Chapter 11
Comparison of HRTF Databases

Abstract Several research institutes have released databases of HRTFs. This chapter introduces representative databases and compares them from the viewpoints of spectral cues and pinna shape.

11.1 Representative HRTF Database

Table 11.1 shows representative publically available sites of HRTF databases.

In this chapter, the databases of the following five research institutes are compared.

1). Acoustics Research Institute (ARI), Austria
2). Center for Image Processing and Integrated Computing Interface Laboratory (CIPIC), U.S.A.
3). Spatial Hearing Laboratory (SHL), Chiba Institute of Technology, Japan
4). Institut de Recherche et Coordination Acoustique/Musique (IRCAM), France
5). Research Institute of Electrical Communication (RIEC), Tohoku University

An outline of these databases is shown in Table 11.2 (Yan et al. 2014). All research institutes measured HRTFs under the blocked-entrance condition (Shaw and Teranishi 1968).

The minimum number of subjects is 45 (CIPIC) and the maximum number of subjects is 105 (RIEC). In the four research institutes other than IRCAM, an array in which multiple loudspeakers are arranged in the vertical direction is installed, and the HRIRs for various three-dimensional directions were measured by rotating the subject or the array in the horizontal direction. At IRCAM, HRIRs were measured by moving one loudspeaker in the vertical direction and then rotating the subject in the horizontal direction.

There are large differences in the lengths of the measured HRIRs, which are 200 samples in CIPIC and 8192 samples in IRCAM. The number of measurement directions is seven to 148 at SHL, 1250 at CIPIC, 1550 at ARI, 187 at IRCAM, and 865 at RIEC.

K. Iida, *Head-Related Transfer Function and Acoustic Virtual Reality*,
https://doi.org/10.1007/978-981-13-9745-5_11

Table 11.1 Publically available sites of HRTF databases

Institute	Country	URL
ARI	Austria	https://www.kfs.oeaw.ac.at/index.php?lang=en
IRCAM	France	http://recherche.ircam.fr/equipes/salles/listen/
CIPIC	USA	http://interface.cipic.ucdavis.edu/sound/hrtf.html
MIT	USA	http://sound.media.mit.edu/resources/KEMAR.html
SHL	Japan	http://www.iida-lab.it-chiba.ac.jp/HRTF/
RIEC	Japan	http://www.ais.riec.tohoku.ac.jp/lab/db-hrtf/index-j.html
Nagoya	Japan	http://www.sp.m.is.nagoya-u.ac.jp/HRTF/index-j.html
ITA	Germany	http://gershwin.akustik.rwth-aachen.de/hrtf/hrtf-lic.php

Table 11.2 Outline of the five HRTF databases considered herein. (Yan et al. 2014)

	ARI	CIPIC	SHL	IRCAM	RIEC
Subject number	82	45	61	50	105
Source signal	ML sequence	MESM	swept–sine	OATSP	log sweep
Data length	256	200	512	8192	512
Sampling frequency (Hz)	48,000	44,100	48,000	44,100	48,000
Number of directions	1550	1250	7–148	187	865
Microphone	KE–4–211–2	ER–7C	WM64AT102	FG3329	FG3329
Manufacturer	Sennhiser	Etymtic	Panasonic	Knowles	Knowles
Loudspeaker	10 BGS	Acoustimass™	FE83E	system600	FE83E
Manufacturer	VIFA	Bose	Fostex	TANNOY	Fostex
Number of pieces	22	5	7	1	35
Data format	mat	mat	bin	mat	SOFA

11.2 Comparison of Spectral Cues

Next, N1, N2, and P1 frequency are compared between databases. The N1, N2 and P1 frequencies were calculated from the HRIRs for the front direction in the five databases using the method described in Sect. 10.2. Their histograms are shown in Fig. 11.1 (Yan et al. 2014).

The histograms for each database are approximately normally distributed. However, the peaks of the ARI histograms are at a higher frequency than the other histograms.

The average, minimum, and maximum frequencies of N1, N2 and P1 for the front direction of each database are shown in Table 11.3. Comparing the databases, the average frequency of RIEC is the lowest, and that of ARI is the highest for N1 and P1. For N2, the average frequency of SHL is the lowest, and that of ARI is the highest.

As such, the N1, N2, and P1 frequencies in Japanese databases are low, and those in ARI are high, as compared to the other databases.

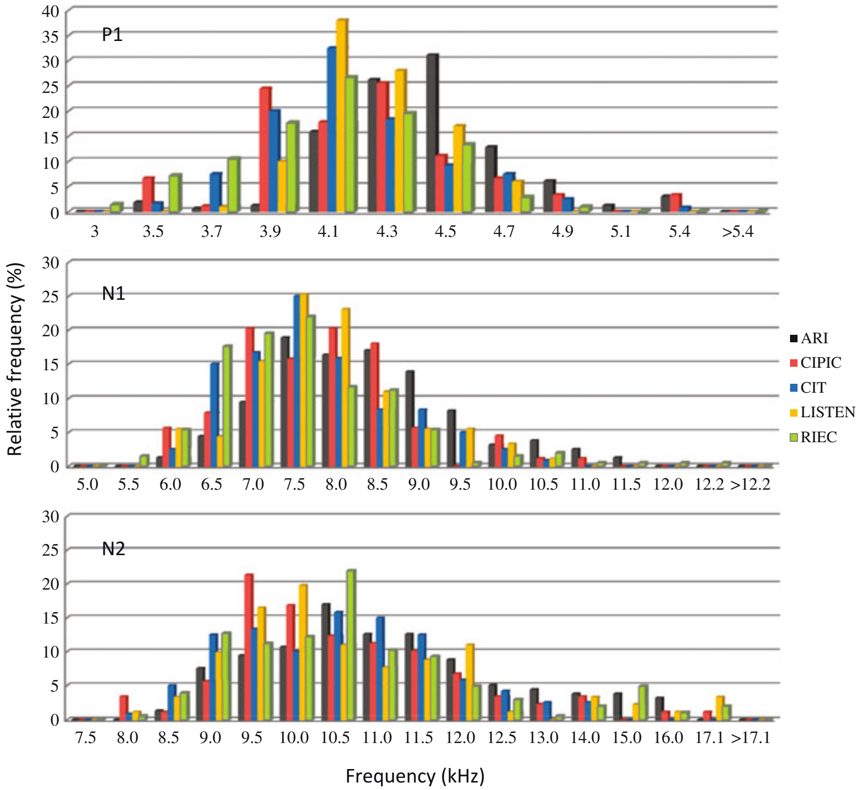

Fig. 11.1 Histograms of P1, N1, and N2 frequencies of HRTFs for front direction for five HRTF databases. (Yan et al. 2014)

Table 11.3 Average, minimum, and maximum frequencies of N1, N2, and P1 for the front direction for five database (Hz). (Yan et al. 2014)

	ARI	CIPIC	SHL	IRCAM	RIEC
P1 Ave	4333	4095	4059	4131	3969
P1 Min	3281	3187	3469	3618	2438
P1 Max	5250	5340	5250	4651	4875
N1 Ave	8101	7545	7481	7585	7301
N1 Min	6000	5771	5531	5685	5063
N1 Max	11,250	10,939	10,031	10,164	12,188
N2 Ave	10,959	10,384	10,287	10,519	10,549
N2 Min	8063	7494	7781	7752	7688
N2 Max	15,938	16,107	13,500	16,882	17,063

Table 11.4 Results of statistical tests for N1, N2, and P1 frequency. (Yan et al. 2014)

N1	ARI	CIPIC	SHL	IRCAM	RIEC
CIPIC	**	–			
CIT	**		–		
LISTEN	**			–	
RIEC	**			*	–
N2	ARI	CIPIC	SHL	IRCAM	RIEC
CIPIC	**	–			
CIT	**		–		
LISTEN	**			–	
RIEC	**				–
P1	ARI	CIPIC	SHL	IRCAM	RIEC
CIPIC	**	–			
CIT	**		–		
LISTEN	**		*	–	
RIEC	**	**	*	**	–

** $p < 0.01$; * $p < 0.05$

Furthermore, statistical tests were performed in order to verify whether there exists a significant difference in the average frequencies among the databases. The results are shown in Table 11.4.

The average frequency of ARI was significantly higher for N1, N2, and P1 compared to the other four databases ($p < 0.01$). For P1, the average frequency of RIEC is significantly lower compared with the other four databases ($p < 0.05$).

11.3 Comparison of Pinna Shape

As mentioned in Chap. 3, since N1, N2, and P1 are caused by the resonance of the cavities in the pinna, the differences in N1, N2, and P1 frequencies among the databases are related to the differences in pinna size.

Among the five HRTF databases, for which the N1, N2, and P1 frequencies were analyzed, detailed pinna anthropometric dimension data for the subjects are available for ARI, CIPIC, and SHL. The number of subjects for ARI, CIPIC, and SHL are 40 (80 ears), 34 (74 ears), and 28 (56 ears), respectively.

Histograms of each pinna anthropometric dimension are shown in Fig. 11.2, and their statistics are shown in Table 11.5. The average pinna anthropometric dimension of SHL is larger than those of ARI and CIPIC, except for x7.

Statistical tests were performed in order to verify whether there exists a significant difference in the average pinna anthropometric dimensions among the databases. The results are shown in Table 11.6. Significant differences ($p < 0.05$) were observed

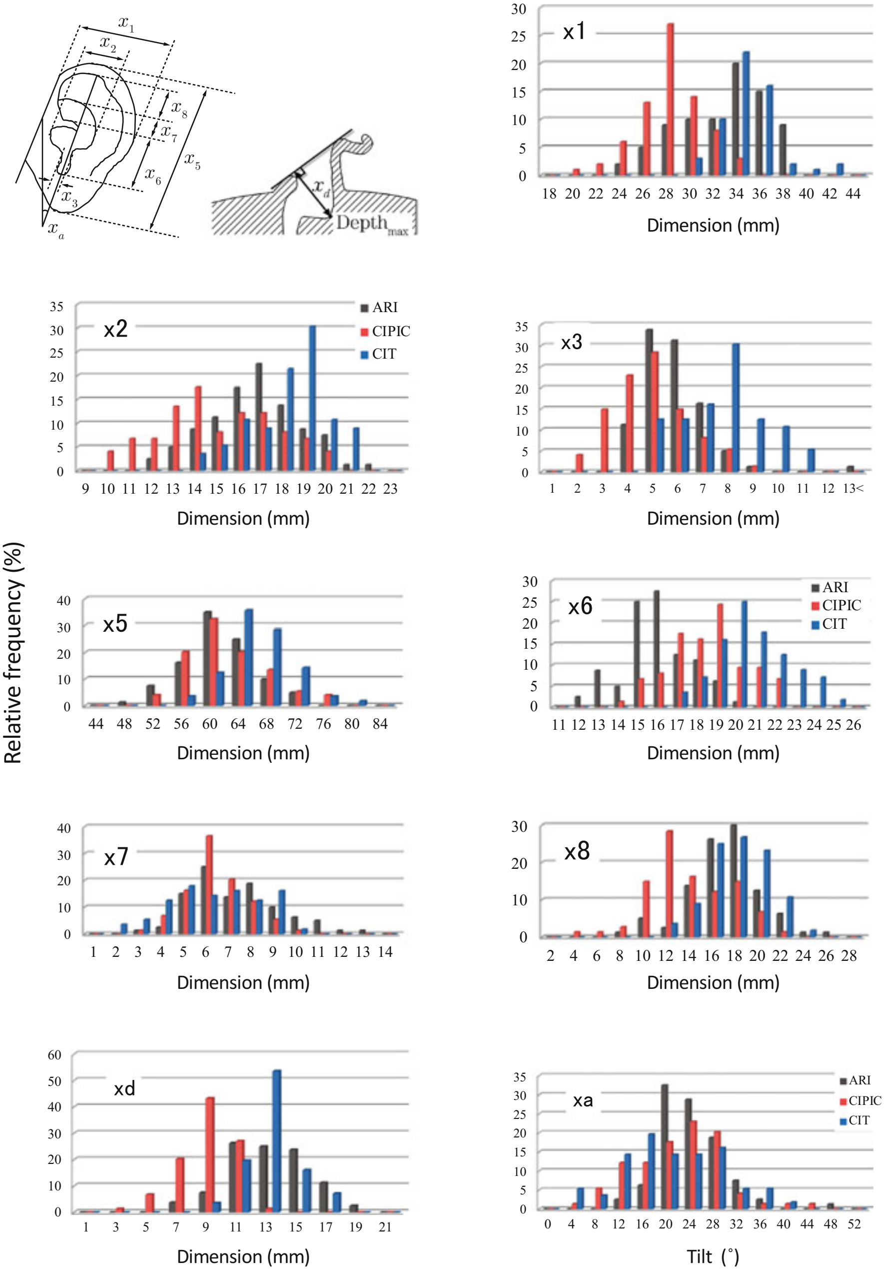

Fig. 11.2 Histograms of pinna anthropometric dimensions. (Yan et al. 2014)

Table 11.5 Statistics for pinna anthropometric dimensions (mm). (Yan et al. 2014)

		x_1	x_2	x_3	x_5	x_6	x_7	x_8	x_d	$x_a[°]$
Mean	ARI	33.48	16.90	6.20	62.43	16.22	7.51	17.72	13.71	25.48
	CIPIC	29.05	15.45	5.40	63.96	18.87	6.83	14.80	9.76	23.29
	SHL	35.48	18.68	8.19	68.26	21.15	6.69	18.92	13.93	21.86
Min	ARI	25.00	12.10	4.00	48.00	12.10	3.80	9.00	7.00	12.00
	CIPIC	21.84	10.37	2.70	54.24	14.32	3.78	5.81	3.65	5.94
	SHL	31.22	14.78	5.34	58.23	17.72	2.58	13.24	9.71	4.00
Max	ARI	39.20	22.00	18.00	74.00	20.00	13.20	26.70	19.20	49.00
	CIPIC	35.34	20.97	9.11	79.55	22.94	10.46	22.37	13.11	44.48
	SHL	43.83	21.84	11.88	83.18	25.09	10.28	24.14	17.60	40.00
Standard deviation	ARI	3.60	2.06	1.67	5.00	1.70	1.99	3.07	2.67	5.57
	CIPIC	2.74	2.58	1.51	5.58	1.96	1.32	3.51	1.79	7.60
	SHL	2.36	1.71	1.65	4.66	1.80	1.99	2.50	1.74	8.37

Table 11.6 Results of statistical tests for pinna anthropometric dimensions. ∗: $p < 0.05$, ∗∗: $p < 0.01$ (Yan et al. 2014)

Pinna anthropometry	Comparison between ARI and CIPIC	ARI and SHL	CIPIC and SHL	Size relationship				
x_1	∗∗	∗∗	∗∗	CIPIC	<	ARI	<	SHL
x_2	∗∗	∗∗	∗∗	CIPIC	<	ARI	<	SHL
x_3	∗∗	∗∗	∗∗	CIPIC	<	ARI	<	SHL
x_5		∗∗	∗∗	ARI	≅	CIPIC	<	SHL
x_6	∗∗	∗∗	∗∗	ARI	<	CIPIC	<	SHL
x_7	∗	∗		SHL	≅	CIPIC	<	ARI
x_8	∗∗	∗	∗∗	CIPIC	<	ARI	<	SHL
x_d	∗∗		∗∗	CIPIC	<	ARI	≅	SHL
x_a	∗	∗∗		SHL	≅	CIPIC	<	ARI

for almost all combinations. In other words, there exists a difference in ear size among the databases.

In the previous section, we found that N1, N2, and P1 frequencies were higher for ARI than for the other four databases. Here, let us consider the reasons. Since N1, N2, and P1 are generated by resonance in the pinna cavities, it is inferred that the pinna dimensions of ARI are smaller than those in other databases. The pinna anthropometric parameter for which the average dimension of ARI was statistically significantly smaller than in other databases was x6 (length of cavity of concha) ($p < 0.01$). As described in Sect. 3.5, the x6 dimension is shown to have a significant effect on the N1, N2, and P1 frequencies. In other words, the fact that x6 is smaller than other databases is considered to be one of the reasons why the N1, N2, and P1 frequencies for ARI are higher than for other databases.

References

Shaw EAG, Teranishi R (1968) Sound pressure generated in an external-ear replica and real human ears by a nearby point source. J Acoust Soc Am 4:240–249

Yan X, Iida K, Ishii Y (2014) Comparison in frequencies of spectral peaks and notches and anthropometric of pinnae between HRTH databases. IEICE Technical Report, EA2014–19. pp. 43–48