Fast Modelling of Pinna Spectral Notches from HRTFs using Linear Prediction Residual Cepstrum

Chaitanya Ahuja and Rajesh M. Hegde
chahuja@iitk.ac.in, rhegde@iitk.ac.in
Indian Institute of Technology, Kanpur

1. Introduction

- Head Related Transfer Functions(HRTFs) are assumed as a linear-system for a given ear.
- HRTF takes into account reflection, resonance and diffraction effects due to the pinna(outer ear), head and torso.
- Accurate Individualized HRTFs are a necessity for reconstruction of accurate spatial audio. [1]
- Measuring HRTFs is a time-consuming job, hence is impractical in the industry.
- Hence, reconstruction of HRTF using ear geometry would be a good start.
- Spectral notches in HRTFs can be linked to the distance of walls of the pinna from the entry of the ear canal [4].
- A new algorithm has been proposed, namely Linear Prediction Residual Cepstrum (LPRC), which provides a more accurate way of extracting Spectral Notches

2. HRTF as an All-Pole Model

- Let $H(r,\theta,\phi,f)$ be the HRTF describing a given ear

$$H(r,\theta,\phi,f) = \frac{\psi(r,\theta,\phi,f)}{\psi_0(f)} \quad (1)$$

where $\psi(r,\theta,\phi,f)$ is sound pressure on right/left ear drum and $\psi_0(f)$ is free-field sound pressure. (r,θ,ϕ) are spherical coordinates denoting the sound source and f is frequency.

- HRTF is approximated using an all pole model as
  - we do not have access to the input sequence
  - it gives a system of equations which can be efficiently solved
- All-Pole Model in time domain is $\hat{x}[n] = \sum_{i=1}^{k} a_i x[n-i]$, where k is the order of approximation (of Linear Prediction(LP) Residual).
  - Order chosen as 12 for all the experiments demonstrated henceforth.
  - The choice of the order does not have a significant effect on the results as long as it is large enough(>8)

3. Estimating HRTF using LP Residual

- Assume n^{th} point of the minimum phase, causal signal $h[n]$ (IFFT of HRTF)is unknown.
- It is modelled as a linear combination of k previous points in the signal, k being the order of the LP Residual
- Expectation of the mean squared error $e[n]$ is minimized

$$e[n] = h[n] - \sum_{i=1}^{k} a_i h[n-i] \quad (2)$$

- LP residual analysis assumes a source filter model and estimates 3 components
  - all-pole model
  - residual, representing excitation of source of sound
  - gain, corresponding to the energy of the signal

4. Linear Prediction Residual Cepstrum

- Windowed signal is transformed using Cepstrum which is defined as

$$c_x[n_q] = \mathcal{R}\mathrm{e}\left(\mathrm{IDCT}\left(\log_{10}\left(|\mathcal{F}\{x[n]\}|\right)\right)\right) \quad (3)$$

where $\mathcal{F}$ is discrete-fourier transform, IDCT is inverse-discrete cosine transform and $\mathcal{R}\mathrm{e}$ is real part of the sequence.

- Cepstrum, by virtue of FFT followed by log function, changes convolution to addition form.
- A half-rectangular lifter eliminates convolutional components of multiple reflections.
- DCT requires fewer coefficients to better approximate the spectrum than FFT.
  - More information can be stored in fewer number of data points.

10. Conclusions

- Linear Prediction Residual Cepstrum (LPRC) proposed as a more accurate algorithm for extraction of spectral notches.
- Lesser number of coefficients are required for storing the information about spectral notches.
- As compared to LPRGD, Mean and Variance in AED of notch distances are significantly smaller for notches extracted using LPRC, which indicates better accuracy of the proposed algorithm.
- Mean of DBR is significantly larger for spectral notches, which indicates sharpness of the valleys in the spectrum.
- Analysis of Variance of the reconstructed HRTF with original HRTF indicate more statistical closeness of HRTF constructed from notches extracted using LPRC.
- Same algorithm can be modified to extract peaks, which are a result of resonance effects.
- Accurate Spectral notch techniques are an essential component for verification of spectral notches (from geometry of the ear) and on-line modelling of the pinna for synthesizing personalized spatial audio.

5. LPRC Algorithm

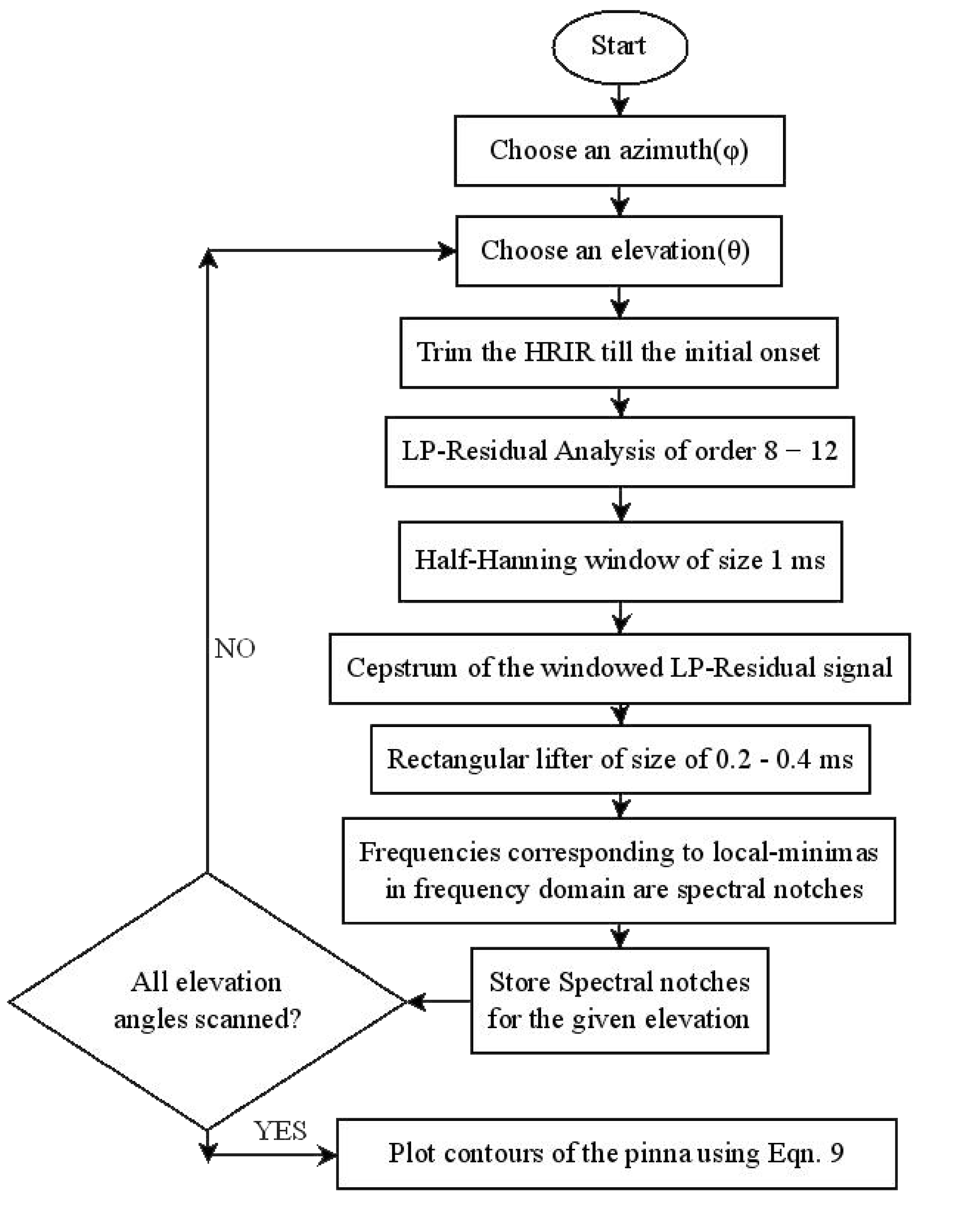


Figure 1: Flowchart of LPRC algorithm

6. Spectral Notch Extraction using LPRC

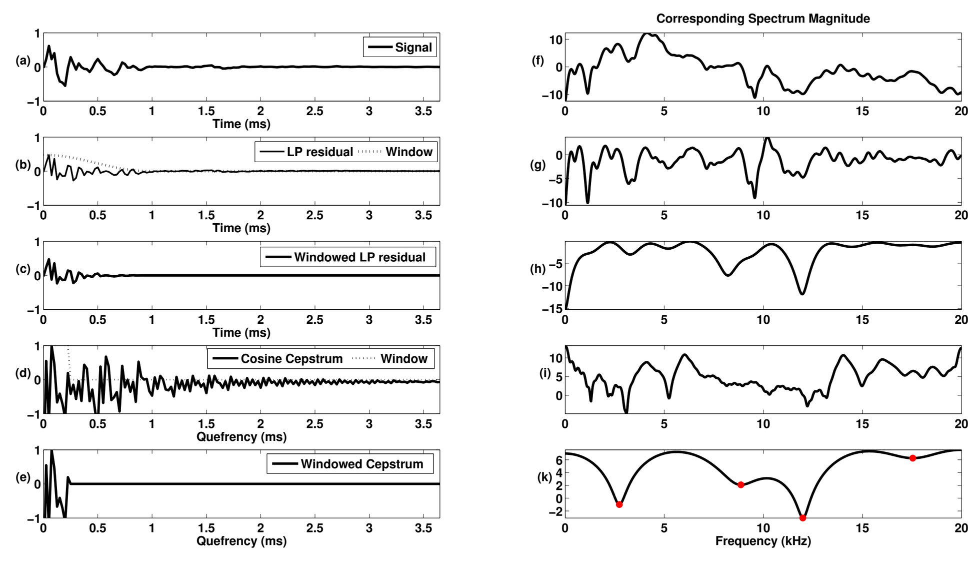


Figure 2: Application of LPRC algorithm for extracting spectral notches for $\theta = 0$ and $\phi = 0$ of subject 119, *Courtesy: CIPIC Database. Figure (a):* Original signal, *Figure (b):* LP residual of original signal, *Figure (c):* Half-hann window of previous signal, *Figure (d):* Cepstrum of windowed signal, *Figure (e):* Rectangular window of previous signal. Figure (f), (g), (h), (i), (k) refer to fourier transforms of Figure (a), (b), (c), (d), (e) respectively. Local minimas in *Figure (k)* correspond to frequencies of spectral notches.

7. Model for Pinna Contour Extraction

- Reflection Model, as described by Batteau [2] and modified by Satarzadeh [3], has been applied to overlay contour of spectral notches on the picture of pinna of a given individual.
- Let a pinna be subjected to a sound wave $x[t]$
- Total signal $y[t]$ received at the ear canal is

$$y[t] = x[t]\,\text{(Direct Signal)} + ax[t - t_d(\theta)]\,\text{(Reflected Signal)} \quad (4)$$

where a is the reflection coefficient and $t_d(\theta)$ is the time delay.

- For destructive superposition of incident and reflected waves we have $t_d(\theta)2\pi f_n(\theta) = (2n+1)\pi \quad \forall n = 0,1,2\ldots$
- For $n = 0$ and $t_d(\theta) = \frac{2d(\theta)}{c}$ we have $f_0(\theta) = \frac{1}{2t_d(\theta)} = \frac{c}{4d(\theta)}$
- Assuming reflection coefficient to be negative (Satarzadeh's argument) we get

$$f_0(\theta) = \frac{c}{2d(\theta)} \quad (5)$$

where c is the speed of sound in air, $d(\theta)$ is the path difference between reflected and direct wave, $f_0(\theta)$ is the frequency of the first spectral notch and θ is the angle of elevation.

8. Results of Pinna Contour extraction

Notches are overlaid on picture using points corresponding to $(d(\theta), \pi + \theta)$ with respect to the ear canal as the origin.

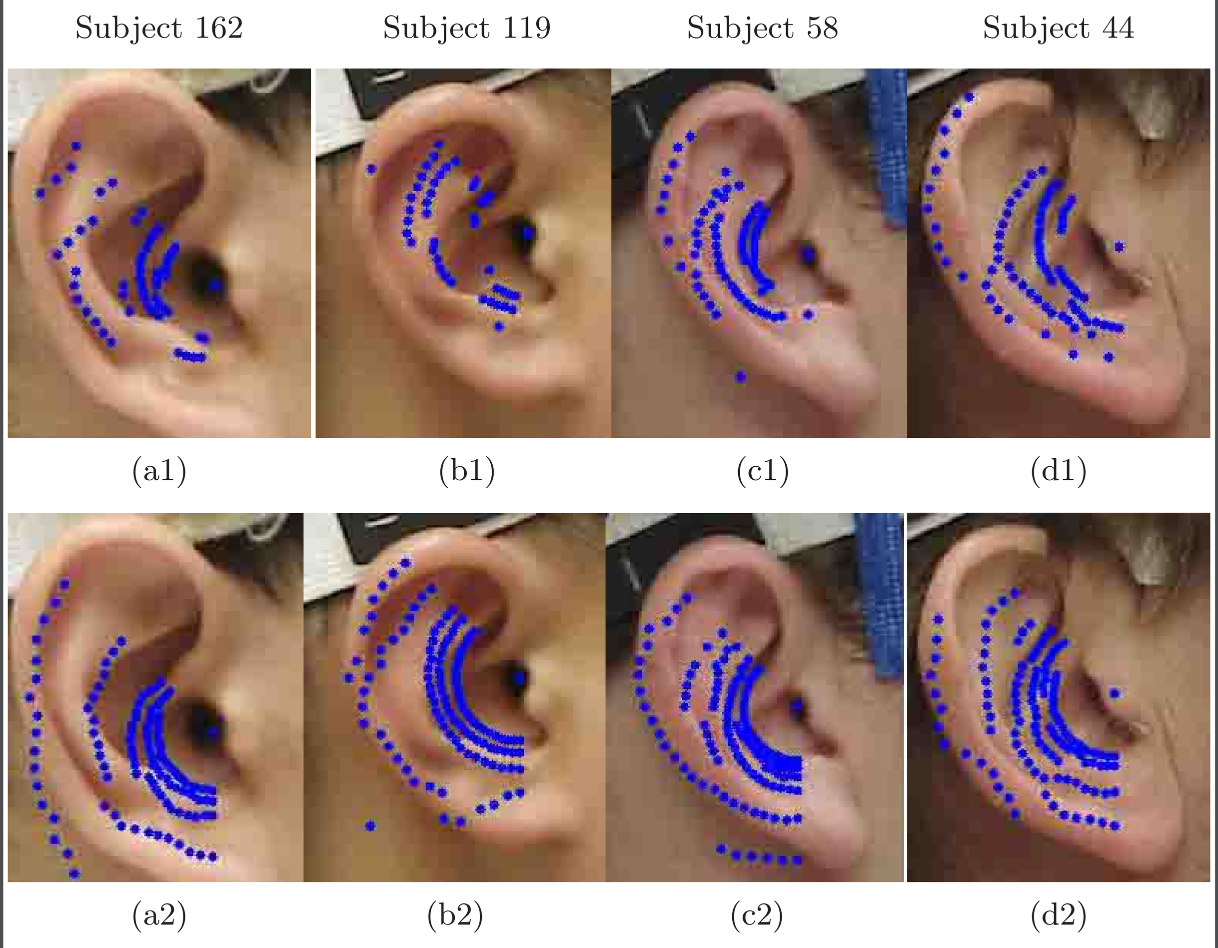


Figure 3: Illustration of pinna images with contours overlaid on them. (a1) through (d1) are generated using LPRGD algorithm [4]. (a2) through (d2) are using LPRC algorithm.

9. Performance Evaluation

9(a). Statistical Analysis

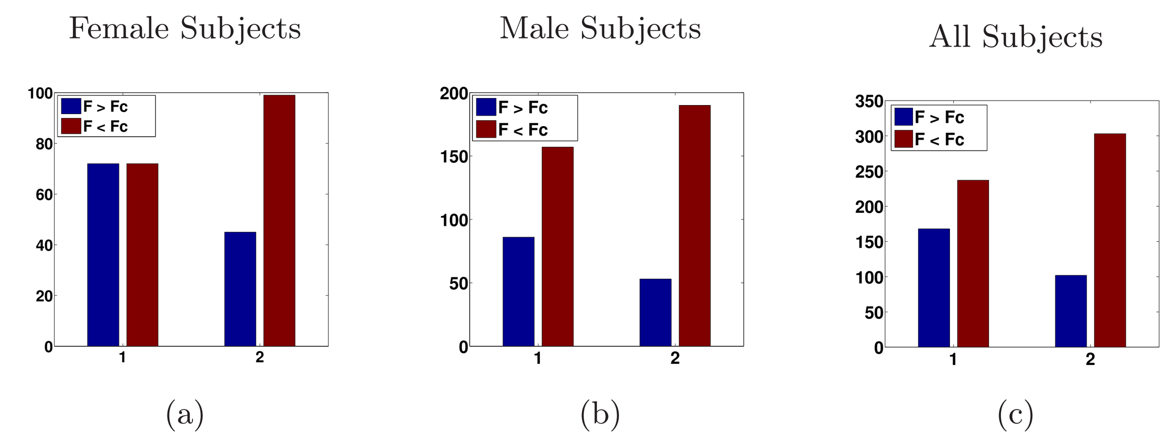
- Publicly available CIPIC Database [5] has been used as the database for testing the algorithm.
- Contours on the pinna were marked manually at discrete angles
- Using Equation 5 frequency of spectral notches were calculated
- Used as reference for calculation of deviation errors
- Average Error Deviation (AED) in notch distances and Mean and Variance of DBR was calculated separately for female and male subjects
- Depth-Bandwidth Ratio DBR $= \frac{\text{Depth}}{\text{3dB Bandwidth}}$ and notch distances were also calculated

	LPRGD			
	AED in Notch Distance		DBR	
	Mean (cm)	Variance (cm)	Mean (dB kHz^{-1})	Variance (dB kHz^{-1})
Female	0.1496	0.1474	2.7600	8.1947
Male	0.1481	0.1375	2.8900	8.5188

	LPRC			
	AED in Notch Distance		DBR	
	Mean (cm)	Variance (cm)	Mean (dB kHz^{-1})	Variance (dB kHz^{-1})
Female	0.0511	0.0848	8.9097	1746.6
Male	0.0349	0.0701	9.9529	1507.0

9(b). Analysis of Variance (ANOVA)

- Frequencies of extracted notches used to synthesise an all-pole filter of fixed bandwidth
- This filter is excited by an impulse train to generate HRIR
- Synthesized HRTF is compared to original spectrum using Analysis of Variance(ANOVA) F-Test
  - Sensitivity $= 5\%$
  - Degrees of freedom of numerator, $n_f = 1$ and denominator, $d_f = 1000$
  - This implies $F_c = 3.85$
  - F-stat values are calculated for all such comparisons of HRTF and plot as a frequency chart
  - Null-Hypothesis is rejected when $F > F_c$



- Bar 1 and Bar 2 represent analysis on LPRGD and LPRC respectively
- Clearly Null-Hypothesis is rejected more prominently in female subjects when using LPRGD Algorithm
- Hence LPRC gives more accurate notches than LPRGD

11. References

[1] Toni Liitola. *Headphone sound externalization.* PhD thesis, Helsinki University of Technology, 2006.

[2] D. W. Batteau. The role of the pinna in human localization. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 168, No.1011:158–180, August 1967.

[3] Patrick Satarzadeh. *A study of physical and circuit models of the human pinnae.* PhD thesis, Citeseer, 2006.

[4] Vikas C. Raykar, Ramani Duraiswami, and B. Yegnanarayana. Extracting the frequencies of the pinna spectral notches in measured head related impulse responses. *The Journal of the Acoustical Society of America*, 118(1):364–374, 2005.

[5] V.R. Algazi, R.O. Duda, D.M. Thompson, and C. Avendano. The cipic hrtf database. In *Applications of Signal Processing to Audio and Acoustics, 2001 IEEE Workshop on the*, pages 99–102, 2001.