statistics," IEEE Trans. Speech Audio Processing, vol. 13, no. 1, pp. 120–134, Jan. 2005.

[36] P. Smaragdis, "Blind separation of convolved mixtures in the frequency domain," Neurocomputing, vol. 22, nos. 1–3, pp. 21–34, Nov. 1998.

[37] H. Sawada, R. Mukai, S. Araki, and S. Makino, "A robust and precise method for solving the permutation problem of frequency-domain blind source separation," IEEE Trans. Speech Audio Processing, vol. 12, no. 5, pp. 530–538,Sept. 2004.

[38] K. Matsuoka, and S. Nakashima, "Minimal distortion principle for blind source separation," in Proc. Independent Component Analysis and Signal Separation, Dec.2001, pp. 722–727.

[39] S. Araki, S. Makino, Y. Hinamoto, R. Mukai, T. Nishikawa, and H. Saruwatari, "Equivalence between frequency domain blind source separation and frequency domain adaptive beamforming for convolutive mixtures," EURASIP J. Appl. Signal Process., vol.2003, no. 11, pp. 1157–1166, Nov. 2003.

[40] Y. Zheng, K. Reindl, and W. Kellermann, "Analysis of dual-channel ICA-based blocking matrix for improved noise estimation," EURASIP J. Appl. Signal Processing, vol.26, 2014. Doi:10.1186/1687-6180-2014-26. [Online]. Available: http://asp.eurasip journals.com/content/2014/1/26

[41] S. Nordholm, I. Claesson, and M. Dahl, "Adaptive microphone array employing calibration signals: An analytical evaluation," IEEE Trans. Speech Audio Processing, vol.7, no. 3, pp. 241–252, May 1999.

[42] T. Gerkmann, C. Breithaupt, and R. Martin, "Improved a posteriori speech presence probability estimation based on a likelihood ratio with fixed priors," IEEE Trans. Audio, Speech, Lang. Processing, vol. 16, no. 5, pp. 910–919, July 2008.

Read more

[43] K. Reindl, S. Markovich-Golan, H. Barfuss, S. Gannot, and W. Kellermann, “Geometrically constrained TRINICON-based relative transfer function estimation in underdetermined scenarios,” in Proc. IEEE Workshop Applications of Signal Processing to Audio and Acoustics (WASPAA), Oct. 2013.

[44] T. Van den Bogaert, S. Doclo, J. Wouters, and M. Moonen, “The effect of multimicrophone noise reduction systems on sound source localization by users of binaural hearing aids,” J. Acoust. Soc. Amer., vol. 124, no. 1, pp. 484–497, July 2008.

[45] G. Grimm, V. Hohmann, and B. Kollmeier, “Increase and subjective evaluation of feedback stability in hearing aids by a binaural coherence-based noise reduction scheme,” IEEE Trans. Audio, Speech, Lang. Processing, vol. 17, no. 7, pp. 1408–1419,
Sept. 2009.

[46] K. Reindl, Y. Zheng, A. Schwarz, S. Meier, R. Maas, A. Sehr, and W. Kellermann, “A stereophonic acoustic signal extraction scheme for noisy and reverberant environments,” Comput. Speech Lang. (CSL), vol. 27, no. 3, pp. 726–745, May 2012.

[47] D. Welker, J. Greenberg, J. Desloge, and P. Zurek, “Microphone-array hearing aids with binaural output–Part II: A two-microphone adaptive system,” IEEE Trans. Speech Audio Processing, vol. 5, no. 6, pp. 543–551, Nov. 1997.

[48] B. Cornelis, S. Doclo, T. Van den Bogaert, J. Wouters, and M. Moonen, “Theoretical analysis of binaural multi-microphone noise reduction techniques,” IEEE Trans. Audio, Speech, Lang. Processing, vol. 18, no. 2, pp. 342–355, Feb. 2010.

[49] E. Hadad, S. Gannot, and S. Doclo, “Binaural linearly constrained minimum variance beamformer for hearing aid applications,” in Proc. Int. Workshop on Acoustic Signal Enhancement (IWAENC), Aachen, Germany,Sept. 2012, pp. 117–120.

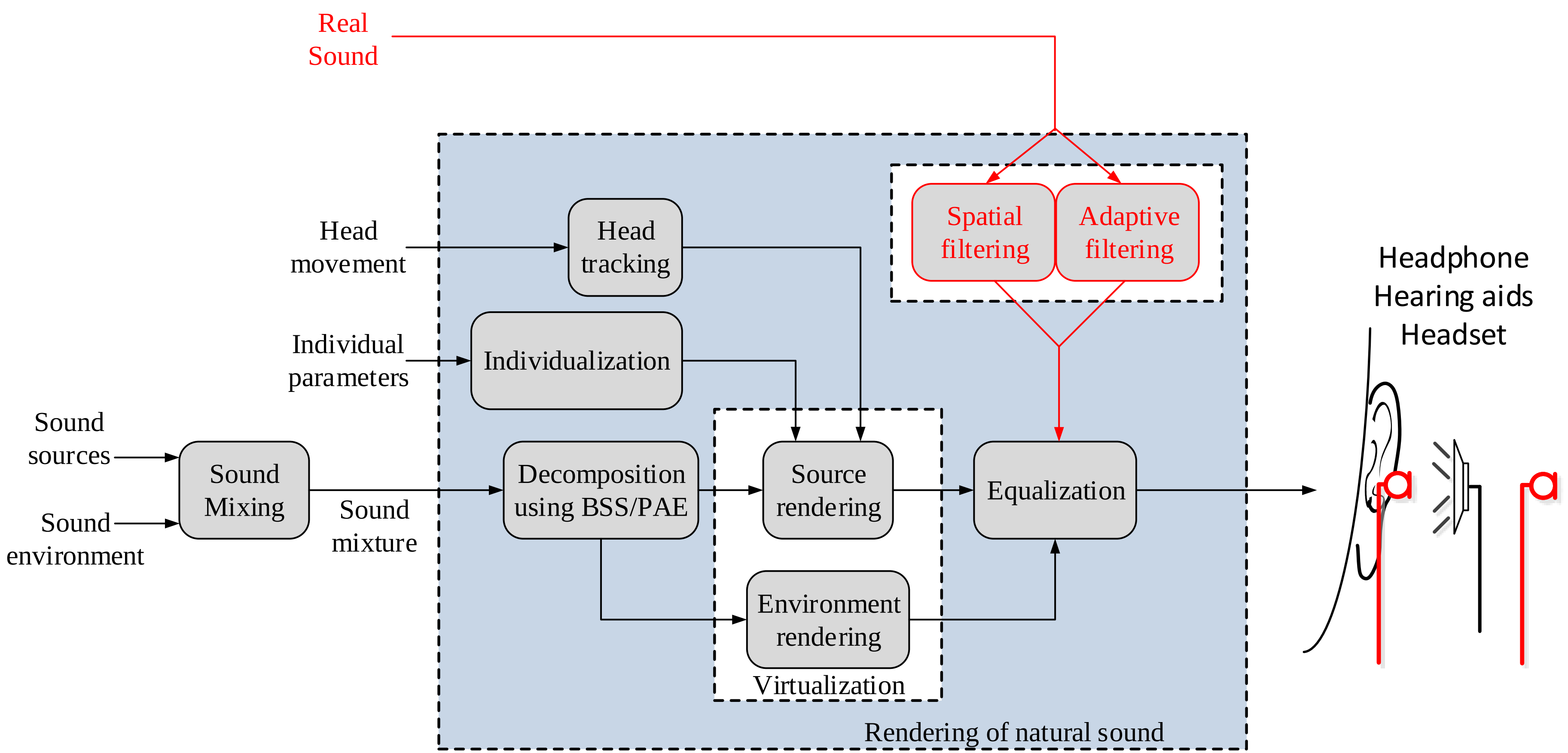
[50] D. Marquardt, V. Hohmann, and S. Doclo, “Perceptually motivated coherence preservation in multi-channel Wiener filtering based noise reduction for binaural hearing aids,” in Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing, Florence, Italy, May 2014, pp. 3688–3692.

[51] Mark Kahrs, and Karlheinz, Bradndenburg, Applications of digital audio processing to audio and acoustics, Kluwer Academic Publishers, 1998.

Module VI
Summary

Aims of Assisted Listening

Summary on Assisted Listening



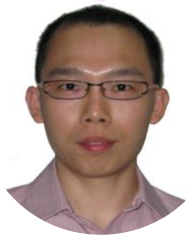
The Future of Assisted Listening

Action/ Technique	You hear		Outcome
	Virtual sound	Real sound	
Put on headphone	Normal	Modified by headphone	Normal listening
Masking	Enhanced	Modified by headphone	Better listening of virtual sound in noisy environment
ANC	Normal	Reduced	
Hear-through	Normal	Recovered	Listening to virtual sound and natural real sound
Acoustic SP	Normal	Personally modified	Listening to virtual sound and enhanced real sound
Acoustic & Audio SP	Natural	Blocked	Virtual Reality
Acoustic & Audio SP	Natural	Recovered	Augmented Reality
Acoustic & Audio SP	Natural	Selective	"Super Reality"

Acknowledgements

Mr. Rishabh Ranjan (Researcher @NTU)

Dr. Kaushik Sunder (Visiting research fellow @NASA)



Dr. Joseph Tan Ee-Leng (CTO @Beijing Sesame World)

Mr. Phyo Ko Ko (Researcher @NTU)

Mr. Nguyen Duy Hai (Researcher @NTU)

And all researchers working in assisted listening, 3D audio, active noise control, virtual reality, augmented reality!

Contact us

Dr. Woon-Seng Gan:
ewsgan@ntu.edu.sg

Mr. Jianjun He:
jhe007@e.ntu.edu.sg

DSP Lab, School of EEE,
Nanyang Technological University,
Singapore