# ON THE MANIFOLDS OF SPATIAL HEARING

VIKAS C. RAYKAR AND RAMANI DURAISWAMI

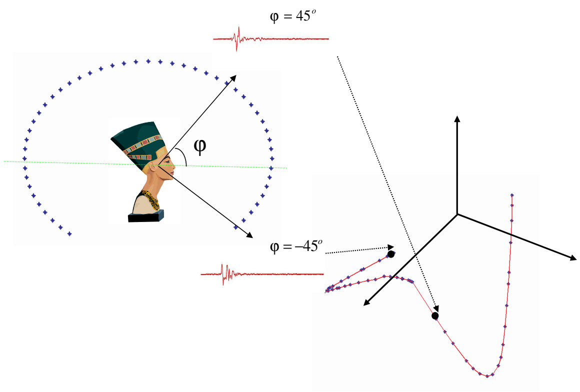


FIGURE 1. Conceptual diagram of a one-dimensional manifold embedded in a higher dimensional space.

Humans have an amazing ability to localize a sound source, *i.e.*, determine the range, elevation, and azimuth angles of the direction of the sound source. The major mechanisms responsible for the directional capability of the human hearing system has been fairly well understood though not completely [2]. One of the primary cues responsible for localization of the sound source is the Interaural Time Difference(ITD) and the Interaural Level Difference(ILD). However ITD and ILD cues alone do not completely explain the source localization mechanism. For example, for all points lying in the hyperboloid of revolution with vertex as the center of the head and passing through a point, the ITD and ILD cues are essentially same. Yet we have the ability to localize the sound source in the vertical plane. This is because there are additional important static and dynamic acoustic cues that arise from the scattering of the sound by the head, torso, and the pinnae. This can be explained in terms of the spectral filtering provided by the torso, head, and the pinnae. This filtering can be described by a complex frequency response function called the Head Related Transfer Function (HRTF). The corresponding impulse response is called the Head Related Impulse Response (HRIR). The spectral features in the HRTF due to pinna diffraction and scattering seem to provide cues for vertical localization, *i.e.*, elevation of the source. By manipulating the cues responsible for the directional hearing capability a virtual audio system which can place the sound to any given location can be built by using just a pair of headphones or only two loud speakers. For a given person the HRTF can be experimentally measured. Figure 2 shows a typical HRIR and the magnitude of the HRTF for both the left and the right ear.

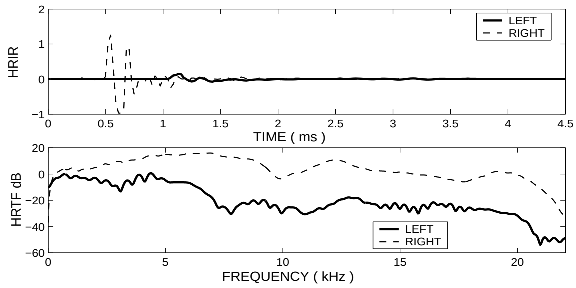


FIGURE 2. HRIR and HRTF for the left and the right ear for a typical person when the source is directly in front of the right ear at a distance of 1m from the center of the head

A HRIR of $N$ samples can be considered as a point in $N$ dimensional space. Consider all the HRIRs in the mid-sagittal plane as shown in Figure 1. Each HRIR corresponding to one elevation is a point in the higher dimensional space. As the elevation is varied smoothly, the points essentially trace out a one-dimensional manifold as shown in Figure 1. Manifolds arise naturally whenever there is a smooth variation of parameters, like the elevation angle in our case. Manifolds encode the perceptual information in a given signal. For all the HRIRs in median plane the dominant perceptual information is the elevation of the source. In the $N$ dimensional Euclidean space of the original HRIRs, two HRIRs corresponding to far apart elevations may be very close to each other. However on the one-dimensional manifold, where we measure the distance between two points as the length of the geodesic on the manifold, they will be far apart. So if we can unfold this low-dimensional manifold we have a good perceptual representation of the signal.

In this talk we will present some exploratory studies using Isomap, Locally Linear Embedding (LLE), and Maximum Variance Unfolding (MVU) to learn the non-linear manifold structure, in the HRTFs, for a given individual. The lower dimensional manifold encodes the perceptual information in the HRIRs, namely the direction of the sound source. Based on this we propose a new method for HRIR interpolation. We also propose a new distance metric between two HRIRs as the geodesic distance on the manifold. This work was first presented in [1]. We think this is an interesting dataset for the attention of the NIPS audience. We will also discuss the current important problems in the field of spatial audio and how techniques from machine learning could be useful.

## REFERENCES

1. R. Duraiswami and V. C. Raykar, *The manifolds of spatial hearing*, In Proceedings of International Conference on Acoustics, Speech and Signal Processing (ICASSP 2005), vol. III, 2005, pp. 285–288.
2. J. C. Middlebrooks and D. M. Green, *Sound localization by human listeners*, Annual Review of Psychology **42** (1991), 135–159.

PERCEPTUAL INTERFACES AND REALITY LAB, UNIVERSITY OF MARYLAND, COLLEGE PARK
*E-mail address*: {vikas,ramani}@umiacs.umd.edu