



UNIVERSITAS INDONESIA

**INDIVIDUALISASI MODEL PARAMETRIK
*HEAD-RELATED TRANSFER FUNCTIONS***

DISERTASI

**Hugeng
0706221136**

**FAKULTAS TEKNIK
PROGRAM STUDI TEKNIK ELEKTRO
UNIVERSITAS INDONESIA
DEPOK
JANUARI 2011**



UNIVERSITAS INDONESIA

**INDIVIDUALISASI MODEL PARAMETRIK
*HEAD-RELATED TRANSFER FUNCTIONS***

DISERTASI

Diajukan sebagai salah satu syarat untuk memperoleh gelar Doktor Teknik

**Hugeng
0706221136**

**FAKULTAS TEKNIK
PROGRAM STUDI TEKNIK ELEKTRO
UNIVERSITAS INDONESIA
DEPOK
JANUARI 2011**

HALAMAN PERNYATAAN ORISINALITAS

**Disertasi ini adalah hasil karya saya sendiri,
dan semua sumber baik yang dikutip maupun dirujuk
telah saya nyatakan dengan benar.**



Nama : Hugeng
NPM : 0706221136
Tanda Tangan : 
Tanggal : 13 Januari 2011

HALAMAN PENGESAHAN

Disertasi ini diajukan oleh :
N a m a : Hugeng
NPM : 0706221136
Judul Disertasi : Individualisasi Model Parametrik
Head-Related Transfer Functions

Telah berhasil dipertahankan di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Doktor pada Program Studi Teknik Elektro, Fakultas Teknik, Universitas Indonesia.

DEWAN PENGUJI

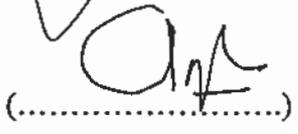
Promotor : Prof. Dr. Ir. Dadang Gunawan, M.Eng. (.....)

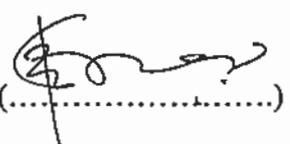
Kopromotor : Ir. Wahidin Wahab, M.Sc., Ph.D. (.....)

Tim Penguji : Prof. Dr.-Ing. Kalamullah Ramli, M.Eng. (.....)

Dr. Ir. Muhamad Asvial, M.Eng. (.....)

Ir. Gunawan Wibisono, M.Sc., Ph.D. (.....)

Dr. Ir. Arman Djohan Diponegoro (.....)

Prof. Dr. Ir. Bambang Riyanto T., M.Eng. (.....)

Ditetapkan di : Depok

Tanggal : Januari 2011

KATA PENGANTAR

Puji dan syukur saya panjatkan kepada Tuhan Yesus Kristus, atas limpahan hikmat dan anugerahNya sehingga disertasi ini dapat diselesaikan. Penulisan disertasi ini dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar Doktor Teknik pada Departemen Teknik Elektro, Fakultas Teknik, Universitas Indonesia. Selama penulisan disertasi ini penulis telah mendapatkan bantuan dan bimbingan banyak pihak. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Prof. Dr. Ir. Dadang Gunawan, M.Eng., selaku Promotor yang telah menyediakan waktu, tenaga dan pikiran selama penyusunan disertasi ini;
2. Ir. Wahidin Wahab, M.Sc., Ph.D., selaku Ko-Promotor atas masukan dan diskusi yang sangat berharga selama penyusunan disertasi ini;
3. Para anggota Dewan Penguji: Prof. Dr. Ir. Bambang Riyanto T., M.Eng., Prof. Dr.-Ing. Kalamullah Ramli, M.Eng., Dr.Ir. Muhamad Asvial, M.Eng., Ir. Gunawan Wibisono, M.Sc., Ph.D., dan Dr. Ir. Arman D. Diponegoro atas saran dan masukannya untuk memperbaiki disertasi ini;
4. Pimpinan penulis di Universitas Trisakti, Jakarta dan rekan-rekan kerja baik staf akademik maupun non-akademik atas bantuan dan dukungan penuh selama penulis menyelesaikan studi;
5. Isteri, ketiga anak, orang tua, serta keluarga besar penulis, yang tanpa dukungannya disertasi ini tidak akan dapat diselesaikan.

Akhir kata, kiranya Tuhan Yesus Kristus sendiri berkenan membalas kebaikan semua pihak yang telah membantu. Semoga disertasi ini membawa manfaat bagi pengembangan ilmu.

Depok, Januari 2011

Penulis

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Universitas Indonesia, saya yang bertanda tangan di bawah ini:

Nama : Hugeng
NPM : 0706221136
Program Studi : Teknik Elektro
Fakultas : Teknik
Jenis Karya : Disertasi

demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Indonesia **Hak Bebas Royalti Noneksklusif (*Non-exclusive Royalty Free Right*)** atas karya ilmiah saya yang berjudul:

Individualisasi Model Parametrik *Head-Related Transfer Functions*

Dengan Hak Bebas Royalti Noneksklusif ini Universitas Indonesia berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat dan memublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Depok
Pada tanggal : Januari 2011

Yang menyatakan



(Hugeng)

ABSTRAK

Nama : Hugeng
Program Studi : Teknik Elektro
Judul : Individualisasi Model Parametrik
Head-Related Transfer Functions

Sintesis sistem audio spasial ditemukan dalam banyak aplikasi krusial, seperti dalam pesawat terbang militer maupun komersial, pengembangan *virtual reality*, termasuk di dalamnya aplikasi dalam permainan-permainan elektronik, maupun sebagai *consumer electronics* yang memberikan efek bunyi spasial dan kenyamanan bagi pengguna dalam mendengarkan musik. Hampir semua sistem yang mengaplikasikan sintesis audio spasial atau tiga dimensi, memerlukan *Head-Related Impulse Response* (HRIR). HRIR merupakan fungsi alih dalam kawasan waktu dari gendang telinga manusia yang melakukan penyaringan bunyi yang datang pada gendang telinga tersebut.

Pemodelan HRIR dengan parameter-parameter yang dapat diatur sesuai dengan antropometri pendengar merupakan salah satu solusi dari masalah yang timbul akibat menggunakan HRIR yang non-individual, seperti bunyi terdengar tidak seperti aslinya, serta bunyi terdengar di dalam kepala jika digunakan *headphones*. Hasil yang sangat bagus diperoleh jika digunakan HRIR individual milik subyek itu sendiri. Tetapi pengukuran HRIR individual ini memerlukan waktu, biaya dan energi yang tidak sedikit.

Di dalam penelitian ini, dimodelkan HRIR fasa minimum, di kawasan waktu, dan *Head-related Transfer Function* (HRTF) *magnitude*, di kawasan frekuensi, dengan menggunakan metode statistik *Principal Components Analysis* (PCA) dengan 10 vektor basis, untuk sumber-sumber bunyi di bidang horisontal dan di bidang median. Sedangkan individualisasi dari model HRIR dan model HRTF tersebut dilakukan berdasarkan *Multiple Linear Regression* (MLR) dari bobot-bobot vektor basis PCA yang dapat diatur, sesuai dengan 8 ukuran antropometris terseleksi. Seleksi 8 dari total 27 ukuran antropometris dilakukan berdasarkan analisis korelasi antara parameter-parameter psikoakustik penting dan bobot-bobot vektor basis PCA terhadap 27 ukuran antropometris.

Individualisasi model HRIR fasa minimum menghasilkan *mean-square error* (MSE) rata-rata yang cukup memuaskan sebesar 22,5% di bidang horisontal dan 20,42% di bidang median. Di kawasan frekuensi, dihasilkan MSE rata-rata lebih kecil untuk individualisasi model HRTF *magnitude*, yaitu 12,17% di bidang horisontal dan 11,21% di bidang median. Model-model HRIR dan HRTF individual dapat mengaproksimasi dengan baik pola-pola dan struktur-struktur detail respon-respon HRIR dan HRTF asli yang bersangkutan. Tes pendengaran dalam menentukan posisi-posisi sumber bunyi di bidang horisontal menghasilkan persentasi kebenaran yang lebih besar jika digunakan model-model HRIR individual daripada jika digunakan HRIR non-individual.

Kata Kunci:

audio spasial, HRTF, HRIR, PCA, MLR, individualisasi HRTF

ABSTRACT

Name : Hugeng
Study Program: Electrical Engineering
Judul : Individualization of Parametric Model of
Head-Related Transfer Functions

Syntheses of spatial audio systems are found in many crucial applications, such as both in military and commercial aeroplanes, development of virtual reality, electronics games, and consumer electronics that have given spatial effects on sound so that the listeners are pleasant in hearing their music. Most of the systems that synthesize spatial audio need head-related impulse responses (HRIRs) of human ears. HRIRs are impulse responses of human's eardrum, in time domain, that are responsible in filtering incident sound to that eardrum.

Modeling HRIRs with tunable parameters that are suitable to anthropometric of a listener is a solution to emerging problems by using non-individualized HRIRs, e.g. sound heard as if it were in head when using headphones and sound heard is not natural. The best results are obtained when the listener uses his/her own individual HRIRs. However, measurements of individual HRIRs require a huge amount of time, cost, and energy.

In this dissertation, minimum phase HRIRs were modeled in time domain and magnitude HRTFs were modeled in frequency domain, by using the statistical method named Principal Components Analysis (PCA) with 10 basis vectors, for sound sources on horizontal plane and on median plane. The individualization method used for the models of minimum phase HRIRs and the models of magnitude HRTFs, was based on Multiple Linear Regression (MLR) between weights of basis vectors obtained from PCA, that could be adjusted due to 8 selected anthropometric measurements. The selection process of 8 out of all 27 anthropometric measurements was performed based on correlations analysis between psychoacoustically crucial parameters; weights of basis vectors of PCA; and 27 anthropometric measurements.

Individualization of models of minimum phase HRIRs resulted in quite satisfying average *mean-square error* (MSE) of 22,5% for sources on horizontal plane and of 20,42% on median plane. In frequency domain, average MSE obtained from individualization of models of magnitude HRTFs was much better, that was 12,17% on horizontal plane and 11,21% on median plane. Individualized HRIRs models and individualized HRTFs models were able to approximate very well the patterns and detail structures found in respective original HRIRs and HRTFs. Listening tests performed for sound sources on horizontal plane resulted in that larger correct rates were obtained, if models of individualized HRIRs were applied, than if models of non-individualized HRIRs were applied.

Keywords:

spatial audio, HRTF, HRIR, PCA, MLR, individualization of HRTF

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERNYATAAN ORISINALITAS	iii
HALAMAN PENGESAHAN	iv
KATA PENGANTAR	v
HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH	vi
ABSTRAK	vii
DAFTAR ISI	ix
DAFTAR GAMBAR	xii
DAFTAR TABEL	xv
DAFTAR SINGKATAN	xvi
DAFTAR NOTASI	xviii
BAB 1. PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Penelitian-penelitian Pemodelan HRTF dan Individualisasi Model HRTF	7
1.3 Identifikasi Masalah	22
1.4 Batasan Masalah	23
1.5 Tujuan Penelitian	24
1.6 Ruang Lingkup Penelitian	24
1.7 Kontribusi Penelitian	28
1.8 Sistematika Penulisan	31
BAB 2. PSIKOAKUSTIK DARI PENDENGARAN SPASIAL	33
2.1 Sifat-sifat Fisik dari Bunyi	33
2.2 Sistem Koordinat yang Digunakan	35
2.3 <i>Azimuth Cues</i>	37
2.4 <i>Elevation Cues</i>	41
2.5 <i>Range Cues</i>	42
2.6 Reverberasi dan Gema-gema	44
2.7 <i>Head-Related Impulse Response</i>	45
2.8 Basis Data yang Digunakan dalam Pemodelan HRTF	49
2.9 Visualisasi dan Analisis Struktur Data HRTF	52
2.9.1 Representasi Data HRTF di Kawasan Frekuensi	53
2.9.2 Representasi Data HRIR di Kawasan Waktu	57
BAB 3. PEMODELAN HEAD-RELATED TRANSFER FUNCTIONS BERDASARKAN PRINCIPAL COMPONENTS ANALYSIS	61
3.1 Jenis-jenis Pengolahan Awal Data HRIR	61
3.2 Aproksimasi Fasa Minimum dari HRTF	63
3.2.1 <i>HRTF Coherence</i>	65
3.2.2 <i>HRTF Group Delay</i>	69
3.3 Pemodelan HRTF Menggunakan PCA	73
3.4 Hasil-hasil Pemodelan PCA dari Berbagai Jenis Data HRIR	75

3.4.1 Analisis Vektor-vektor Basis PCA dari Model HRIR Fasa Minimum	77
3.4.2 Analisis Kesalahan Model HRIR Fasa Minimum	79
3.5 Hasil-hasil Pemodelan PCA dari Berbagai Jenis Data HRTF	81
3.5.1 Analisis Vektor-vektor Basis PCA dari Model HRTF <i>Magnitude</i>	83
3.5.2 Analisis Kesalahan Model HRTF <i>Magnitude</i>	84
BAB 4. METODE SELEKSI UKURAN-UKURAN ANTROPOMETRIS	87
4.1 Analisis Korelasi ITD Maksimum Terhadap Antropometri	89
4.2 Analisis Korelasi ILD Maksimum Terhadap Antropometri	91
4.3 Analisis Korelasi Frekuensi Lembah Telinga Terhadap Antropometri	92
4.4 Analisis Korelasi Bobot-Bobot Vektor Basis Terhadap Antropometri	94
BAB 5. INDIVIDUALISASI MODEL <i>HEAD-RELATED TRANSFER FUNCTIONS</i> DENGAN <i>MULTIPLE LINEAR REGRESSION</i>	98
5.1 Model MLR dari Bobot-bobot PCA dengan Prediktor Antropometri Subyek	99
5.2 Metode Individualisasi Model HRIR Fasa Minimum yang Diusulkan	99
5.3 Analisis Hasil Individualisasi Model HRIR Fasa Minimum	101
5.3.1 Hasil Simulasi Obyektif	102
5.3.2 Diagram <i>Pole-Zero</i> Model HRIR Fasa Minimum Individual	106
5.3.3 Hasil Penentuan Posisi Bunyi	109
5.4 Metode Individualisasi Model HRTF <i>Magnitude</i> yang Diusulkan	111
5.5 Analisis Hasil Individualisasi Model HRTF <i>Magnitude</i>	114
BAB 6. KESIMPULAN	120
DAFTAR ACUAN	122
LAMPIRAN A <i>PRINCIPAL COMPONENTS ANALYSIS</i>	A1
LAMPIRAN B <i>MULTIPLE LINEAR REGRESSION</i>	B1
LAMPIRAN C <i>PERHITUNGAN HRIR FASA MINIMUM</i>	C1

LAMPIRAN D	MODEL-MODEL HRIR FASA MINIMUM INDIVIDUAL SUBYEK 003 DI BIDANG HORIZONTAL	D1
LAMPIRAN E	MODEL-MODEL HRIR FASA MINIMUM INDIVIDUAL SUBYEK 003 DI BIDANG MEDIAN	E1
LAMPIRAN F	MODEL-MODEL HRTF <i>MAGNITUDE</i> INDIVIDUAL SUBYEK 003 DI BIDANG HORIZONTAL	F1
LAMPIRAN G	MODEL-MODEL HRTF <i>MAGNITUDE</i> INDIVIDUAL SUBYEK 003 DI BIDANG MEDIAN	G1



DAFTAR GAMBAR

Gambar 1.1	Diagram Alir Penelitian	25
Gambar 1.2	Berbagai Pengolahan Awal dalam Pemodelan	26
Gambar 2.1	Gelombang Bunyi yang Ditransmisikan ke Telinga Pendengar	33
Gambar 2.2	Peristiwa-peristiwa Fisik dari Gelombang Bunyi	34
Gambar 2.3	Sistem Koordinat Segi Empat yang Berpusat pada Kepala	35
Gambar 2.4	Sistem-sistem Koordinat Bola [44]	36
Gambar 2.5	Ilustrasi untuk Menentukan ITD [44]	37
Gambar 2.6	ITD Subyek 028 pada Beberapa Bidang dengan Elevasi Tetap	39
Gambar 2.7	ILD Subyek 028 pada Bidang Horisontal dan Bidang Median	40
Gambar 2.8	Respon Frekuensi untuk Kedatangan Sinyal Bunyi dari Dua Arah yang Berbeda [44]	41
Gambar 2.9	Deskripsi Penyaringan Sumber Bunyi	46
Gambar 2.10	Sistem Audio Spasial Berbasis HRTF: Convolvotron™	48
Gambar 2.11	Ukuran-ukuran Antropometris Subyek [1]	51
Gambar 2.12	Seluruh HRTF <i>Magnitude</i> dari Subyek 028 pada Bidang Horisontal Depan [2]	55
Gambar 2.13	Seluruh HRTF <i>Magnitude</i> dari Subyek 028 pada Bidang Median [2]	56
Gambar 2.14	Grafik 3D Seluruh HRIR dari Subyek 028 pada Bidang Horisontal	58
Gambar 2.15	<i>Image</i> Seluruh HRIR dari Subyek 028 pada Bidang Horisontal	58
Gambar 2.16	Grafik 3D Seluruh HRIR dari Subyek 028 pada Bidang Median	59
Gambar 2.17	<i>Image</i> Seluruh HRIR dari Subyek 028 pada Bidang Median	60
Gambar 3.1	Model Fasa Minimum dari HRTF	63
Gambar 3.2	Filter <i>Allpass</i> Orde-1	68
Gambar 3.3	<i>Retained Coherence</i> untuk Seorang Subyek: HRTF Diukur dengan Mikrofon <i>Open Meatus</i> [57]	69
Gambar 3.4	<i>Retained Coherence</i> untuk Seorang Subyek: HRTF Diukur dengan Mikrofon <i>Blocked Meatus</i> [57]	69

Gambar 3.5	<i>Excess Group Delay</i> dan Respon <i>Magnitude</i> [57]	71
Gambar 3.6	<i>Retained Coherence</i> Subyek 003 untuk Semua Posisi Bunyi	71
Gambar 3.7	Respon Impuls dan Respon Fasa dari HRIR Asli dan HRIR Fasa Minimum pada Posisi Bunyi ($0^{\circ}, 0^{\circ}$)	72
Gambar 3.8	Sepuluh Vektor Basis Pertama dari Model DIR Fasa Minimum	79
Gambar 3.9	Persentase MSE antara Model HRIR _{mp} dan HRIR _{mp} Asli	80
Gambar 3.10	Sepuluh Vektor Basis Pertama dari Model DTF <i>Magnitude</i>	84
Gambar 3.11	Persentase MSE antara Model HRTF <i>Magnitude</i> dan HRTF <i>Magnitude</i> Asli	85
Gambar 4.1	Bobot-bobot Vektor Basis Model HRIR Fasa Minimum dari Subyek 003 di Bidang Horisontal Depan	94
Gambar 4.2	Korelasi antara Ukuran Kepala dan <i>Torso</i> Terhadap Bobot $w_1(\theta)$ pada Bidang Horisontal Depan	96
Gambar 5.1	Metode Individualisasi Model HRIR Fasa Minimum yang Diusulkan	100
Gambar 5.2	Kesalahan Model HRIR _{mp} Individual Subyek 003 dan Subyek 163 di Bidang Horisontal Depan	103
Gambar 5.3	Kesalahan Model HRIR _{mp} Individual Subyek 003 dan Subyek 163 di Bidang Median	104
Gambar 5.4	Model HRIR _{mp} Individual dan HRIR _{mp} Asli dari Subyek 003 di Bidang Horisontal Depan	105
Gambar 5.5	Model HRIR _{mp} Individual dan HRIR _{mp} Asli dari Subyek 003 di Bidang Median	106
Gambar 5.6	Diagram <i>Pole-Zero</i> dari HRIR _{mp} 200 Sampel dan HRIR _{mp} 67 Sampel Beserta <i>Magnitude</i> dari Nilai-nilai <i>Zero</i> yang Bersangkutan	107
Gambar 5.7	Diagram <i>Pole-Zero</i> dari Model-model HRIR _{mp} Individual 67 Sampel dari Telinga Kiri Subyek 003	108
Gambar 5.8	Panel untuk Sesi Latihan dan Sesi Tes dari Percobaan Penentuan Posisi Bunyi	110
Gambar 5.9	Metode Individualisasi Model HRTF <i>Magnitude</i> yang Diusulkan	112
Gambar 5.10	Perbandingan HRIR Asli dan Model HRIR dari Rekonstruksi Model-model HRTF <i>Magnitude</i>	114
Gambar 5.11	Kesalahan Model HRTF <i>Magnitude</i> Individual Subyek 003 dan Subyek 163 di Bidang Horisontal Depan	116

Gambar 5.12 Kesalahan Model HRTF <i>Magnitude</i> Individual Subyek 003 dan Subyek 163 di Bidang Median	117
Gambar 5.13 Model HRTF <i>Magnitude</i> Individual dan HRTF <i>Magnitude</i> Asli dari Subyek 003 di Bidang Horisontal Depan	118
Gambar 5.14 Model HRTF <i>Magnitude</i> Individual dan HRTF <i>Magnitude</i> Asli dari Subyek 003 di Bidang Median	118



DAFTAR TABEL

Tabel 2.1	Statistik Antropometri Basis Data	51
Tabel 3.1	Persentase MSE Rata-rata dari Berbagai Jenis Data HRIR	77
Tabel 3.2	Persentase Varians dari Vektor Basis Model DIR Fasa Minimum	78
Tabel 3.3	Persentase MSE Rata-rata dari Berbagai Jenis Data HRTF	82
Tabel 3.4	Persentase Varians dari Vektor Basis Model DTF <i>Magnitude</i>	83
Tabel 4.1	Koefisien Korelasi antara ITD_{max} dan Ukuran-ukuran Antropometris	90
Tabel 4.2	Koefisien Korelasi antara ILD_{max} dan Ukuran-ukuran Antropometris	92
Tabel 4.3	Koefisien Korelasi antara f_{pn} dan Ukuran-ukuran Antropometris	93

DAFTAR SINGKATAN

AES	<i>Audia Engineering Society</i>
ACM	<i>Association for Computing Machinery</i>
ASA	<i>Acoustical Society of America</i>
BEM	<i>Boundary Element Method</i>
DIR	<i>Direct Impulse Response</i>
DIR _{nc}	<i>DIR Response that is normalized by its energy</i>
DIR _{nz}	<i>DIR that is normalized by its standard deviation</i>
DFT	<i>Discrete Fourier Transform</i>
DTF	<i>Direct Transfer Function</i>
DTF _{nc}	<i>DTF that is normalized by its energy</i>
DTF _{nz}	<i>DTF that is normalized by its standard deviation</i>
FFT	<i>Fast Fourier Transform</i>
FIR	<i>Finite Impulse Response</i>
IEEE	<i>Institute of Electrical and Electronics Engineers</i>
IIR	<i>Infinite Impulse Response</i>
ILD	<i>Interaural Level Difference</i>
ILD _{max}	<i>maximum Interaural Level Difference</i>
ITD	<i>Interaural Time Difference</i>
ITD _{max}	<i>maximum Interaural Time Difference</i>
GA	<i>Genetic Algorithm</i>
HRIR	<i>Head-Related Impulse Response</i>
HRIR _{dp}	<i>direct-pulse HRIR</i>
HRIR _{mp}	<i>minimum phase HRIR</i>
HRIR _{nc}	<i>HRIR that is normalized by its energy</i>
HRIR _{nz}	<i>HRIR that is normalized by its standard deviation</i>
HRTF	<i>Head-Related Transfer Function</i>
HRTF _{nc}	<i>HRTF that is normalized by its energy</i>
HRTF _{nz}	<i>HRTF that is normalized by its standard deviation</i>
KEMAR	<i>Knowles Electronics Manikin for Acoustic Research</i>
LS	<i>Least Square</i>

LTI	<i>Linear Time-Invariant</i>
MLR	<i>Multiple Linear Regression</i>
MSE	<i>Mean square error</i>
OLS	<i>Ordinary Least Square</i>
PC	<i>Principal Component</i>
PCA	<i>Principal Components Analysis</i>
PCW	<i>Principal Component's Weight</i>



DAFTAR NOTASI

θ	sudut azimut dari posisi sumber bunyi
ϕ	sudut elevasi dari posisi sumber bunyi
(θ, ϕ)	pasangan sudut azimut dan sudut elevasi yang menunjukkan posisi sumber bunyi
f_{pn}	<i>pinna notch frequency</i>
$x_m(n)$	sinyal bunyi monaural
$h_l(n)$	HRIR telinga kiri
$h_r(n)$	HRIR telinga kanan
$y_l(n)$	sinyal keluaran yang didengar telinga kiri
$y_r(n)$	sinyal keluaran yang didengar telinga kanan
h	vektor HRIR / matriks yang terdiri dari sepasang HRIR
y	sinyal keluaran pada kedua telinga
$H(e^{j\omega})$	HRTF
$X_m(e^{j\omega})$	spektrum frekuensi dari sinyal $x_m(n)$
$Y_l(e^{j\omega})$	spektrum frekuensi dari sinyal $y_l(n)$
$Y_r(e^{j\omega})$	spektrum frekuensi dari sinyal $y_r(n)$
$H_{ap}(e^{j\omega})$	HRTF dengan fasa <i>allpass</i>
$H_{mp}(e^{j\omega})$	HRTF dengan fasa minimum
$F_D\{\}$	Transformasi Fourier Diskrit
x_1	<i>head width</i>
x_2	<i>head height</i>
x_3	<i>head depth</i>
x_4	<i>pinna offset down</i>
x_5	<i>pinna offset back</i>
x_6	<i>neck width</i>
x_7	<i>neck height</i>
x_8	<i>neck depth</i>
x_9	<i>torso top width</i>
x_{10}	<i>torso top height</i>
x_{11}	<i>torso top depth</i>

x_{12}	<i>shoulder width</i>
x_{13}	<i>head offset forward</i>
x_{14}	<i>height</i>
x_{15}	<i>seated height</i>
x_{16}	<i>head circumference</i>
x_{17}	<i>shoulder circumference</i>
d_1	<i>cavum concha height</i>
d_2	<i>cymba concha height</i>
d_3	<i>cavum concha width</i>
d_4	<i>fossa height</i>
d_5	<i>pinna height</i>
d_6	<i>pinna width</i>
d_7	<i>intertragal incisure width</i>
d_8	<i>cavum concha depth</i>
θ_1	<i>pinna rotation angle</i>
θ_2	<i>pinna flare angle</i>
$\mathcal{H}\{\}$	transformasi Hilbert
$h_{mp}(n)$	HRIR fasa minimum
$win(n)$	fungsi jendela dalam perhitungan <i>real cepstrum</i>
$F_D^{-1}\{\}$	invers transformasi Fourier
$\exp\{\}$	fungsi eksponensial
v_i	vektor basis / vektor eigen / <i>principal component</i> yang ke-i
λ_i	nilai eigen yang ke-i
μ	HRIR rata-rata dari seluruh data HRIR HRTF rata-rata dari seluruh data HRTF
S	matriks kovarians dari seluruh HRIR / HRTF
$E\{\}$	nilai ekspektasi
I	matriks identitas
$\ v\ $	<i>magnitude</i> dari vektor v
h_i	vektor HRIR/HRTF yang ke-i
H	matriks yang terdiri dari semua HRIR/HRTF
\hat{h}_i	vektor model HRIR/HRTF ke-i

\hat{H}	matriks yang terdiri dari semua model HRIR/HRTF
d_i	vektor DIR / DTF ke-i
D	matriks yang terdiri dari seluruh DIR / DTF
\hat{d}_i	vektor model DIR/DTF ke-i
\hat{D}	matriks yang terdiri dari model-model DIR/DTF
V	matriks vektor-vektor eigen yang telah berurutan berkaitan dengan nilai eigen terbesar pertama, kedua, dan seterusnya.
V_L	matriks yang hanya terdiri dari L vektor eigen pertama
w	vektor bobot-bobot hasil transformasi PCA
$w_i(\theta)$	vektor bobot-bobot dari vektor basis ke-i pada azimuth θ di bidang horisontal dari seluruh subyek
W	matriks yang terdiri dari vektor-vektor bobot w
Λ	matriks diagonal dari nilai-nilai eigen berurutan
$g[m]$	konten energi kumulatif dari vektor eigen pertama sampai vektor eigen ke-m
β_i	koefisien regresi ke-i
$\hat{\beta}_i$	estimasi koefisien regresi ke-i
β	vektor koefisien-koefisien regresi
$\hat{\beta}$	vektor estimasi-estimasi koefisien regresi
e_i	kesalahan estimasi MLR ke-i
e	vektor kesalahan-kesalahan estimasi MLR
X	matriks yang terdiri dari vektor-vektor ukuran antropometris
ρ	koefisien korelasi
$\varphi_{mp}(\omega)$	fasa minimum
$\eta(\omega)$	<i>excess phase response</i>
$\tau(\omega)$	<i>group delay</i>
$\gamma(\omega)$	<i>excess group delay</i>
$\psi_{xy}(l)$	koherensi silang antara sinyal x dan sinyal y dengan lag l
$\kappa(\theta, \phi)$	<i>retained coherence</i>

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Bagi kalangan masyarakat tertentu, kebutuhan pokok bukan merupakan masalah lagi. Pada gilirannya kebutuhan untuk mendapatkan hiburan dirasakan penting untuk dipenuhi, terutama bagi masyarakat perkotaan. Salah satu bentuk hiburan sekarang ini adalah mendengarkan musik dan lagu kesukaan mereka dengan nyaman. Dalam dekade terakhir ini, dikembangkan satu teknologi baru dalam menghasilkan musik yang terdengar alami, nyaman dan menarik bagi seseorang, yang menghasilkan efek-efek bunyi tiga dimensi atau spasial. Teknologi baru ini mengimplementasikan fungsi-fungsi alih dari kedua telinga pendengar dalam reproduksi musik. Fungsi alih telinga ini dikenal sebagai *Head-Related Transfer Function* (HRTF). Karena persepsi bunyi setiap orang berbeda dengan persepsi orang lain, maka HRTF dari masing-masing telinga seseorang bersifat sangat khas. Dengan kata lain, pasangan HRTF seseorang berbeda dengan pasangan HRTF orang lain, oleh karena itu dikatakan HRTF bersifat individual. Musik terdengar nyaman dan alami jika musik tersebut terdengar memenuhi ruang spasial dan bunyi-bunyi yang berbeda terdengar berasal dari posisi-posisi yang sebenarnya. Jika pasangan HRTF dari kedua telinga seseorang tersedia pada semua posisi di ruang spasial, maka semua HRTF tersebut dapat digunakan untuk menghasilkan efek-efek spasial yang lengkap dalam reproduksi musik.

Dalam beberapa dekade belakangan ini, telah dikembangkan sistem-sistem elektronik yang menyerupai sistem pendengaran alami manusia dengan menggunakan kedua telinganya; dimana seseorang dapat menentukan posisi sumber bunyi yang berada dalam ruang spasial dengan menggunakan kedua telinganya. Sistem-sistem ini disebut juga dengan sistem audio spasial. Sintesis sistem audio spasial ditemukan dalam banyak aplikasi yang krusial, seperti digunakan dalam percakapan pada pesawat terbang militer maupun komersial, pengembangan *virtual reality*, termasuk di dalamnya aplikasi-aplikasi dalam permainan-permainan elektronik, maupun sebagai *consumer electronics* untuk

mendengarkan musik yang memberikan kenyamanan dan efek-efek spasial bagi pengguna.

Sistem audio spasial sederhana terdiri dari tiga kelas dasar, yaitu stereo, multikanal dan *binaural recordings*. Kedua kelas yang disebutkan pertama terbatas hanya menghasilkan bunyi spasial pada posisi azimut atau horisontal saja. Reproduksi bunyi pada posisi elevasi ditentukan oleh posisi vertikal yang tetap dari penguat-penguat suara dan tidak tergantung dari bunyi yang disimulasikan.

Pada mulanya dikenal dua jenis sistem reproduksi bunyi, yaitu sistem reproduksi bunyi mono dan sistem reproduksi bunyi stereo. Sistem reproduksi bunyi mono menghasilkan bunyi kanal tunggal. Di sini digunakan satu mikrofon untuk merekam bunyi dan satu penguat suara atau lebih untuk mereproduksi bunyi dari mikrofon tadi. Sistem ini tidak menghasilkan informasi spasial dari bunyi yang dikeluarkan. Sebaliknya dalam sistem reproduksi stereo, terdapat dua kanal independen baik pada sisi perekaman maupun pada sisi reproduksi. Bunyi yang dihasilkan mengandung informasi spasial dan memiliki kualitas subyektif yang lebih baik, sehingga bunyi tersebut terdengar lebih natural. Hal ini terjadi karena efek stereo tergantung pada faktor-faktor psikoakustik yang dapat dimengerti oleh sistem otak manusia.

Efek stereo ditentukan oleh perbedaan-perbedaan penting di antara kedua sinyal dari kedua kanal bunyi yang dihasilkan walaupun jarak antara kedua mikrofon cukup kecil. Dalam industri hiburan audio, sistem stereo adalah produk komersial pertama yang sukses yang melibatkan bunyi spasial. Konsep dasar stereo adalah menempatkan bunyi di sisi kiri dan mengirimkan sinyalnya ke penguat suara kiri, demikian juga menempatkan bunyi lain di sisi kanan, lalu mengirimkan sinyalnya ke penguat suara kanan. Suatu *phantom source* terjadi pada suatu titik di tengah-tengah antara kedua penguat suara jika sinyal yang sama dikirimkan ke kedua penguat suara yang dihubungkan sefasa di dalam ruangan yang cukup beraturan secara akustik. Dengan melakukan *crossfading* terhadap sinyal dari satu penguat ke penguat lainnya, maka seseorang dapat menghasilkan impresi pergerakan bunyi secara kontinu di antara kedua penguat suara. *Crossfading* yang sederhana tidak akan menghasilkan impresi bahwa suatu

sumber berada di luar segmen garis antara kedua pengeras suara. Impresi tersebut hanya dapat dilakukan dengan stereo yang *crosstalk-cancelled*.

Cara lain untuk menghasilkan bunyi pada posisi-posisi tertentu adalah dengan memberikan sinyal pada kanal-kanal terpisah untuk setiap posisi yang diinginkan, termasuk di atas dan di bawah kepala pengguna. Cara ini sebenarnya yang diaplikasikan pada sistem-sistem teater seperti *Dolby Pro Logic Surround Sound*. Dalam lingkungan reverberansi tertentu, asal suatu bunyi dapat ditentukan dengan menggunakan efek Franssen. Efek Franssen adalah ilusi audio yang membuat pendengarnya salah menentukan posisi sumber bunyi. Sistem multikanal menggunakan pengeras-pengeras kecil di posisi-posisi yang diinginkan, kecuali satu pengeras besar atau *subwoofer* yang menghasilkan bunyi frekuensi rendah yang memancar ke segala arah. Walaupun sistem-sistem multikanal menghasilkan efek-efek spasial yang impresif, sistem-sistem ini jelas mahal dan rumit sehingga sistem-sistem tersebut kurang memiliki peranan dalam pengembangan sistem audio spasial selanjutnya.

Sejak lama diketahui, bahwa untuk menciptakan suatu bunyi spasial yang meyakinkan, tidak diperlukan banyak kanal, tetapi dua kanal saja sudah cukup. Cara yang digunakan adalah dengan mereproduksi tekanan-tekanan bunyi pada gendang telinga kiri dan gendang telinga kanan yang terjadi jika seorang pendengar seolah-olah benar-benar hadir di sekitar sumber bunyi. Pendekatan konseptual yang sederhana dilakukan dengan menempatkan dua mikrofon pada kanal-kanal telinga dari suatu manikin akustik atau pada kedua kanal telinga manusia dan merekam bunyi yang ditangkap kedua mikrofon tersebut. Ketika sinyal kiri dan kanan diumpankan pada *headphones* kiri dan kanan, maka seolah-olah pendengar hadir di sekeliling sumber bunyi. Jika manikin dan pendengar mempunyai ukuran dan bentuk kepala yang sama, informasi *interaural time difference* (ITD) dan *interaural level difference* (ILD) yang sama akan diperoleh, demikian juga jika manikin dan pendengar mempunyai ukuran dan bentuk daun telinga yang sama, faktor-faktor elevasi juga akan muncul. Perekaman dengan cara ini disebut *binaural recordings* yang dapat menghasilkan bunyi spasial yang lebih alami daripada kedua sistem lainnya. Selain keuntungan *binaural recordings* dari segi ekonomi dan keefektifannya, perekaman jenis ini mengandung beberapa

kelemahan. Pertama, diperlukan *headphones* untuk reproduksi. Kedua, sistem ini tidak interaktif melainkan harus direkam terlebih dahulu. Ketiga, jika pendengar bergerak, bunyi juga ikut bergerak. Keempat, sumber-sumber bunyi tepat di depan biasanya kelihatan terlalu dekat. Terakhir, karena bentuk-bentuk daun telinga berbeda bagi setiap orang, efek-efek elevasi tidak dapat diandalkan. Pengembangan lebih lanjut dari *binaural recordings* memerlukan pengertian tentang HRTF.

Kemampuan manusia untuk melokalisasi suatu sumber bunyi lateral berhubungan dengan kenyataan bahwa manusia mendengar dengan dua telinga dan secara sadar membandingkan sinyal-sinyal akustik yang diterima keduanya. Jika suatu gelombang bunyi datang dari suatu posisi yang tidak berada dalam bidang simetri vertikal (bidang median), maka salah satu telinga sedikit banyak dibayangi oleh kepala, sementara itu telinga lain terbuka penuh terhadap kedatangan bunyi. Sebagai tambahan pada perbedaan amplitudo yang disebabkan efek ini, perbedaan-perbedaan dalam waktu-waktu transit dari kedua sinyal telinga juga terjadi. Gelombang datang dihamburkan oleh kepala dan kedua daun telinga. Hasil dari proses ini berbeda untuk kedua telinga jika bunyi datang dari arah samping. Proses ini secara kuantitatif dideskripsikan oleh HRTF, yaitu fungsi alih yang mengacu pada transmisi sinyal bunyi dari medan jauh menuju pangkal kanal telinga. Fungsi alih ini dapat dibayangkan sebagai fungsi filter yang mengubah spektrum sinyal bunyi yang datang. Fungsi alih telinga kiri seseorang berbeda dengan fungsi alih telinga kanannya untuk suatu posisi sumber bunyi yang disebabkan oleh difraksi dan refleksi sinyal bunyi pada permukaan kepala dan daun telinga.

HRTF mencakup semua faktor fisik dari penentuan posisi sumber bunyi. Jika HRTF untuk telinga kiri dan HRTF untuk telinga kanan diperoleh, suatu sumber monaural dapat disintesis dengan menggunakan kedua HRTF tadi menjadi sinyal-sinyal binaural. HRTF merupakan fungsi yang rumit dari empat variabel: tiga koordinat ruang dan frekuensi. Dalam koordinat bola, jika jarak lebih dari satu meter, sumber bunyi dikatakan berada dalam medan jauh dan HRTF bertambah kecil dengan bertambahnya jarak. Kebanyakan pengukuran HRTF dilakukan di dalam medan jauh, yang pada intinya mengakibatkan HRTF menjadi

suatu fungsi dari azimut, elevasi dan frekuensi. Pada jarak yang lebih dekat, nilai-nilai ILD yang diamati di antara kedua telinga akan menjadi cukup besar, bahkan pada daerah frekuensi rendah, dimana dalam medan jauh diamati nilai-nilai ILD yang dapat diabaikan. Pasangan transformasi Fourier dari HRTF di kawasan waktu disebut *head-related impulse response* (HRIR). Pengukuran HRIR dilakukan para peneliti di CIPIC *Interface Laboratory* di *University of California at Davis* dengan mengambil jarak satu meter antara sumber bunyi dan pusat kepala subyek. Mereka menghasilkan basis data HRTF yang diberi nama CIPIC HRTF *Database*. Dalam penelitian ini, digunakan basis data tersebut, yang terdiri dari HRIR-HRIR dari 45 subyek untuk 1250 posisi sumber bunyi per telinga [1,2]. Jadi terdapat seluruhnya 112.500 HRIR, ditambah data ITD, waktu-waktu *Onset* dan ukuran-ukuran antropometris setiap subyek. Sedangkan setiap HRIR terdiri dari 200 respon impuls dengan frekuensi cuplik 44.100 Hz. Jadi, dapat dibayangkan betapa rumit analisis data HRIR tersebut dan betapa rumit pemodelan HRIR/HRTF dengan menggunakan basis data tadi. Penjelasan lebih rinci dari CIPIC HRTF *Database* dapat ditemukan pada subbab 2.8.

Reproduksi bunyi yang alami melalui *headphones* maupun penguas suara membutuhkan HRTF telinga kiri dan HRTF telinga kanan dari seorang pendengar sendiri. Pengukuran HRTF milik sendiri tentu saja membutuhkan banyak tenaga dan biaya, misalnya diperlukan pengukuran HRTF untuk posisi-posisi ruang yang sebanyak mungkin untuk merepresentasikan posisi ruang yang kontinu, belum lagi peralatan dan ruang kedap suara yang diperlukan menghabiskan banyak biaya. Tetapi penggunaan HRTF individual menyebabkan hasil-hasil yang sangat bagus. Untuk menyiasati kelemahan penggunaan HRTF individual, terdapat tiga pendekatan. Pendekatan pertama adalah dengan menggunakan suatu kompromi yaitu berupa sepasang HRTF standar. Metode ini memberikan hasil-hasil elevasi yang agak lemah untuk beberapa bagian dari populasi, tetapi metode ini praktis untuk sistem-sistem yang tidak mahal. Saat ini instituti-institusi dunia bidang audio baik *Institute of Electrical and Electronics Engineers* (IEEE), *Association for Computing Machinery* (ACM), *Audio Engineering Society* (AES) maupun *Acoustical Society of America* (ASA) belum ada yang mendefinisikan sepasang HRTF standar.

Pendekatan kedua adalah dengan menggunakan sepasang HRTF dari sekumpulan pasangan HRTF standar. Metode ini membutuhkan pengukuran HRTF untuk sejumlah kecil masyarakat yang merepresentasikan seluruh kelompok dengan karakteristik yang berbeda-beda dalam populasi. Kemudian diperlukan suatu cara sederhana agar seorang pendengar dapat memilih sepasang HRTF yang paling sesuai dengannya, misalnya digunakan dalam produk komersial berupa amplifier *headphones* HEARO dari perusahaan *AKG-Acoustics*. Walaupun telah banyak diusulkan, belum ada suatu kumpulan pasangan HRTF standar yang sudah dibakukan. Pendekatan pertama dan kedua memiliki kelemahan berkaitan dengan kenyataan bahwa jika seseorang mendengarkan bunyi spasial hasil reproduksi dengan menggunakan pasangan HRTF milik orang lain atau pasangan HRTF standar, gangguan-gangguan persepsi akan terjadi, misalnya bunyi terdengar tidak seperti bunyi aslinya, bunyi terdengar berada di dalam kepala jika digunakan *headphones* dan jarak yang terdengar juga tidak sesuai dengan yang sebenarnya.

Untuk mengatasi kelemahan pada HRTF individual, sepasang HRTF standar dan kumpulan pasangan HRTF standar, seperti misalnya karena tingginya biaya dan usaha yang diperlukan untuk mengukur HRTF individual, diperlukan suatu model HRTF atau HRIR yang bersifat parametrik, yang berarti bahwa model tersebut terdiri dari beberapa parameter yang dapat diatur berdasarkan ukuran-ukuran antropometris dari *torso*, kepala dan daun telinga dari pendengar, sehingga model yang telah ditala tersebut sesuai dengan atau mendekati HRTF dari si pendengar. Para peneliti di dunia telah menganalisis beberapa cara untuk memodelkan HRTF [3, 4]. Sampai saat ini terdapat tiga jenis pemodelan umum yang sudah banyak diteliti.

Pemodelan pertama berupa fungsi rasional atau model *pole-zero*. Dalam hal ini masalah pemodelan dipandang sebagai identifikasi sistem yang telah memiliki solusi-solusi klasik. Akan tetapi, model ini menggunakan koefisien-koefisien yang biasanya merupakan fungsi yang sedemikian rumit dari azimuth dan elevasi sehingga harus ditabulasi, yang berarti tidak ditemukan suatu cara yang sederhana untuk mengatur koefisien-koefisien tersebut agar diperoleh HRTF yang

cocok untuk seorang pendengar. Hal ini pada akhirnya menghancurkan manfaat dari model itu sendiri.

Pemodelan kedua merupakan model-model struktural. Di sini diusahakan untuk menghasilkan fungsi-fungsi alih sederhana yang memperhitungkan mekanisme fisik, seperti difraksi-difraksi daun telinga, bayangan kepala, refleksi-refleksi pundak dan *torso* dan sebagainya. Pemodelan ini telah dikembangkan secara mendalam oleh Genuit [5] dan dilanjutkan oleh kelompok peneliti di *CIPIC Interface Laboratory*. Pemodelan struktural memiliki beberapa masalah yang belum terpecahkan, yaitu pemodelan yang terlalu sederhana dan proses integrasi yang sulit dari beberapa model yang berkaitan dengan ukuran antropometris tertentu yang dilakukan secara terpisah.

Pemodelan ketiga berupa ekspansi-ekspansi deret (*series expansions*) atau model statistik. Di dalam ekspansi-ekspansi deret Karhunen-Loève, juga dikenal sebagai *Principal Components Analysis (PCA)*, HRTF dimodelkan sebagai jumlah terbobot (*weighted sum*) dari fungsi-fungsi basis yang lebih sederhana. Salah satu kelemahan dalam model ini adalah kompleksitas perhitungan dalam pemodelan. Kelemahan tersebut dapat diatasi karena pemodelan ini dilakukan terlebih dulu dengan menggunakan komputer. Pemodelan PCA memanfaatkan basis data HRTF yang cukup besar dari hasil pengukuran HRIR sejumlah besar subyek dan dilakukan secara statistik. Oleh karena itu, diharapkan kombinasi fungsi-fungsi basis hasil ekstraksi dari basis data, dapat merepresentasikan model HRTF dari populasi manusia. Selanjutnya hasil pemodelan HRTF dengan PCA diharapkan dapat diatur menggunakan ukuran-ukuran antropometris seorang pendengar agar sesuai dengan HRTF individual orang tersebut.

1.2 Penelitian-penelitian Pemodelan HRTF dan Individualisasi Model HRTF

Seperti telah disebutkan di atas bahwa pemodelan HRTF mencakup tiga pendekatan umum. Pemodelan pertama berupa fungsi rasional atau model *pole-zero*. Pemodelan kedua merupakan model-model struktural. Pemodelan ketiga berupa ekspansi-ekspansi deret atau model-model statistik. Di dalam bagian-bagian berikut dibahas mengenai kelebihan dan kelemahan penelitian-penelitian

yang telah dilakukan dalam pemodelan HRTF. Kemudian dibahas juga individualisasi model HRTF yang dilakukan oleh para peneliti sebelumnya.

Beberapa peneliti berusaha memodelkan HRTF individu dengan menggunakan model *pole-zero* [6,7]. Durant dan Wakefield [7] menggunakan *genetic algorithm* (GA) untuk membangkitkan secara efisien aproksimasi-aproksimasi *pole-zero* orde rendah untuk HRTF dari respon-respon impuls yang didapat dari pengukuran, yaitu dengan cara meminimasi suatu kriteria kesalahan logaritmik. Pendekatan ini lebih sederhana dan lebih unggul dalam efisiensi daripada algoritma-algoritma pencarian lainnya. Penelitian mereka merupakan pengembangan dari penelitian sebelumnya dalam aproksimasi HRTF orde rendah [6]. Blommer dan Wakefield sebelumnya menganalisis HRTF dan berbagai jenis aproksimasi orde rendah lain. Mereka memperkenalkan suatu algoritma desain *pole-zero* yang menggunakan teknik-teknik pencarian (*search*) gradien untuk meminimasi baik kesalahan *log-magnitude* maupun kesalahan respon fasa. Model *pole-zero* HRTF ini memberikan hasil yang baik dalam mengaproksimasi HRTF orde rendah, tetapi para peneliti di sini tidak mencoba untuk membuat model yang parametrik sehingga tidak dapat dicari hubungan antara parameter-parameter model dengan ukuran-ukuran antropometris subyek.

Kelompok peneliti yang lain melakukan pemodelan HRTF jenis kedua menggunakan metode *boundary element method* (BEM). Model ini didasarkan pada pipa-pipa yang digabung-gabungkan (*concatenated ducts*) dan menggunakan metode-metode numerik untuk menghitung HRTF untuk frekuensi-frekuensi sampai dengan 6 kHz [8-10]. Dalam [8], BEM digunakan untuk menghitung fungsi-fungsi alih dari daun telinga dan basis data HRTF dari suatu kepala buatan (*dummy head*). Suatu pemindai laser (*laser scanner*) digunakan untuk mendapatkan model-model geometri dengan keakuratan tinggi, yang dikonversi menjadi jaring-jaring (*meshes*) BEM yang *valid* dengan resolusi-resolusi poligon yang berbeda-beda untuk jangkauan frekuensi yang berbeda-beda pula. Penelitian ini memiliki tingkat kesulitan tinggi dan membutuhkan peralatan yang akurat dan mahal, sehingga diperoleh juga hasil pemodelan HRTF yang memuaskan. Di sini juga tidak diusahakan untuk membuat model parametrik HRTF yang berdasarkan antropometri subyek.

Metode pemodelan fungsi rasional HRTF yang ketiga adalah menggunakan *beamforming* yang diusulkan oleh Chen et al. [11]. Model ini merupakan suatu model fungsional untuk karakteristik-karakteristik *filtering* spasial dan temporal dari daun telinga. Keluaran dari suatu *beamformer* adalah suatu kombinasi yang diboboti dari data yang diterima pada suatu deretan (*array*) sensor-sensor yang terdistribusi secara spasial. Bobot-bobot *beamformer* tersebut dan geometri deret menentukan karakteristik-karakteristik *filtering* spasial dan temporalnya. Bobot-bobot tersebut dipilih sedemikian rupa sehingga *mean-squared error* antara respon *beamformer* dan respon yang diukur dari daun telinga diminimasi. Respon dari model merupakan *inner product* dari vektor bobot dan suatu vektor respon deret. Vektor respon deret menyatakan secara eksplisit geometri medan dan posisi sumber bunyi. Metode ini memodelkan karakteristik-karakteristik transformasi dari daun telinga sebagai hubungan masukan-keluaran. Penelitian ini juga membutuhkan peralatan yang akurat dan mahal serta lebih sulit untuk dilakukan.

Model yang keempat dari fungsi rasional HRTF adalah pemodelan atau aproksimasi HRTF menggunakan filter *Infinite Impulse Response* (IIR) [12]. Seperti diketahui sampel-sampel dari HRIR juga merupakan koefisien-koefisien dari filter *Finite Impulse Response* (FIR). Setiap HRIR dalam CIPIC HRTF *Database* memiliki 200 sampel, sehingga filter FIR yang digunakan untuk merepresentasikan HRIR tersebut memiliki 200 koefisien. Tujuan aproksimasi HRTF dengan filter IIR adalah agar diperoleh filter yang lebih sederhana dengan jumlah koefisien yang lebih sedikit. Filter IIR yang didesain berupa kaskade dari beberapa filter IIR orde dua. Skenario pertama adalah menggunakan tiga tahap filter IIR orde dua untuk mensimulasikan HRTF. Skenario kedua menggunakan 4 sampai 7 tahap filter IIR orde dua. Sinyal-sinyal binaural menggunakan HRTFs hasil aproksimasi filter IIR skenario pertama dan kedua disimulasikan menggunakan pengeras suara. Percobaan lokalisasi sumber bunyi dilakukan pada bidang horisontal. Hasil dari percobaan ini dibandingkan dengan hasil percobaan menggunakan HRTFs hasil simulasi dengan filter FIR. Dari perbandingan tadi dihasilkan bahwa HRTFs yang diaproksimasi dengan skenario kedua memiliki akurasi yang sama dengan HRTFs yang disimulasi dengan filter FIR. Penelitian

Hasegawa et al. [12] di atas menghasilkan aproksimasi HRTFs yang cukup memuaskan. Mereka menggunakan 16 sampai 28 koefisien filter IIR agar diperoleh akurasi yang sama seperti pada filter FIR. Penelitian ini memberikan inspirasi untuk menghasilkan aproksimasi HRTF yang lebih sederhana. Sayangnya, mereka tidak mencoba menghubungkan koefisien-koefisien filter IIR yang dihasilkan dengan ukuran-ukuran antropometris agar diperoleh aproksimasi HRTF dengan filter IIR yang dapat diindividualisasi.

Secara umum, model *pole-zero* atau fungsi rasional dari HRTF memberikan banyak parameter yang sulit untuk dihubungkan dengan antropometri manusia. Selain itu model BEM dan *beamforming* membutuhkan peralatan yang canggih dan mahal. Di lain pihak, pemodelan HRTF dengan pendekatan fungsi rasional memberikan hasil yang memuaskan.

Model struktural pertama dikembangkan tahun 1960-an yang dikenal dengan nama model kawasan waktu (*time-domain model*), yang pertama kali diusulkan oleh Batteau [13]. Batteau mengukur dan melakukan studi terhadap respon impuls (*impulse response*) dari suatu replika daun telinga yang diperbesar dan menyimpulkan bahwa struktur fisik dari daun telinga menyebabkan dua impuls penting di samping bunyi *original*. Satu impuls yang bervariasi dengan perubahan posisi azimut dari sumber bunyi mempunyai waktu tunda dalam jangkauan 0 - 80 μ s, sementara itu impuls yang lain yang bervariasi terhadap elevasi memiliki waktu tunda antara 100 sampai 300 μ s. Batteau menduga bahwa daun telinga dapat dimodelkan sebagai suatu *three-hole-coupler*, yang ekuivalen dengan suatu sistem *two-delay-and-add*.

Studi-studi selanjutnya menunjukkan bahwa respon frekuensi dari suatu sistem *two-delay-and-add* tersebut menyerupai respon daun telinga untuk sudut-sudut kedatangan dalam bidang median. Keuntungan dari model Batteau adalah kesederhanaannya. Model ini hanya terdiri dari empat parameter a_1 , a_2 , $\tau_1(\theta)$ dan $\tau_2(\phi)$, dimana $\tau_1(\theta)$ mengacu pada waktu tunda impuls azimut dan $\tau_2(\phi)$ adalah waktu tunda impuls elevasi; a_1 dan a_2 merupakan dua konstanta yang menyatakan koefisien-koefisien refleksi. Posisi sumber bunyi yang ditentukan oleh sudut azimut θ dan sudut elevasi ϕ dikodekan secara eksplisit dalam model ini.

Kelemahan utama dalam model Batteau adalah aproksimasi yang kurang baik yang disebabkan oleh penggunaan suatu model waktu tunda untuk mendeskripsikan sistem fisik yang terdistribusi. Hanya hasil-hasil kualitatif yang diperoleh dengan model ini. Batteau menyadari masalah ini dan mencoba untuk menggeneralisasi modelnya menjadi suatu fungsi alih daun telinga. Tetapi, model ini lalu kehilangan kesederhanaannya dan menjadi terlalu umum untuk secara eksplisit mengindikasikan bagaimana posisi sumber bunyi dinyatakan.

Pemodelan struktural berikutnya dilakukan di kawasan frekuensi oleh Shaw pada tahun 1977 [14]. Shaw memodelkan daun telinga sebagai suatu silinder terbuka. Dimensi dari silinder dipilih agar menyebabkan silinder tersebut beresonansi kira-kira pada frekuensi 4,5 kHz, yang ekuivalen dengan resonansi rata-rata telinga manusia. Eksperimen-eksperimen yang dilakukan menunjukkan bahwa peralatan mekanik sederhana ini dapat mereproduksi suatu puncak mayor dalam fungsi alih daun telinga. Dengan menempelkan elemen-elemen tambahan yang mengacu pada daun telinga terhadap silinder ini, Shaw dapat menduplikasi secara aproksimasi, puncak-puncak resonansi tambahan dalam fungsi alih daun telinga untuk frekuensi-frekuensi di bawah 5 kHz.

Kemudian Shaw memperbaiki modelnya dengan menambahkan suatu repreneisasi dari kanal telinga dan suatu impedansi akhir untuk mensimulasikan beban yang disebabkan oleh gendang telinga. Dengan menggunakan sumber-sumber gelombang progresif sebagai stimuli, Shaw mengamati delapan resonansi antara 2,6 sampai 15,7 kHz dalam modelnya. Percobaan lanjutan mengindikasikan bahwa frekuensi-frekuensi resonansi ini bervariasi terhadap sudut datang dari sumber bunyi, seperti halnya resonansi-resonansi dalam fungsi alih daun telinga. Sayangnya, karena bentuk yang rumit dari struktur tulang rawan daun telinga, ukuran dari setiap resonator yang diperlukan untuk mendapatkan resonansi-resonansi pada frekuensi-frekuensi yang telah ditentukan sebelumnya, sangatlah sulit untuk dikalkulasi. Demikian juga, tidak ada ekspresi analitis yang mendeskripsikan keluaran model jika diberikan posisi sumber dan sinyal bunyi itu sendiri.

Suatu model parametrik dari HRTF dilaporkan oleh Genuit [5]. Yang dimaksudkan dengan model parametrik adalah model yang memiliki parameter-

parameter yang berhubungan dengan ukuran-ukuran antropometris. Genuit menyatakan bahwa dengan menggunakan integral-integral difraksi Kirrchoff, dimungkinkan untuk mendeskripsikan, secara pendekatan, fungsi alih daun telinga yang dipengaruhi oleh kepala, *torso*, pundak dan daun telinga. Dalam modelnya, 16 kanal dari kombinasi-kombinasi waktu tunda filter merepresentasikan kontribusi akustik dari kepala, *torso*, pundak dan daun telinga kepada respon keseluruhan daun telinga. Semua *delay* ini dan koefisien-koefisien filter merupakan fungsi-fungsi matematis eksplisit dari ukuran-ukuran telinga dan posisi sumber bunyi. Model Genuit adalah model pertama yang berusaha untuk mendeskripsikan semua fungsi akustik daun telinga dan membangun suatu hubungan matematis antara ukuran-ukuran telinga dan parameter-parameter dari model. Genuit menunjukkan bahwa spektrum amplitudo daun telinga yang diprediksi terletak antara amplitudo-amplitudo minimum dan maksimum yang diperoleh dari keenam pengukuran dari suatu respon daun telinga dari seorang subyek dalam satu posisi tunggal sumber bunyi. Studi-studi berikutnya yang memvalidasi keakuratan dari model ini di dalam suatu jangkauan posisi lainnya belum dilaporkan.

Peneliti-peneliti dari CIPIC *Interface Laboratory* [15-19] mengembangkan model struktural HRTF sebagai sekumpulan fungsi frekuensi dari bayangan kepala dan ITD, pantulan-pantulan pundak serta efek-efek daun telinga. Model ini didasarkan pada deskripsi kawasan waktu yang disederhanakan dari sifat-sifat fisik propagasi dan difraksi gelombang. Mereka membuat fungsi-fungsi frekuensi atau filter-filter sederhana yang merupakan respon dari masing-masing antropometri yang berkaitan dengan gelombang bunyi yang datang. Desain filter-filter tersebut dilakukan secara terpisah, kemudian digabungkan untuk memperoleh respon lengkap dari subyek. Model yang didapat menghasilkan efek-efek vertikal dan horisontal yang terkendali dengan baik. Salah satu keunggulan dari model struktural ini adalah bahwa komponen-komponen model memiliki suatu korespondensi satu-satu dengan sumber-sumber fisik dari difraksi, *delay* dan refleksi bunyi. Walaupun diperoleh model yang sederhana, tetapi ada beberapa parameter yang sulit untuk diatur agar diperoleh model HRTF untuk individu tertentu. Seperti dikemukakan oleh Satarzadeh et al. [19], bahwa meskipun

mungkin untuk mengestimasi parameter-parameter model dari beberapa ukuran geometri daun telinga, tetapi estimasi langsung dari antropometri daun telinga secara umum masih merupakan masalah yang belum terselesaikan.

Suatu dekomposisi HRTF menjadi kombinasi dari filter-filter, dimana setiap filter merepresentasikan efek-efek struktur anatomis secara terisolasi, merupakan suatu aproksimasi yang heuristik. Terdapat banyak interaksi antara gelombang-gelombang yang dihamburkan dari berbagai struktur yang berbeda. Sekarang ini belum ditemukan bahwa apakah mungkin untuk menyusun kembali HRTF lengkap dari HRTF-HRTF untuk bagian-bagian berbeda dari tubuh yang diukur secara terisolasi. Demikian pula, model struktural menghasilkan aproksimasi HRTF yang sederhana, dalam arti model hanya mengikuti pola umum dari HRTF, tetapi tidak merepresentasikan HRTF secara lebih rinci.

Yang dimaksud dengan model statistikal HRTF di sini adalah model HRTF yang dibangun menggunakan pendekatan statistik dengan memanfaatkan kekayaan data di dalam basis data. Model statistikal HRTF disebut juga dengan ekspansi-ekspansi deret yang merepresentasikan HRTF. Penelitian ini difokuskan pada model statistikal HRTF menggunakan PCA. Sebelum mengemukakan alasannya, terlebih dulu dibahas penelitian-penelitian terdahulu mengenai model statistikal HRTF.

Model statistikal pertama dilakukan menggunakan pendekatan GA dalam pemodelan HRTF oleh Cheung dan Trautman [20]. Basis data yang digunakan diperoleh dari pengukuran HRTF menggunakan obyek *Knowles Electronics Manikin for Acoustic Research (KEMAR)* dari MIT Media Laboratory. Seluruh HRIR basis data (710 posisi sumber bunyi) diubah menjadi respon amplitudo HRTF 207 titik cuplik dari 13,75 Hz sampai 22.050 Hz. Setiap respon amplitudo dikurangi dengan rata-rata respon amplitudo dari seluruh 710 posisi, fungsi ini disebut juga sebagai *Directional Transfer Function (DTF)*. GA didesain untuk memilih spektrum-spektrum basis yang representatif dari seluruh DTF. Fungsi *fitness* yang berorientasi pada persepsi telinga dikembangkan untuk menuntun proses pencarian. Tiga DTF terpilih oleh GA sebagai spektrum-spektrum basis yang merupakan DTF pada posisi (θ, ϕ) : $(-20^\circ, 0^\circ)$, $(60^\circ, 270^\circ)$ dan $(10^\circ, 5^\circ)$. Untuk

DTF dengan posisi-posisi yang jauh dari DTF spektrum basis, dihasilkan model-model HRTF yang kurang baik.

Model statistik HRTF berusaha secara statistik melakukan parameterisasi seluruh kumpulan HRTF (digunakan nilai *logarithmic magnitude* HRTF [21-23], nilai kompleks HRTF [24], atau nilai amplitudo linier HRTF [25,26]) dari suatu basis data dalam kawasan frekuensi atau parameterisasi seluruh set HRIR dalam kawasan waktu [27] dengan menggunakan metode statistik yang dikenal dengan nama PCA, *singular value decomposition*, atau *Karhunen-Loève expansion*. Metode-metode statistik ini pada dasarnya berasal dari ide yang sama, yaitu diuraikan sekumpulan spektrum *magnitude* dari HRTF atau respon impuls dari HRIR menjadi fungsi-fungsi basis dengan cara yang mirip seperti cara analisis Fourier menguraikan suatu bentuk gelombang menjadi komponen-komponen sinus dan cosinus. Fungsi-fungsi basis dapat diasumsikan sebagai bentuk-bentuk spektral dasar dari mana setiap spektrum dalam kumpulan HRTF dibangun, atau dapat juga diasumsikan sebagai bentuk-bentuk respon impuls dasar dari mana setiap respon impuls dalam kumpulan HRIR dibangun.

Data masukan bagi metode-metode ini adalah sekumpulan grafik daun telinga di dalam kawasan waktu yang disebut HRIR, atau di dalam kawasan frekuensi yang disebut HRTF. Kumpulan HRTF ini berjumlah seluruhnya N spektrum HRTF; misalnya terdapat 250 posisi sumber dan diukur dari 10 subyek untuk kedua daun telinga, maka $N = 250 \times 10 \times 2 = 5000$ spektrum. Setiap spektrum HRTF terdiri dari M komponen frekuensi. Suatu matriks kovarians spektrum S yang berukuran $(M \times M)$ dihitung dengan menggunakan N spektrum HRTF tadi. Dari matriks kovarians S , dapat dihitung dengan menyelesaikan persamaan eigen, sebanyak M vektor eigen yang berurutan dimulai dari vektor eigen dengan nilai eigen yang terbesar sampai dengan vektor eigen dengan nilai eigen terkecil. HRTF individual dapat direpresentasikan secara sempurna dan merupakan jumlah terbobot dari M vektor eigen. Untuk mengurangi ukuran data dan perhitungan, hanya digunakan sebanyak L vektor eigen yang mewakili sebagian besar (biasanya lebih dari 90%) variasi yang terdapat dalam basis data HRTF. Vektor eigen dengan nilai eigen terbesar menyumbang bagian terbesar variasi yang terdapat dalam basis data HRTF dalam merepresentasi setiap HRTF. Bobot-bobot

yang tergantung posisi bunyi menyatakan kontribusi setiap vektor eigen yang digunakan dalam suatu respon impuls atau dalam suatu spektrum. Dekomposisi seperti ini berarti bahwa komponen-komponen arah dalam bentuk bobot-bobot dan komponen-komponen frekuensi dalam bentuk vektor-vektor eigen terpisah satu sama lain. Bobot-bobot tadi juga tergantung dari data antropometri dari setiap subyek dalam penelitian.

Wightman dan Kistler [21] mengusulkan suatu model matematis untuk merepresentasikan HRTF sebagai suatu kombinasi linier dari sejumlah kecil spektrum basis. Mereka menggunakan nilai *log-magnitude* dari HRTF dari 5300 HRIR. Rata-rata dari semua *log-magnitude* HRTF tadi dihitung dan kemudian dikurangi dari masing-masing *log-magnitude* HRTF. Fungsi hasil pengurangan tersebut dikenal sebagai DTF. Keseluruhan DTF lalu diproses dalam PCA untuk menghitung spektrum-spektrum basis dan bobot-bobotnya. Setiap model *log-magnitude* HRTF dinyatakan sebagai hasil penjumlahan spektrum-spektrum basis yang diboboti ditambah dengan rata-rata *log-magnitude* HRTF. Fasa HRTF dimodelkan dengan menganggap bahwa HRTF adalah fungsi fasa minimum dan bahwa perbedaan-perbedaan fasa antar telinga dapat diaproksimasi dengan suatu waktu tunda yang sederhana. Mereka menggunakan lima spektrum basis untuk memodelkan 5300 DTF secara efektif; di sini diperoleh reduksi data sebanyak 30 kali. Dalam model matematis ini, masalah utama adalah bagaimana mencari spektrum-spektrum basis yang representatif. Mereka melihat bahwa terdapat hubungan antara bobot-bobot dan posisi bunyi serta hubungan antara bobot-bobot dan data antropometri subyek. Tetapi hubungan-hubungan ini tidak dinyatakan secara kuantitatif untuk menghasilkan individualisasi model HRTF.

Leung dan Carlile [28] membandingkan berbagai teknik *preprocessing* (HRIR linier [40], HRTF *log-magnitude* [21-23], HRTF amplitudo linier [25,26], HRTF pasangan riil-kompleks [24]) dalam hal efisiensi kompresi berdasarkan suatu analisis varians. Dalam kawasan frekuensi, HRTFs diberikan skala linier dan logaritmik. Baik dalam skala linier maupun skala logaritmik, HRTF amplitudo linier memberikan varians yang optimal, yakni mendekati varians 100% jika HRTFs direkonstruksi dari 10 vektor basis PCA. Hal yang menarik adalah bahwa dengan hanya 5 vektor basis PCA, konten varians mendekati 95%.

Perbedaan-perbedaan antara HRTF asli dan bentuk filter yang direkonstruksi oleh PCA dalam kawasan frekuensi dikuantifikasi dengan MSE, sama seperti dalam disertasi ini. Baik dalam skala linier dan skala logaritmik, MSE berkurang dengan tajam untuk HRTF amplitudo linier, yang juga memiliki deviasi standar yang terkecil. HRTF amplitudo linier yang disebutkan di sini adalah sama dengan HRTF *magnitude* yang digunakan dalam disertasi ini. Jadi hasil-hasil dalam [28] mengkonfirmasi bahwa HRTF *magnitude* adalah jenis data HRTF yang terbaik dalam pemodelan PCA di kawasan frekuensi dibandingkan jenis-jenis data HRTF lain. Tetapi Leung dan Carlile menggunakan basis data yang berbeda daripada yang digunakan dalam disertasi ini.

Percobaan-percobaan *psychophysical* kemudian dilakukan untuk menginvestigasi akurasi perseptual dari HRTF yang diregenerasi dengan berbagai tingkat kompresi. Percobaan-percobaan ini dilakukan dalam medan bebas dan *virtual auditory space* (VAS) dalam suatu ruang kedap bunyi. Stimuli adalah *Gaussian noise bursts* selebar 150 ms yang dikombinasikan dengan berbagai HRTF yang dikompresi dengan berbagai jumlah vektor basis PCA. Setiap *trial* terdiri dari 76 lokasi target dan diulangi 5 kali dengan total sebanyak 380. Semua subyek memberikan tingkat mislokalisasi yang lebih tinggi jika HRTFs diregenerasi dengan hanya 5 vektor basis PCA dibandingkan jika HRTFs diregenerasi dengan 10 vektor basis PCA. Jumlah kesalahan depan-belakang juga meningkat dengan HRTFs yang mengandung 5 vektor basis. Leung dan Carlile tidak melakukan individualisasi terhadap model-model HRTF amplitudo linier yang diperoleh dengan PCA.

Dalam penelitian dari Pec dan Strumillo [29], pemodelan HRTFs dengan FIR orde rendah berdasarkan PCA dianalisis. Metode-metode yang dilaporkan mengizinkan untuk mereduksi panjang filter FIR menjadi 48 orde tanpa penurunan resolusi lokalisasi dari sumber-sumber bunyi. Di sini PCA dilakukan terhadap HRTF *magnitude*, sama seperti dalam disertasi ini. Model HRTF *magnitude* direkonstruksi menggunakan 6 vektor basis PCA, menjadi filter FIR 48 orde berdasarkan algoritma *least square* terhadap respon-respon frekuensi. HRTFs yang digunakan dalam tes subyektif lokalisasi sumber bunyi, diukur secara individual untuk setiap subyek (terdiri dari 2 perempuan dan 3 laki-laki).

Setiap kumpulan HRTF dimodelkan dengan: a) suatu filter FIR 128-koefisien; b) suatu filter FIR 48-koefisien tanpa *preprocessing* apapun; dan c) suatu filter FIR 48-koefisien yang direkonstruksi dengan 6 vektor basis PCA. Sumber-sumber bunyi virtual diletakkan pada lokasi-lokasi dengan 12 azimut (arah-arah penunjuk waktu pada permukaan jam) dan 3 elevasi (-27° , 0° , 27°). Hasil-hasil tes pendengaran menunjukkan bahwa model-model filter orde rendah HRTF tanpa tahap *preprocessing* (skenario b) secara signifikan mengalami penurunan akurasi lokalisasi sumber bunyi. Tetapi penggunaan rekonstruksi HRTF dengan PCA, mengizinkan reduksi orde filter (skenario c), yang menyebabkan penurunan akurasi lokalisasi sumber bunyi yang tidak berarti, jika dibandingkan dengan akurasi lokalisasi sumber bunyi menggunakan suatu filter lengkap 128-koefisien. Di sini juga tidak dilakukan individualisasi pada model-model filter FIR orde 48-koefisien yang diperoleh dengan PCA.

Fokus penelitian Sodnik et al. [26] adalah studi mengenai variasi-variasi dari bobot-bobot PCA. Tujuannya adalah untuk menemukan beberapa keteraturan dalam variasi-variasi tersebut untuk posisi-posisi sumber bunyi spasial. Variasi-variasi tersebut dinyatakan dengan fungsi-fungsi matematik sederhana untuk posisi-posisi sumber bunyi di bidang horisontal depan dan lalu rekonstruksi data asli dilakukan. Jadi, di sini hanya diobservasi variasi bobot-bobot untuk azimut-azimut yang berbeda dengan elevasi tetap (0°). Basis data yang digunakan dalam penelitian mereka adalah MIT Media Lab HRTF *Library*. Tahap selanjutnya adalah rekonstruksi data dengan model yang diusulkan dan validasinya dengan beberapa tes lokalisasi sumber bunyi. Di sini dilakukan pemodelan HRTF amplitudo linier dengan PCA (dinamakan HRTF *magnitude* dalam disertasi ini). Rekonstruksi model-model HRTF amplitudo linier dilakukan dengan 4 vektor basis PCA beserta 4 bobotnya.

Variasi dari bobot PCA pertama, w_1 , bobot paling penting yang berkaitan dengan nilai eigen terbesar, dapat secara akurat dideskripsikan dengan suatu fungsi sinus setengah periode terhadap azimut. Fungsi sinus yang sesuai tersebut didapatkan dari suatu algoritma spesial yang berdasarkan pada kesalahan *least mean square* antara data asli dan data hasil rekonstruksi. Bobot-bobot kedua, w_2 , sampai bobot-bobot keempat, w_4 , diaproksimasi dengan suku-suku polinomial

azimut dengan orde yang berbeda. Variasi dari bobot yang lebih tinggi memerlukan orde polinomial yang lebih tinggi. Bobot kedua, w_2 , diaproksimasi dengan suku-suku polinomial azimut orde-8, bobot ketiga, w_3 , dengan suku-suku polinomial azimut orde-9 dan bobot keempat, w_4 , dengan suku-suku polinomial azimut orde-10.

Tes lokalisasi sumber bunyi dilakukan dengan mengukur waktu lokalisasi rata-rata yang diperlukan untuk menentukan azimut oleh subyek-subyek yang terlibat. Serangkaian tes dilakukan dengan menggunakan HRIR asli dan HRIR hasil rekonstruksi dengan fungsi-fungsi bobot-bobot dan vektor-vektor basis PCA. Hasil yang diperoleh adalah bahwa diperlukan waktu lokalisasi rata-rata yang hampir sama, saat digunakan HRIR hasil rekonstruksi PCA dan saat digunakan HRIR asli. Walaupun dicari hubungan antara bobot-bobot PCA sebagai fungsi azimut di bidang horisontal depan, tetapi di sini tidak dicari hubungan antara bobot-bobot tersebut dengan ukuran-ukuran antropometris subyek, atau dengan kata lain, model yang dihasilkan di sini tidak diindividualisasi.

Dalam penelitian ini, digunakan model statistik HRTF, dimana diharapkan dalam model statistik HRTF dapat dilakukan parameterisasi seluruh kumpulan HRTF dalam kawasan frekuensi, dan parameterisasi seluruh kumpulan HRIR dalam kawasan waktu. Dari hasil observasi basis data yang digunakan, diperoleh bahwa jumlah HRIR yang diukur sangat banyak pada posisi-posisi yang merepresentasikan suatu ruang dalam koordinat bola. Selain itu pengukuran HRIR dilakukan terhadap 45 subyek dengan memperhatikan data statistik dari ukuran-ukuran antropometris semua subyek. Subyek-subyek dipilih supaya mereka sedikit banyak merepresentasikan populasi manusia. Hal ini terlihat dari variasi yang memadai dari ukuran-ukuran antropometris yang termasuk di dalam basis data. Dengan kelengkapan data di atas, diharapkan model yang dihasilkan dapat merepresentasikan HRIR dari populasi manusia. Alasan kedua adalah bahwa jumlah fungsi basis yang terlibat dalam pemodelan dapat dipilih sesuai dengan tingkat akurasi yang diinginkan. Alasan ketiga bahwa bobot-bobot dari fungsi basis dapat dihubungkan dengan antropometri subyek agar diperoleh model HRTF individual. Alasan keempat adalah karena penggunaan PCA dapat

mengurangi jumlah memori yang diperlukan untuk implementasi HRIR individual bagi seorang pengguna dalam suatu rangkaian terintegrasi.

HRTF merupakan representasi proses fisik dari interaksi antara gelombang bunyi yang datang dan telinga, kepala dan *torso*. Oleh karena itu dapat dibuat hipotesis bahwa struktur HRTF berhubungan dengan bagian tubuh tempat penghamburan, dimensi tubuh dan orientasinya. Beberapa studi telah dilakukan untuk individualisasi HRTF yang dikaitkan dengan data antropometri subyek. Beberapa studi seperti pengelompokan (*clustering*) HRTF dan seleksi HRTF dari beberapa HRTF yang paling representatif [30], representasi fungsional dari HRTF menggunakan model *spatial feature extraction and regularization* [24], suatu model struktural untuk komposisi dan dekomposisi HRTF dengan memasukkan parameter-parameter yang berhubungan dengan daun telinga, kepala, *torso* dan pundak dalam model mereka [15-19] dan khususnya percobaan-percobaan dengan pengembangan suatu model fungsional yang menghubungkan ukuran-ukuran morfologikal dengan HRTF [31] dan dengan penskalaan HRTF [32,33] telah disarankan. Hipotesis yang dibuat dalam beberapa hal berlaku, walaupun suatu lokalisasi sempurna yang ekuivalen dengan lokalisasi HRTF orang itu sendiri, tidak diperoleh dengan HRTF orang lain yang dimodifikasi.

Penelitian dari Middlebrooks [32,33] berdasarkan pada penskalaan frekuensi. Middlebrooks mengamati bahwa jika telinga diperbesar, HRTF akan mempertahankan bentuknya tetapi akan tergeser menuju frekuensi-frekuensi lebih rendah pada sumbu frekuensi. Karena pendengar dianggap mengindera elevasi sumber bunyi dari posisi-posisi puncak dan lembah dalam spektrum bunyi yang datang, penggunaan HRTF dari telinga yang diperbesar akan mengakibatkan bias sistematis dalam estimasi elevasi. Pendengaran oleh telinga orang-orang yang berbeda akan jauh berbeda daripada sekedar hanya suatu penskalaan sederhana. Perubahan kecil yang kelihatannya tidak signifikan dalam struktur telinga akan menyebabkan perubahan-perubahan dramatis dalam HRTF.

Suatu pendekatan tengah yang dilakukan oleh Zotkin et al. [34] merupakan suatu usaha untuk menyeleksi HRTF yang paling cocok dengan HRTF pendengar dari suatu basis data HRTF yang ada dan menggunakannya untuk sintesis audio virtual; yaitu dengan membuat HRTF semi personal. Mereka

melakukan studi tentang hipotesis bahwa struktur HRTF berhubungan dengan ukuran-ukuran telinga. HRTF yang paling cocok terhadap seorang pendengar dicari dengan cara memotret telinga pendengar tadi, lalu mengukur ukuran-ukuran telinga dari foto yang diperoleh. Kemudian minimasi ukuran jarak tertentu antara ukuran-ukuran telinga pendengar terhadap ukuran-ukuran telinga subyek-subyek dalam basis data dilakukan, untuk menemukan HRTF yang paling sesuai dari dalam basis data HRTF dari 43 subyek. Metode individualisasi HRTF ini memerlukan sistem pencitraan tambahan untuk memotret telinga seorang pendengar dan menghitung ukuran-ukuran telinga dari citra fotonya. Tetapi HRTF yang digunakan tetap non individual, yakni HRTF milik seorang subyek dalam CIPIC HRTF Database.

Shin dan Park [35] telah mengusulkan metode individualisasi HRIR berdasarkan pengaturan subyektif terhadap respon-respon daun telinga dari HRIR pulsa langsung dalam kawasan waktu. Pemodelan HRIR di sini dilakukan dengan metode statistikal PCA dengan 4 atau 5 fungsi basis terhadap CIPIC HRTF Database. Proses individualisasi untuk memperoleh respon daun telinga yang sesuai adalah dengan mengizinkan seorang subyek mengatur bobot pada setiap fungsi basis. Mereka menunjukkan suatu peningkatan kinerja untuk persepsi elevasi dengan serangkaian tes pendengaran subyektif. Penelitian mereka hanya berfokus pada respon daun telinga (hanya mengambil 0,2 ms dari sekitar 2 ms HRIR). Walaupun respon-respon daun telinga penting bagi pendengar untuk mengindera posisi bunyi, respon-respon pundak dan torso juga mempengaruhi faktor-faktor bunyi direksional terutama perbedaan posisi atas dan bawah.

Pendekatan lain dalam individualisasi model HRTF menggunakan PCA dilakukan oleh Hwang et al. [36,37]. Mereka mengusulkan suatu metode individualisasi HRIR pulsa langsung dengan pengaturan bobot-bobot dari tiga fungsi basis dominan berdasarkan tiga deviasi standar bobot yang terbesar pada setiap elevasi, secara subyektif dan manual. Hasil-hasil tes pendengaran subyektif menunjukkan keterbalikan depan-belakang dan persepsi vertikal dapat ditingkatkan dengan HRIR pulsa langsung yang diindividualisasi. Di sini seluruh HRIR pulsa langsung di bidang median dari CIPIC HRTF Database (kecuali 45 HRIR pada 230,625°) digunakan dalam analisis tunggal. Jadi semua HRIR pulsa

langsung di bidang median menggunakan bersama satu kumpulan fungsi basis yang sama. Fungsi-fungsi basis tidak hanya mencakup variasi antar-individu tetapi juga antar-elevasi. Respon-respon sepanjang 1,5 ms (67 sampel dengan $f_s = 44,1$ kHz) dari HRIR asli basis data sejak kedatangan pulsa langsung dalam HRIR, yang mengandung efek-efek daun telinga, kepala, pundak dan *torso*, dimasukkan dalam PCA. Model-model HRIR pulsa langsung direkonstruksi dengan menggunakan 12 fungsi basis PCA.

Hu et al. [38] memodelkan HRTF dengan menggunakan PCA dan mengikuti pendekatan yang dilakukan oleh Wightman dan Kistler [21]. Mereka menggunakan 10 fungsi basis untuk merekonstruksi DTF. Dalam penelitian mereka, dibangun suatu model *multiple linear regression* (MLR) antara parameter-parameter karakteristik DTF, berupa bobot-bobot PCA, dan ukuran-ukuran antropometris pada posisi-posisi bunyi di bidang horisontal. Kemudian koefisien-koefisien regresi MLR dicari dengan estimasi *least-square* sebagai fungsi regresi dari ukuran-ukuran antropometris dan bobot-bobot PCA. Penelitian lanjutan dari Hu et al. [39] adalah mengubah frekuensi-frekuensi tengah dua lembah pinna yang signifikan dari estimasi respon *log-magnitude* dalam [38], dengan frekuensi-frekuensi tengah yang diestimasi dari ukuran-ukuran pinna subyek.

Dalam penelitian mereka [38,39], ukuran-ukuran antropometris yang berjumlah 27 dipilih dengan menggunakan analisis korelasi dalam dua tahap, sehingga pada akhir proses terpilih 8 ukuran sebagai variabel bebas dari model MLR. Pada tahap pertama, pemilihan ukuran-ukuran antropometris untuk individualisasi dilakukan dengan analisis korelasi antara sejumlah besar DTF dan seluruh 27 ukuran antropometris, sehingga terseleksi 18 ukuran antropometris. Pada tahap kedua, analisis korelasi dilakukan di antara 18 ukuran terseleksi, sehingga pada akhirnya terseleksi 8 ukuran antropometris yang berkorelasi kuat satu sama lain, yaitu x_1 , x_3 , x_{12} , d_1 , $d_3 - d_6$. Definisi dari ukuran-ukuran antropometris dapat dilihat dari Tabel 2.1 dan Gambar 2.11.

Parameter untuk mengukur kinerja model HRTF adalah nilai distorsi spektral. Parameter ini menunjukkan jarak antara HRTF basis data dan model HRTF *log-magnitude*. Nilai-nilai distorsi spektral dari model-model subyek 003

dan subyek 163 dari CIPIC HRTF *Database* tidak melebihi 3,5 dB. Kinerja dari model HRTF juga diukur dari tes pendengaran subyektif dalam menentukan posisi sumber bunyi. Hasil tes pendengaran dari lima subyek menunjukkan bahwa *correct rate* yang lebih besar diperoleh jika stimuli disintesis menggunakan model HRTF berdasarkan ukuran-ukuran antropometris miliknya sendiri daripada menggunakan HRTF non-individual. Dari hasil pengamatan, ternyata pemilihan ukuran-ukuran antropometris untuk individualisasi model HRTF dapat dilakukan dengan analisis korelasi antara sejumlah DTF dan ukuran-ukuran antropometris.

1.3 Identifikasi Masalah

Berdasarkan fakta-fakta dalam latar belakang, dapat disimpulkan bahwa terdapat beberapa permasalahan pada pemodelan HRTF/HRIR dan individualisasi model HRTF/HRIR yang dapat dikemukakan di sini, yaitu:

1. Terdapat beberapa pendekatan untuk memodelkan HRTF dalam kawasan frekuensi maupun HRIR dalam kawasan waktu. Setiap pendekatan memiliki masing-masing kelebihan dan kelemahan sendiri, sehingga perlu dipilih pendekatan pemodelan dengan performansi yang memadai.
2. Individualisasi model-model HRIR dalam usaha memperoleh HRIR individual masih sulit dilakukan karena variabilitas HRIR terhadap antropometri subyek yang masih belum terungkap secara detail. Di dalam metode pemodelan statistikal PCA untuk pemodelan HRIR, hubungan antara bobot-bobot fungsi-fungsi basis baik terhadap posisi sumber bunyi (azimut dan elevasi) maupun terhadap antropometri subyek juga belum tereksploitasi dengan menyeluruh.
3. Kelemahan individualisasi HRTF/HRIR sebelumnya:
 - a) Model *Clustering* dan pemilihan HRIR dari basis data hanya menggunakan HRIR non-individual, hanya mencari subyek yang cocok antropometrinya dengan pendengar [30,34].
 - b) Model HRTF semi personal dari Zotkin et al. [34] menyeleksi HRTF dari semua HRTF basis data berdasarkan ukuran-ukuran daun telinga antara seorang pendengar dengan subyek-subyek basis data. Jadi HRTF yang digunakan adalah milik seorang subyek dari basis data atau HRTF non individual.

- c) Model struktural dari CIPIC *Interface Laboratory* terlalu sederhana memodelkan HRTF sebagai fungsi dari antropometri. Efek-efek psikoakustik penting akan hilang dalam model HRTF sederhana. Selain itu, desain filter-filter yang merepresentasikan efek setiap struktur antropometri dilakukan secara terpisah, sehingga mengabaikan interaksi antara gelombang-gelombang yang dihamburkan dari berbagai struktur yang berbeda [15-19].
- d) Model penskalaan frekuensi HRTF sesuai dengan ukuran telinga [32,33] memiliki kelemahan, yaitu bahwa pendengaran oleh telinga orang-orang yang berbeda akan jauh berbeda daripada sekedar hanya suatu penskalaan sederhana serta perubahan kecil yang kelihatannya tidak signifikan dalam struktur telinga akan menyebabkan perubahan-perubahan dramatis dalam HRTF.
- e) Model HRIR dengan PCA dari Shin dan Park [35] hanya memodelkan respon daun telinga, tidak mencakup pengaruh kepala dan *torso*. Pemodelan dilakukan dengan sampel-sampel yang merepresentasikan respon daun telinga pada respon awal HRIR dan hanya pada sumber bidang median. Pengaturan bobot-bobot PCA dilakukan secara subyektif.
- f) Hwang et al. [36,37] memodelkan HRIR pulsa langsung dengan PCA hanya pada bidang median, diambil 67 sampel (1,5 ms) dengan *delay* dibuang. Pengaturan bobot-bobot PCA dilakukan secara subyektif.
- g) Hu et al. [38,39] memodelkan nilai *log-magnitude* HRTF ($20 \log_{10}$ dari HRTF *magnitude*) menggunakan PCA di bidang horisontal. Individualisasi model HRTF tadi dilakukan menggunakan MLR antara bobot-bobot PCA dan 8 ukuran antropometris. Pemilihan ukuran-ukuran antropometris untuk individualisasi dilakukan dengan analisis korelasi antara sejumlah besar DTF dan seluruh ukuran antropometris. Tidak diungkapkan parameter-parameter apa dari DTFs yang diamati.

1.4 Batasan Masalah

Di dalam pengolahan data HRTF seperti yang telah dirumuskan pada identifikasi masalah diperlukan batasan-batasan sebagai berikut :

- a. Bahasa pemrograman yang digunakan adalah Matlab versi 7.0.
- b. Basis data HRTF yang digunakan adalah CIPIC HRTF *Database* dari CIPIC *Interface Laboratory* di *University of California at Davis*, Amerika Serikat. Basis data ini diunduh dari situs http://interface.cipic.ucdavis.edu/CIL_html/CIL_HRTF_database.htm.
- c. Ukuran-ukuran antropometris subyek yang digunakan berdasarkan definisi dari CIPIC *Interface Laboratory*.
- d. Dalam individualisasi model HRTF/HRIR hanya digunakan beberapa ukuran antropometris yang dominan dalam persepsi bunyi yang alami.

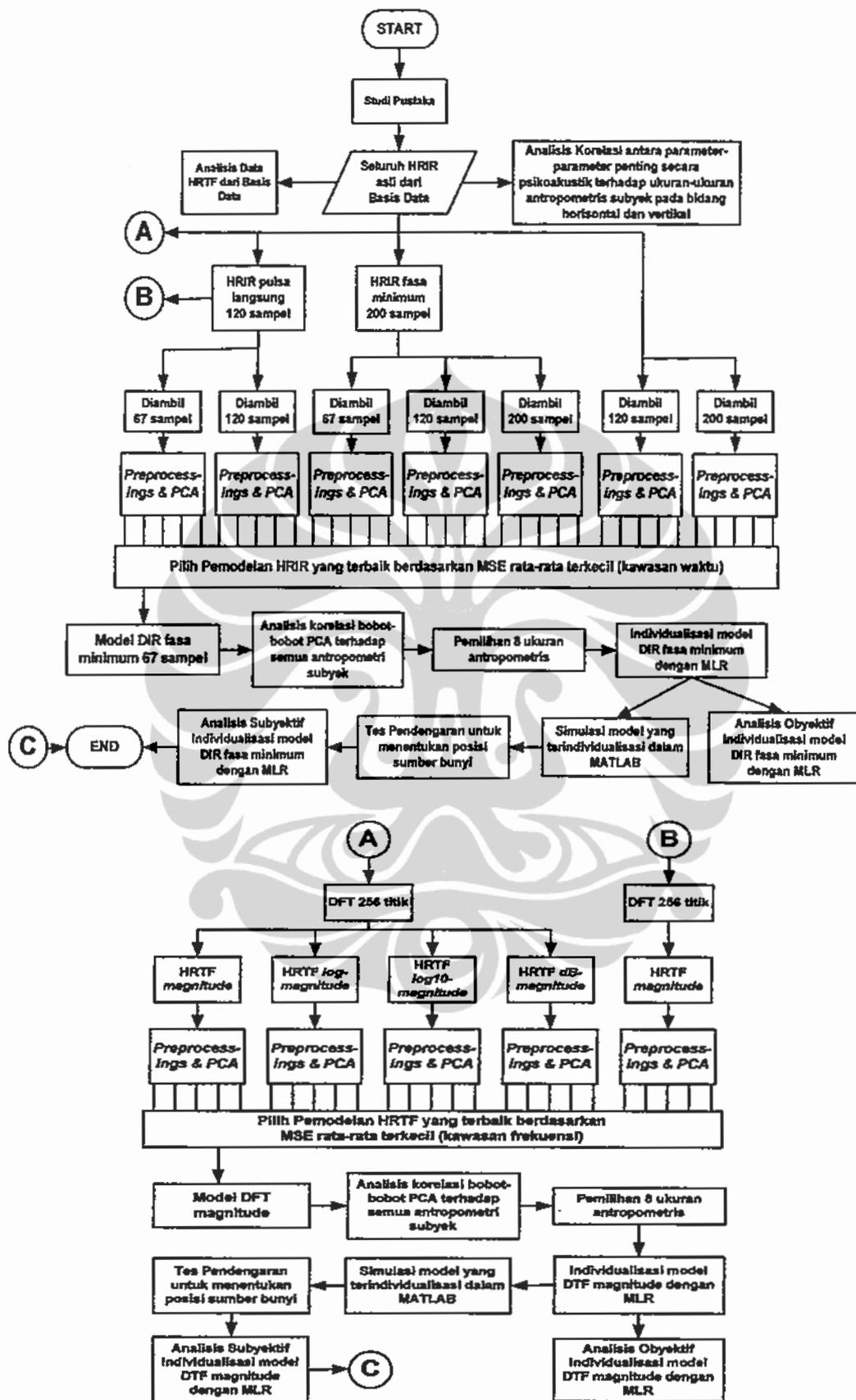
1.5 Tujuan Penelitian

Tujuan penelitian adalah untuk memodelkan HRTF dan HRIR menggunakan PCA, kemudian mengindividualisasi model HRTF dan HRIR tersebut. Individualisasi model-model dilakukan dengan mengatur bobot-bobot PCA yang berjumlah sesedikit mungkin secara otomatis, berdasarkan ukuran-ukuran antropometris subyek, sehingga dihasilkan efek-efek psikoakustik penting untuk memberikan kesan bunyi alami.

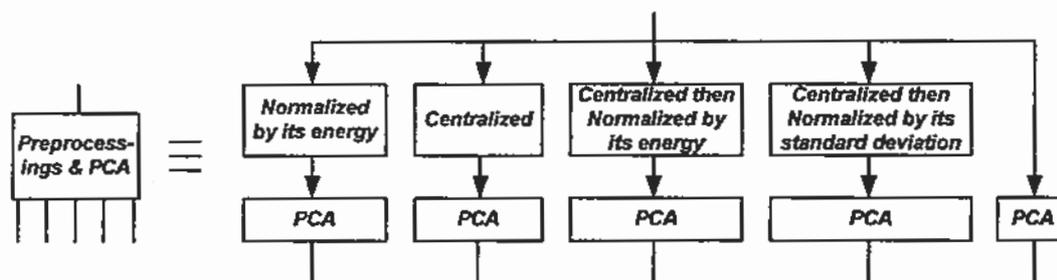
Kinerja sistem yang diteliti dinyatakan dengan parameter *mean-square error* (MSE) secara obyektif serta secara subyektif dilakukan tes pendengaran dalam menentukan posisi sumber bunyi.

1.6 Ruang Lingkup Penelitian

Penelitian ini difokuskan pada individualisasi model HRIR di kawasan waktu maupun model HRTF di kawasan frekuensi menggunakan PCA dan MLR. Deskripsi lengkap mengenai penelitian ini dapat dilihat pada Gambar 1.1. Pemodelan HRIR menggunakan PCA diimplementasikan dengan menggunakan semua HRIR asli dari CIPIC HRTF *Database*. HRIR asli mengandung waktu tunda pulsa langsung yang diperlukan bunyi untuk merambat dari sumber bunyi ke gendang telinga. Dua pengolahan awal terhadap HRIR asli yaitu menghilangkan waktu tunda tersebut dari HRIR asli dengan hasil berupa HRIR pulsa langsung / *direct pulse* HRIR ($HRIR_{dp}$) dan merekonstruksi HRIR fasa minimum / *minimum-phase* HRIR ($HRIR_{mp}$) dari HRIR asli menggunakan analisis *cepstral*.



Gambar 1.1. Diagram Alir Penelitian



Gambar 1.2. Berbagai Pengolahan Awal dalam Pemodelan

Berikutnya dilakukan penentuan jumlah sampel (67, 120 atau 200 sampel) dari masing-masing jenis HRIR yang diinginkan untuk pemodelan HRIR yang bersangkutan. Jumlah sampel HRIR berkaitan dengan respon masing-masing bagian tubuh yang terlibat dalam menghasilkan HRIR. Sampel-sampel awal HRIR merupakan respon yang diperoleh sebagai akibat difraksi gelombang bunyi oleh daun telinga. Semakin banyak sampel HRIR, semakin banyak bagian tubuh yang terlibat dalam menghasilkan HRIR, yakni dimulai dari respon bagian tubuh yang terdekat dengan gendang telinga yaitu daun telinga sampai dengan respon dari *torso*, bagian tubuh yang terjauh yang tercakup dalam HRIR.

Kemudian terhadap setiap jenis HRIR di atas dilakukan 4 jenis pengolahan awal tahap berikutnya, sehingga dihasilkan sekumpulan HRIR berikut:

- a. HRIR : HRIR asli;
- b. $HRIR_{nc}$: HRIR yang dinormalisasi oleh energinya;
- c. *Direct Impulse Response* (DIR): respon impuls yang diperoleh dari pengurangan *global mean* seluruh HRIR asli dari setiap HRIR asli;
- d. DIR_{nc} : DIR yang dinormalisasi oleh energinya;
- e. DIR_{sz} : DIR yang dinormalisasi oleh deviasi standar HRIR.

Setiap kelompok HRIR dari butir a sampai butir e digunakan masing-masing untuk menghitung matriks kovarians S di dalam pemodelan PCA, dimana pemodelan masing-masing kelompok HRIR menggunakan 10 vektor basis. Semua jenis pengolahan awal ini dapat dilihat dalam Gambar 1.2 bagian atas.

Langkah berikutnya adalah perhitungan MSE rata-rata dari 35 kumpulan hasil pemodelan HRIR dalam kawasan waktu dengan PCA. Pengolahan awal dan jenis data HRIR yang menghasilkan MSE rata-rata terkecil dipilih sebagai kumpulan data pemodelan HRIR untuk digunakan lebih lanjut dalam

individualisasi model HRIR dengan menggunakan MLR. Dari sini terpilih DIR fasa minimum yang terdiri dari 67 sampel [40]. Hasil perhitungan nilai-nilai MSE dari 35 kumpulan hasil pemodelan HRIR dapat dilihat pada Tabel 3.1.

Individualisasi model DIR fasa minimum 67 sampel dilakukan untuk kumpulan data HRIR di bidang horisontal [41], dengan menggunakan metode individualisasi MLR. Dalam [41], dianalisis korelasi bobot-bobot PCA dari 10 vektor basis terhadap semua ukuran antropometris subyek. Pada awal penelitian juga telah dilakukan analisis korelasi antara parameter-parameter yang penting secara psikoakustik yaitu, ITD_{max} , ILD_{max} , dan *pinna notch frequency* (f_{pn}); masing-masing terhadap semua ukuran antropometris subyek. Berdasarkan hasil dari keempat jenis analisis korelasi ini, dipilih 8 ukuran antropometris yang digunakan dalam MLR. Kemudian dilakukan pemodelan bobot-bobot PCA sebagai fungsi 8 ukuran antropometris oleh MLR. MLR membangkitkan koefisien-koefisien regresi untuk memperoleh estimasi bobot-bobot dari masing-masing fungsi basis dalam PCA, berdasarkan 8 ukuran antropometris yang terseleksi. Dengan cara yang serupa, individualisasi model DIR fasa minimum 67 sampel dilakukan untuk kumpulan data HRIR di bidang median [43]. Untuk menguji kinerja individualisasi model HRIR secara subyektif, dilakukan tes pendengaran untuk menentukan posisi-posisi sumber bunyi.

Individualisasi model HRTF juga dilakukan di kawasan frekuensi. Untuk memperoleh HRTF, dilakukan transformasi Fourier diskrit (*Discrete Fourier Transform / DFT*) 256-titik terhadap HRIR asli dan HRIR pulsa langsung. Keluaran DFT dari suatu HRIR berupa HRTF kompleks. Pertama-tama, dihitung respon *magnitude* dari HRTF kompleks hasil DFT dari HRIR pulsa langsung. Tetapi beberapa pengolahan awal dilakukan terhadap HRTF kompleks hasil DFT HRIR asli. Pertama dihitung respon *magnitude* dari HRTF kompleks tersebut. Pengolahan awal kedua adalah dihitung nilai logaritma natural dari respon *magnitude* di atas, yang disebut HRTF *log-magnitude*. Pengolahan awal ketiga adalah dihitung harga logaritma-10 dari respon *magnitude* HRTF di atas, yang disebut dengan HRTF *log10-magnitude*. Terakhir, dihitung harga dB ($20 \cdot \log_{10} |HRTF|$) dari HRTF kompleks, yang disebut HRTF dB. Fasa dari semua bentuk HRTF dimodelkan sebagai fasa minimum. Sebelum dimasukkan ke dalam

PCA, terhadap kelima jenis HRTF di atas dilakukan pengolahan awal tahap berikutnya, seperti pada data HRIR, yaitu:

- a. HRTF asli : pasangan DFT dari HRIR asli;
- b. $HRTF_{nc}$: HRTF yang dinormalisasi oleh energinya;
- c. *Direct Transfer Function* (DTF): fungsi alih yang diperoleh dari pengurangan *global mean* seluruh HRTF asli dari setiap HRTF asli;
- d. DTF_{nc} : DTF yang dinormalisasi oleh energinya;
- e. DTF_{nz} : DTF yang dinormalisasi oleh deviasi standar HRTF.

Setiap kelompok HRTF dari butir a sampai butir e digunakan masing-masing untuk menghitung matriks kovarians S di dalam pemodelan PCA. Seluruh bentuk data HRTF ini masing-masing menghasilkan 10 vektor basis beserta bobot-bobotnya masing-masing. Pemodelan masing-masing kelompok HRTF menggunakan 10 vektor basis dan bobot-bobot tersebut.

Dari pemodelan PCA menggunakan beberapa bentuk data HRTF tadi, dipilih bentuk mana yang menghasilkan kesalahan rata-rata pemodelan terkecil di seluruh posisi sumber bunyi dan subyek. Parameter yang digunakan secara obyektif untuk mengukur kesalahan pemodelan juga sama seperti pada kawasan waktu, yaitu MSE. Kumpulan data HRTF yang menghasilkan pemodelan PCA dengan MSE rata-rata terkecil adalah DTF *magnitude* (pengolahan awal butir c) [40]. Setelah diperoleh model terbaik dengan kombinasi berbagai pengolahan awal dan bentuk HRTF, dilakukan pemodelan bobot-bobot PCA sebagai fungsi antropometri subyek. Pemodelan bobot-bobot tersebut diperoleh dengan metode MLR, seperti pada HRIR. Simulasi pemodelan HRTF dan HRIR dilakukan dengan menggunakan perangkat lunak MATLAB.

1.7 Kontribusi Penelitian

Kontribusi pertama dari penelitian ini adalah mendapatkan kombinasi terbaik dari beberapa pengolahan awal terhadap HRIR dan HRTF, sehingga dihasilkan jenis data HRIR terbaik dan jenis data HRTF terbaik dalam pemodelan PCA yang optimal dibandingkan dengan jenis data HRIR dan HRTF lain, seperti yang dijelaskan dalam subbab 1.6. Dalam penelitian ini, jenis data HRIR yang menghasilkan pemodelan PCA terbaik dengan 10 vektor basis adalah DIR fasa minimum dengan 67 sampel seperti dijelaskan dalam subbab 3.4 dan [40]. Hwang

et al. [36,37] memodelkan DIR pulsa langsung 67 sampel menggunakan PCA dengan 1–20 vektor basis di bidang median. Pemodelan mereka dengan menggunakan 10 vektor basis di bidang median menghasilkan MSE rata-rata sebesar sekitar 6,67% [36]. Dengan menggunakan data DIR fasa minimum 67 sampel dalam penelitian ini daripada DIR pulsa langsung 67 sampel dalam [36,37] dan dengan menggunakan *setup* PCA yang sama seperti dalam [36], maka model-model DIR fasa minimum 67 sampel, menggunakan juga 10 vektor basis di bidang median dalam penelitian ini, memberikan MSE rata-rata sebesar 5,31% [40] yang lebih baik daripada hasil Hwang et al.

Jenis data HRTF yang menghasilkan pemodelan PCA terbaik adalah DTF *magnitude* dengan 256 komponen frekuensi, seperti dijelaskan dalam subbab 3.5 dan [40]. Seperti dijelaskan dalam subbab 1.2, Leung dan Carlile [28] membandingkan berbagai teknik pengolahan awal (HRIR linier, HRTF *log-magnitude*, HRTF amplitudo linier, HRTF pasangan riil-kompleks) dalam hal efisiensi kompresi berdasarkan suatu analisis varians. HRTF amplitudo linier menghasilkan nilai-nilai MSE yang terkecil dari antara jenis data yang lain. HRTF amplitudo linier yang disebutkan di sini adalah sama dengan HRTF *magnitude* yang digunakan dalam penelitian ini. Jadi hasil-hasil dalam [28] mengkonfirmasi bahwa HRTF *magnitude* adalah jenis data HRTF yang terbaik dalam pemodelan PCA di kawasan frekuensi dibandingkan jenis-jenis data HRTF lain. Tetapi Leung dan Carlile menggunakan basis data yang berbeda daripada yang digunakan dalam penelitian ini.

Kontribusi kedua dari penelitian ini adalah mendapatkan metode yang sistematis untuk menyeleksi ukuran-ukuran antropometris terbaik yang memberikan efek-efek psikoakustik penting dalam persepsi bunyi alami dan memberikan individualisasi model HRIR dan model HRTF yang optimal [41]. Hu et al. [38,39] melakukan seleksi ukuran-ukuran antropometris dalam dua tahap. Tahap pertama adalah analisis korelasi antara sejumlah besar DTF dan ukuran-ukuran antropometris. Tahap kedua adalah analisis korelasi di antara 18 ukuran antropometris yang terseleksi dari tahap pertama. Seleksi ukuran-ukuran antropometris dalam penelitian ini dilakukan secara lebih spesifik daripada Hu et al., yakni dengan memperhatikan parameter-parameter psikoakustik yang penting

dalam mempengaruhi pendengaran manusia, yaitu ITD_{max} , ILD_{max} , f_{pn} , dimana parameter-parameter tersebut diekstrak dari respon-respon HRIR dan spektrum-spektrum HRTF yang relevan. Kemudian dilakukan analisis korelasi antara parameter-parameter psikoakustik penting tadi, masing-masing terhadap ukuran-ukuran antropometris yang didefinisikan dalam Tabel 2.1. Delapan ukuran antropometris yang memiliki korelasi-korelasi yang kuat dengan parameter-parameter tersebut dipilih untuk digunakan dalam individualisasi. Seleksi 8 ukuran antropometris juga didasarkan pada analisis korelasi antara bobot-bobot vektor basis PCA terhadap ukuran-ukuran antropometris, yang belum pernah dilakukan peneliti-peneliti lain. Analisis-analisis korelasi ini berbeda dengan analisis-analisis korelasi yang dilakukan Hu et al. Delapan ukuran antropometris yang digunakan oleh Hu et al. juga berbeda dengan 8 ukuran antropometris yang digunakan dalam penelitian ini. Dengan menggunakan metode individualisasi dan model HRTF yang sama seperti dalam [38], tetapi digunakan 8 ukuran antropometris dalam penelitian ini, dihasilkan kinerja yang sebanding daripada jika digunakan 8 ukuran antropometris dari Hu et al.

Kontribusi ketiga dari penelitian ini adalah mendapatkan metode baru dalam individualisasi model-model HRIR fasa minimum yang lebih sederhana dan metode baru dalam individualisasi model-model HRTF *magnitude*, untuk posisi-posisi sumber bunyi baik di bidang horisontal maupun di bidang median, berdasarkan MLR antara bobot-bobot PCA yang bersangkutan dan 8 ukuran antropometris yang terseleksi [41-43].

Shin dan Park [35] hanya memodelkan respon-respon pinna (HRIR pulsa langsung 0,2 ms) di bidang median menggunakan PCA dengan 4 atau 5 fungsi basis. Individualisasi model HRIR yang dihasilkan dilakukan dengan pengaturan setiap bobot secara subyektif dan manual. Hwang et al. [36,37] memodelkan HRIR pulsa langsung 1,5 ms di bidang median menggunakan PCA dengan 12 fungsi basis. Individualisasi model HRIR mereka dilakukan dengan pengaturan tiga bobot PCA yang dominan berdasarkan tiga deviasi standar bobot yang terbesar untuk setiap elevasi, secara subyektif dan manual. Terlihat jelas perbedaan dari kedua penelitian di atas dengan metode individualisasi dalam penelitian ini. Di dalam penelitian ini, dilakukan suatu individualisasi model

HRIR fasa minimum secara otomatis, berdasarkan 8 ukuran antropometris dengan metode MLR, tetapi Shin dan Park serta Hwang et al. melakukan individualisasi model HRIR pulsa langsung secara subyektif dan manual. Mereka hanya mengindividualisasi model mereka untuk posisi-posisi bunyi di bidang median, tetapi dalam penelitian ini di bidang median dan di bidang horisontal.

Hu et al. [38,39] mengindividualisasi model HRTF *log-magnitude* (20 log10 dari HRTF *magnitude*) dengan MLR. Tetapi dalam penelitian ini diindividualisasi model HRTF *magnitude* dengan MLR. Berdasarkan perbandingan pemodelan PCA dari berbagai jenis data HRTF[40], didapatkan bahwa HRTF *magnitude* dalam penelitian ini, menghasilkan pemodelan PCA yang lebih baik daripada HRTF *log-magnitude* dari Hu et al. Selain itu HRTF *magnitude* dalam penelitian ini lebih mudah dihitung daripada HRTF *log-magnitude* dari Hu et al. Individualisasi model HRIR fasa minimum dalam bagian lain penelitian ini juga lebih sederhana daripada individualisasi model HRTF *log-magnitude* dari Hu et al., karena tidak melibatkan proses DFT dan invers DFT dalam merekonstruksi model HRIR individual bagi seorang pendengar.

1.8 Sistematika Penulisan

Sistematika penulisan dari proposal penelitian ini adalah sebagai berikut:

BAB 1. Pendahuluan

Dalam bagian ini dibahas latar belakang pemilihan topik penelitian, penelitian-penelitian pemodelan HRTF dan individualisasi model HRTF, identifikasi masalah, batasan masalah, tujuan penelitian, ruang lingkup penelitian, kontribusi penelitian, serta sistematika penulisan.

BAB 2. Psikoakustik dari Pendengaran Spasial

Di dalam bab ini dibahas sifat-sifat fisik bunyi, sistem koordinat yang digunakan, *azimuth cues*, *elevation cues*, *range cues*, reverberasi dan gema-gema, pengertian HRIR dan HRTF, basis data yang digunakan dalam pemodelan HRIR dan HRTF, serta visualisasi dan analisis struktur data HRTF.

BAB 3. Pemodelan *Head-Related Transfer Functions* Berdasarkan *Principal Components Analysis*

Di sini dipaparkan jenis-jenis pengolahan awal data HRIR, aproksimasi fasa minimum dari HRTF, metode pemodelan HRTF menggunakan PCA, hasil-hasil pemodelan PCA dari berbagai jenis data HRIR, analisis vektor-vektor basis PCA dari model HRIR fasa minimum, analisis kesalahan model HRIR fasa minimum, hasil-hasil pemodelan PCA dari berbagai jenis data HRTF, analisis vektor-vektor basis PCA dari model HRTF *magnitude*, serta analisis kesalahan model HRTF *magnitude*.

BAB 4. Metode Seleksi Ukuran-ukuran Antropometris

Di dalam bab ini, dibahas mengenai metode seleksi ukuran-ukuran antropometris yang dilakukan berdasarkan analisis korelasi antara ITD maksimum dan 27 ukuran antropometris, analisis korelasi antara ILD maksimum dan 27 ukuran antropometris, analisis korelasi antara frekuensi lembah pinna, f_{pn} , dan 10 ukuran pinna, serta analisis korelasi antara bobot-bobot vektor basis PCA dan 17 ukuran kepala dan *torso*.

BAB 5. Individualisasi Model *Head-Related Transfer Functions* Berdasarkan *Multiple Linear Regression*

Di dalam bab ini, dibahas model MLR dari bobot-bobot vektor basis PCA dengan prediktor antropometri subyek, metode individualisasi model HRIR fasa minimum yang diusulkan, analisis hasil individualisasi model HRIR fasa minimum, yang terdiri dari: hasil simulasi obyektif, diagram *pole-zero* model HRIR fasa minimum individual, hasil penentuan posisi bunyi; metode individualisasi model HRTF *magnitude* yang diusulkan, analisis hasil individualisasi model HRTF *magnitude*; dan analisis hasil individualisasi model HRTF *magnitude*.

BAB 6. Kesimpulan

Bab ini merupakan penutup dari pembahasan dan analisis yang dilakukan; berisi kesimpulan dan opsi pengembangan lanjut dari disertasi ini.

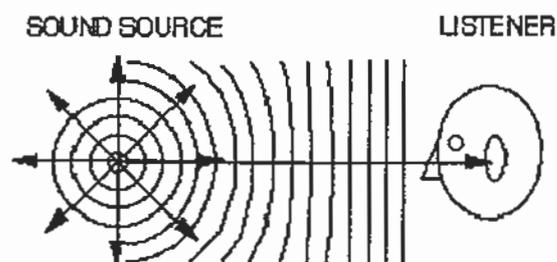
BAB 2

PSIKOAKUSTIK DARI PENDENGARAN SPASIAL

Semua komputer memberikan *output* berupa semacam bunyi peringatan ‘*beep*’ dan komputer yang dilengkapi multimedia dewasa ini mampu memberikan bunyi stereo dengan kualitas *compact disc*. Stereo konvensional dapat secara mudah menempatkan suatu bunyi pada setiap titik antara pengeras bunyi kiri dan kanan. Akan tetapi, dengan bunyi tiga dimensi yang sebenarnya, sumber dapat diletakkan di sebarang lokasi, yaitu kiri atau kanan, atas atau bawah, serta dekat atau jauh. Bab ini berisi penjelasan bagaimana hal ini dapat terjadi. Topik-topik yang dibahas adalah sifat-sifat fisik dari bunyi, sistem koordinat yang digunakan, *azimuth cues*, *elevation cues*, *range cues*, reverberasi dan gema-gema, *head-related impulse response*, basis data yang digunakan dalam pemodelan HRTF, serta visualisasi dan analisis struktur data HRTF.

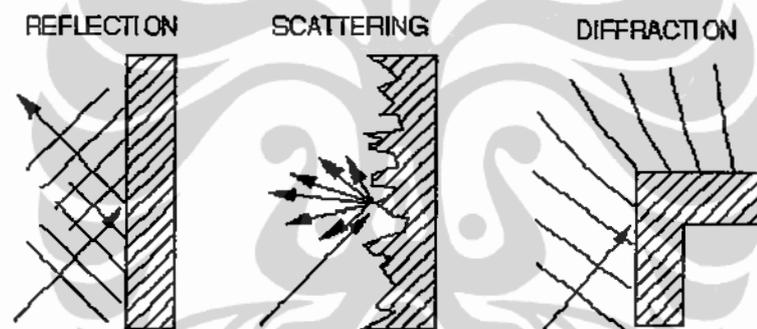
2.1 Sifat-sifat Fisik dari Bunyi

Dari segi pandang ilmu fisika, bunyi adalah suatu gelombang akustik yang dihasilkan jika suatu sumber yang bervibrasi, seperti pita suara manusia, mengganggu suatu medium elastis seperti udara. Ketika suatu gelombang bunyi mencapai gendang telinga pendengar, vibrasi-vibrasi tersebut ditransmisikan ke telinga bagian dalam atau *cochlea*, dimana perpindahan mekanik dikonversikan menjadi pulsa-pulsa syaraf yang kemudian dikirimkan ke otak dan akhirnya menghasilkan sensasi bunyi. Gambar 2.1 mengilustrasikan peristiwa ini.



Gambar 2.1. Gelombang Bunyi yang Ditransmisikan ke Telinga Pendengar

Pada Gambar 2.2 di bawah dapat dilihat peristiwa-peristiwa yang terjadi pada gelombang bunyi jika mengenai suatu permukaan dari medium padat. Dalam beberapa segi, gelombang-gelombang bunyi sangat mirip dengan gelombang elektromagnetik dan cahaya. Dalam suatu medium homogen, gelombang bunyi merambat dengan kecepatan konstan sebesar c (kecepatan bunyi dalam udara sebesar 343 ms^{-1}). Suatu sumber titik yang seragam meradiasikan gelombang-gelombang *spheric* yang memiliki amplitudo yang berkurang dengan bertambahnya jarak dari sumber bunyi. Gelombang-gelombang ini dipantulkan oleh permukaan-permukaan yang padat dan dihamburkan oleh permukaan-permukaan yang kasar. Suatu permukaan dikatakan halus jika ukuran ketidakteraturan (*irregularities*) adalah relatif kecil daripada panjang gelombang dan dikatakan kasar jika sebaliknya.



Gambar 2.2. Peristiwa-peristiwa Fisik dari Gelombang Bunyi

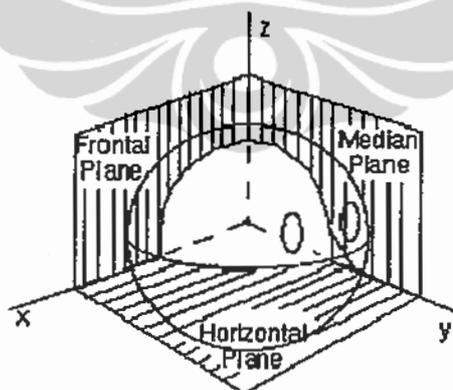
Gelombang-gelombang bunyi dibiarkan di sekitar obyek-obyek yang berada di dekatnya. Jika obyek itu relatif kecil terhadap panjang gelombang, obyek itu memiliki efek yang sangat kecil, dengan kata lain gelombang hanya melewatinya tanpa terganggu. Jika obyek itu besar, suatu bayangan bunyi (*sound shadow*) timbul di belakang obyek tersebut dan sejumlah energi yang signifikan dipantulkan kembali menuju sumber bunyi. Kemudian jika obyek itu hampir sama ukurannya dengan panjang gelombang, maka hal ini menjadi rumit dan pola-pola difraksi yang menarik muncul.

Telah diketahui bahwa manfaat dari telinga adalah untuk menuntun mata. Kemampuan dari sistem auditori untuk melokalisasi sumber-sumber bunyi adalah hanya satu komponen dari sistem penginderaan yang memiliki nilai ketahanan yang tinggi dan makhluk hidup telah menemukan banyak cara untuk mengekstrak

informasi arah dari bunyi. Walaupun misteri-misteri penginderaan masih tersisa, faktor-faktor utama telah diketahui sejak waktu yang lama, studi-studi psikologis yang cermat telah membangun seberapa akurat manusia dapat menentukan lokasi. Setiap orang yang ingin membangkitkan bunyi spasial, harus mengetahui hal-hal apa saja yang mempengaruhi sistem auditori manusia. Bagian ini membahas faktor-faktor utama yang mempengaruhi pendengaran spasial. Topik-topik yang dibahas selanjutnya adalah sistem koordinat yang digunakan, *azimut cues*, *elevation cues*, *range cues*, serta reverberasi dan gema.

2.2 Sistem Koordinat yang Digunakan

Untuk menyatakan lokasi suatu sumber bunyi relatif terhadap pendengar, suatu sistem koordinat diperlukan. Sebuah pilihan yang alami adalah sistem koordinat segi empat yang berpusat pada kepala seperti ditunjukkan Gambar 2.3 di bawah ini. Dalam sistem koordinat ini, sumbu-x hampir melewati telinga kanan, sumbu-y menunjukkan tepat ke depan dan sumbu-z adalah yang vertikal. Susunan koordinat seperti ini mendefinisikan tiga bidang standar yakni bidang horisontal atau xy, bidang frontal atau xz dan bidang median atau yz (juga sering disebut bidang *mid-sagittal*). Terlihat jelas bahwa bidang horisontal memisahkan bagian atas dan bawah, bidang frontal memisahkan bagian depan dan belakang serta bidang median memisahkan bagian kiri dan kanan.



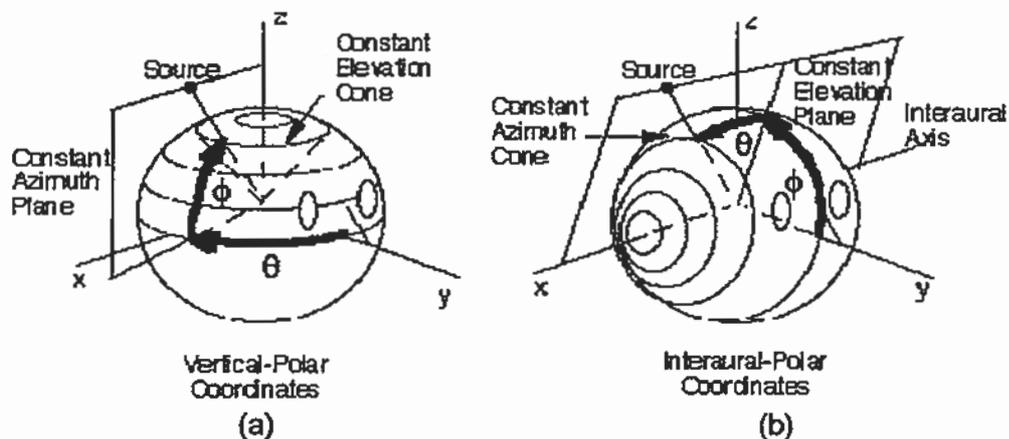
Gambar 2.3. Sistem Koordinat Segi Empat yang Berpusat pada Kepala

Karena kepala hampir berbentuk bola, suatu sistem koordinat bola biasanya digunakan. Di sini koordinat standar adalah azimut, elevasi dan jarak. Akan tetapi ada lebih dari satu cara untuk mendefinisikan koordinat-koordinat ini dan orang-orang yang berbeda mendefinisikannya secara berlainan. Sistem

koordinat *vertical-polar* yang terlihat pada Gambar 2.4a adalah yang paling populer. Di sini pertama-tama diukur azimuth θ yaitu sudut dari bidang median ke suatu bidang vertikal yang mengandung sumber bunyi dan sumbu-z dan kemudian baru diukur elevasi ϕ yakni sudut dari bidang horisontal. Dengan pilihan ini, permukaan-permukaan dengan azimuth konstan adalah bidang-bidang melalui sumbu-z dan permukaan-permukaan dengan elevasi konstan adalah kerucut-kerucut yang konsentris di sekitar sumbu-z.

Suatu alternatif yang penting adalah sistem koordinat *interaural-polar* seperti terlihat pada Gambar 2.4b. Di sini pada mulanya diukur elevasi sebagai sudut dari bidang horisontal ke suatu bidang yang melalui sumber bunyi dan sumbu-x, yang merupakan sumbu interaural; azimuth kemudian diukur sebagai sudut yang melalui bidang median. Dengan pilihan ini, permukaan-permukaan dengan elevasi konstan adalah bidang-bidang yang melalui sumbu interaural dan permukaan-permukaan dengan azimuth konstan adalah kerucut-kerucut yang konsentris dengan sumbu interaural.

Sistem *vertical-polar* jelas lebih tepat dalam menggambarkan sumber-sumber yang berada pada bidang horisontal, karena azimuth haruslah sudut antara -180° dan $+180^\circ$. Dengan sistem *interaural-polar*, azimuth selalu terletak antara -90° dan $+90^\circ$; hal yang mengejutkan di sini adalah bahwa perbedaan depan/belakang harus dispesifikasikan oleh elevasi, yang adalah 0° untuk sumber-sumber di depan bidang horisontal dan 180° atau -180° untuk sumber-sumber di belakang. Walaupun kurang lazim, sistem *interaural-polar* lebih memudahkan dalam mengekspresikan perbedaan-perbedaan interaural pada semua elevasi.

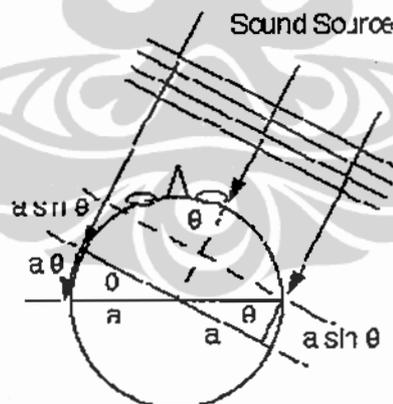


Gambar 2.4. Sistem-sistem Koordinat Bola [44]

2.3 Azimut Cues

Salah seorang pelopor dalam penelitian pendengaran spasial adalah John Strutt yang lebih dikenal sebagai Lord Rayleigh. Kurang lebih 100 tahun yang lalu, Lord Rayleigh mengembangkan Teori Duplex [45]. Menurut teori ini, terdapat dua *cue* utama untuk azimuth yakni selisih waktu *interaural* (ITD) dan selisih level *interaural* (ILD).

ITD merupakan perbedaan waktu antara kedatangan pulsa pertama dari telinga kiri dan kedatangan pulsa pertama dari telinga kanan, pada gendang telinga masing-masing dari posisi sumber bunyi yang sama. Posisi-posisi sumber bunyi pada bidang median memiliki ITD mendekati nol, dimana untuk kepala yang simetris sempurna tidak ada ITD pada bidang tersebut. Semakin jauh posisi sumber bunyi dari bidang median, maka semakin besar ITD yang diperoleh pada posisi tersebut. ITD maksimum, ITD_{max} , terjadi pada posisi-posisi ekstrem persis di sebelah salah satu telinga dengan elevasi $\phi=0^\circ$. Dengan demikian ITD merupakan fungsi dari azimuth pada bidang-bidang dengan elevasi tetap. ITD dapat juga dihitung dari waktu tunda hasil korelasi silang maksimum antara sepasang HRIR.



Gambar 2.5. Ilustrasi untuk Menentukan ITD [44]

Lord Rayleigh memiliki suatu penjelasan yang sederhana untuk ITD. Bunyi merambat dengan kecepatan sebesar c sekitar 343 ms^{-1} . Bayangkan suatu gelombang bunyi dari suatu sumber yang jauh mengenai suatu kepala *spheric* dengan jari-jari a dari suatu arah yang ditentukan oleh sudut azimuth θ . Jelaslah bahwa bunyi datang pada telinga kanan sebelum telinga kiri, karena gelombang itu harus merambat jarak ekstra sebesar $(a\theta + a \sin \theta)$ untuk mencapai telinga kiri.

Jika jarak ekstra itu dibagi dengan kecepatan bunyi c , formula sederhana untuk ITD berikut ini yang juga akurat diperoleh:

$$\text{ITD} = \frac{a}{c} (\theta + \sin \theta), \quad -90^\circ \leq \theta \leq 90^\circ \quad (2.1)$$

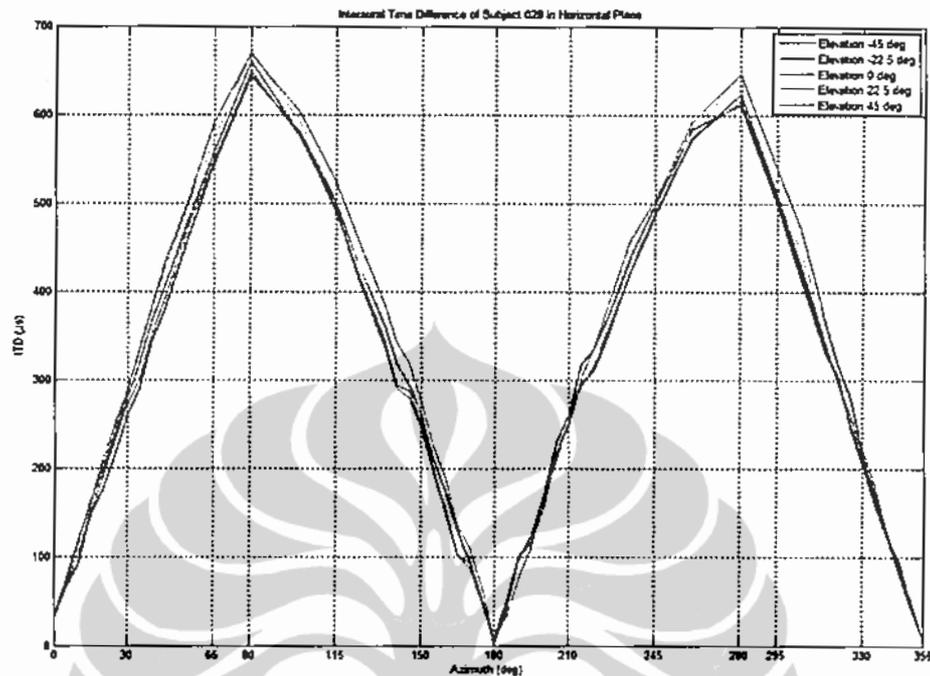
Jadi ITD adalah nol ketika sumber bunyi berada tepat di depan dan maksimum sebesar $(a/c)(\pi/2 + 1)$ ketika sumber bunyi berada pada satu sisi kepala. Hal ini menyatakan suatu selisih waktu kedatangan sekitar 0,7 ms untuk ukuran tipikal kepala manusia dan dapat dengan mudah diindera [45].

Perlu diperhatikan bahwa jika sistem koordinat *interaural-polar* digunakan dan memegang azimuth konstan, maka suatu nilai konstan untuk ITD diperoleh. Jadi terdapat suatu korespondensi satu-satu antara ITD dan kerucut dengan azimuth konstan, yang kadang-kadang disebut *cone of confusion*. Hal ini tidak terjadi untuk sistem *vertical-polar*. Akhirnya perhatikan bahwa ITD sendiri hanya menunjukkan bahwa sumber berada di suatu tempat pada *cone of confusion*. Informasi tentang ITD tidak cukup untuk menentukan apakah sumber terletak di depan atau belakang kepala.

Data ITD dari subyek 028 dari CIPIC HRTF *Database* pada beberapa bidang dengan elevasi tetap ($\phi = -45^\circ; -22,5^\circ; 0^\circ; 22,5^\circ; 45^\circ$) diilustrasikan pada Gambar 2.6 sebagai fungsi dari azimuth. Bidang-bidang tersebut berbentuk lingkaran yang terletak masing-masing di bawah bidang horisontal, pada bidang horisontal dan di atas bidang horisontal. Secara umum terlihat bahwa perubahan ITD terhadap azimuth menyerupai bentuk sinus. Jadi persamaan ITD sekali lagi dibuktikan di sini. ITD_{\max} pada setiap bidang terlihat pada azimuth $\theta = 80^\circ$ dan $\theta = 280^\circ$. Walaupun kelihatan di sini, bahwa urutan ITD_{\max} dari yang terkecil sampai terbesar diperoleh pada bidang $\phi = -45^\circ, -22,5^\circ, 0^\circ, 22,5^\circ$ dan 45° , tetapi urutan sistematis seperti ini tidak berlaku secara umum untuk semua elevasi. Terlihat bahwa perbedaan ITD per azimuth antar elevasi-elevasi yang berbeda adalah sangat kecil dan tidak melebihi 2 μs .

Lord Rayleigh juga mengamati bahwa gelombang-gelombang bunyi yang datang dibelokkan oleh kepala. Persamaan gelombang benar-benar diselesaikan untuk menunjukkan bagaimana suatu gelombang bidang dibelokkan oleh bola keras. Solusinya menunjukkan bahwa sebagai tambahan untuk selisih waktu,

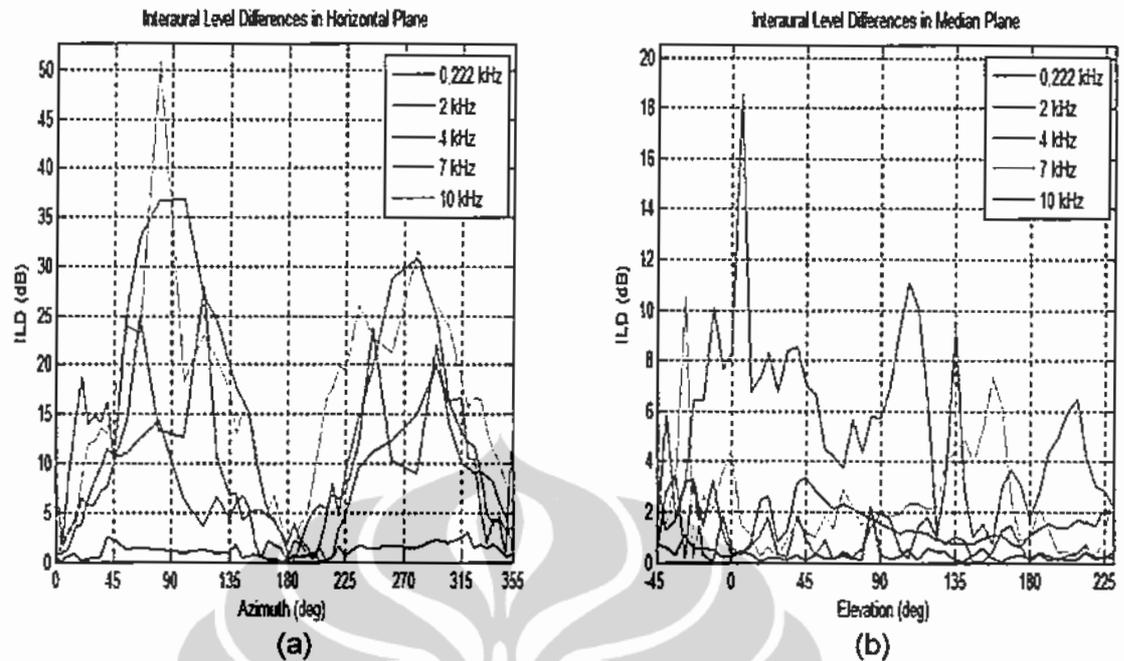
terdapat juga suatu selisih signifikan antara level-level sinyal pada kedua telinga yaitu ILD.



Gambar 2.6. ITD Subyek 028 pada Beberapa Bidang dengan Elevasi Tetap [2]

ILD didefinisikan sebagai perbedaan *level* atau *magnitude* dalam satuan dB di kawasan frekuensi antara sepasang HRTF *magnitude* yang berasal dari kedua telinga. Untuk suatu posisi sumber bunyi, diperoleh nilai ILD dari masing-masing komponen frekuensi (0 – 22050 Hz). Seperti yang diharapkan, ILD sangat tergantung terhadap frekuensi. Pada frekuensi-frekuensi rendah, dimana panjang gelombang bunyinya relatif panjang terhadap diameter kepala, terdapat selisih yang sangat kecil dalam tekanan bunyi pada kedua telinga. Tetapi pada frekuensi-frekuensi tinggi, dimana panjang gelombangnya pendek, akan terdapat suatu selisih sebesar 20 dB atau lebih. Hal ini disebut efek bayangan kepala (*head-shadow effect*), dimana telinga jauh berada di dalam bayangan bunyi dari kepala.

ILD biasanya dianalisis untuk komponen frekuensi tertentu di sepanjang bidang horisontal dan di sepanjang bidang median. Gambar 2.7 di bawah menunjukkan nilai-nilai ILD dari empat komponen frekuensi yang berubah pada bidang horisontal dan bidang median.



Gambar 2.7. ILD Subyek 028 pada Bidang Horisontal dan Bidang Median [2]

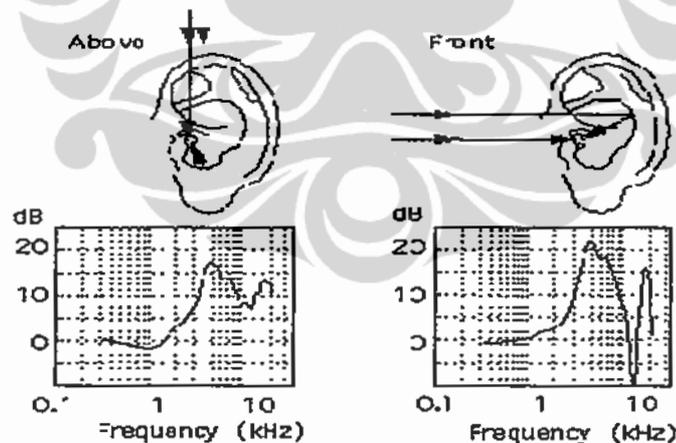
Seperti halnya perubahan ITD terhadap azimuth di bidang horisontal, ILD dari setiap frekuensi bertambah besar ketika bunyi mendekati kedua telinga. ILD_{max} pada bidang horisontal terjadi di sekitar azimuth 90° and 270° . Di sini dapat diamati bahwa ILD merupakan fungsi dari azimuth di bidang horisontal. Karena efek difraksi dan bayangan kepala, pada komponen-komponen frekuensi lebih tinggi pada umumnya terjadi ILD yang lebih besar.

Pada bidang median, perubahan ILD terhadap elevasi menjadi bertambah kompleks dengan bertambahnya frekuensi. Di sepanjang bidang median, nilai-nilai ILD dari frekuensi 10 kHz paling banyak bervariasi daripada variasi dari frekuensi-frekuensi lainnya. Dan sebaliknya nilai-nilai ILD dari 222 Hz paling stabil. Variasi nilai-nilai ILD yang cukup besar dari komponen-komponen frekuensi berasal dari pantulan-pantulan daun telinga. Dari sini dapat disimpulkan daun telinga memiliki peranan penting dalam persepsi elevasi dari posisi sumber bunyi. Ketika nilai-nilai ITD mendekati nol pada bidang median (terlihat pada azimuth 0° dan 180° dari semua bidang pada Gambar 2.6), nilai-nilai ILD memegang peranan penting dalam membedakan elevasi dari sumber bunyi; terlihat dari Gambar 2.7b, yaitu nilai-nilai ILD pada frekuensi-frekuensi tertentu berubah-ubah sebagai fungsi elevasi di bidang median. Observasi ini sesuai dengan teori Duplex seperti dijelaskan berikut ini.

Teori Duplex menambahkan bahwa ILD dan ITD saling melengkapi. Pada frekuensi-frekuensi rendah (di bawah sekitar 1,5 kHz), ada sedikit informasi ILD, tetapi ITD menggeser gelombang itu sebagian dari satu periode, yang dengan mudah dapat dideteksi. Pada frekuensi-frekuensi tinggi (di atas kira-kira 1,5 kHz), terdapat kebingungan dalam ITD karena ada pergeseran sebanyak beberapa periode, tetapi ILD menutupi kebingungan arah ini. Teori Duplex dari Rayleigh mengatakan bahwa ILD dan ITD yang diambil bersama-sama memberikan informasi lokalisasi sepanjang jangkauan frekuensi yang terdengar.

2.4 Elevation Cues

Sementara *cue* utama untuk azimuth adalah binaural, *cue* utama untuk elevasi sering disebut monaural. Faktor-faktor ini berasal dari fakta bahwa daun telinga manusia atau pinna berfungsi sebagai suatu antena akustik. Rongga-rongga resonansinya memperkuat beberapa frekuensi dan bentuk geometrinya menyebabkan efek-efek interferensi yang meredam frekuensi-frekuensi lain. Lebih jauh lagi bahwa respon frekuensinya tergantung pada arah datang gelombang bunyi.



Gambar 2.8. Respon Frekuensi untuk Kedatangan Sinyal Bunyi dari Dua Arah yang Berbeda [44]

Gambar 2.8 di atas menunjukkan respon-respon frekuensi yang diukur untuk dua arah kedatangan sinyal bunyi yang berbeda. Dalam setiap kasus terdapat dua jalur dari sumber ke kanal telinga, satu jalur langsung dan satu jalur lebih panjang yang berasal dari refleksi dari pinna. Pada frekuensi-frekuensi rendah yang moderat, pinna pada dasarnya mengumpulkan energi bunyi tambahan

dan sinyal-sinyal dari kedua jalur datang dengan fasa yang sama. Tetapi pada frekuensi-frekuensi tinggi, sinyal yang tertunda berlawanan fasa dengan sinyal langsung sehingga terjadi interferensi yang destruktif. Interferensi terbesar terjadi ketika selisih panjang jalur, d , adalah setengah panjang gelombang yaitu ketika $f = c/2d$ [44]. Dalam contoh di atas, hal ini menghasilkan suatu *pinna notch* di sekitar 10 kHz. Dengan nilai-nilai tipikal untuk d , frekuensi *notch* biasanya dalam jangkauan 6 kHz sampai 16 kHz [4].

Karena *pinna* merupakan suatu reflektor yang lebih efektif untuk bunyi yang datang dari depan daripada bunyi dari atas, *notch* yang dihasilkan lebih jelas untuk sumber-sumber di depan daripada untuk sumber-sumber di atas kepala. Perlu ditambahkan bahwa selisih panjang jalur berubah dengan sudut elevasi, sehingga frekuensi *notch* juga bergerak dengan elevasi. Walaupun masih terdapat perdebatan mengenai fitur-fitur mana yang paling penting yang dapat diindera [46], tetapi telah banyak disepakati bahwa *pinna* memberikan faktor-faktor utama untuk elevasi.

2.5 *Range Cues*

Pada saat melokalisasi suatu sumber, orang paling baik dalam mengestimasi azimut, terbaik selanjutnya dalam mengestimasi elevasi dan yang terburuk dalam mengestimasi jarak (*range*). Dengan alasan yang serupa, *azimuth cues* lebih banyak dimengerti, *elevation cues* lebih kurang dimengerti dan *range cues* paling kurang dimengerti. *Range cues* terdiri dari faktor-faktor berikut ini:

1. Kekerasan Bunyi (*Loudness*);
2. Gerakan paralaks (*Motion parallax*);
3. ILD yang berlebihan (*excess ILD*);
4. Rasio dari bunyi yang langsung dan yang dipantulkan.

Dasar fisik untuk *cue* kekerasan bunyi jelasnya berasal dari fakta bahwa energi bunyi yang tertangkap yang datang langsung dari sumber mengalami penurunan bersamaan dengan meningkatnya kuadrat jarak. Jadi ketika suatu sumber dengan energi konstan mencapai seorang pendengar, kekerasan bunyi meningkat. Juga telah jelas bahwa energi yang diterima sebanding dengan energi yang dipancarkan sumber dan bahwa tidak akan ada relasi satu-satu antara

kekerasan bunyi dan jarak. Dengan hanya memutar bunyi pada suatu level volume yang rendah tidak serta merta menyebabkan sumber terdengar jauh. Untuk menggunakan kekerasan bunyi sebagai *cue* untuk jarak, karakteristik-karakteristik sumber juga harus dimengerti. Dalam kasus bunyi pembicaraan manusia, setiap orang mengetahui dari pengalaman kualitas bunyi yang berbeda yang berkaitan dengan berbisik, berbicara normal dan berteriak, tidak peduli level bunyi. Kombinasi antara kekerasan bunyi dan pengetahuan tentang sumber memberikan informasi berguna untuk penentuan jarak.

Pergerakan paralaks mengacu pada fakta bahwa jika seorang pendengar menggerakkan kepalanya, perubahan dalam azimut akan tergantung pada jarak. Untuk sumber-sumber yang sangat dekat, suatu pergeseran kecil menyebabkan perubahan besar dalam azimut, sementara untuk sumber-sumber jauh pada dasarnya tidak ada perubahan azimut. Sebagai tambahan, ketika sumber bunyi terletak sangat dekat dengan kepala, ILD akan meningkat. Peningkatan ini menjadi sangat terasa untuk jarak-jarak di bawah sekitar satu meter. Suatu kasus ekstrim adalah ketika terdapat serangga yang ada dekat salah satu telinga, atau ketika seseorang berbisik pada salah satu telinga. Pada umumnya bunyi-bunyi yang terdengar hanya dalam satu telinga akan sangat mengganggu dan tidak nyaman untuk didengarkan. Hal-hal ini khususnya perlu diperhatikan dalam merancang sistem interaksi manusia dan komputer untuk *headphones*. Akan terlihat bahwa untuk memberikan pemikiran pada pendengar bahwa bunyi berada pada satu sisi telinga, tidak diperlukan sama sekali untuk memberikan bunyi hanya pada telinga tersebut dan tidak ada bunyi apapun pada telinga lain.

Cue terakhir adalah rasio dari bunyi langsung dan yang dipantulkan. Seperti disebutkan di atas, energi yang diterima secara langsung dari sumber bunyi menurun berbanding terbalik dengan kuadrat jarak. Tetapi dalam ruangan-ruangan biasa, bunyi dipantulkan dan terhambur beberapa kali dari permukaan-permukaan di sekelilingnya. Energi pantulan yang mencapai telinga tidak terlalu banyak berubah dengan jarak dari sumber ke pendengar. Jadi rasio dari energi bunyi langsung dan yang dipantulkan adalah suatu *cue* utama untuk jarak. Pada jarak dekat, rasio ini sangat besar, sementara pada jarak jauh rasio ini sangat kecil.

Akan tetapi rasio ini adalah *cue* yang relatif mudah dan efektif untuk dimanipulasi untuk aplikasi-aplikasi sistem audio spasial.

2.6 Reverberasi dan Gema-gema

Sering tidak disadari bahwa bunyi yang terdengar datang dari refleksi-refleksi dari permukaan-permukaan di lingkungan sekitar. Bahkan di lingkungan luar sejumlah energi signifikan dipantulkan oleh tanah dan oleh struktur-struktur dan tumbuh-tumbuhan sekitar. Refleksi-refleksi ini hanya terasa ketika waktu tunda menjadi lebih panjang daripada kira-kira 30 sampai 50 ms (ambang batas gema) dan disebut gema-gema [47]. Ruangan-ruangan spesial yang disebut *anechoic chambers* dibangun untuk menyerap energi bunyi, sehingga hanya energi yang langsung diradiasikan yang mencapai telinga. Pada saat pertama kali memasuki suatu ruang kedap bunyi, kebanyakan orang terkagum oleh bagaimana lembut dan tumpulnya semua bunyi terdengar.

Jika bunyi yang dipantulkan menjadi begitu umum dalam lingkungan akustik biasa, seseorang akan heran mengapa refleksi-refleksi ini tidak berinterferensi dengan kemampuan manusia untuk melokalisasi sumber-sumber bunyi. Jawabannya adalah bahwa manusia secara cepat beradaptasi dengan suatu lingkungan baru dan sistem auditori manusia hanya menggunakan sebagian dari mekanisme-mekanisme yang dimengerti untuk menekan efek-efek dari refleksi-refleksi dan reverberasi. Fakta bahwa manusia menentukan lokasi berdasarkan sinyal-sinyal yang meraih kedua telinga pada mulanya disebut sebagai *precedence effect* atau *Law of the First Wavefront* [48]. Di sini tidak dikatakan bahwa seseorang tidak merasakan refleksi-refleksi berikutnya dan informasi ini digunakan untuk mengestimasi jarak. Jika reverberasi tidak besar, refleksi-refleksi hanya memiliki pengaruh relatif kecil pada kemampuan manusia dalam menglokalisasi bunyi.

Bagaimanapun *precedence effect* memaksa agar Teori Duplex dari Rayleigh dimodifikasi. Dalam suatu ruang tipikal, refleksi-refleksi mulai timbul beberapa millidetik setelah bunyi awal. Untuk bunyi dengan frekuensi rendah yang memiliki periode lebih panjang daripada waktu yang diperlukan untuk timbulnya refleksi (misalnya di bawah 250 Hz), refleksi-refleksi mulai

berdatangan bahkan sebelum satu siklus sinyal terlengkapi. Pada saat beberapa siklus telah datang dan sistem auditori dapat mulai mengestimasi warna bunyi/*pitch*, pola bunyi dalam ruang adalah sekumpulan gelombang-gelombang berdiri dan sekarang sistem auditori tidak mungkin untuk mengestimasi ITD. Dengan demikian dalam ruang yang *reverberant*, informasi frekuensi rendah pada dasarnya tidak berguna untuk lokalisasi.

Tetapi hal ini tidak berarti bahwa selisih-selisih waktu interaural tidak penting. Informasi waktu penting datang dari Selisih Selubung Interaural / *Interaural Envelope Difference*. Hal ini secara jelas didemonstrasikan oleh efek Franssen. Jika suatu gelombang sinus tiba-tiba diaktifkan dan suatu versi yang disaring frekuensi tingginya dikirim ke pengeras suara A sementara versi yang disaring frekuensi rendahnya dikirim ke pengeras suara B, kebanyakan pendengar akan melokalisasi bunyi pada pengeras suara A. Hal ini juga benar bahkan jika frekuensi gelombang sinus cukup rendah sehingga pada keadaan mantap sebagian besar energi datang dari pengeras suara B. Pada dasarnya, transien permulaan memberikan informasi lokalisasi yang jelas, sementara sinyal keadaan mantap sangat sulit untuk dilokalisasi, dan dalam hal ini sistem auditori mengabaikan saja informasi yang membingungkan. Dengan resiko terlalu sederhana, dapat disimpulkan dan dikatakan bahwa dalam lingkungan-lingkungan yang reverberan energi frekuensi tinggi dan bukan energi frekuensi rendah yang penting untuk lokalisasi.

2.7 *Head-Related Impulse Response*

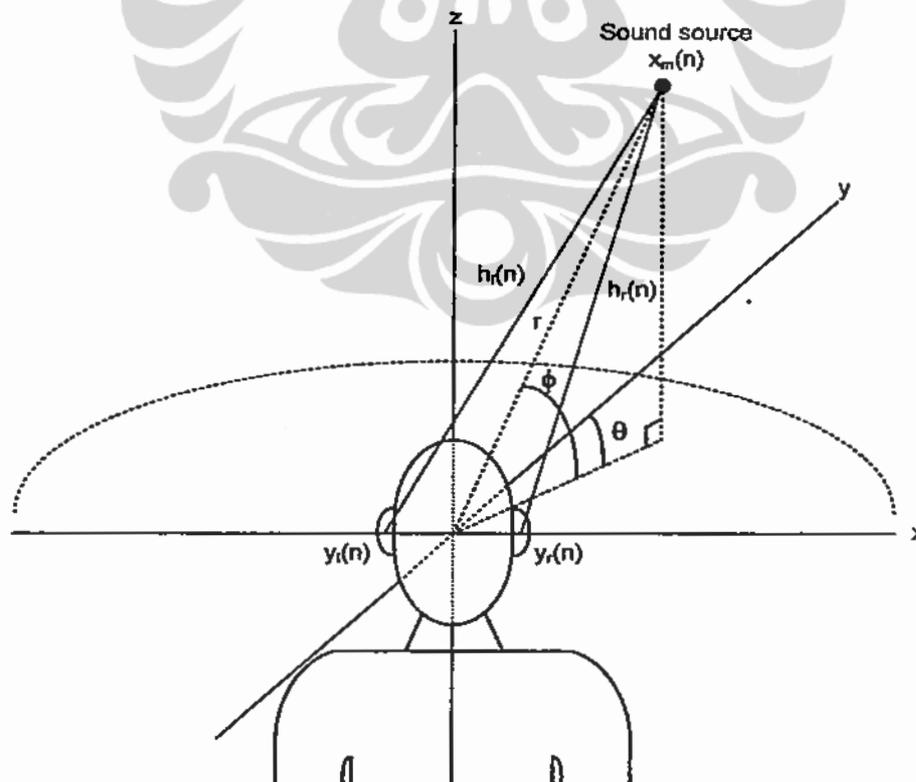
Untuk menemukan tekanan bunyi yang dihasilkan oleh suatu sumber sembarang pada gendang telinga manusia, diperlukan respon impuls dari sumber tersebut sampai ke gendang telinga. Respon impuls ini disebut dengan *Head-Related Impulse Response (HRIR)*. Jadi, definisi dari HRIR adalah suatu respon impuls gendang telinga manusia yang berfungsi untuk menyaring bunyi yang datang padanya dari suatu posisi sehingga dihasilkan tekanan bunyi tertentu pada gendang telinga tadi. Proses sintesis binaural dapat dirumuskan dalam suatu sistem *Linear Time-Invariant (LTI)*. Jika sinyal masukan monaural disimbolkan

dengan $x_m(n)$, sinyal-sinyal binaural yang dihasilkan pada kedua telinga manusia diperoleh dari konvolusi berikut:

$$\mathbf{y} = \mathbf{h} * x_m(n) \quad (2.2)$$

$$\mathbf{y} = \begin{bmatrix} y_l(n) \\ y_r(n) \end{bmatrix} \text{ dan } \mathbf{h} = \begin{bmatrix} h_l(n) \\ h_r(n) \end{bmatrix} \quad (2.3)$$

dimana \mathbf{y} adalah suatu vektor kolom dari sinyal-sinyal binaural dan \mathbf{h} adalah suatu vektor kolom sepasang HRIR yang digunakan dalam sintesis binaural. Persamaan (2.3) menunjukkan bahwa $h_l(n)$ dan $h_r(n)$ adalah masing-masing HRIR telinga kiri dan HRIR telinga kanan serta $y_l(n)$ dan $y_r(n)$ adalah masing-masing tekanan bunyi yang sampai pada gendang telinga kiri dan gendang telinga kanan dari seorang subyek. Jadi $y_l(n)$ dan $y_r(n)$ merupakan hasil penyaringan sumber bunyi monaural $x_m(n)$ yang dilakukan oleh masing-masing HRIR telinga kiri, $h_l(n)$ dan HRIR telinga kanan, $h_r(n)$. Deskripsi penyaringan ini diilustrasikan oleh Gambar 2.9 berikut. Terlihat dari gambar tersebut, sumber bunyi berada pada posisi (θ, ϕ) tertentu, dimana θ adalah sudut azimuth dan ϕ adalah sudut elevasi sumber bunyi terhadap pusat kepala.



Gambar 2.9. Deskripsi Penyaringan Sumber Bunyi

Pasangan Transformasi Fourier dari HRIR dikenal sebagai *Head-Related Transfer Function* (HRTF). HRTF didefinisikan sebagai fungsi alih sistem penyaringan sumber bunyi oleh gendang telinga kiri dan gendang telinga kanan dalam kawasan frekuensi. Jika HRIR dilambangkan $h(n)$ maka HRTF, $H(e^{j\omega})$, dinyatakan oleh:

$$H(e^{j\omega}) = F_D\{h(n)\}, \quad (2.4)$$

dimana $F_D\{\}$ adalah Transformasi Fourier Diskrit (*Discrete Fourier Transform*/DFT). Representasi kawasan frekuensi diskrit dari persamaan (2.2) (representasi kawasan waktu diskrit) dapat dituliskan sebagai:

$$\mathbf{Y} = \mathbf{H} X_m(e^{j\omega}) \quad (2.5)$$

$$\mathbf{Y} = \begin{bmatrix} Y_l(e^{j\omega}) \\ Y_r(e^{j\omega}) \end{bmatrix} \text{ dan } \mathbf{H} = \begin{bmatrix} H_l(e^{j\omega}) \\ H_r(e^{j\omega}) \end{bmatrix} \quad (2.6)$$

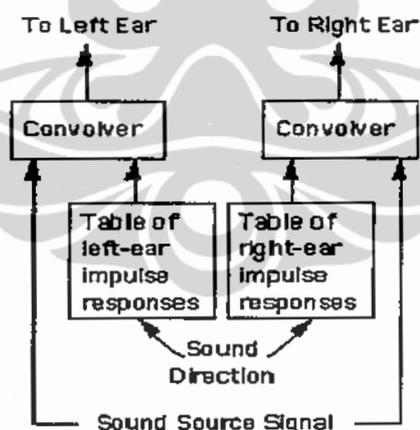
dimana $Y_l(e^{j\omega})$ dan $Y_r(e^{j\omega})$ adalah masing-masing tekanan bunyi yang timbul pada gendang telinga kiri dan gendang telinga kanan pada kawasan frekuensi, sedangkan $X_m(e^{j\omega})$, $H_l(e^{j\omega})$, $H_r(e^{j\omega})$ berturut-turut adalah sinyal *input* monaural dalam kawasan frekuensi, HRTF telinga kiri dan HRTF telinga kanan dari seorang subyek.

Biasanya HRTF diukur dari suatu posisi bunyi pada waktu tertentu sepanjang frekuensi yang diinginkan, dengan menggunakan teknik-teknik sistem identifikasi yang berbasis *fast Fourier transform* (FFT), sehingga $H_{l,r}(e^{j\omega})$ dicuplik secara tidak teratur dalam posisi spasial tetapi dicuplik seragam di dalam frekuensi. Pengukuran langsung HRTF menggunakan prosedur-prosedur identifikasi sistem yang sesuai untuk diimplementasikan pada manusia dan manikin.

Mikrofon-mikrofon sangat kecil diletakkan pada kanal telinga kiri dan kanan. Lalu deretan penguat-penguat suara digunakan untuk mengumpangkan stimulus pada kedua telinga dari beberapa posisi yang telah diketahui. Berdasarkan jenis sinyal stimulus yang digunakan, prosedur identifikasi sistem yang berbeda dapat digunakan dalam menghubungkan sinyal yang diukur dan sinyal yang diumpangkan, untuk menghasilkan estimasi-estimasi HRTF pada posisi-posisi penguat suara yang bersangkutan.

Sistem-sistem berbasis HRTFs dapat menghasilkan efek-efek elevasi dan jarak sebaik efek-efek azimut. Hal ini pada dasarnya berarti bahwa sistem-sistem tersebut dapat menciptakan impresi bahwa suatu bunyi berada dalam suatu posisi spasial yang diinginkan – kiri atau kanan, atas atau bawah, dekat atau jauh. Dalam prakteknya, karena perbedaan-perbedaan antar individu dan keterbatasan komputasi, azimut jauh lebih mudah dikendalikan daripada elevasi atau jarak. Sekarang ini sistem-sistem berbasis HRTF telah dengan cepat menjadi baku bagi *advanced 3D audio interfaces*.

Salah satu sistem berbasis HRTF adalah Convolvotron™, suatu sistem audio spasial yang terkenal dan efektif, seperti terlihat pada Gambar 2.10. Convolvotron, dikembangkan untuk NASA dan dibuat oleh *Crystal River Engineering*, memberikan suatu cara yang sederhana untuk menggunakan HRTFs dalam audio spasial. Tersirat dari namanya, sistem ini terdiri dari dua mesin konvolusi, dimana setiap mesin konvolusi ini dapat mengkonvolusi aliran masukan audio yang sama dengan suatu segmen terbatas dari sebuah HRIR yang dipanggil dari tabel yang berisi nilai-nilai hasil pengukuran. Kedua keluaran dari Convolvotron menuju *headphones* yang dikenakan pendengar melalui *amplifier*.



Gambar 2.10. Sistem Audio Spasial Berbasis HRTF: Convolvotron™

Jika HRIRs individual dari pendengar tadi cukup dekat dengan HRIRs yang digunakan oleh mesin konvolusi, maka bunyi yang dihantarkan ke kedua telinga pendengar akan mengandung semua *cue* spasial yang tepat, sehingga *sound image* dapat dengan tepat pula dilokalisasi. Lokasi ini dapat ditentukan oleh azimut, elevasi dan jarak tertentu yang digunakan untuk *indexing* tabel-tabel yang disimpan.

Ide dasar dari Convolvotron dapat dielaborasi lebih jauh dalam beberapa cara, yaitu:

1. Sumber-sumber banyak dapat diakomodasikan dengan menduplikasi perangkat keras dan menjumlahkan keluarannya untuk setiap telinga;
2. Gerakan kepala dapat diperhitungkan dengan menggabungkan lokasi absolut sumber dengan keluaran-keluaran dari suatu *head tracker* untuk memilik HRIRs yang tepat;
3. Tabel-tabel dapat diberi nomor dengan azimut dan elevasi saja, dengan menggunakan jarak dari sumber bunyi ke setiap telinga untuk menghasilkan efek-efek amplitudo;
4. Jumlah HRIR yang disimpan dalam tabel dapat dikurangi dengan menggunakan *sampling* yang kasar dari lokasi spasial dan kemudian menginterpolasi secara tepat antar titik-titik yang berdekatan;
5. Gema-gema dan reverberasi ruang dapat ditambahkan dengan memasukkan suatu model simulasi ruang;
6. Sistem ini dapat diindividualisasi untuk pendengar tertentu dengan mengukur dan menggunakan HRIR miliknya.

Walau bagaimanapun sistem tersebut diperluas, konsep dasar tetap sama yaitu untuk menciptakan sinyal-sinyal telinga kiri dan telinga kanan yang tepat dengan cara konvolusi waktu nyata antara masukan bunyi monaural dengan respon-respon impuls yang ditabulasi.

2.8 Basis Data yang Digunakan dalam Pemodelan HRTF

Kebanyakan sistem komersial mengkonvolusi sinyal input dengan „satu pasang HRIR standar“ dan beberapa studi telah mempergunakan data publik dari *Knowles Electronic Manikin for Acoustic Research (KEMAR)* [46,49]. Seperti telah diketahui bahwa HRTF berubah secara signifikan untuk setiap individu, sehingga diperlukan suatu basis data hasil pengukuran HRTF yang memadai dalam pemodelan HRTF.

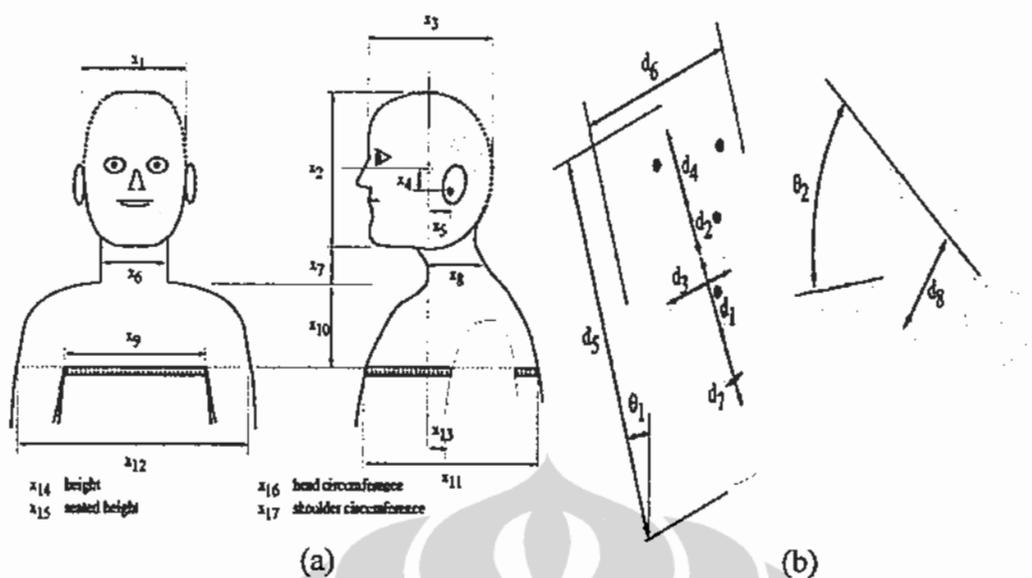
CIPIC Interface Laboratory di U.C. Davis, Amerika Serikat, telah mengukur HRTF dengan resolusi spasial tinggi bagi lebih dari 90 subyek. *CIPIC HRTF Database Release 1.2*, yang adalah *subset* basis data untuk 45 subyek,

tersedia untuk diakses dari situs web mereka. Basis data CIPIC tidak hanya terdiri dari respon impuls untuk 1250 posisi untuk setiap telinga dan setiap subyek, tetapi juga termasuk sekumpulan ukuran antropometris untuk setiap telinga dari setiap subyek [2]. Penelitian ini menggunakan basis data di atas.

Subyek untuk pengukuran respon impuls berjumlah 43 orang, 27 pria dan 16 perempuan dan KEMAR. Para subyek merupakan para mahasiswa dan para pengunjung di CIPIC *Interface Laboratory*. Semua HRTF diukur dengan subyek duduk di pusat lingkaran dengan jari-jari satu meter. Posisi kepala subyek tidak dibuat tetap, tetapi subyek dapat memantau posisi kepalanya. Pengeras suara *Bose AcoustimassTM* (diameter *cone* 5,8 cm) dipasang pada posisi-posisi yang berbeda di *hoop* setengah lingkaran. Suatu sistem *SnapshotTM* yang dimodifikasi dari *Crystal River Engineering* membangkitkan sinyal-sinyal kode Golay. Kanal telinga dari subyek ditutup dan mikrofon-mikrofon *probe Etymotic Research ER-7C* digunakan untuk menangkap sinyal-sinyal kode Golay. Keluaran dari mikrofon dicuplik dengan frekuensi 44.100 Hz, resolusi 16 bit dan diproses oleh fungsi *Snapshot's oneshot* untuk menghasilkan suatu *raw HRIR*. Suatu jendela Hanning yang dimodifikasi digunakan pada pengukuran *raw HRIR* untuk menghilangkan refleksi ruang dan sebagai hasilnya dikompensasi *free-field* untuk memperbaiki karakteristik spektral dari transduser-transduser yang digunakan. Panjang dari setiap *HRIR* adalah 200 sampel, yang berdurasi sekitar 4,5 ms.

Posisi sumber bunyi ditentukan oleh sudut azimut θ dan sudut elevasi ϕ dalam koordinat kutub-interaural. Elevasi disampel dalam $360/64 = 5,625^\circ$ langkah dari -45° sampai $+230,625^\circ$, sedangkan azimut disampel pada -80° , -65° , -55° , dari -45° sampai $+45^\circ$ dalam langkah 5° , pada $55^\circ, 65^\circ, 80^\circ$. Jadi diperoleh sampel spasial pada 1250 titik.

Basis data CIPIC yang dipublikasikan tersebut mengandung antropometri setiap subyek. Ukuran-ukuran antropometris, walaupun tidak sempurna, memungkinkan investigasi tentang kemungkinan korespondansi atau korelasi antara dimensi fisik dan fitur-fitur HRTF. Dalam basis data CIPIC terdapat 27 ukuran antropometris, yakni 17 parameter untuk kepala dan *torso*; dan 10 parameter untuk daun telinga, yang mengikuti pendekatan yang diusulkan oleh Genuit [5], seperti diilustrasikan dalam Gambar 2.11.



Gambar 2.11. Ukuran-ukuran Antropometris Subyek, (a) kepala dan torso, (b) daun telinga [1].

Tabel 2.1. Statistik Antropometri Basis Data, $\% = 2\sigma/\mu$ [1]

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
x_9	torso top width	31.50	3.19	20
x_{10}	torso top height	13.42	1.85	28
x_{11}	torso top depth	23.84	2.95	25
x_{12}	shoulder width	45.90	3.78	16
x_{13}	head offset forward	3.03	2.29	151
x_{14}	height	172.43	11.61	13
x_{15}	seated height	88.83	5.53	12
x_{16}	head circumference	57.33	2.47	9
x_{17}	shoulder circumference	109.43	10.30	19
d_1	cavum concha height	1.91	0.18	19
d_2	cymba concha height	0.68	0.12	35
d_3	cavum concha width	1.58	0.28	35
d_4	fossa height	1.51	0.33	44
d_5	pinna height	6.41	0.51	16
d_6	pinna width	2.92	0.27	18
d_7	intertragal incisure width	0.53	0.14	51
d_8	cavum concha depth	1.02	0.16	32
θ_1	pinna rotation angle	24.01	6.59	55
θ_2	pinna flare angle	28.53	6.70	47

Secara umum, histogram dari ukuran-ukuran individu pada dasarnya mengindikasikan suatu distribusi nilai-nilai secara normal. Tabel 2.1 memberikan ukuran-ukuran antropometris yang didefinisikan dalam basis data yang digunakan, beserta masing-masing parameter statistik. Di sini jarak-jarak diukur dalam cm dan sudut-sudut dalam derajat dan persentase variasi adalah $2\sigma / \mu$ persen, dimana μ dan σ masing-masing adalah nilai rata-rata dan deviasi standar dari ukuran antropometris yang bersangkutan. Sebagai contoh, rata-rata dari lebar kepala adalah 14,49 cm dan dengan anggapan suatu distribusi normal, 95% dari kasus berada dalam $\pm 13\%$ dari nilai rata-rata. Tanpa memperhatikan ukuran-ukuran offset x_4, x_5, x_{13} , dimana persentase deviasi tidak berarti, terlihat persentase deviasi rata-rata adalah $\pm 26\%$. Jadi terdapat sejumlah variasi dalam ukuran-ukuran dan bentuk-bentuk dari subyek dalam basis data.

2.9 Visualisasi dan Analisis Struktur Data HRTF

Visualisasi data HRTF memegang peranan penting dalam menemukan fitur-fitur utama HRTF yang berhubungan dengan faktor-faktor yang menentukan persepsi posisi sumber bunyi. Puncak-puncak (*peaks*) dan lembah-lembah (*notches*) dari respon-respon absolut HRTF berhubungan dengan azimuth dan elevasi. Visualisasi data HRTF juga penting dalam perbandingan data HRTF antar subyek. Pengamatan lebih mendalam mengenai variasi karakteristik HRTF tertentu dapat dilakukan antar beberapa set data HRTF yang berasal dari beberapa subyek. Tentu saja tidak efisien untuk menggambarkan dan menganalisis semua data HRTF dalam basis data yang sangat besar. Dalam penelitian ini digunakan suatu basis data HRTF publik dengan resolusi spasial tinggi yaitu CIPIC HRTF *Database*. Seperti diuraikan dalam subbab 2.8, satu HRIR dari basis data terdiri dari 200 respon impuls. Basis data ini terdiri dari HRIR untuk 45 subyek pada 25 azimuth berbeda dan 50 elevasi berbeda. Resolusi dari azimuth dan elevasi hampir sebesar 5° . Jadi terdapat seluruhnya 1250 HRIR untuk satu telinga.

Perbandingan data HRTF antar subyek dilakukan lebih efisien dengan menggunakan suatu set yang lebih kecil, misalnya set HRTF pada bidang horisontal dan set HRTF pada bidang median. Dalam visualisasi selanjutnya, digunakan data HRTF dari subyek 028 dari basis data karena dari hasil pengamatan dari data HRTF seluruh subyek, diperoleh bahwa variasi HRTF dari

subyek ini pada bidang horisontal dan bidang median pada umumnya mewakili variasi HRTF subyek-subyek lain. Di dalam basis data ini, tercakup juga ukuran-ukuran antropometris dari semua subyek.

Di dalam subbab 2.9 ini digambarkan dan kemudian dianalisis struktur data HRTF baik berupa representasi HRIR di kawasan waktu maupun representasi HRTF *magnitude* di kawasan frekuensi. Dengan menganalisis struktur data HRTF, maka dapat terlihat beberapa karakteristik HRTF. Dalam bagian-bagian berikutnya, diuraikan analisis-analisis representasi data HRTF di kawasan frekuensi dan analisis-analisis representasi HRIR di kawasan waktu.

2.9.1 Representasi Data HRTF di Kawasan Frekuensi

Data HRTF yang dianalisis di sini adalah data HRTF *magnitude* dalam satuan dB pada bidang horisontal dan pada bidang median dari subyek 028 basis data; dimana HRTF *magnitude* $H(f) = 20 \log\{|F_D[h(n)]|\}$. Dalam representasi ini dibahas variasi HRTF terhadap variasi azimuth ($\theta = -80^\circ$ sampai $+80^\circ$) dengan elevasi tetap ($\phi = 0^\circ$) pada bidang horisontal depan serta variasi HRTF terhadap variasi elevasi ($\phi = -45^\circ$ sampai $+230^\circ$) dengan azimuth tetap ($\theta=0^\circ$) pada bidang median. Tanda negatif azimuth menunjukkan posisi sumber bunyi di sebelah kiri dan tanda positif di sebelah kanan. Elevasi negatif adalah elevasi posisi sumber bunyi di bawah, sedangkan elevasi positif adalah elevasi posisi sumber bunyi di atas. Acuan posisi sumber bunyi adalah pusat kepala.

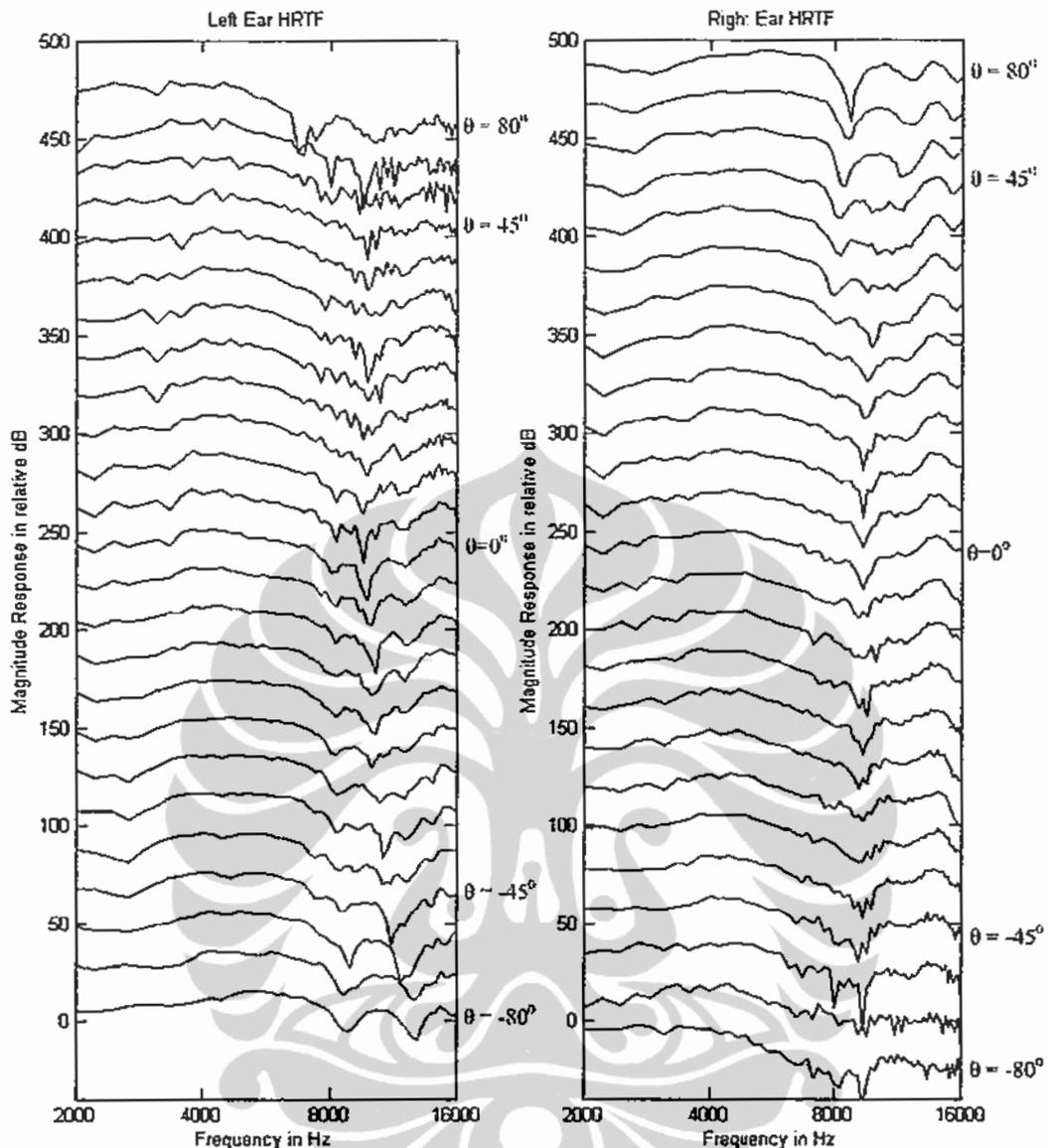
Representasi data HRTF *magnitude* di kawasan frekuensi dapat dilihat pada Gambar 2.12 untuk posisi-posisi sumber bunyi di bidang horisontal depan dan Gambar 2.13 untuk posisi-posisi sumber bunyi di bidang median. Pada kedua gambar tersebut, HRTF *magnitude* paling bawah adalah HRTF *magnitude* sebenarnya, tetapi setiap HRTF *magnitude* lainnya ditambah 20 dB dari HRTF *magnitude* di bawahnya. HRTF *magnitude* dari hasil pengukuran merupakan fungsi yang rumit dari posisi sumber bunyi dan frekuensi. Terdapat beberapa puncak dan lembah dalam spektrum *magnitude* HRTF tersebut, dimana fitur-fitur ini berkaitan dengan penyaringan gelombang bunyi oleh daun telinga, kepala dan torso.

Dari Gambar 2.12 terlihat bahwa secara umum *magnitude* HRTF dari telinga yang terletak pada sisi yang sama dengan sumber bunyi (*ipsilateral*), lebih

besar daripada HRTF *magnitude* dari telinga yang terletak pada sisi yang berlawanan dengan sumber bunyi (*contralateral*). Hasil ini menunjukkan bahwa energi HRTF *magnitude* dari telinga *ipsilateral* lebih besar daripada energi HRTF *magnitude* dari telinga *contralateral*. Terlihat juga bahwa HRTF *magnitude* pada sisi *ipsilateral* secara umum lebih halus daripada HRTF *magnitude* pada sisi *contralateral*. Kedua hal tersebut menunjukkan efek bayangan kepala karena telinga *ipsilateral* menerima langsung gelombang bunyi yang datang sedangkan telinga *contralateral* dibayangi oleh kepala pada saat menerima gelombang bunyi tersebut.

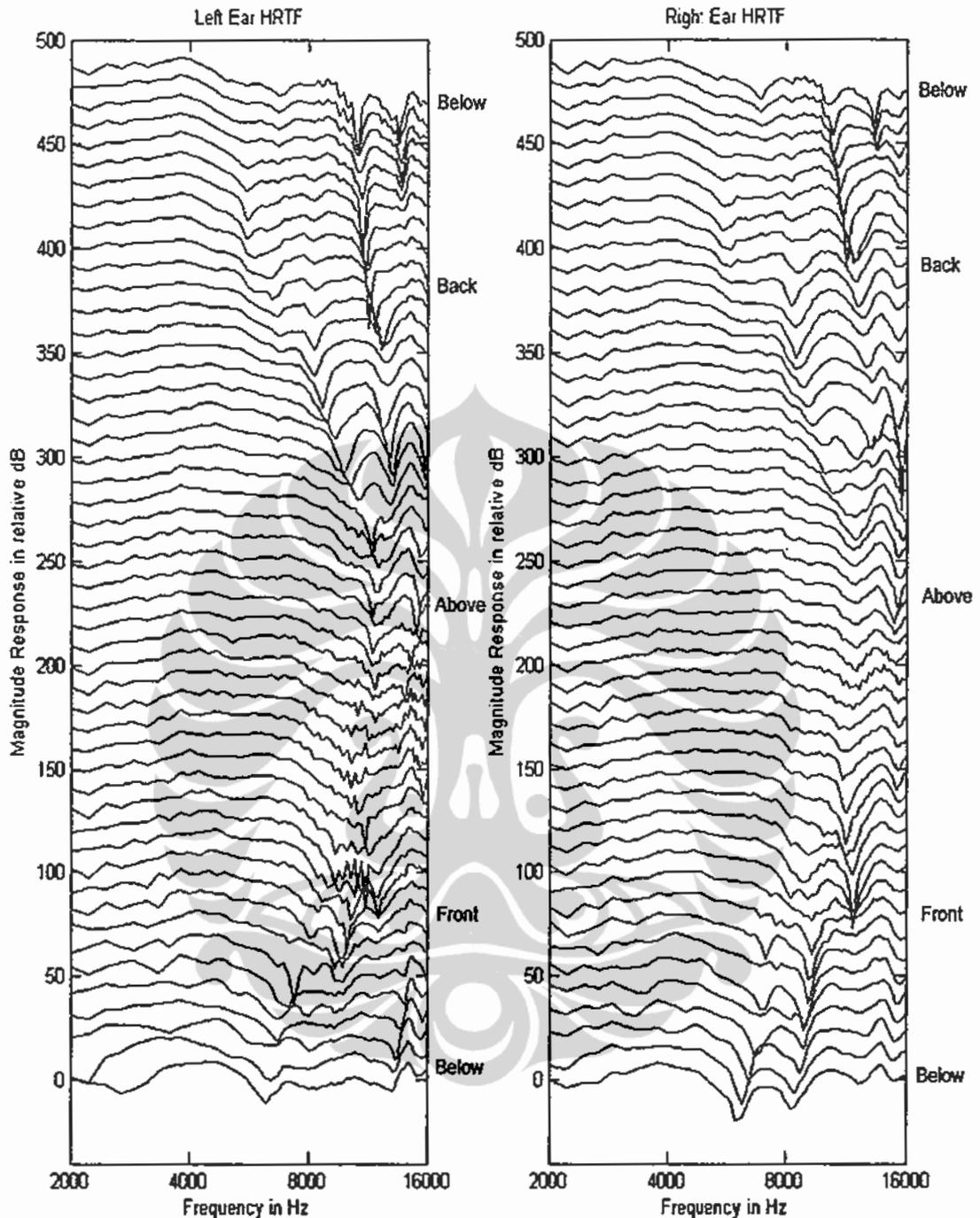
Terdapat dua jenis efek struktural yang terjadi pada HRTF. Kedua jenis efek ini menyebabkan struktur-struktur spesifik dalam HRTF. Efek jenis pertama adalah efek difraksi kepala yang mempengaruhi HRTF *contralateral*. Untuk beberapa frekuensi dan sudut datang, kepala memiliki efek amplifikasi pada suatu gelombang datar yang datang pada titik-titik tertentu di dekat kepala. Pada posisi-posisi bunyi di sisi *contralateral* kepala (azimut $0^\circ - 80^\circ$ untuk telinga kiri), komponen-komponen frekuensi tinggi lebih dari 1,5 kHz, diatenuasi lebih banyak daripada komponen-komponen frekuensi rendah. Hal ini terjadi karena kepala menutupi sebagian atau seluruh telinga *contralateral*. Sebaliknya, pada posisi-posisi bunyi di sisi *ipsilateral* telinga (azimut $-80^\circ - 0^\circ$ untuk telinga kiri), komponen-komponen frekuensi tinggi diperkuat (*boosting*). Fenomena-fenomena ini disebut efek difraksi kepala atau efek bayangan kepala. Efek difraksi kepala bertambah besar untuk bertambahnya frekuensi (berkurangnya panjang gelombang). Misalnya, suatu gelombang sinus 3 kHz pada azimut 90° akan diatenuasi sekitar 10 dB, suatu gelombang 6 kHz akan diatenuasi sekitar 20 dB, dan suatu gelombang 10 kHz akan diatenuasi sekitar 35 dB [50].

Efek kedua adalah efek elevasi daun telinga yang mempengaruhi HRTF *ipsilateral*. Detil-detil dari spektrum yang berhubungan dengan elevasi disebabkan oleh daun telinga. Frekuensi-frekuensi antara 6 – 8 kHz (panjang gelombang $\lambda = 4,1 - 5,5$ cm) adalah penting dalam persepsi elevasi, karena frekuensi-frekuensi ini berhubungan dengan panjang gelombang λ yang hampir sama dengan panjang daun telinga. Terdapat fitur-fitur di sekitar frekuensi-frekuensi ini yang telah dibuktikan secara psikofisik, berhubungan dengan persepsi elevasi [48,50].



Gambar 2.12. Seluruh HRTF *Magnitude* dari Subyek 028 pada Bidang Horisontal Depan [2]

Efek difraksi kepala dapat dilihat pada Gambar 2.12. HRTF *magnitude* telinga kiri dari azimuth 45° sampai 80° dan HRTF *magnitude* telinga kanan dari azimuth -45° sampai -80° masing-masing memiliki lobus utama (*main lobe*) dengan *magnitude* kecil pada frekuensi rendah. Lobus utama ini merepresentasikan efek amplifikasi kepala pada frekuensi-frekuensi rendah yang disebabkan difraksi gelombang bunyi pada sisi *contralateral*. Efek amplifikasi lain pada HRTF *magnitude ipsilateral* juga terlihat dari penguatan frekuensi-frekuensi tinggi yang disebabkan oleh pantulan gelombang bunyi pada daun telinga.



Gambar 2.13 Seluruh HRTF *Magnitude* dari Subyek 028 pada Bidang Median [2]

Efek elevasi daun telinga juga dapat dilihat pada HRTF *magnitude* di bidang median. Gambar 2.13 mengilustrasikan deretan kurva yang menunjukkan respon frekuensi HRTF *magnitude* berubah-ubah ketika sumber bunyi bergerak di sekeliling bidang median. Resonansi kanal telinga yang lebar sekitar 4 kHz tidak berubah untuk semua respon frekuensi [4]. Efek elevasi daun telinga ditunjukkan

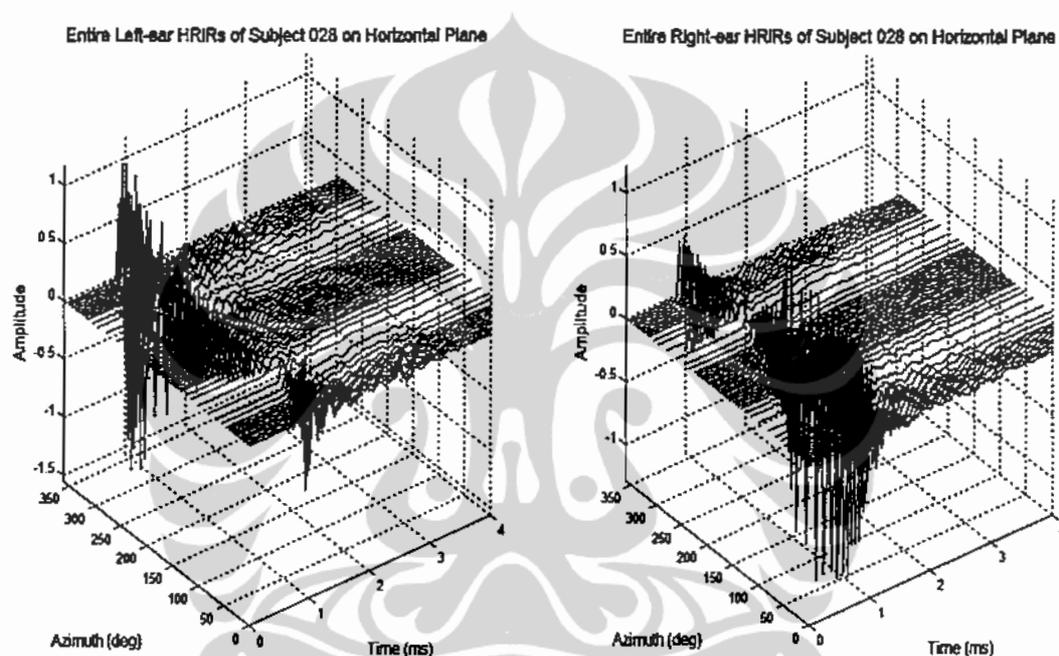
oleh frekuensi lembah daun telinga, f_{pn} . Frekuensi ini berubah secara signifikan terhadap elevasi. Terlihat frekuensi ini berubah mulai dari sedikit di bawah 6 kHz pada elevasi rendah sampai 10 kHz pada posisi di atas kepala. Ketika sumber bunyi berada tepat di atas kepala, lembah daun telinga sulit dilihat karena respon frekuensi relatif rata. Lembah daun telinga kembali muncul ketika sumber bunyi bergerak di sekitar belakang dan kembali ke lantai.

2.9.2 Representasi Data HRIR di Kawasan Waktu

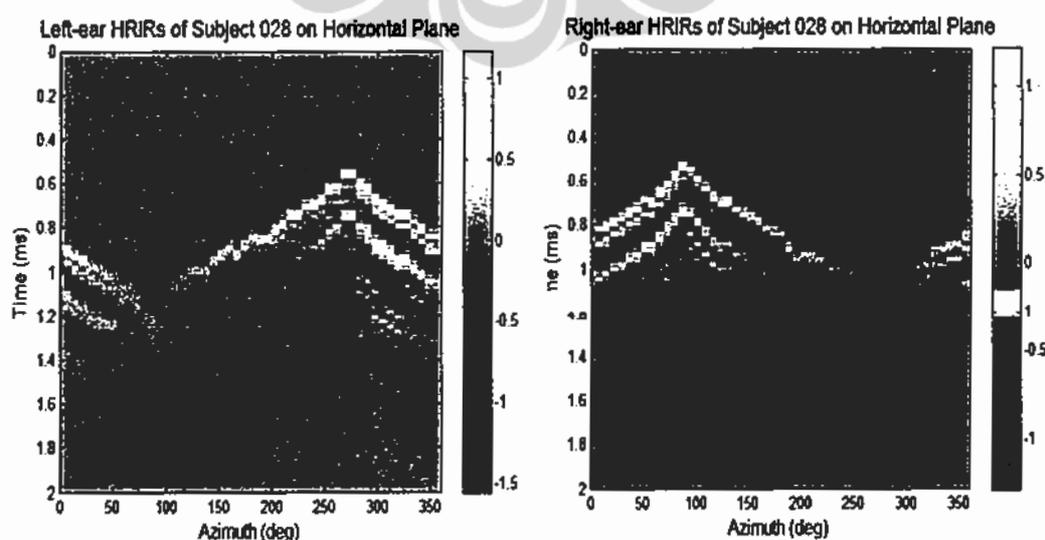
Gambar 2.14 menunjukkan grafik-grafik tiga dimensi dari seluruh HRIR subyek 028 pada bidang horisontal. Gambar 2.15 merupakan representasi *image* dari seluruh HRIR yang sama seperti pada Gambar 2.14, dimana intensitas HRIR ditunjukkan oleh derajat kecerahan titik pada gambar tersebut. Kedua gambar ini memberikan interpretasi yang serupa. Terlihat dari kedua gambar, bahwa gelombang bunyi paling kuat dan tiba paling cepat ketika gelombang itu datang dari sisi *ipsilateral*, yaitu $\theta = 270^\circ$ untuk telinga kiri dan $\theta = 90^\circ$ untuk telinga kanan. Hal ini terlihat dari Gambar 2.14 sebagai respon impuls yang memiliki waktu tunda paling kecil, sedangkan dari Gambar 2.15 terlihat dari garis putih paling atas dengan waktu tunda paling kecil. Demikian juga, gelombang itu paling lemah dan datang paling akhir ketika datang dari sisi *contralateral* ($\theta = 90^\circ$ untuk telinga kiri dan $\theta = 270^\circ$ untuk telinga kanan). Waktu kedatangan gelombang bunyi bervariasi dengan azimuth di sepanjang bidang horisontal dalam bentuk yang menyerupai gelombang sinus. Ternyata waktu kedatangan ini mengkonfirmasi cukup baik persamaan ITD, misalnya perbedaan waktu kedatangan terpendek dan terpanjang adalah sekitar 0,7 ms seperti yang diprediksi oleh Teori Duplex. Terlihat dari Gambar 2.14 dan Gambar 2.15, bahwa waktu-waktu kedatangan sebagai fungsi azimuth di bidang horisontal untuk telinga kiri (panel kiri) merupakan pencerminan dari waktu-waktu kedatangan sebagai fungsi azimuth di bidang horisontal untuk telinga kanan (panel kanan).

Fitur-fitur yang terlihat dalam Gambar 2.15 dapat dijelaskan dengan mengingat fisik yang terlibat. Sebagai contoh, pita-pita terang gelap disebabkan oleh refleksi-refleksi daun telinga. Puncak impuls yang datang sekitar 0,4 ms setelah puncak awal disebabkan oleh refleksi pundak. Respon impuls ketika

sumber bunyi terletak di depan agak mirip dengan respon impuls ketika sumber bunyi berada di belakang. Terlihat dari Gambar 2.15 (panel kanan), perbedaan-perbedaan yang tampak menunjukkan simetri yang tidak sempurna di sekitar garis vertikal $\theta=90^{\circ}$. Hal ini terlihat dari ,lembah' gelap setelah pulsa awal yang lebih dominan untuk sumber bunyi di depan daripada di belakang. Manusia sering memiliki masalah dalam membedakan sumber bunyi di depan atau di belakang dan memecahkan masalah ini dengan menggerakkan kepala.

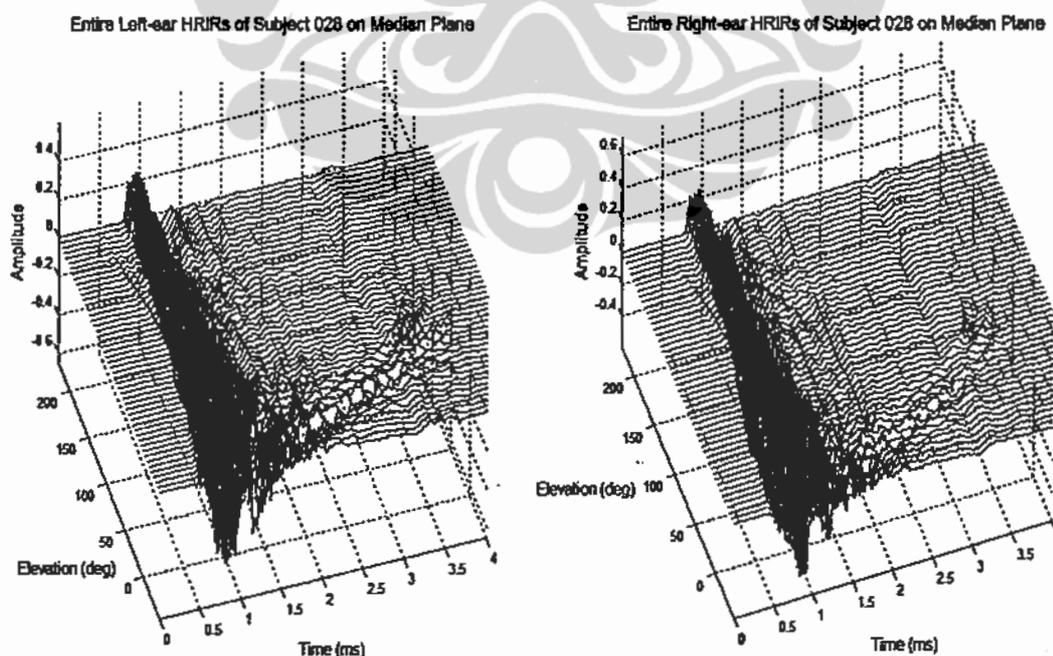


Gambar 2.14. Grafik 3D Seluruh HRIR dari Subyek 028 pada Bidang Horisontal

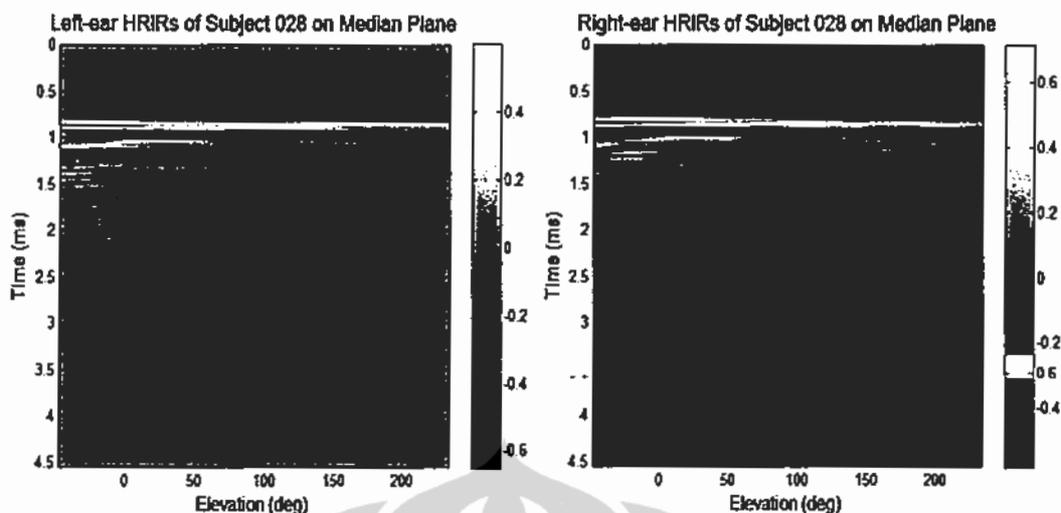


Gambar 2.15. *Image* Seluruh HRIR dari Subyek 028 pada Bidang Horisontal

Gambar 2.16 menunjukkan grafik-grafik tiga dimensi dari seluruh HRIR subyek 028 pada bidang median, sedangkan Gambar 2.17 merupakan representasi *image* dari seluruh HRIR yang sama seperti pada Gambar 2.16, dimana intensitas HRIR ditunjukkan oleh derajat kecerahan titik pada gambar tersebut. Kedua gambar ini juga memberikan interpretasi yang serupa. Terlihat dari kedua gambar, bahwa ketika sumber bunyi bergerak mengelilingi kepala pada bidang median, maka variasi-variasi HRIR jauh lebih kecil. Hal ini terlihat dari Gambar 2.16 (panel kiri dan panel kanan) bahwa semua HRIR di bidang median hampir sama jika ditinjau dari segi struktur detail, amplitudo dan waktu kedatangan gelombang bunyi. Waktu kedatangan juga kurang lebih sama untuk semua HRIR seperti yang diharapkan. Hal ini disebabkan karena pada bidang median, jarak antara sumber bunyi dengan setiap telinga adalah sama (dengan anggapan kepala manusia bersifat simetri sempurna). Variasi-variasi utama HRIR terletak pada waktu-waktu kedatangan relatif dan intensitas dari refleksi-refleksi daun telinga. Hal ini tampak pada HRTF *magnitude* di kawasan frekuensi sebagai suatu lembah dengan komponen frekuensi yang bervariasi menurut elevasi, seperti tampak pada Gambar 2.13.



Gambar 2.16 Grafik 3D Seluruh HRIR dari Subject 028 pada Bidang Median



Gambar 2.17 *Image* Seluruh HRIR dari Subject 028 pada Bidang Horizontal [2]

Perbedaan antara HRIR untuk sumber bunyi di depan dan HRIR untuk sumber bunyi di belakang ditunjukkan oleh ketidaksempurnaan simetri kecil tetapi kelihatan jelas di sekitar suatu garis vertikal pada elevasi $\phi=90^{\circ}$, seperti terlihat dari Gambar 2.17. Dari kedua panel Gambar 2.17, diamati bahwa kedua *image* hampir sama, dimana setiap *image* merepresentasikan semua HRIR untuk telinga kiri dan untuk telinga kanan. Garis samar yang menurun pada sekitar $\phi=30^{\circ}$, pada bagian atas Gambar 2.17, telah diselidiki [18] sebagai pantulan lantai. Tentu saja, sangat sulit memperoleh data ekperimental yang bebas dari cacat.

BAB 3

PEMODELAN *HEAD-RELATED TRANSFER FUNCTIONS* BERDASARKAN *PRINCIPAL COMPONENTS ANALYSIS*

Di dalam bab ini dibahas mengenai metode pemodelan HRTF menggunakan PCA beserta hasil-hasilnya. Pemodelan dilakukan baik terhadap berbagai jenis data HRIR di kawasan waktu maupun terhadap berbagai jenis data HRTF di kawasan frekuensi, dengan maksud agar dipilih suatu model terbaik di kawasan waktu dan satu model terbaik di kawasan frekuensi. Berbagai jenis data HRIR dan HRTF diperoleh dari beberapa pengolahan awal seperti yang dijelaskan dalam subbab 1.6. Di dalam bagian-bagian berikutnya dari bab ini, dijelaskan mengenai jenis-jenis pengolahan awal data HRIR, aproksimasi fasa minimum dari HRTF, metode pemodelan HRTF menggunakan PCA, hasil-hasil pemodelan PCA dari berbagai jenis data HRIR, analisis vektor-vektor basis PCA dari model HRIR fasa minimum, analisis kesalahan model HRIR fasa minimum, analisis vektor-vektor basis PCA dari model HRTF *magnitude*, dan analisis kesalahan model HRTF *magnitude*.

3.1 Jenis-jenis Pengolahan Awal Data HRIR

Dalam kawasan waktu, digunakan jenis data berikut dalam pemodelan HRIR: HRIR asli dengan jumlah sampel 120 dan 200 (langsung dari CIPIC HRTF *Database*, tanpa pengolahan awal); HRIR pulsa langsung ($HRIR_{dp}$) dengan jumlah sampel 67 dan 120, $HRIR_{dp}$ adalah HRIR dengan waktu tunda inisial dihilangkan; dan HRIR fasa minimum ($HRIR_{mp}$) dengan jumlah sampel 67, 120 dan 200. Kemudian terhadap setiap jenis HRIR di atas dilakukan empat jenis pengolahan awal lanjutan seperti ditunjukkan daalam Gambar 1.1 dan Gambar 1.2 pada subbab 1.6. Sedangkan dalam kawasan frekuensi, jenis-jenis data yang digunakan dalam pemodelan PCA adalah HRTF *magnitude* dari DFT 256-titik terhadap $HRIR_{dp}$; dengan menggunakan DFT 256-titik terhadap HRIR asli, diperoleh HRTF *magnitude*, HRTF *log-magnitude*, HRTF *log10-magnitude* dan HRTF dB. Penjelasan yang rinci dari pengolahan awal yang dilakukan untuk

mendapatkan jenis-jenis data HRTF di atas, dapat dilihat pada subbab 1.6. Sebelum dimasukkan ke dalam PCA, terhadap kelima jenis HRTF di atas dilakukan pengolahan awal lanjutan seperti pada data HRIR. Pengolahan awal lanjutan ini dapat dilihat pada Gambar 1.2 dan dijelaskan secara rinci dalam subbab 1.6.

Jenis-jenis data HRIR yang disebutkan di atas juga digunakan dalam penelitian-penelitian sebelumnya dalam pemodelan HRIR menggunakan PCA. HRIR pulsa langsung digunakan oleh Shin dan Park [35]; serta Hwang et al. [36,37]. HRIR fasa minimum dimodelkan oleh Hugeng et al. [41]. Wu et al. menggunakan HRIR yang dinormalisasi [27]. HRTF *magnitude* digunakan dalam [25,26,52]; HRTFs *log-magnitude* dalam [21,23,38,39], dan HRTFs *log-magnitude* yang dinormalisasi digunakan dalam [53]. Setiap pengolahan awal yang terlibat dalam model-model PCA di atas, dijelaskan lebih rinci dalam referensi-referensi yang diberikan.

Penelitian-penelitian di atas dalam pemodelan HRIR menggunakan PCA, baik dalam kawasan waktu maupun kawasan frekuensi, mengimplementasikan basis data HRIR yang berbeda, yang diperoleh dari pengukuran-pengukuran HRIR oleh institusi-institusi yang berbeda dan menggunakan subyek-subyek berupa manusia, kucing hidup yang dianestesi dan manikin akustik. Para peneliti tersebut menggunakan jumlah vektor basis PCA yang berbeda yakni 4 - 10 vektor basis. Demikian juga kinerja dari pemodelan diuji dengan parameter-parameter kesalahan yang berbeda, yakni *Spectral Distortion score* (nilai SD) dan *Mean-square Error* (MSE). Karena subyek-subyek, basis data HRIR, dan jumlah vektor basis yang berbeda yang digunakan dalam PCA, suatu perbandingan yang baik dan obyektif antara model-model di atas sangatlah sulit diperoleh. Oleh sebab itu, dilakukan perbandingan yang komprehensif dari berbagai pengolahan awal yang disebutkan terdahulu dalam memodelkan HRIR berdasarkan PCA dengan menggunakan 10 vektor basis dan CIPIC HRTF *Database*. Tujuan dari perbandingan ini adalah untuk mendapatkan suatu pengolahan awal yang efektif dalam pemodelan HRIR menggunakan PCA [40].

Seperti diusulkan dalam [36,37], HRIR pulsa langsung diperoleh dengan cara menghilangkan waktu tunda inisial dan mengambil sampel-sampel awal yang

berdurasi 1,5 ms atau 67 sampel dengan frekuensi cuplik, f_s , sebesar 44.100 Hz, sejak kedatangan pulsa langsung. Waktu tunda inisial dari HRIR menunjukkan waktu propagasi bunyi dari sumbernya ke gendang telinga manusia. Jika nanti waktu tunda ini diperlukan, maka waktu tunda ini dapat disisipkan kemudian. Respon impuls sebesar 1,5 ms termasuk efek-efek daun telinga, kepala dan torso manusia. Seperti dijelaskan dalam lampiran C, HRIR fasa minimum dapat diperoleh melalui perhitungan *real cepstrum* dari HRIR asli, yang memiliki fasa sembarang. Dapat dikatakan bahwa HRIR fasa minimum merupakan versi HRIR dengan waktu tunda inisial yang dihilangkan, mirip seperti HRIR pulsa langsung. HRIR asli dan HRIR fasa minimum yang dihitung dari HRIR asli tersebut, memiliki spektrum *magnitude* yang sama dalam kawasan frekuensi [55].

3.2 Aproksimasi Fasa Minimum dari HRTF

Dalam subbab ini dibahas sifat fasa minimum dari HRTF. Untuk sintesis audio binaural, suatu HRTF pada umumnya diimplementasikan sebagai kaskade antara suatu waktu tunda murni (*pure delay*) dan filter fasa minimum, seperti ditunjukkan Gambar 3.1. Penggunaan model ini mengurangi biaya dari sintesis binaural dengan cara memperpendek panjang filter FIR yang digunakan dan mengurangi jumlah komponen yang diperlukan dalam metode-metode dekomposisi linier. Ditambah lagi, interpolasi yang halus HRIR pada posisi-posisi diskrit, yang diperlukan dalam simulasi sumber bunyi yang bergerak, sangat diuntungkan dari representasi fasa minimum.



Gambar 3.1. Model Fasa Minimum dari HRTF

Studi-studi yang menginvestigasi HRTF sebagai fungsi fasa minimum telah dilakukan oleh banyak peneliti. Mehrgardt dan Mellet [56] menemukan bahwa HRTFs mendekati sistem fasa minimum untuk frekuensi-frekuensi sampai dengan 10 kHz. Wightman dan Kistler [21], dalam memodelkan HRTF dengan PCA, menganggap bahwa HRTF hasil pengukuran adalah fungsi-fungsi dengan perbedaan-perbedaan fasa antar-telinga yang diaproksimasi dengan suatu waktu

tunda murni. Kulkarni et al. [22] memeriksa validitas dari model fasa minimum dengan tes pendengaran subyektif dengan menggunakan spektrum-spektrum fasa yang dimodifikasi, termasuk fasa minimum-*plus-delay*, fasa linier, dan fasa terbalik-*plus-delay*. Mereka juga menganalisis similaritas antara suatu HRTF dan komponen fasa minimumnya. Walaupun studi-studi ini mempergunakan metode-metode yang berbeda untuk mengestimasi ITD, kesimpulan umum dari percobaan-percobaan subyektif adalah bahwa model fasa minimum tidak dapat dipisahkan dari HRTF asalnya, jika ITD yang tepat digunakan. Pertanyaan yang muncul adalah apakah kesimpulan tadi muncul karena pendengaran manusia yang tidak sensitif, atau apakah HRTF itu sendiri pada intinya bersifat fasa minimum. Kulkarni et al. [22] dan Nam et al. [57] memeriksa fasa dari HRTF dan membuktikan bahwa HRTF pada intinya adalah sistem fasa minimum.

Untuk mengevaluasi konten fasa minimum dari HRTF, dihitung koherensi-koherensi silang (*cross-coherences*) antara HRTF dan versi fasa minimumnya. Secara teoritik, jika HRTF adalah suatu waktu tunda murni diikuti dengan suatu sistem fasa minimum, koherensi silang akan bernilai 1,0 (satu) karena kedua filter adalah versi pergeseran waktu (*time shifted version*) satu sama lain. Di sini koherensi silang maksimum antara HRTF asli dan HRTF fasa minimum disebut sebagai *retained coherence*. Sebagian besar HRTF yang diselidiki dalam [57] memiliki *retained coherence* lebih besar dari 0,9; yang mengindikasikan bahwa sebagian besar energi HRTF terletak pada bagian fasa minimum dari HRTF tersebut. Kulkarni et al. [22] menghitung *retained coherence* untuk sejumlah HRTF dari dua subyek dan mencatat bahwa mayoritas HRTF yang diselidiki memiliki nilai *retained coherence* di atas 0,9.

Pendekatan lain untuk mengevaluasi konten fasa minimum dari HRTF adalah dengan memeriksa *excess group delay* [57]. Jika nilai *excess group delay* adalah fungsi konstan dari frekuensi, maka HRTF tersebut adalah suatu waktu tunda murni yang diikuti suatu filter fasa minimum. Dalam semua HRTF dari pengukuran, dua jenis *excess group delay* yang tidak konstan ditemukan, yaitu satu jenis yang berkaitan dengan daerah-daerah energi spektral moderat dan jenis lain berkaitan dengan lembah-lembah spektral yang terisolasi; kedua jenis tersebut

biasanya ditemukan pada frekuensi-frekuensi di atas 10 kHz. Selanjutnya dijelaskan HRTF *retained coherence* dan HRTF *group delay*.

3.2.1 HRTF Coherence

HRTF pada frekuensi ω untuk posisi bunyi dengan azimut θ dan elevasi ϕ , (θ, ϕ) dituliskan sebagai $H(\omega; \theta, \phi)$. Tanpa menuliskan posisi bunyi, HRTF dapat ditulis dalam bentuk *magnitude-phase* sebagai berikut:

$$H(\omega) = |H(\omega)| e^{j\varphi(\omega)}, \quad (3.1)$$

dimana $|H(\omega)|$ adalah respon *magnitude* dan $\varphi(\omega)$ adalah respon fasa, yang dapat diuraikan ke dalam respon fasa minimum $\varphi_{mp}(\omega)$ dan respon *excess phase* $\eta(\omega)$,

$$\varphi(\omega) = \varphi_{mp}(\omega) + \eta(\omega). \quad (3.2)$$

Group delay $\tau(\omega)$ dari $H(\omega)$ didefinisikan sebagai nilai negatif dari derivatif dari respon fasa,

$$\tau(\omega) = -\text{Im} \left\{ \frac{d}{d\omega} \log H(\omega) \right\} = -\frac{d}{d\omega} \varphi(\omega). \quad (3.3)$$

Khususnya, *excess group delay* $\gamma(\omega)$ adalah :

$$\gamma(\omega) = -\frac{d}{d\omega} \eta(\omega). \quad (3.4)$$

Karena fungsi alih fasa minimum adalah:

$$H_{mp}(\omega) = |H(\omega)| e^{j\varphi_{mp}(\omega)}, \quad (3.5)$$

maka HRTF dapat ditulis sebagai produk dari fungsi alih fasa minimum dan respon *excess phase*,

$$H(\omega) = H_{mp}(\omega) e^{j\eta(\omega)}. \quad (3.6)$$

Seperti diketahui, dalam representasi standar HRTF yang digunakan dalam sintesis binaural, respon *excess phase* $\eta(\omega)$ adalah suatu fasa linier, $\eta(\omega) = \omega\tau$.

Koherensi silang $\psi_{xy}(l)$ antara sinyal $x(n)$ dan $y(n)$ adalah:

$$\psi_{xy}(l) = \frac{\sum_{n=0}^{\infty} x(n-l)y(n)}{\sqrt{\sum_{n=0}^{\infty} x^2(n) \sum_{n=0}^{\infty} y^2(n)}}, \quad (3.7)$$

dimana l adalah pergeseran waktu antara kedua sinyal pada mana koherensi dievaluasi. Terlihat dari persamaan (3.7) bahwa koherensi antara dua sinyal $x(n)$ dan $y(n)$ adalah korelasi silang antara kedua sinyal tersebut yang dinormalisasi

oleh rata-rata geometrik dari energi-energi sinyal. Koherensi silang maksimum sepanjang *lag* l ,

$$\psi_{\max} = \max_l \{\psi_{xy}(l)\}, \quad (3.8)$$

merupakan suatu indikasi similaritas antara sinyal-sinyal $x(n)$ dan $y(n)$.

Di sini digunakan koherensi silang maksimum antara HRTF dan versi fasa minimumnya sebagai suatu cara mengkuantisasi konten fasa minimum HRTF. Koherensi maksimum ini disebut di sini sebagai *retained coherence*, yang dilambangkan dengan $\kappa(\theta, \phi)$, sebagai berikut:

$$\kappa(\theta, \phi) = \max_l \left\{ \frac{\sum_{n=0}^{\infty} h_{mp}(n-l)h(n)}{\sum_{n=0}^{\infty} h^2(n)} \right\} \quad (3.9)$$

dimana $h(n)$ adalah HRIR asli yang dievaluasi pada azimut θ , dan elevasi ϕ , $h_{mp}(n)$ adalah HRIR fasa minimum. Jika $h(n)$ terdiri dari suatu waktu tunda murni dan deret fasa minimum, *retained coherence* $\kappa(\theta, \phi)$ akan bernilai 1,0. Dengan demikian, jika $\kappa(\theta, \phi)$ mendekati 1,0, maka $h(n)$ hampir identik bentuknya dengan $h_{mp}(n)$.

Untuk mengerti sifat-sifat dari sistem *allpass*, maka diamati suatu filter *allpass* orde-1, $G(z)$, yang terdiri dari suatu *pole* pada $z = r$ dan suatu *zero* $z = 1/r$,

$$G(z) = \frac{z^{-1} - r}{1 - rz^{-1}}. \quad (3.10)$$

Dengan mensubstitusi $z = e^{j\omega}$, maka diperoleh DFT dari $G(z)$, sebagai berikut:

$$G(e^{j\omega}) = \frac{e^{-j\omega} - r}{1 - re^{-j\omega}}, \quad (3.11)$$

dimana $e^{-j\omega}$ dapat ditulis dalam bentuk rasio polinomial-polinomial dalam $s=j\omega$ sebagai berikut:

$$e^{-j\omega} = \frac{1 - \frac{s}{2} + \frac{s^2}{8} - \dots}{1 + \frac{s}{2} + \frac{s^2}{8} + \dots} \approx \frac{1 - \frac{s}{2}}{1 + \frac{s}{2}}. \quad (3.12)$$

Persamaan (3.12) dan $s=j\omega$ disubstitusi ke dalam persamaan (3.11) untuk memberikan:

$$G(e^{j\omega}) = \frac{(1-r) - j\frac{\omega}{2}(1+r)}{(1-r) + j\frac{\omega}{2}(1+r)}. \quad (3.13)$$

Dari persamaan (3.13) dapat dihitung *magnitude* dari $G(e^{j\omega})$ seperti berikut ini:

$$|G(e^{j\omega})|=1, \quad (3.14)$$

dan fasa dari $G(e^{j\omega})$,

$$\eta(\omega) = -2 \arctan \frac{\frac{\omega}{2}(1+r)}{1-r}. \quad (3.15)$$

Dari persamaan (3.14) dan (3.15) di atas, dapat dituliskan filter *allpass* orde-1, $G(e^{j\omega})$, sebagai berikut:

$$G(e^{j\omega}) = e^{j\eta(\omega)}. \quad (3.16)$$

Terlihat jelas dari persamaan (3.16) bahwa filter *allpass* orde-1, $G(e^{j\omega})$, merupakan waktu tunda murni dengan fasa sebesar $\eta(\omega)$; dimana $\eta(\omega) = \omega\tau$. Hal ini membuktikan persamaan (3.6) yang menyatakan bahwa suatu HRTF adalah kaskade antara filter *allpass*, yang berbentuk waktu tunda murni $e^{j\eta(\omega)}$, dan filter fasa minimum $H_{mp}(\omega)$.

Untuk mengerti sifat-sifat dari *retained coherence* dalam kehadiran fasa linier dan waktu tunda tidak murni, perhatikan suatu filter *allpass* orde-1, $G(z)$, dalam persamaan (3.10), dengan suatu *pole* $z = r$ dan suatu *zero* $z=1/r$. Gambar 3.2a menunjukkan respon-respon impuls yang berkaitan dengan $G(z)$ untuk radius *pole* r yang berbeda. Pada kasus-kasus ekstrim ($r=0$ dan $r=0,99$), respon impuls sangat mirip fungsi impuls dan *retained coherence* sangat mendekati 1,0. Sebaliknya, respon-respon impuls yang berkaitan dengan radius-radius di antaranya ($0 < r < 0,99$) cenderung mengalami dispersi, yang mengurangi *retained coherences*. Untuk filter *allpass* orde-1 $G(z)$, *retained coherence* diberikan oleh:

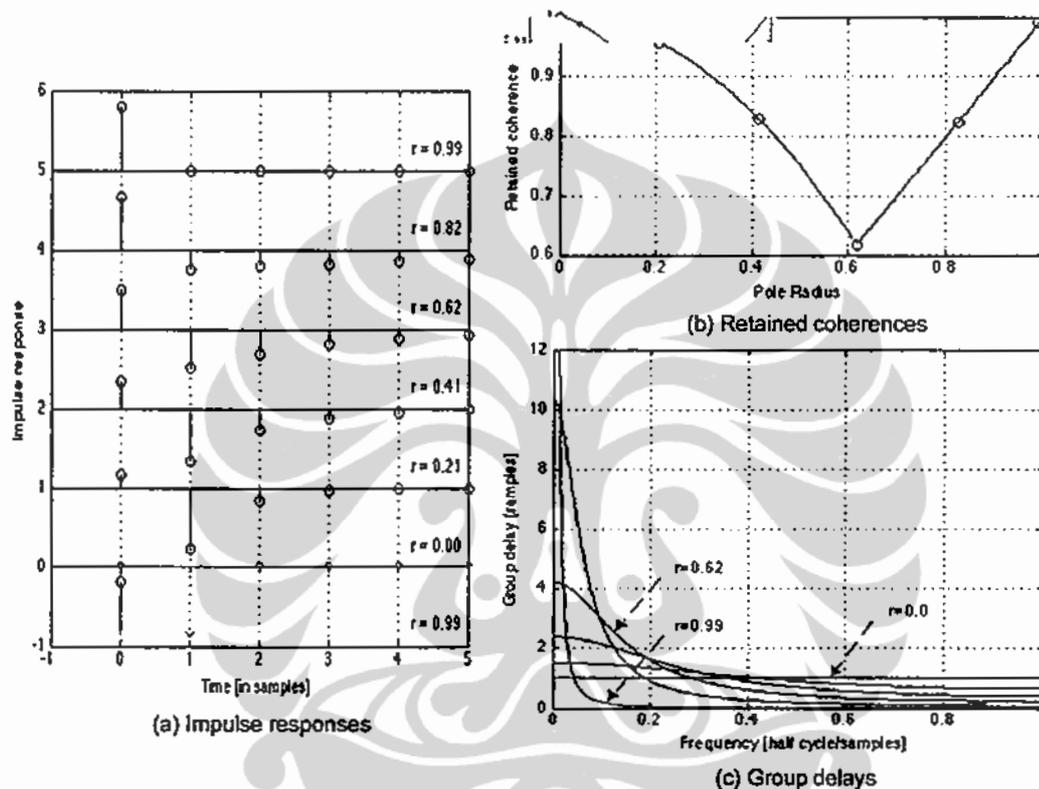
$$\kappa = \max(|r|, 1 - r^2), \quad (3.17)$$

dan *group delay* yang berkaitan adalah:

$$\tau(\omega) = \frac{1-r^2}{1-2r\cos\omega+r^2}. \quad (3.18)$$

Retained Coherence ditunjukkan dalam Gambar 3.2b, dan *group delay* yang berkaitan ditunjukkan Gambar 3.2c. Pada $r=0$, *group delay* adalah suatu fungsi

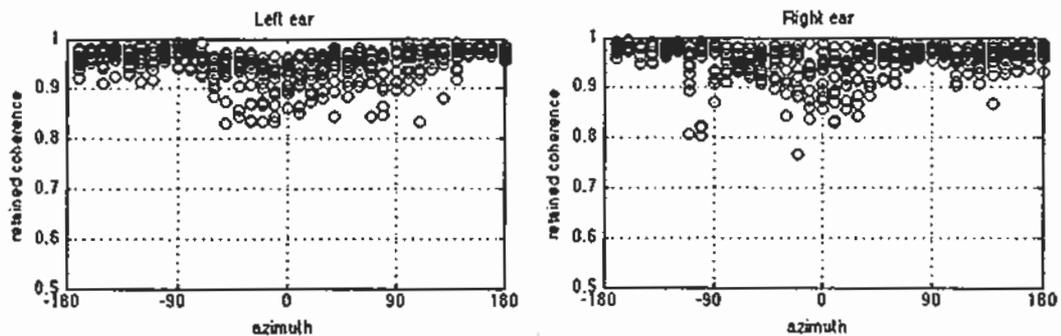
konstan, yang dengan jelas memiliki *retained coherence* sebesar 1,0. Pada $r=0,99$, walaupun *group delay* memiliki suatu puncak tajam, tetapi *group delay* ini rata sepanjang mayoritas pita frekuensi, dan *retained coherence* bernilai hampir 1,0. Tetapi puncak-puncak moderat dalam *group delay* seperti pada $r=0,62$ berkaitan dengan suatu *retained coherence* yang kecil; yang menunjukkan bahwa fitur-fitur *group delay* yang non-konstan seperti itu mengurangi *retained coherence*.



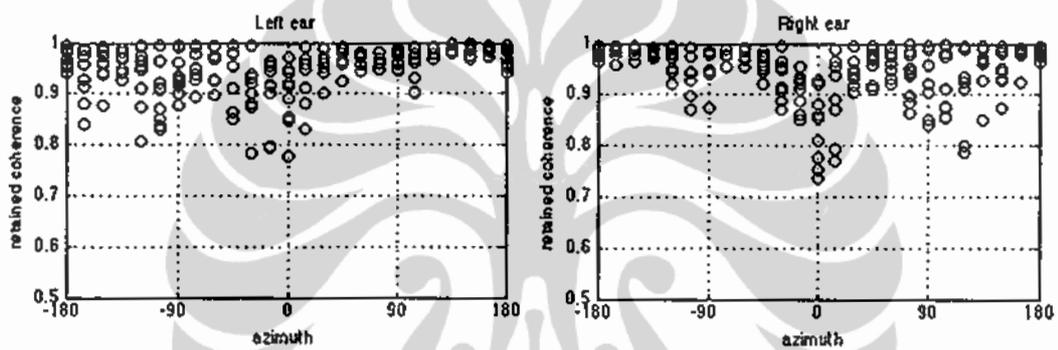
Gambar 3.2. Filter Allpass Orde-1

Tiga basis data HRTF digunakan untuk mengeksplorasi fasa HRTF oleh Nam et al.[57]. Satu basis data, "CCRMA HRTFs", terdiri dari 8 subyek yang diukur di dalam studio rekaman CCRMA, menggunakan mikrofon-mikrofon *blocked meatus*. HRTF diukur setiap 15° azimuth dan setiap 10° elevasi, dengan jangkauan dari -40° sampai $+40^\circ$. Basis data kedua, "Snapshot HRTFs", melibatkan 23 subyek dan HRTF diukur menggunakan mikrofon-mikrofon *blocked meatus*. Respon-respon impuls diukur setiap 30° azimuth dan setiap 18° elevasi, dengan jangkauan dari -36° sampai $+54^\circ$, menggunakan kode-kode Golay untuk mengukur *raw HRIR*. Basis data ketiga, "WK HRTFs", dari Wightman dan

Kistler di *University of Wisconsin-Madison*, terdiri dari 5 subyek, yang diukur pada 505 posisi bunyi dengan menggunakan mikrofon-mikrofon *open meatus*.



Gambar 3.3. *Retained Coherence* untuk Seorang Subyek: HRTF Diukur dengan Mikrofon *Open Meatus* [57]



Gambar 3.4. *Retained Coherence* untuk Seorang Subyek: HRTF Diukur dengan Mikrofon *Blocked Meatus*[57]

Contoh dari *retained coherences* untuk seorang subyek WK dan seorang subyek CCRMA masing-masing dapat dilihat pada Gambar 3.3 dan Gambar 3.4. Perhatikan bahwa kebanyakan *retained coherence* dari HRTF bernilai di atas 0,9, dan bahwa *retained coherence* dari HRTF bernilai di bawah 0,9 terkonsentrasi pada posisi-posisi depan dan sisi *ipsilateral*. *Retained coherence* yang bernilai di bawah 0,8 sangat jarang ditemui.

3.2.2 HRTF Group Delay

Retained Coherence dapat dinyatakan sebagai suatu jumlah yang diboboti dari kesalahan-kesalahan fasa:

$$\kappa = \max_i \left[\frac{F_D^{-1} \{H(\omega)H_{mp}(\omega)\}}{\sum_{\omega} |H(\omega)|^2} \right] \quad (3.19)$$

$$\kappa = \max_i \left\{ \sum_{\omega} V(\omega) \cos(\omega l - \eta(\omega)) \right\}, \quad (3.20)$$

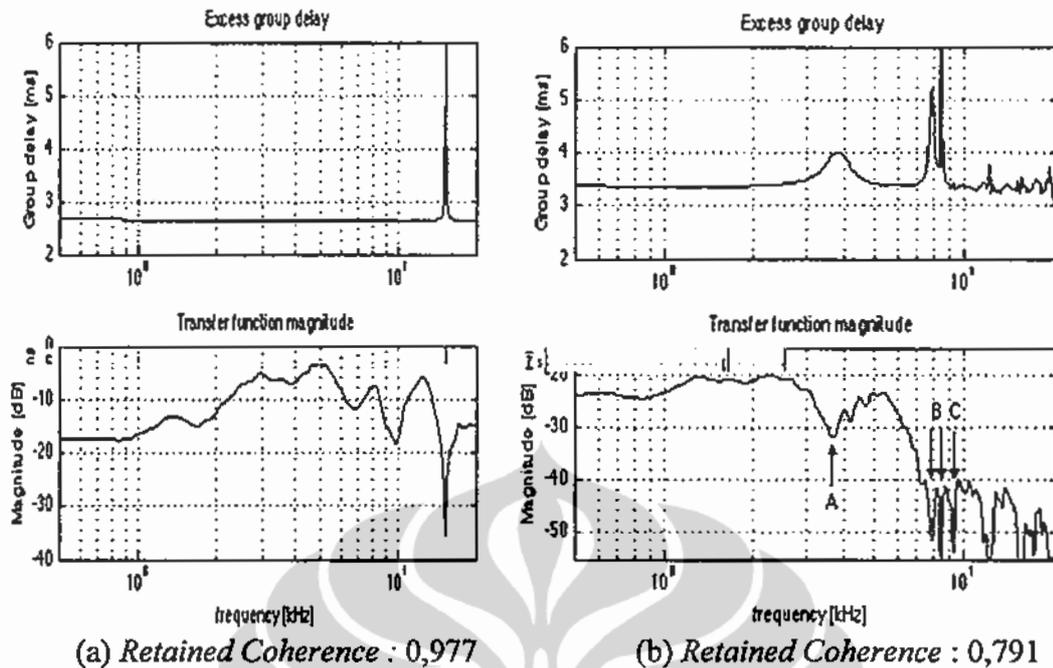
dimana bobot $V(\omega)$ sebanding dengan daya HRTF pada frekuensi ω ,

$$V(\omega) = \frac{|H(\omega)|^2}{\sum_{\omega} |H(\omega)|^2}. \quad (3.21)$$

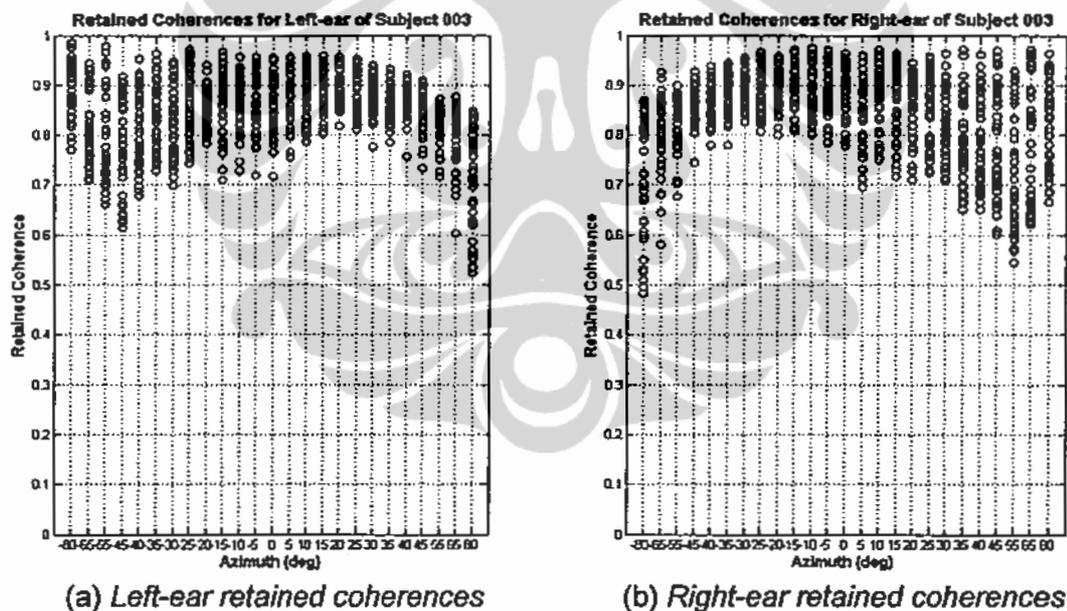
Perhatikan bahwa *excess-phase* adalah suatu fungsi linier dari frekuensi $\eta(\omega) = \omega l$ dan penjumlahan dalam persamaan (3.20) mencapai nilai maksimum dengan $l = \tau$, sehingga suatu *retained coherence* sebesar 1,0 didapatkan. Pada umumnya, *retained coherence* berkurang dari 1,0 dalam kasus-kasus dimana *excess phase* terdeviasi dari suatu karakteristik linier, dan HRTF memiliki energi pada frekuensi-frekuensi dimana deviasi-deviasi terjadi. Hal ini mengarah pada ide bahwa similaritas antara HRTF dan pasangan fasa minimumnya dapat ditemukan dengan memeriksa *excess group delay* dan respon *magnitude*.

Nilai-nilai *excess group delay* dari HRTF dihitung untuk HRTF hasil pengukuran. Banyak HRTF memiliki *excess group delay* yang tidak konstan, terutama HRTF pada azimuth-azimut *contralateral*. Gambar 3.5 menunjukkan dua *excess group delay* tipikal dan respon-respon *magnitude*, satu pada suatu azimuth *ipsilateral* dan yang lain pada azimuth *contralateral*. *Excess group delay* dalam Gambar 3.5a dicirikan oleh suatu puncak tinggi dengan lebar pita sempit. Terdapat suatu lembah spektral pada frekuensi yang sama, dan karena itu *retained coherence* masih mendekati 1,0. Fitur lebar pita sempit ini mungkin berasal dari suatu *zero non fasa minimum* yang terisolasi.

Daerah-daerah yang serupa dari *excess group delay* tidak konstan dapat dilihat pada Gambar 3.5b. Puncak-puncak tajam dari *group delay* berhubungan dengan lembah-lembah spektral, seperti lembah "B" dalam Gambar 3.5b, dan karena itu hanya sedikit berpengaruh pada *retained coherence*. Lembah "C" berhubungan dengan suatu daerah dengan *group delay* konstan dan dengan suatu *zero fasa minimum*. Jenis fitur lain yang kadang-kadang muncul adalah suatu puncak *group delay* lebar yang berhubungan dengan suatu daerah dari energi spektral moderat, seperti fitur yang berlabel "A" dalam Gambar 3.5b. Dalam hal ini, *retained coherence* berkurang dari 1,0.



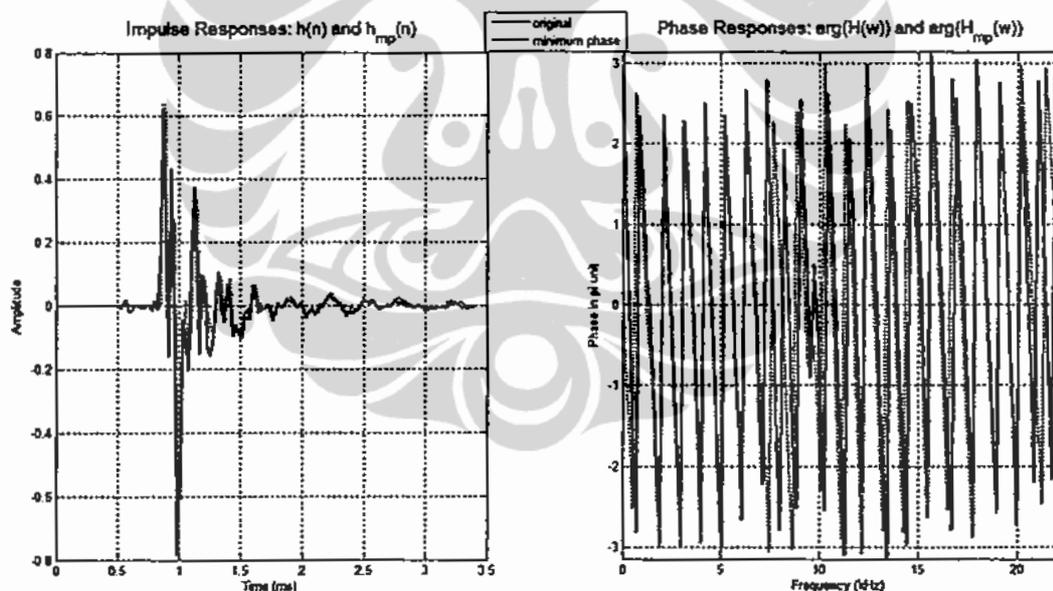
Gambar 3.5. Excess Group Delay dan Respon Magnitude pada
(a) azimuth= 90° , elevasi= 0° ; (b) azimuth= 255° , elevasi= 0° [57]



Gambar 3.6. Retained Coherences Subyek 003 untuk Semua Posisi Bunyi

Dalam penelitian ini, dihitung koherensi silang maksimum (disebut *retained coherence*) untuk telinga kiri dan telinga kanan dari subyek 003 antara respon-respon impuls HRIR asli yang berasal dari CIPIC HRTF Database dan respon-respon impuls HRIR fasa minimum yang berasal dari rekonstruksi fasa minimum dari HRIR asli. Gambar 3.6a menunjukkan nilai-nilai *retained*

coherence untuk semua posisi sumber bunyi dari telinga kiri subyek 003. Terlihat dari Gambar 3.6a, bahwa *retained coherence* kurang dari 0,6 terjadi untuk sumber-sumber bunyi pada bidang vertikal dengan azimut 80° (sisi *contralateral* dari telinga kiri). Persentase jumlah nilai *retained coherence* yang lebih besar dari 0,8 untuk telinga kiri subyek 003 adalah 77,84%. Dari Gambar 3.6b, ditunjukkan nilai-nilai *retained coherence* untuk semua posisi sumber bunyi dari telinga kanan subyek 003. Fenomena yang sama terjadi dalam kasus nilai-nilai *retained coherence* kurang dari 0,6 terjadi pada posisi-posisi bunyi pada bidang vertikal di sisi *contralateral* dari telinga kanan (azimut -80°). Dihitung dari Gambar 3.6b, persentase jumlah nilai *retained coherence* yang lebih besar dari 0,8 untuk telinga kanan subyek 003 adalah 73,92%. Secara keseluruhan subyek 003, persentase jumlah nilai *retained coherence* yang lebih dari 0,8 adalah sebesar 75,88%. Hal ini berarti bahwa sebagian besar (75,88%) HRIR asli subyek 003 menyerupai versi waktu tunda murni yang diikuti HRIR fasa minimum, yang dihitung dari HRIR asli subyek 003.



Gambar 3.7. Respon Impuls dan Respon Fasa dari HRIR Asli dan HRIR Fasa Minimum pada Posisi Bunyi ($0^\circ, 0^\circ$)

Respon-respon impuls dari HRIR asli hampir sama dengan versi waktu tunda murni yang diikuti respon-respon HRIR fasa minimum, seperti terlihat dari Gambar 3.7. Pada panel kiri Gambar 3.7, ditunjukkan respon-respon impuls yang bersangkutan dari telinga kiri subyek 003 pada posisi bunyi tepat di depan pusat

kepala, $(0^{\circ}, 0^{\circ})$, sedangkan pada panel kanan terlihat respon-respon fasanya. Kedua jenis respon pada masing-masing panel hampir sama satu sama lain.

3.3 Pemodelan HRTF Menggunakan PCA

Penjelasan mengenai pemodelan HRTF menggunakan PCA di bawah ini hanya diberikan bagi HRTF *magnitude* dengan jumlah vektor basis, jumlah subyek dan jumlah posisi sumber bunyi yang sama seperti pemodelan jenis-jenis data lain. Penjelasan dari pemodelan PCA dari jenis-jenis data lain, baik di kawasan frekuensi ataupun di kawasan waktu, adalah hampir sama seperti penjelasan untuk HRTF *magnitude*. Hanya saja, pemodelan HRIR menggunakan sampel-sampel dari HRIR sebagai masukan PCA, sedangkan pemodelan HRTF menggunakan komponen-komponen frekuensi sebagai masukan PCA.

Hanya 128 komponen frekuensi (dari 256 komponen frekuensi) dari suatu HRTF *magnitude*, $|H(\omega)|$, diambil untuk dianalisis karena sifat simetri dari suatu spektrum frekuensi. Semua HRTF *magnitude* dihitung dari seluruh HRIR dari telinga kiri dan telinga kanan dari 45 subyek, yang berkaitan dengan seluruh posisi sumber bunyi sebanyak 1250 posisi di ruang spasial. Terdapat 25 azimuth dan 50 elevasi dalam sistem koordinat *interaural-polar* dari posisi-posisi sumber bunyi untuk setiap telinga dari seorang subyek. Suatu matriks yang terdiri dari *Direct Transfer Functions (DTFs) magnitude* diperlukan oleh PCA agar data masukan menjadi tersentralisasi. Matriks data asli, \mathbf{H} ($M \times N$), terdiri dari semua HRTF *magnitude*, dimana setiap vektor kolom dari matriks \mathbf{H} , \mathbf{h}_i ($i=1,2,\dots,N$), merepresentasikan suatu HRTF *magnitude* dari suatu telinga seorang subyek pada posisi sumber bunyi tertentu. Jumlah HRTF *magnitude* dari seorang subyek adalah 2500 (2 telinga x 1250 posisi untuk satu telinga), sehingga jumlah total HRTF *magnitude* dari 45 subyek adalah 112.500. Jadi ukuran matriks \mathbf{H} adalah 128×112.500 ($M=128$ dan $N=112.500$). Vektor rata-rata empiris, $\boldsymbol{\mu}$ ($M \times 1$), dari seluruh HRTF *magnitude* diberikan oleh:

$$\boldsymbol{\mu} = (1/N) \sum_{i=1}^N \mathbf{h}_i \quad (3.22)$$

Matriks DTF *magnitude*, \mathbf{D} , berasal matriks \mathbf{H} dengan rata-rata empiris, $\boldsymbol{\mu}$, yang dikurangkan dan diberikan oleh:

$$\mathbf{D} = \mathbf{H} - \boldsymbol{\mu} \cdot \mathbf{y}, \quad (3.23)$$

dimana y adalah suatu vektor baris ($1 \times N$) dengan semua komponen bernilai 1. Langkah berikutnya adalah menghitung suatu matriks kovarians S dari data DTF *magnitude*. Matriks kovarians dapat dihitung dari persamaan berikut:

$$S = D.D^*/(N-1) \quad (3.24)$$

dimana $*$ menandai operator konjugasi transpos. Vektor-vektor basis atau *principal components* (PCs), v_i ($i=1,2,\dots,L$), adalah L vektor eigen dari matriks kovarians, S , yang berkaitan dengan L nilai eigen yang terbesar. Jika $L=M$, maka semua DTF *magnitude* dapat direkonstruksi secara sempurna oleh suatu kombinasi linier terbobot dari M vektor basis. Tetapi, pada umumnya L diambil kurang dari M karena tujuan PCA adalah untuk mereduksi dimensi dari data asli. Suatu estimasi dari kumpulan data asli diperoleh di sini hanya dengan menggunakan 10 vektor basis (dipilih $L=10$), yang mencakup lebih dari 90% variasi yang ada dalam data asli D (tepatnya 94,30% variasi dari data DTF *magnitude*). Dengan menggunakan 10 vektor basis untuk memodelkan HRTF *magnitude*, diharapkan diperoleh model-model yang memadai. Matriks vektor basis, $V = [v_1 \ v_2 \ \dots \ v_M]$, yang terdiri dari seluruh vektor basis atau vektor eigen, dapat diperoleh dengan memecahkan persamaan eigen berikut ini:

$$S V = \Lambda V \quad (3.25)$$

dimana $\Lambda = \text{diag}\{\lambda_1, \dots, \lambda_M\}$, adalah suatu matriks diagonal yang terbentuk dari 128 nilai eigen ($M=128$) yang berkaitan dengan 128 komponen frekuensi HRTF *magnitude* yang dianalisis dalam PCA. Setiap nilai eigen, λ_i , merepresentasikan varians sampel dari DTF *magnitude* yang diproyeksikan pada vektor eigen ke- i , v_i .

Kemudian bobot-bobot dari setiap vektor basis (*Principal Component's Weight / PCW*), W ($L \times N$), yang berkaitan dengan seluruh DTF *magnitude*, D , dapat diperoleh dari:

$$W = V^*.D, \quad (3.26)$$

dimana matriks V ($M \times M$) sekarang direduksi menjadi $V = [v_1 \ v_2 \ \dots \ v_L]$. PCW merepresentasikan kontribusi dari setiap vektor basis pada satu DTF *magnitude*. PCW mengandung fitur-fitur spasial dan perbedaan-perbedaan antar individu dari DTF *magnitude*. Jadi, matriks yang terdiri dari model-model HRTF *magnitude*, \hat{H} , diberikan oleh:

$$\hat{\mathbf{H}} = \mathbf{V} \cdot \mathbf{W} + \boldsymbol{\mu} \cdot \mathbf{y}. \quad (3.27)$$

Kinerja dari model-model HRTF *magnitude*, yang dihasilkan oleh PCA, dievaluasi dengan persentase dari perbandingan *mean-square error* dari simpangan antara model HRTF *magnitude* dan HRTF *magnitude* asli yang dihitung dari HRIR basis data, terhadap *mean-square error* dari HRTF *magnitude* asli. Parameter di atas disebut secara sederhana sebagai MSE [21,27,37], yang dihitung dengan persamaan berikut:

$$e_j(\theta, \varnothing) = 100 \% \times \frac{\| \mathbf{h}_j(\theta, \varnothing) - \hat{\mathbf{h}}_j(\theta, \varnothing) \|^2}{\| \mathbf{h}_j(\theta, \varnothing) \|^2} \quad (3.28)$$

dimana $\mathbf{h}_j(\theta, \varnothing)$ adalah HRTF *magnitude* asli yang ke- j , $j=1,2,\dots,N$, pada posisi dengan azimuth θ dan elevation \varnothing , serta $\hat{\mathbf{h}}_j(\theta, \varnothing)$ adalah model HRTF *magnitude* yang bersangkutan. Semakin besar MSE, maka kinerja model HRTF *magnitude* menjadi semakin buruk. Sebaliknya, kinerja yang baik ditandai dengan MSE yang kecil. MSE ini juga digunakan untuk semua jenis data HRIR dan HRTF yang lain dalam menghitung kesalahan pemodelan jenis data yang bersangkutan.

3.4 Hasil-hasil Pemodelan PCA dari Berbagai Jenis Data HRIR

Pada awalnya, dari total 200 sampel di dalam satu HRIR asli, dipilih jumlah sampel HRIR asli yang digunakan langsung dalam pemodelan PCA sebesar 120 dan 200 sampel; di sini tidak dimodelkan HRIR asli dengan 67 sampel karena sebagian besar dari 67 sampel (maksimum sekitar 55 sampel) hanya berisi waktu tunda sebelum pulsa langsung. Lalu HRIR asli dengan 200 sampel digunakan untuk pengolahan awal lanjutan.

Pertama, diambil sampel-sampel mulai dari pulsa langsung (*direct pulse*) sampai pulsa terakhir dari HRIR asli; hasil pengolahan ini disebut HRIR pulsa langsung (HRIR_{dp}). HRIR_{dp} dengan 67 sampel dan 120 sampel dimodelkan masing-masing oleh PCA; di sini tidak dimodelkan HRIR_{dp} dengan 200 sampel karena alasan berikut. Karena jumlah sampel HRIR_{dp} adalah kurang dari 200 (200 sampel HRIR asli dikurangi sampel-sampel waktu tunda), maka untuk memperoleh 200 sampel HRIR_{dp} , dilakukan penambahan pulsa nol (disebut *zero padding*) ke dalam bagian akhir dari HRIR_{dp} sehingga didapatkan HRIR_{dp} dengan 200 sampel. Terlihat di sini bahwa pulsa-pulsa akhir dari HRIR_{dp} 200 sampel pada dasarnya adalah pulsa-pulsa yang mendekati nol dan pulsa-pulsa nol.

Kedua, dilakukan pengolahan data HRIR asli dengan 200 sampel dengan *real cepstrum* untuk menghasilkan 200 sampel HRIR fasa minimum (HRIR_{mp}). Jumlah sampel HRIR_{mp} yang terlibat dalam pemodelan PCA adalah 67, 120 dan 200. Jadi, terdapat tiga jenis data HRIR yang diumpangkan pada PCA, yaitu HRIR asli, HRIR pulsa langsung dan HRIR fasa minimum. Kemudian setiap jenis HRIR tadi dilakukan empat jenis pengolahan awal lanjutan seperti dijelaskan Gambar 1.2 pada subbab 1.6.

MSE rata-rata pemodelan PCA dari seluruh subyek dan seluruh posisi sumber bunyi dari masing-masing data HRIR ditunjukkan oleh Tabel 3.1. Pemodelan dari sekumpulan HRIR asli (no.1 – 5 Tabel 3.1) menghasilkan MSE rata-rata yang mengecewakan. Hal ini disebabkan karena HRIR asli mengandung pulsa-pulsa yang sangat bervariasi dan sangat detail, termasuk di dalamnya waktu tunda kedatangan gelombang bunyi, sehingga kurang baik jika dimodelkan hanya dengan 10 vektor basis. Sedangkan dari kumpulan HRIR pulsa langsung (no. 6 – 10 Tabel 3.1), diperoleh nilai-nilai MSE rata-rata yang kecil hanya ditemui pada jenis data DIR (no.8 Tabel 3.1). Hasil-hasil yang serupa dengan kumpulan HRIR pulsa langsung adalah hasil-hasil dari kumpulan HRIR fasa minimum (no. 11 – 15 Tabel 3.1). Jumlah sampel HRIR fasa minimum tidak mempengaruhi secara signifikan MSE rata-rata dan nilai-nilai MSE rata-rata yang terkecil juga dijumpai pada jenis data DIR. Seperti yang diharapkan, bahwa kedua hasil ini mirip karena kemiripan dari kedua jenis data HRIR yaitu HRIR dengan waktu tunda yang dihilangkan. Hasil-hasil yang baik yang ditemukan dari no. 8 dan no. 13 dapat dipahami terutama karena tidak terlibatnya waktu tunda dalam pemodelan PCA.

Terlihat dari no. 6 dan no. 11 Tabel 3.1, bahwa nilai-nilai MSE rata-rata lebih dari 100%. Nilai ini dimungkinkan oleh definisi MSE pada persamaan (3.28). Hal ini terjadi karena jumlah kuadrat dari perbedaan-perbedaan antara sampel-sampel HRIR dengan sampel-sampel modelnya, yaitu pembilang dari persamaan (3.28), melebihi jumlah kuadrat dari sampel-sampel HRIR tersebut, yaitu penyebut dari persamaan (3.28). Kesalahan-kesalahan pemodelan yang besar terjadi dalam hal ini karena tidak dilakukan *mean centering* pada data HRIR seperti yang disyaratkan oleh PCA sehingga vektor eigen pertama yang dihasilkan tidak merepresentasikan arah varians maksimum.

Tabel 3.1. Persentase MSE Rata-rata dari Berbagai Jenis Data HRIR

No.	Jenis HRIR	Pengolahan Awal I	Jumlah Sample		
			67	120	200
1.	HRIR	Tidak ada	-	47.71	47.89
2.	HRIR _{nc}	Tidak ada	-	44.51	44.79
3.	DIR	Tidak ada	-	46.73	46.92
4.	DIR _{nc}	Tidak ada	-	34.6	35.15
5.	DIR _{nz}	Tidak ada	-	62.6	72.62
6.	HRIR	<i>Direct pulse</i>	100.53	100.68	-
7.	HRIR _{nc}	<i>Direct pulse</i>	11.57	12.24	-
8.	DIR	<i>Direct pulse</i>	8.91	9.59	-
9.	DIR _{nc}	<i>Direct pulse</i>	91.51	87.17	-
10.	DIR _{nz}	<i>Direct pulse</i>	54.02	66.15	-
11.	HRIR	Fasa minimum	100.33	100.16	100.16
12.	HRIR _{nc}	Fasa minimum	7.69	8.36	8.40
13.	DIR	Fasa minimum	7.26	7.93	7.97
14.	DIR _{nc}	Fasa minimum	90.66	86.36	86.68
15.	DIR _{nz}	Fasa minimum	51.92	61.44	69.7

Dari hasil-hasil di atas, maka dalam penelitian ini dipilih model HRIR fasa minimum dengan 67 sampel untuk diindividualisasi dengan MLR, karena model HRIR fasa minimum dengan 67 sampel menghasilkan MSE rata-rata terkecil sebesar 7,26% dibandingkan jenis-jenis data HRIR dalam kawasan waktu lainnya. Model HRIR fasa minimum 67 sampel dianalisis lebih lanjut di dalam subbab-subbab berikut mengenai vektor-vektor basis PCA dan nilai-nilai MSE yang didapatkan dari seorang subyek untuk seluruh posisi sumber bunyi.

3.4.1 Analisis Vektor-vektor Basis PCA dari Model HRIR Fasa Minimum

Pada subbab ini dianalisis vektor-vektor basis yang diperoleh dari pemodelan PCA dengan data DIR fasa minimum sebanyak 67 sampel sebagai data masukan PCA, yaitu data nomor 13 pada Tabel 3.1 di atas. Metode pemodelan PCA yang dilakukan dijelaskan ringkas sebagai berikut. Mula-mula setiap HRIR asli basis data (terdapat seluruhnya 112.500 HRIR) diubah menjadi HRIR fasa minimum (HRIR_{mp}), dengan menggunakan perhitungan dalam lampiran C. Lalu dihitung nilai rata-rata HRIR_{mp} dari semua HRIR_{mp} yang diperoleh. Setiap HRIR_{mp} dikurangi nilai rata-rata HRIR_{mp} menghasilkan DIR fasa minimum (DIR_{mp}). Semua DIR_{mp} digunakan untuk menghitung matriks kovarians sampel S, seperti ditunjukkan oleh persamaan (3.24). Dengan menyelesaikan persamaan

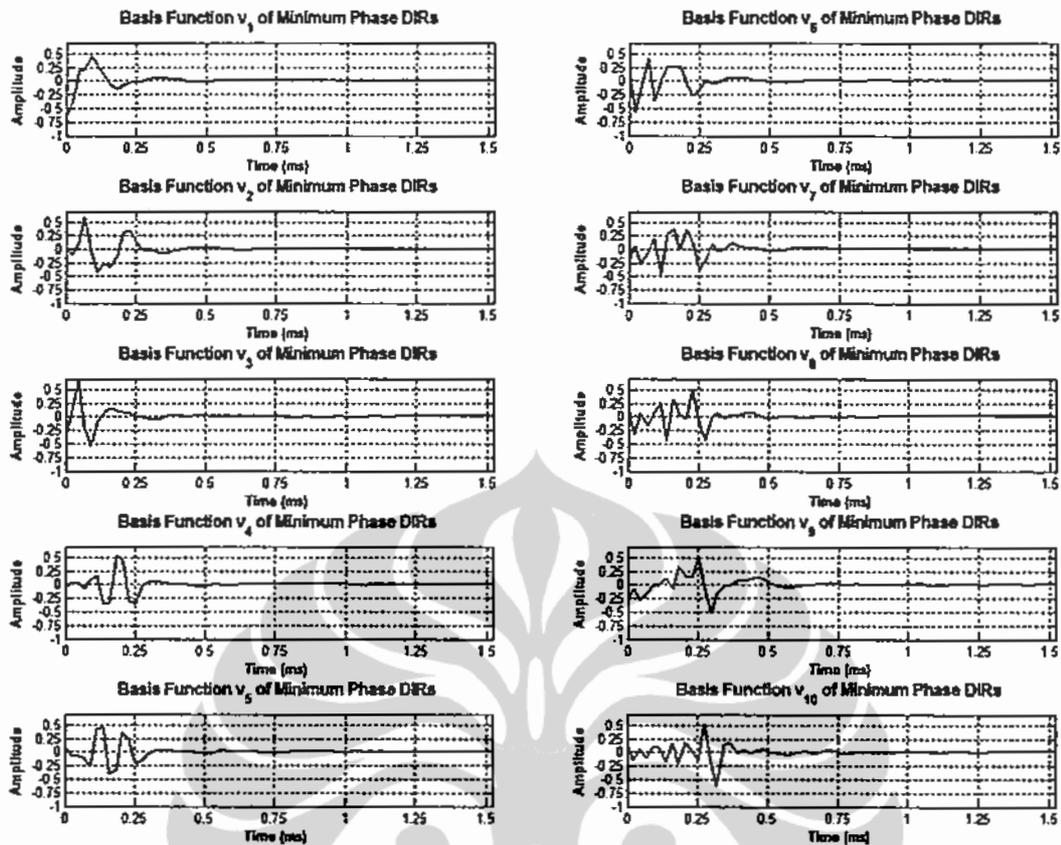
(3.25) yang melibatkan S , diperoleh matriks basis V , yang terdiri dari 67 vektor basis v_i , $i=1,2,\dots,67$, dimana $V = [v_1 v_2 \dots v_{67}]$. Pemodelan DIR_{mp} yang sempurna diperoleh dengan menggunakan 67 vektor basis. Tetapi di sini digunakan hanya 10 vektor basis pertama yang mencakup 90,4% varians dari seluruh DIR_{mp} .

Tabel 3.2 memberikan persentase varians yang tercakup dari masing-masing $v_1 - v_{10}$ (kolom kedua) dan persentase varians kumulatif dari vektor-vektor basis sebelumnya sampai vektor basis ke- i , $i=1,2,\dots,10$ (kolom ketiga). Dari Tabel 3.2, terlihat bahwa vektor basis pertama, v_1 , mencakup % varians terbesar dari seluruh data asli, yaitu sebesar 45%. Kemudian vektor basis kedua, v_2 , mencakup % varians terbesar kedua dan seterusnya. Vektor-vektor basis $v_1 - v_{10}$ mencakup %variens kumulatif sebesar 90,40% dari seluruh data asli.

Gambar 3.8 menunjukkan vektor-vektor basis v_1 sampai v_{10} yang digunakan dalam pemodelan DIR_{mp} dengan metode PCA. Vektor-vektor basis $v_1 - v_{10}$ berkaitan dengan 10 nilai eigen terbesar secara berurutan mulai dari yang terbesar sampai yang terkecil ($\lambda_1 > \lambda_2 > \dots > \lambda_{10}$). Terlihat bahwa v_1 memiliki bentuk paling sederhana. Jika nilai eigen semakin kecil, vektor basis yang bersangkutan memiliki bentuk yang semakin detail. Vektor basis $v_1 - v_5$ memiliki amplitudo tidak nol sampai dengan waktu sekitar 0,3 ms, sedangkan $v_6 - v_{10}$ memiliki amplitudo tidak nol sampai dengan waktu sekitar 0,6 ms. Semua amplitudo vektor-vektor basis terletak pada interval yang sama. Setiap vektor basis merefleksikan sebagian dari struktur-struktur data DIR fasa minimum.

Tabel 3.2 Persentase Varians dari Vektor Basis Model DIR Fasa Minimum

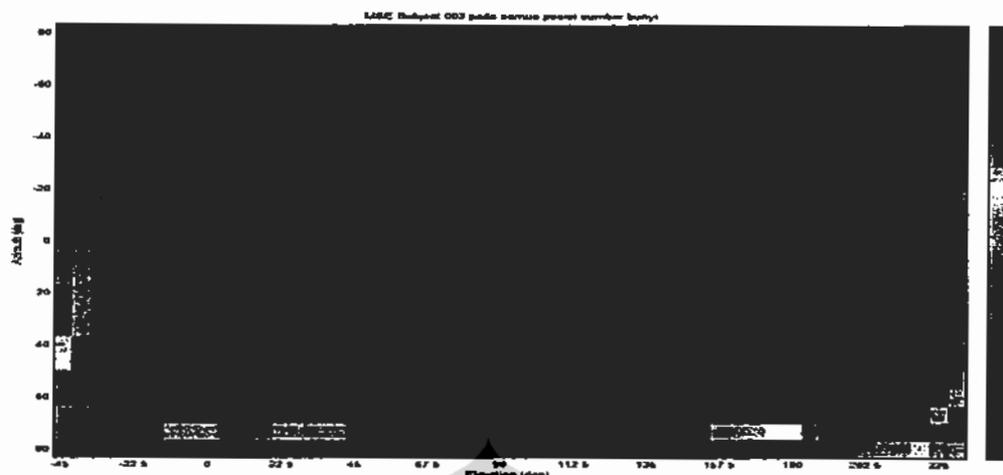
Vektor Basis	% Varians	% Varians Kumulatif
v_1	45,00	45,00
v_2	11,98	56,98
v_3	10,19	67,17
v_4	5,50	72,66
v_5	5,23	77,89
v_6	4,67	82,56
v_7	2,50	85,06
v_8	2,19	87,25
v_9	1,77	89,02
v_{10}	1,39	90,40



Gambar 3.8. Sepuluh Vektor Basis Pertama dari Model DIR Fasa Minimum

3.4.2 Analisis Kesalahan Model HRIR Fasa Minimum

Pada subbab ini dibahas tentang kesalahan pemodelan PCA dari jenis data HRIR nomor 13 pada Tabel 3.1. Dengan menggunakan PCA, diperoleh model-model $HRIR_{mp}$ dengan 10 vektor basis beserta bobot-bobot mereka. Di bagian ini dibahas mengenai sampai seberapa jauh kemiripan atau keakuratan antara model $HRIR_{mp}$ dan $HRIR_{mp}$ asli. Di sini akurasi dari model $HRIR_{mp}$ dihitung menggunakan suatu ukuran jarak antara dua vektor yang dinyatakan dengan MSE. Perhitungan MSE dapat dilihat dari persamaan (3.28). Jika model $HRIR_{mp}$ yang diperoleh sama dengan $HRIR_{mp}$ asli, maka dihasilkan $MSE = 0$. Semakin tidak akurat model $HRIR_{mp}$, maka semakin besar nilai MSE yang dihasilkan.



Gambar 3.9. Persentase MSE antara Model HRIR_{mp} dan HRIR_{mp} Asli

Hasil perhitungan MSE untuk subyek 003 (seorang subyek basis data), pada semua posisi sumber bunyi untuk telinga kiri dapat dilihat pada Gambar 3.9; di sini nilai MSE dinyatakan oleh warna-warna degradasi pada *colorbar* di sisi kanan gambar. Dari hasil pengamatan diperoleh bahwa secara umum terjadi kesalahan-kesalahan yang mirip pada kedua telinga pada daerah-daerah istimewa yang akan dijelaskan selanjutnya. Jadi di sini hanya dibahas kesalahan-kesalahan yang terjadi untuk telinga kiri. Secara umum, kesalahan-kesalahan kecil terjadi untuk sebagian besar posisi bunyi, terutama pada daerah tengah-tengah pusat kepala, yaitu daerah di sekitar bidang median (bidang-bidang dengan azimuth -20° sampai 20° , pada Gambar 3.9 terlihat dari nilai-nilai MSE mendatar dengan azimuth -20° sampai 20°) dan bidang-bidang di atas kepala dengan elevasi $67,5^\circ - 135^\circ$ (nilai-nilai MSE vertikal dengan elevasi $67,5^\circ - 135^\circ$). MSE rata-rata telinga kiri dari posisi-posisi bunyi di bidang median adalah 4,23% dan di bidang di atas kepala dengan elevasi 90° adalah 2,03%. Untuk posisi-posisi bunyi pada bidang horisontal (elevasi 0°) dihasilkan kesalahan-kesalahan dengan MSE rata-rata telinga kiri 5,15%.

Daerah lain yang menghasilkan kesalahan-kesalahan pemodelan kecil adalah daerah *ipsilateral* dari sumber bunyi (sumber bunyi terletak satu sisi dengan telinga yang mendengar). Sisi *ipsilateral* dari telinga kiri adalah daerah-daerah dengan azimuth $-90^\circ < \theta < 0^\circ$. MSE rata-rata telinga kiri dari posisi-posisi bunyi paling dekat telinga kiri pada bidang dengan azimuth -80° adalah 2,65% (berasal dari nilai-nilai MSE mendatar paling atas dengan azimuth -80°).

Pemodelan HRIR_{mp} telinga kiri yang kurang baik terjadi pada daerah bawah subyek yaitu bidang-bidang pada elevasi -45° (bawah depan kepala) dan $230,6^\circ$ (bawah belakang kepala) dengan MSE rata-rata masing-masing 13,18% dan 12,62%. Pemodelan terburuk terjadi untuk bidang *contralateral* (sumber bunyi terletak pada sisi berlawanan dari telinga yang mendengar) dengan azimuth 65° untuk telinga kiri; dihasilkan MSE rata-rata 15,33 %. MSE rata-rata di sepanjang bidang *contralateral* dengan azimuth 80° adalah 6,86%.

Secara keseluruhan untuk seluruh posisi sumber bunyi, dihasilkan MSE rata-rata untuk telinga kiri dan kanan dari subyek 003 sebesar 5,07% dan 5,39%.

3.5 Hasil-hasil Pemodelan PCA dari Berbagai Jenis Data HRTF

Hasil-hasil pemodelan PCA di sini menggunakan berbagai jenis data HRTF seperti dijelaskan dalam subbab 1.6. Secara umum, pemodelan HRTF menghasilkan MSE rata-rata di antara semua posisi bunyi dan semua subyek, yang jauh lebih baik, sekitar 5% seperti terlihat dari Tabel 3.3, dibandingkan MSE rata-rata pada pemodelan HRIR dalam subbab 3.4, seperti terlihat dari Tabel 3.1. Jenis-jenis data HRTF yang dimodelkan dapat dibagi menjadi lima kumpulan, seperti terlihat pada Tabel 3.3.

Kumpulan pertama adalah sekumpulan HRTF *magnitude* hasil DFT 256-titik dari HRIR asli (jenis HRTF nomor 1 – nomor 5). Kumpulan kedua adalah sekumpulan HRTF *magnitude* hasil DFT 256-titik dari HRIR pulsa langsung (jenis HRTF nomor 6 – nomor 10). Kumpulan ketiga sampai kelima masing-masing adalah sekumpulan HRTF hasil logaritma natural dari kumpulan pertama (no.11 – no. 15), sekumpulan HRTF hasil logaritma-10 dari kumpulan pertama (no. 16 – no. 20), dan sekumpulan HRTF hasil kelipatan 20 dari kumpulan keempat (no. 21 – no.25).

Dari pemodelan PCA semua jenis data HRTF, diperoleh bahwa kumpulan HRTF yang berasal dari kumpulan HRIR yang berbeda menghasilkan MSE rata-rata yang berbeda (kumpulan satu dan kumpulan kedua). Sebaliknya, kumpulan HRTF yang merupakan kelipatan dari kumpulan HRTF yang lain (kumpulan ketiga sampai kumpulan kelima), memberikan MSE rata-rata yang sama. Hasil-hasil ini dapat dimengerti jika dilihat definisi MSE yang digunakan pada

persamaan (3.28) yaitu perbandingan antara MSE dari selisih HRTF basis data dan model HRTF terhadap MSE HRTF basis data. Hasil-hasil buruk pemodelan HRTF dengan PCA ditemukan untuk pengolahan awal normalisasi DTF dengan deviasi standar HRTF, DTF_{nz} , pada masing-masing jenis HRTF (no. 5, 10, 15, 20 dan 25 pada Tabel 3.3).

Tabel 3.3. Persentase MSE rata-rata dari Berbagai Jenis HRTF

No.	Jenis HRTF	Pengolahan Awal	MSE rata-rata(%)
1.	HRTF	<i>magnitude</i>	3,46
2.	HRTF _{nc}	<i>magnitude</i>	3,41
3.	DTF	<i>magnitude</i>	3,30
4.	DTF _{nc}	<i>magnitude</i>	3,40
5.	DTF _{nz}	<i>magnitude</i>	11,88
6.	HRTF	<i>direct-pulse magnitude</i>	4,58
7.	HRTF _{nc}	<i>direct-pulse magnitude</i>	3,84
8.	DTF	<i>direct-pulse magnitude</i>	3,53
9.	DTF _{nc}	<i>direct-pulse magnitude</i>	3,79
10.	DTF _{nz}	<i>direct-pulse magnitude</i>	13,68
11.	HRTF	<i>log-magnitude</i>	5,26
12.	HRTF _{nc}	<i>log-magnitude</i>	5,26
13.	DTF	<i>log-magnitude</i>	5,17
14.	DTF _{nc}	<i>log-magnitude</i>	5,22
15.	DTF _{nz}	<i>log-magnitude</i>	18,90
16.	HRTF	<i>log10-magnitude</i>	5,26
17.	HRTF _{nc}	<i>log10-magnitude</i>	5,26
18.	DTF	<i>log10-magnitude</i>	5,17
19.	DTF _{nc}	<i>log10-magnitude</i>	5,22
20.	DTF _{nz}	<i>log10-magnitude</i>	18,90
21.	HRTF	dB	5,26
22.	HRTF _{nc}	dB	5,26
23.	DTF	dB	5,17
24.	DTF _{nc}	dB	5,22
25.	DTF _{nz}	dB	18,90

Hasil pemodelan PCA yang terbaik diperoleh dari jenis data DTF *magnitude* (no.3). Dalam subbab-subbab selanjutnya, dianalisis vektor-vektor basis PCA yang diperoleh dari model DTF *magnitude* dan dianalisis kesalahan-kesalahan pemodelan jenis data HRTF ini dari seorang subyek yang mencakup seluruh posisi bunyi. Dalam Bab 4, dibahas individualisasi model ini menggunakan MLR, kemudian dianalisis MSE rata-rata yang diperoleh setelah

dilakukan individualisasi. Demikian juga dijelaskan cara rekonstruksi model tersebut menjadi model HRIR di kawasan waktu agar dapat digunakan dalam menghasilkan efek bunyi spasial, yaitu dengan cara konvolusi antara model HRIR tadi dengan stimuli bunyi dari berbagai posisi di ruang spasial.

3.5.1 Analisis Vektor-vektor Basis PCA dari Model HRTF *Magnitude*

Gambar 3.10 menunjukkan 10 vektor basis pertama yang digunakan untuk memodelkan HRTF *magnitude* menggunakan PCA. Sepuluh vektor basis ini merepresentasikan secara bersama-sama 94,30% varians kumulatif dari data DTF *magnitude*. Pada umumnya, 10 vektor basis dapat dikatakan konstan mendekati nol pada frekuensi-frekuensi di bawah sekitar 2-3 kHz karena tidak ada ketergantungan antara posisi bunyi dengan variasi-variasi dari HRTF *magnitude* dalam daerah frekuensi ini. Frekuensi-frekuensi rendah dari gelombang bunyi memancar dari sumber bunyi ke segala arah di ruang spasial (bersifat omnidireksional). Penjumlahan kombinasi linier dari kesepuluh vektor basis di daerah frekuensi rendah dari semua posisi sumber bunyi mendekati nol.

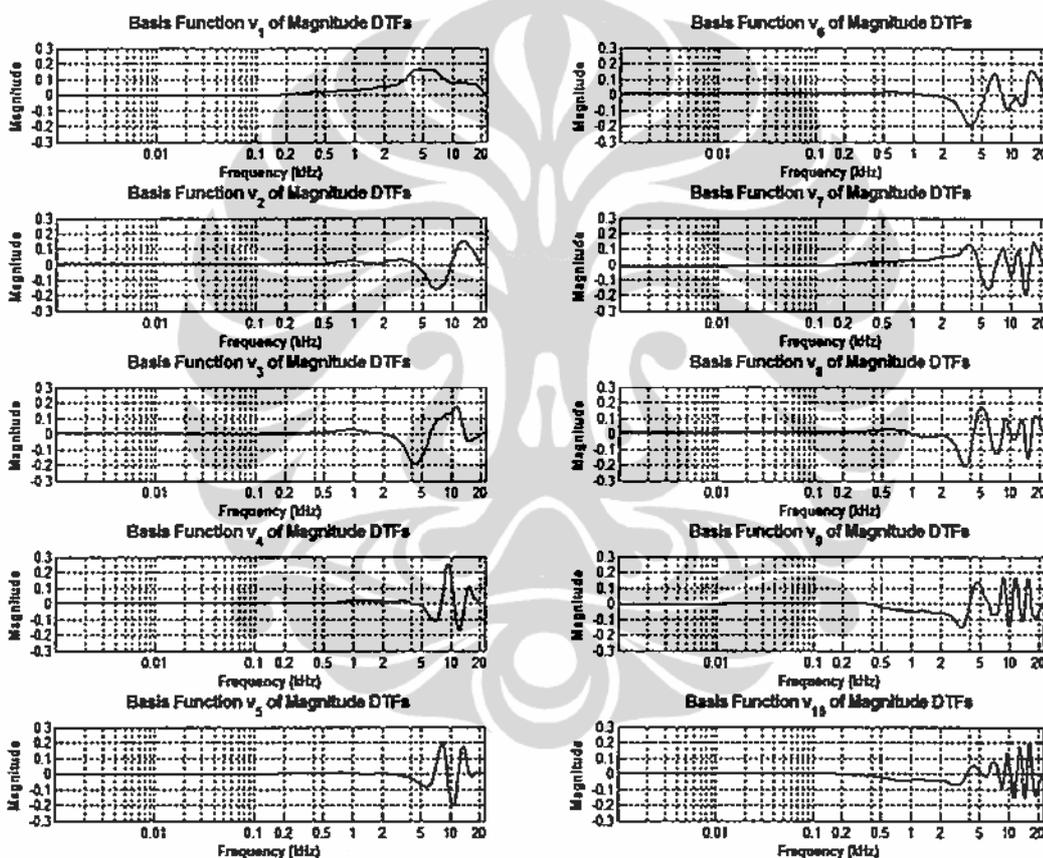
Tabel 3.4. Persentase Varians dari Vektor Basis Model DTF *Magnitude*

Vektor Basis	% Varians	% Varians Kumulatif
v_1	63,37	63,37
v_2	9,24	72,62
v_3	7,62	80,24
v_4	3,54	83,77
v_5	3,17	86,94
v_6	2,91	89,85
v_7	1,56	91,40
v_8	1,20	92,60
v_9	0,93	93,53
v_{10}	0,77	94,30

Di atas kira-kira 3 kHz, semua vektor basis memiliki *magnitude* tidak nol. Variasi-variasi vektor basis pada frekuensi-frekuensi lebih tinggi (kecuali dalam vektor basis pertama) merepresentasikan puncak-puncak dan lembah-lembah frekuensi tinggi yang banyak tergantung dari posisi sumber bunyi. Seperti terjadi pada model HRIR fasa minimum, vektor basis pertama dalam model HRTF *magnitude* juga memiliki bentuk spektrum *magnitude* yang paling sederhana. Vektor basis pertama dapat dikatakan sebagai faktor penguatan dalam daerah kira-

kira 3 kHz sampai dengan 22,05 kHz dan mencakup 63,37% varians (terlihat dari Tabel 3.4) dari data DTF *magnitude*.

Spektrum *magnitude* yang lebih detail ditemui pada vektor basis yang berhubungan dengan nilai eigen yang lebih kecil. Terlihat dari Gambar 3.10, vektor basis kesepuluh memiliki spektrum *magnitude* paling detail dibandingkan vektor-vektor basis pertama sampai vektor basis ke-9. Pada umumnya, spektrum-spektrum *magnitude* dari seluruh vektor basis memiliki level-level dalam interval yang hampir sama. Setiap vektor basis ini mencerminkan sebagian dari struktur-struktur detil dari data DTF *magnitude*.

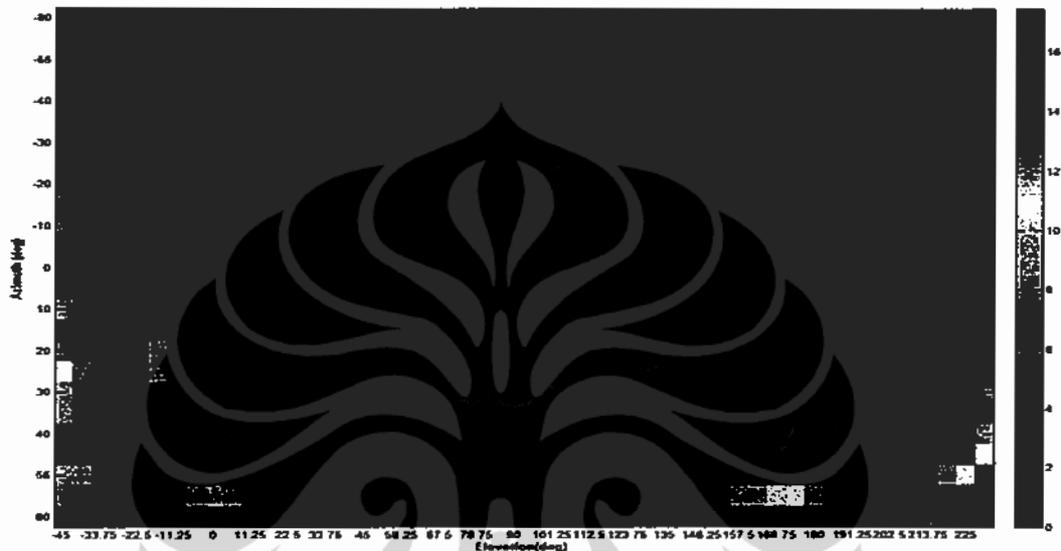


Gambar 3.10. Sepuluh Vektor Basis Pertama dari Model DTF *Magnitude*

3.5.2 Analisis Kesalahan Model HRTF *Magnitude* Menggunakan PCA

Pada subbab ini dibahas mengenai kinerja dari pemodelan PCA terhadap model-model HRTF *magnitude* dari telinga kiri subyek 003 pada semua posisi sumber bunyi di ruang spasial. Dengan menggunakan PCA, diperoleh model-model HRTF *magnitude* dari kombinasi linier 10 vektor basis beserta bobot-bobot mereka. Dalam subbab ini dibahas mengenai sampai seberapa jauh kemiripan atau

keakuratan antara model *HRTF magnitude* dan *HRTF magnitude* asli. Di sini akurasi dari model *HRTF magnitude* dihitung menggunakan suatu ukuran jarak antara dua vektor yang disebut dengan MSE dari persamaan (3.28). Jika model *HRTF magnitude* yang diperoleh identik dengan *HRTF magnitude* asli, maka dihasilkan $MSE = 0$. Semakin tidak akurat model *HRTF magnitude*, maka semakin besar nilai MSE.



Gambar 3.11. Persentase MSE antara Model *HRTF Magnitude* dan *HRTF Magnitude* Asli

Dari hasil pengamatan diperoleh bahwa secara umum terjadi kesalahan-kesalahan yang mirip pada kedua telinga pada daerah-daerah istimewa, seperti akan dijelaskan berikut ini. Jadi di sini hanya dibahas kesalahan-kesalahan yang terjadi untuk telinga kiri. Seperti terlihat dari Gambar 3.11 di atas, kinerja yang sangat baik ditunjukkan oleh model-model *HRTF magnitude* telinga kiri subyek 003 di hampir seluruh posisi sumber bunyi dengan MSE sekitar 5% (nilai MSE ditunjukkan oleh *colorbar* pada Gambar 3.11). Daerah posisi-posisi sumber bunyi dengan nilai-nilai MSE kecil adalah daerah tengah di sekitar pusat kepala, yaitu daerah di sekitar bidang median (bidang dengan azimuth 0°). MSE rata-rata dari posisi-posisi sepanjang bidang median adalah 1,95%, sedangkan sepanjang bidang horisontal (bidang dengan elevasi 0°) adalah 2,39%.

Dihitung dari nilai-nilai MSE pada Gambar 3.11, didapatkan MSE rata-rata pada posisi-posisi di bidang vertikal dengan azimuth -80° (bidang pada sisi *ipsilateral* dan paling dekat dengan telinga kiri) adalah 1,47%, sedangkan MSE

rata-rata pada posisi-posisi di bidang vertikal yang terjauh dari telinga kiri (bidang vertikal dengan azimuth 80°) adalah 2,62%. MSE rata-rata terkecil diperoleh dari posisi-posisi bidang horisontal di atas kepala dengan elevasi 90° , yaitu sebesar 1,15%. Pada bidang mendatar depan bawah dari pusat kepala dengan elevasi -45° dan belakang bawah dari pusat kepala dengan elevasi $230,6^\circ$, diperoleh MSE rata-rata cukup besar yaitu masing-masing 6,30% dan 5,05%. Jadi, MSE rata-rata dari semua posisi sumber di ruang spasial untuk telinga kiri subyek 003 adalah 2,44% sedangkan untuk telinga kanan adalah 2,53%.



BAB 4

METODE SELEKSI UKURAN-UKURAN ANTROPOMETRIS

Model-model terbaik dari HRIR dan HRTF hasil pemodelan PCA, seperti dijelaskan pada Bab 3, diindividualisasi dengan menggunakan MLR, dijelaskan pada Bab 5, yaitu dengan cara mengestimasi bobot-bobot vektor basis PCA (PCWs) berdasarkan 8 ukuran antropometris yang diseleksi dalam bab ini. Seleksi ukuran-ukuran antropometris dilakukan berdasarkan analisis korelasi antara ITD maksimum dan 27 ukuran antropometris, analisis korelasi antara ILD maksimum dan 27 ukuran antropometris, analisis korelasi antara frekuensi lembah pinna, f_{pn} , dan 10 ukuran pinna, serta analisis korelasi antara PCWs dan 17 ukuran kepala dan *torso*. Dalam penelitian ini, metode seleksi ukuran-ukuran antropometris di atas dan 8 ukuran yang terseleksi berbeda dengan metode seleksi dan 8 ukuran antropometris yang diseleksi oleh Hu et al. [38,39]. Mereka menyeleksi ukuran-ukuran antropometris hanya berdasarkan analisis-analisis korelasi antara ukuran-ukuran antropometris dan sejumlah DTF.

CIPIC HRTF *Database* yang digunakan dalam penelitian ini, terdiri dari sejumlah besar HRIR hasil pengukuran dan ukuran-ukuran antropometris dari 43 orang dan 2 manikin KEMAR dengan pinna kecil dan pinna besar. Tabel 2.1 berisi nama-nama 27 ukuran antropometris yang didefinisikan oleh CIPIC *Interface Laboratory*, beserta data statistik setiap ukuran. Gambar 2.11 menunjukkan diagram-diagram skematik dari ukuran-ukuran kepala, *torso* dan pinna. Dalam penelitian ini, telah diukur 8 ukuran antropometris, yang diseleksi seperti dijelaskan nanti pada subbab di bawah, untuk keperluan rekonstruksi model-model HRIR individual dari dua orang subyek [41]. Pada mulanya, diambil foto-foto dari kedua orang subyek, dengan sebuah penggaris yang diletakkan dekat ukuran-ukuran yang diinginkan, sebagai suatu acuan. Foto tampak depan dan tampak sisi kiri, seperti pada Gambar 2.11a, diambil untuk mengukur x_1 , x_3 , x_6 , x_{12} dan ukuran-ukuran pinna kiri. Foto tampak sisi kanan diambil untuk mengukur ukuran-ukuran pinna kanan. Jumlah *pixel* dari 1 cm digunakan sebagai

acuan panjang. Kemudian, setiap ukuran diukur dari titik-titik acuan, seperti ditunjukkan Gambar 2.11. Panjang suatu ukuran antropometris dinyatakan dalam jumlah *pixel*. Perbandingan antara jumlah *pixel* suatu ukuran dan jumlah *pixel* dari 1 cm menentukan panjang ukuran antropometris tersebut. Prosedur pengukuran yang sama dilakukan oleh Zotkin et al. [40].

Model HRIR individual untuk seorang pendengar melalui ukuran-ukuran antropometrisnya berpengaruh langsung terhadap keandalan dan kompleksitas sistem pemodelan dan individualisasi. Tentu saja tidak disarankan untuk menggunakan semua 27 ukuran antropometris di dalam individualisasi model HRIR tersebut. Beberapa informasi penting akan hilang yang disebabkan oleh ukuran-ukuran yang tidak diperlukan, yang menyebabkan suatu model regresi yang lebih buruk. Di samping itu, banyak ukuran yang sulit untuk diukur secara sederhana dengan benar, misalnya x_4 , x_5 , x_{13} , θ_1 , θ_2 , dan d_8 . Alasannya adalah karena sulit untuk menentukan titik pusat kepala yang benar sebagai acuan ukuran-ukuran lain dan pengukuran sudut-sudut pinna yang sulit.

Terdapat tiga parameter yang secara psikoakustik krusial dalam persepsi bunyi alami, yaitu ITD, ILD, dan f_{pn} . ITD adalah selisih waktu antara kedatangan pulsa pertama dari sumber bunyi pada posisi tertentu, pada gendang telinga kiri dan pada gendang telinga kanan. Pada posisi-posisi bunyi di bidang median, nilai-nilai ITD mendekati nol. Jika kepala simetris sempurna, maka tidak terdapat ITD pada bidang median. Karena itu ITD merupakan suatu fungsi azimuth pada bidang-bidang dengan elevasi tetap. ITD dapat dihitung dari waktu tunda dari korelasi silang maksimum antara HRIR telinga kiri dan HRIR telinga kanan untuk posisi bunyi tertentu.

ILD adalah selisih level atau *magnitude* dalam dB, di kawasan frekuensi, antara sepasang HRTF *magnitude* dari telinga kiri dan telinga kanan, untuk posisi bunyi tertentu. ILD dapat dihitung dari setiap komponen frekuensi dalam jangkauan 0 – 22.050 Hz dari pasangan HRTF *magnitude*. Nilai-nilai ILD biasanya dianalisis untuk komponen frekuensi tertentu untuk sumber bunyi di bidang horisontal dan di bidang median. Parameter psikoakustik penting lainnya adalah frekuensi lembah pinna, yang terjadi karena adanya difraksi dan refleksi gelombang bunyi pada pinna. Frekuensi lembah pinna adalah suatu lembah

minimum pada frekuensi tertentu dalam spektrum *magnitude* HRTF. Frekuensi-frekuensi lembah pinna bervariasi sesuai dengan posisi-posisi bunyi.

ITD dan ILD penting dalam persepsi azimuth dari sumber bunyi. Banyak variasi HRIR di bidang horisontal dipengaruhi kedua parameter ini. Tetapi ILD dan f_{pn} adalah parameter-parameter penting dalam persepsi elevasi dari sumber bunyi dan mempengaruhi variasi HRIR di bidang median. Tetapi, pada umumnya, penentuan karakteristik dari berbagai variasi HRIR di antara subyek-subyek, masih bersifat problematik.

Untuk menyeleksi beberapa ukuran antropometris yang paling berpengaruh secara psikoakustik dalam persepsi bunyi, maka analisis-analisis korelasi dilakukan antara ITD_{max} , ILD_{max} , f_{pn} , dan bobot-bobot vektor basis, dari seluruh subyek, terhadap ukuran-ukuran antropometris.

4.1 Analisis Korelasi ITD Maksimum Terhadap Antropometri

Nilai-nilai ITD hampir nol untuk sumber-sumber bunyi di bidang median. Untuk sumber-sumber bunyi yang bergeser menjauhi bidang median, nilai-nilai ITD yang didapatkan menjadi semakin besar. ITD mencapai nilai maksimum ketika sumber bunyi terletak tepat di depan salah satu telinga di bidang horisontal (azimuth -90° untuk telinga kiri dan azimuth 90° untuk telinga kanan). Di dalam CIPIC HRTF *Database* tidak terdapat HRIR yang diukur untuk posisi-posisi bunyi tersebut. HRIR hanya diukur dari lokasi-lokasi bunyi ekstrim dengan azimuth -80° di sisi kiri dan azimuth 80° di sisi kanan. Jadi, dalam hal ini nilai-nilai maksimum ITD terjadi untuk posisi-posisi di dua bidang vertikal dengan kedua azimuth ekstrim tetap dan semua elevasi, seperti yang ditunjukkan Gambar 2.6. Untuk seorang subyek, nilai-nilai ITD_{max} di setiap bidang vertikal dirata-rata, sehingga didapatkan dua nilai rata-rata ITD_{max} dari kedua bidang vertikal tadi. Nilai yang terbesar dari kedua nilai rata-rata ITD_{max} diambil sebagai nilai ITD_{max} yang mewakili subyek tersebut.

Prosedur yang mirip dilakukan untuk menentukan nilai ITD_{max} dari semua subyek. Jadi di sini didapatkan 45 nilai ITD_{max} untuk 45 subyek. Nilai-nilai ini dikumpulkan dalam suatu vektor. Vektor lain diperoleh dari 45 nilai ukuran antropometris tertentu dari 45 subyek. Regresi-regresi linier sederhana dilakukan

antara vektor ITD_{max} dan vektor setiap ukuran antropometris, sehingga dihasilkan korelasi-korelasi antara ITD_{max} dan 27 ukuran antropometris.

Tabel 4.1. Koefisien Korelasi antara ITD_{max} dan Ukuran-ukuran Antropometris

Anthropometric Measurements	Correlation Coefficient, ρ	Anthropometric Measurements	Correlation Coefficient, ρ
x1 head width	0.736	x9 and x11 and x12	0.777
x2 head height	0.199	x10 and x11 and x12	0.805
x3 head depth	0.709	x13 head offset forward	0.222
x1 and x2	0.737	x14 height	0.624
x1 and x3	0.853	x15 seated height	0.518
x2 and x3	0.709	x14 and x15	0.626
x1 and x2 and x3	0.855	x16 head circumference	0.626
x4 pinna off.down	0.161	x17 shoulder circumfer.	0.600
x5 pinna off.back	0.098	x16 and x17	0.693
x4 and x5	0.180	Cavum c. height d1L	0.373
x6 neck width	0.726	Cymba c. height d2L	0.173
x7 neck height	0.061	Cavum c. width d3L	0.213
x8 neck depth	0.665	Fossa height d4L	0.494
x6 and x7	0.731	Pinna height d5L	0.607
x6 and x8	0.755	Pinna width d6L	0.242
x7 and x8	0.669	Inter.incisure width d7L	0.191
x6 and x7 and x8	0.760	Cavum c.depth d8L	0.446
x9 torso top width	0.641	Pinna rot.angle θ 1L	0.243
x10 torso top height	0.605	Pinna flare angle θ 2L	0.284
x11 torso top depth	0.440	Cavum c.height d1R	0.363
x12 shoulder width	0.768	Cymba c.height d2R	0.283
x9 and x10	0.725	Cavum c.width d3R	0.173
x9 and x11	0.677	Fossa height d4R	0.463
x9 and x12	0.770	Pinna height d5R	0.608
x10 and x11	0.628	Pinna width d6R	0.228
x10 and x12	0.804	Inter. incisure width d7R	0.203
x11 and x12	0.775	Cavum c.depth d8R	0.397
x9 and x10 and x11	0.732	Pinna rot.angle θ 1R	0.090
x9 and x10 and x12	0.807	Pinna flare angle θ 2R	0.141

Korelasi-korelasi yang terjadi di antara ITD_{max} dan 27 ukuran antropometris dapat dilihat pada Tabel 4.1 di atas. Nilai-nilai pada tabel tersebut adalah koefisien-koefisien korelasi, ρ , yang bersangkutan. Dari seluruh subyek di dalam basis data, ITD_{max} terdistribusi hampir mendekati normal, dengan nilai rata-rata 646 μ s dan deviasi standar 33 μ s, yang berkaitan dengan $\pm 10,3$ % variasi [1]. ITD_{max} tentu saja berkorelasi kuat dengan ukuran-ukuran kepala dan *torso* dan dapat diestimasi cukup akurat dengan regresi linier sederhana dengan algoritma

least square. Prediktor tunggal yang paling baik dari ITD_{max} adalah lebar kepala x_1 dengan koefisien korelasi $\rho = 0,736$. Pasangan prediktor yang terbaik adalah lebar kepala x_1 dan tebal kepala x_3 dengan $\rho = 0,853$.

Dari Tabel 4.1, diseleksi 4 ukuran kepala dan *torso*, yang memiliki korelasi-korelasi yang paling kuat terhadap ITD_{max} ; yaitu lebar kepala x_1 dengan $\rho = 0,736$, tebal kepala x_3 dengan $\rho = 0,709$, lebar leher x_6 dengan $\rho = 0,726$, dan lebar pundak x_{12} dengan $\rho = 0,768$. Ukuran-ukuran ini diperlukan untuk individualisasi model-model HRIR dan HRTF dengan MLR.

4.2 Analisis Korelasi ILD Maksimum Terhadap Antropometri

Seperti terlihat dari Gambar 2.7a dan Gambar 2.7b, bahwa ILD dari komponen-komponen frekuensi yang lebih tinggi bervariasi lebih besar daripada variasi-variasi ILD dari frekuensi-frekuensi lebih rendah, baik di bidang horisontal (Gambar 2.7a) maupun di bidang median (Gambar 2.7b). Di sini dipilih frekuensi 10 kHz untuk menghitung ILD karena mulai dari 10 kHz terjadi puncak-puncak dan lembah-lembah *magnitude* HRTF yang signifikan. Untuk seorang subyek, nilai-nilai ILD di frekuensi 10 kHz di bidang horisontal dihitung dari respon-respon HRIR yang bersangkutan. Dari nilai-nilai ILD tersebut dicari nilai maksimum ILD yang digunakan sebagai nilai ILD_{max} yang mewakili subyek tadi. Perhitungan-perhitungan selanjutnya dilakukan untuk mencari 45 nilai ILD_{max} dari 45 subyek. Kemudian regresi linier sederhana dilakukan antara vektor ILD_{max} , terdiri dari 45 nilai ILD_{max} tadi, dan vektor yang berisi 45 ukuran antropometris tertentu dari 45 subyek.

Tabel 4.2 berisi koefisien-koefisien korelasi antara ILD_{max} dan setiap ukuran antropometris. Seperti terlihat dari Tabel 4.2, ILD_{max} tidak berkorelasi dengan ukuran-ukuran antropometris sekuat seperti ITD_{max} . Prediktor tunggal terbaik adalah tinggi tubuh x_{14} dengan $\rho = 0,619$ diikuti dengan tinggi tubuh ketika duduk x_{15} dengan $\rho = 0,589$. Sedangkan pasangan prediktor terbaik adalah x_{14} dan x_{15} , dengan $\rho = 0,637$. Terlihat dari Tabel 4.2, bahwa hasil seleksi 4 ukuran kepala dan *torso* berdasarkan ITD_{max} dikonfirmasi dengan korelasi-korelasi yang lebih kuat antara ILD_{max} dan 4 ukuran tersebut, daripada korelasi-korelasi ILD_{max} dan ukuran-ukuran kepala dan *torso* lainnya. Seperti terlihat dari

Tabel 4.1 dan Tabel 4.2, ILD_{max} dan ITD_{max} memiliki korelasi yang kuat dengan ukuran-ukuran antropometris horisontal seperti lebar dan keliling.

Tabel 4.2. Koefisien Korelasi antara ILD_{max} dan Ukuran-ukuran Antropometris

Anthropometric Measurements	Correlation Coefficient ρ	Anthropometric Measurements	Correlation Coefficient ρ
x1 head width	0.270	x9 and x11 and x12	0.517
x2 head height	0.063	x10 and x11 and x12	0.480
x3 head depth	0.459	x13 head offset forward	0.265
x1 and x2	0.271	x14 height	0.619
x1 and x3	0.466	x15 seated height	0.589
x2 and x3	0.460	x14 and x15	0.637
x1 and x2 and x3	0.468	x16 head circumference	0.436
x4 pinna off. down	0.063	x17 shoulder circumfer.	0.385
x5 pinna off. back	0.005	x16 and x17	0.467
x4 and x5	0.063	Cavum c. height d1L	0.281
x6 neck width	0.454	Cymba c. height d2L	0.172
x7 neck height	0.405	Cavum c. width d3L	0.162
x8 neck depth	0.266	Fossa height d4L	0.199
x6 and x7	0.598	Pinna height d5L	0.240
x6 and x8	0.461	Pinna width d6L	0.200
x7 and x8	0.481	Inter. incisure width d7L	0.047
x6 and x7 and x8	0.603	Cavum c. depth d8L	0.121
x9 torso top width	0.491	Pinna rot. angle θ 1L	0.074
x10 torso top height	0.255	Pinna flare angle θ 2L	0.141
x11 torso top depth	0.331	Cavum c. height d1R	0.320
x12 shoulder width	0.458	Cymba c. height d2R	0.203
x9 and x10	0.492	Cavum c. width d3R	0.117
x9 and x11	0.517	Fossa height d4R	0.304
x9 and x12	0.495	Pinna height d5R	0.378
x10 and x11	0.348	Pinna width d6R	0.177
x10 and x12	0.459	Inter. incisure width d7R	0.061
x11 and x12	0.479	Cavum c. depth d8R	0.178
x9 and x10 and x11	0.519	Pinna rotation angle θ 1R	0.056
x9 and x10 and x12	0.459	Pinna flare angle θ 2R	0.226

4.3 Analisis Korelasi Frekuensi Lembah Telinga Terhadap Antropometri

Pengaruh dari pinna terhadap persepsi bunyi lebih besar di bidang median daripada di bidang horisontal [48,50]. Tetapi secara keseluruhan ukuran-ukuran pinna mempengaruhi persepsi bunyi dari segala arah. Untuk menyeleksi beberapa ukuran pinna yang sederhana dan memegang peranan penting dalam persepsi bunyi, maka analisis-analisis korelasi dilakukan antara f_{pn} dan 10 ukuran pinna. Frekuensi-frekuensi f_{pn} dari telinga kiri dan telinga kanan diperoleh masing-

masing dari HRTF *magnitude* telinga kiri dan HRTF *magnitude* telinga kanan, untuk posisi bunyi tepat di depan kepala ($0^{\circ}, 0^{\circ}$). Regresi linier sederhana dilakukan antara f_{pn} dan 10 ukuran pinna yang bersangkutan.

Secara statistik, f_{pn} terdistribusi hampir normal, dengan $\mu = 7600$ Hz dan $\sigma = 1050$ Hz, yang berkaitan dengan variasi yang agak besar $\pm 28\%$ [1]. Seperti yang diharapkan, f_{pn} berkorelasi dengan ukuran-ukuran daun telinga, tetapi hubungannya tidak kuat dan regresi linier sederhana kurang berhasil dalam mengestimasi f_{pn} dari ukuran-ukuran antropometris. Dari Tabel 4.1 dan Tabel 4.3 terlihat, bahwa analisis-*analisis* korelasi antara f_{pn} dan 10 ukuran pinna memberikan korelasi-korelasi yang lebih lemah daripada korelasi-korelasi dari ITD_{max} yang bersangkutan. Prediktor tunggal terbaik dari f_{pn} adalah tinggi cavum concha d_{1L} dengan $\rho = 0,435$, seperti tampak pada Tabel 4.3. Secara mengejutkan pasangan prediktor terbaik adalah kedua sudut, θ_1 dan θ_2 , dengan $\rho = 0,42$. Tiga prediktor terbaik adalah kedua sudut dan tinggi fossa d_4 dengan $\rho = 0,51$ [1].

Tabel 4.3. Koefisien Korelasi antara f_{pn} dan Ukuran-ukuran Antropometris

Anthropometric Measurements	Correlation Coefficient ρ	Anthropometric Measurements	Correlation Coefficient ρ
Cavum c. height d_{1L}	0.435	Cavum c. height d_{1R}	0.029
Cymba c. height d_{2L}	0.150	Cymba c. height d_{2R}	0.166
Cavum c. width d_{3L}	0.003	Cavum c. width d_{3R}	0.360
Fossa height d_{4L}	0.108	Fossa height d_{4R}	0.070
Pinna height d_{5L}	0.204	Pinna height d_{5R}	0.053
Pinna width d_{6L}	0.031	Pinna width d_{6R}	0.280
Inter.incisure width d_{7L}	0.158	Inter. incisure width d_{7R}	0.197
Cavum c. depth d_{8L}	0.060	Cavum c. depth d_{8R}	0.031
Pinna rot. angle θ_{1L}	0.233	Pinna rot. angle θ_{1R}	0.112
Pinna flare angle θ_{2L}	0.206	Pinna flare angle θ_{2R}	0.044

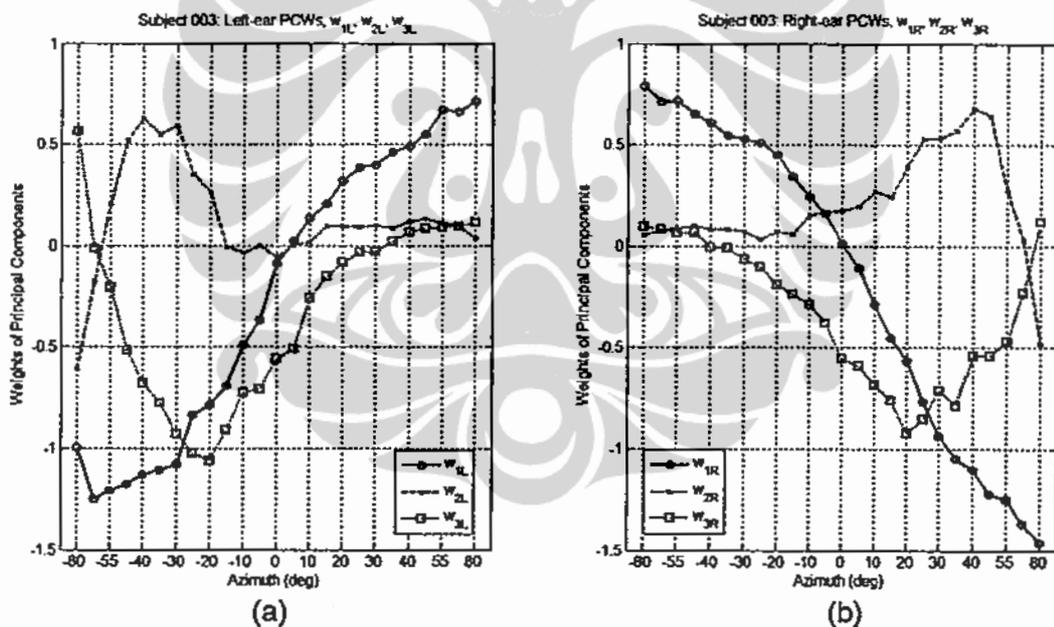
Di sini diseleksi 4 ukuran daun telinga, berdasarkan koefisien-koefisien korelasi terbesar antara f_{pn} dan ukuran-ukuran pinna pada Tabel 4.3. Keempat ukuran ini adalah tinggi *cavum concha* d_1 dengan $\rho = 0,435$, lebar *cavum concha* d_3 dengan $\rho = 0,360$, tinggi pinna d_5 dengan $\rho = 0,204$, and lebar pinna d_6 $\rho = 0,280$. Keempat ukuran pinna tersebut ternyata mudah diukur dan merepresentasikan faktor tinggi dan lebar. Dari Tabel 4.1, ITD_{max} berkorelasi lebih kuat dengan tinggi pinna d_5 daripada ukuran-ukuran pinna lainnya.

Demikian juga dari Tabel 4.2, ILD_{max} berkorelasi lebih kuat dengan tinggi *cavum concha* d_1 , tinggi pinna d_5 dan lebar pinna d_6 .

Hasil-hasil analisis-analisis korelasi di atas menunjukkan bahwa penghamburan gelombang-gelombang datang oleh daun telinga, kepala dan *torso*, adalah proses kompleks yang berkaitan dengan fitur-fitur detail. Namun analisis-analisis regresi sederhana antara ITD_{max} , ILD_{max} dan f_{pn} , terhadap semua ukuran antropometris dapat mengidentifikasi 8 ukuran antropometris yang penting.

4.4 Analisis Korelasi Bobot-bobot Vektor Basis Terhadap Antropometri

Gambar 4.1 menunjukkan bobot-bobot vektor basis (PCW) pertama sampai ketiga, w_1 , w_2 , dan w_3 dari model-model HRIR fasa minimum, sebagai fungsi dari azimuth, di bidang horisontal depan untuk subyek 003. Metode individualisasi model HRIR fasa minimum dijelaskan dalam subbab 5.2.



Gambar 4.1. Bobot-bobot Vektor Basis Model HRIR Fasa Minimum dari Subyek 003 di Bidang Horisontal Depan

Terlihat dari Gambar 4.1, bahwa perubahan nilai suatu PCW terhadap azimuth untuk telinga kiri merupakan pencerminan perubahan nilai PCW tersebut untuk telinga kanan. Dari ketiga jenis PCW tersebut, PCW pertama, w_1 , mengalami perubahan nilai yang paling besar sebagai fungsi azimuth dibandingkan kedua jenis bobot lainnya; diikuti berikutnya adalah perubahan nilai PCW kedua,

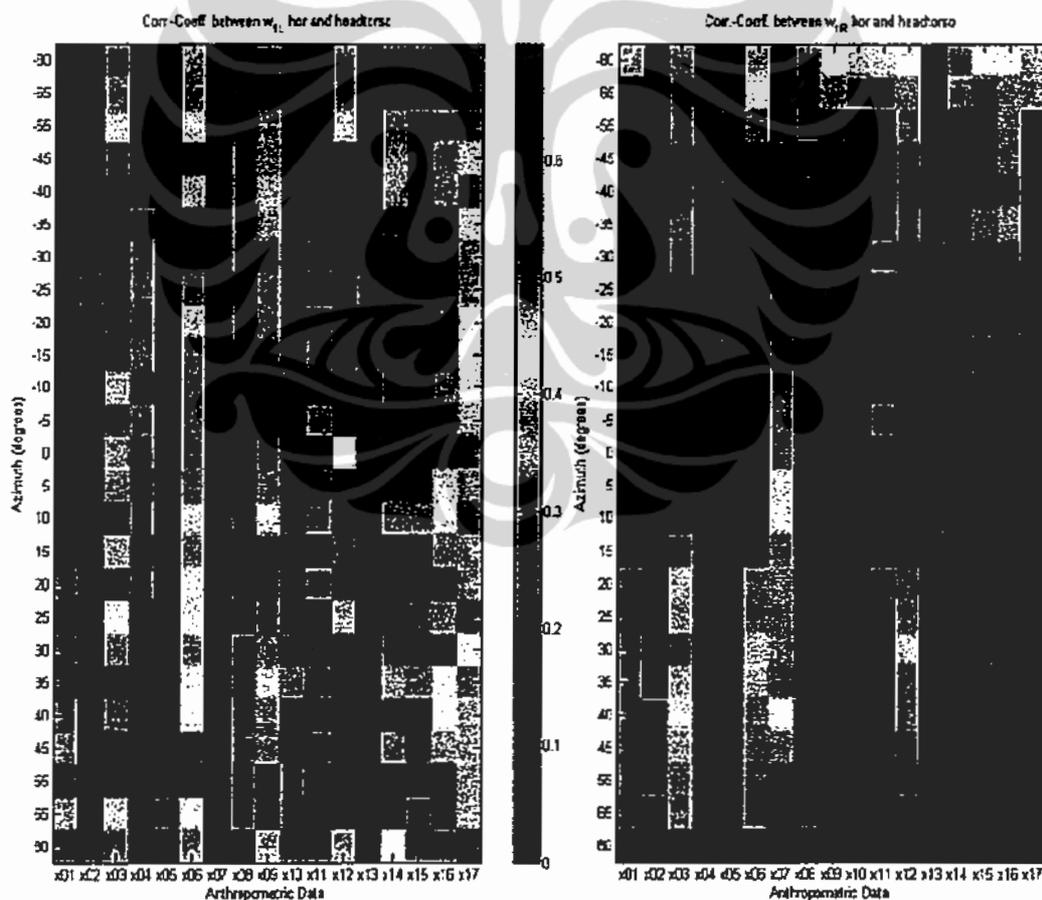
w_2 , dan perubahan terkecil dalam Gambar 4.1 adalah nilai-nilai PCW ketiga, w_3 . Perubahan nilai-nilai jenis PCW berikutnya akan semakin kecil untuk jenis PCW dengan orde lebih tinggi ($\Delta w_4 > \Delta w_5 > \dots > \Delta w_{10}$).

Diamati dari nilai-nilai PCW ke-2 sampai PCW ke-10, terlihat bahwa nilai-nilai dari 9 jenis PCW tersebut memiliki variabilitas yang lebih besar di sisi *ipsilateral* daripada di sisi *contralateral*. Dimulai dari azimuth 0° di sisi *contralateral*, nilai-nilai masing-masing PCW adalah konstan mendekati nol. Vektor basis dengan orde yang lebih tinggi memiliki perubahan nilai-nilai bobot dengan pola yang semakin rata untuk sumber-sumber bunyi di bidang horisontal depan. Juga ditemui bahwa pola-pola dari perubahan nilai-nilai dari 10 jenis PCW sebagai fungsi azimuth di bidang horisontal depan adalah sama di antara kedua telinga dan di antara semua subyek.

Analisis korelasi dilakukan antara nilai-nilai PCW pertama sampai PCW ke-10 ($w_1 - w_{10}$) dan 17 ukuran kepala dan *torso* di bidang horisontal depan [41]. Diamati secara umum bahwa di bidang horisontal depan, setiap ukuran antropometris tidak memiliki korelasi yang konsisten terhadap salah satu PCW. Terdapat tiga ukuran antropometris, yaitu tebal kepala x_3 , lebar leher x_6 dan lebar pundak x_{12} memiliki korelasi-korelasi kuat dengan $w_1(\theta)$ sampai $w_7(\theta)$ di sepanjang bidang horisontal depan. Vektor $w_i(\theta)$, $i=1,2,\dots,10$, berisi nilai-nilai PCW ke- i yang berkaitan dengan posisi bunyi pada azimuth θ di bidang horisontal depan, dari suatu telinga, yang dikumpulkan dari semua subyek. Regresi linier sederhana, antara $w_1(\theta)$ telinga kiri di bidang horisontal depan dan tebal kepala x_3 , lebar leher x_6 dan lebar pundak x_{12} menghasilkan koefisien korelasi rata-rata $\bar{\rho}$ berturut-turut sebesar 0,479; 0,428 dan 0,508. Nilai-nilai $\bar{\rho}$ ini dihitung dari nilai-nilai koefisien korelasi yang besar di bidang horisontal depan untuk ketiga ukuran kepala dan *torso*, seperti yang terlihat dari kolom ke-3, kolom ke-6, dan kolom ke-12 dari Gambar 4.2. Sumbu mendatar dari Gambar 4.2 adalah 17 ukuran kepala dan *torso*, sedangkan sumbu vertikal adalah sudut-sudut azimuth yang menentukan posisi bunyi di bidang horisontal depan. Besarnya nilai koefisien korelasi dinyatakan oleh warna-warna, yang ditunjukkan oleh *colorbar* di tengah-tengah Gambar 4.2.

Hubungan antara $w_2(\theta)$ telinga kiri di bidang horisontal depan dan tebal kepala x_3 , lebar leher x_6 dan lebar pundak x_{12} memberikan masing-masing $\bar{\rho}$ sebesar 0,422; 0,369 dan 0,409. Vektor $w_3(\theta)$ telinga kiri di bidang horisontal depan berkorelasi dengan masing-masing x_3 , x_6 dan x_{12} berturut-turut dengan $\bar{\rho}$ sebesar 0,300; 0,331 dan 0,351. Nilai-nilai $\bar{\rho}$ untuk telinga kanan adalah lebih kecil daripada nilai-nilai $\bar{\rho}$ yang bersangkutan untuk telinga kiri. Tetapi, nilai-nilai koefisien korelasi rata-rata, $\bar{\rho}$, untuk ketiga ukuran antropometris di atas dari masing-masing telinga adalah nilai-nilai yang terbesar daripada untuk ukuran-ukuran kepala dan *torso* lainnya.

Hasil-hasil di atas mengkonfirmasi seleksi tebal kepala x_3 , lebar leher x_6 dan lebar pundak x_{12} berdasarkan ITD_{max} , seperti dijelaskan dalam subbab 4.1, untuk digunakan sebagai sebagian prediktor dari PCWs dengan metode MLR.



Gambar 4.2. Korelasi antara Ukuran Kepala dan Torso Terhadap Nilai-Nilai Bobot $w_1(\theta)$ pada Bidang Horisontal Depan

Dari penjelasan subbab-subbab sebelumnya dalam bab ini, maka dihasilkan 8 ukuran antropometris yang diseleksi, yaitu x_1 , x_3 , x_6 , x_{12} , d_1 , d_3 , d_5 , dan d_6 untuk mengestimasi PCWs model-model DIR fasa minimum dan PCWs model-model DTF *magnitude*, untuk setiap posisi bunyi dalam bidang horisontal dan bidang median dengan menggunakan metode MLR.



BAB 5

INDIVIDUALISASI MODEL *HEAD-RELATED TRANSFER FUNCTIONS* DENGAN *MULTIPLE LINEAR REGRESSION*

Model-model terbaik dari HRIR dan HRTF hasil pemodelan PCA selanjutnya diindividualisasi dengan MLR, yaitu dengan cara mengestimasi bobot-bobot vektor basis PCA (PCWs) berdasarkan 8 ukuran antropometris yang diseleksi dalam Bab 4. Di dalam bab ini dibahas tentang model MLR dari PCWs dengan prediktor antropometri subyek, metode individualisasi model HRIR fasa minimum yang diusulkan, analisis hasil individualisasi model HRIR fasa minimum beserta diagram-diagram *pole-zero* yang relevan, metode individualisasi model HRTF *magnitude* yang diusulkan, dan analisis hasil individualisasi model HRTF *magnitude*.

Tujuan dari penelitian ini adalah untuk membangun suatu metode baru dalam individualisasi model-model HRTF dan model-model HRIR baik di bidang horisontal maupun di bidang median. Dari Bab 3 ditemukan jenis data HRIR yang terbaik untuk pemodelan menggunakan PCA adalah DIR fasa minimum. Hasil-hasil dari Bab 3 mengenai penentuan jenis data HRIR dan jenis data HRTF yang terbaik untuk pemodelan PCA telah dipublikasikan dalam jurnal internasional [40]. Penelitian ini dilakukan secara komprehensif yang meliputi individualisasi model HRIR fasa minimum baik di bidang horisontal [41] maupun di bidang median [42] dan individualisasi model HRTF *magnitude* baik di bidang horisontal dan di bidang median [43].

Metode individualisasi yang digunakan adalah MLR antara PCWs dari DIR fasa minimum dan DTF *magnitude*, terhadap 8 ukuran antropometris yang diseleksi dalam Bab 4. Beberapa ukuran antropometris hasil seleksi tidak diberikan dalam CIPIC HRTF *Database* untuk 8 subyek. Oleh karena itu pemodelan PCA dari HRIR fasa minimum dan HRTF *magnitude* di bidang horisontal dan di bidang median serta individualisasi model-model tersebut hanya menggunakan data HRIR dari 37 subyek dari total 45 subyek dalam basis data.

5.1 Model MLR dari Bobot-bobot Vektor Basis PCA dengan Prediktor Antropometri Subyek

Hubungan antara PCWs, $w_i(\theta, \phi)$, pada posisi sumber bunyi (θ, ϕ) dari persamaan (3.26) dan ukuran-ukuran antropometris subyek dapat dijelaskan oleh persamaan berikut:

$$w_i(\theta, \phi) = \mathbf{X} \beta_i(\theta, \phi) + e_i(\theta, \phi) \quad (5.1)$$

dimana $w_i(\theta, \phi)$ adalah vektor kolom ($n \times 1$) dari PCW ke- i , $i=1,2,\dots,10$, pada posisi sumber bunyi tertentu, (θ, ϕ) , untuk suatu telinga dari setiap subyek dari total n subyek. $\beta_i(\theta, \phi)$ adalah vektor kolom ($[p+1] \times 1$) koefisien-koefisien regresi yang digunakan untuk mengestimasi $w_i(\theta, \phi)$; \mathbf{X} adalah matriks ukuran-ukuran antropometris, dimana setiap baris ke- j dari \mathbf{X} , $\mathbf{x}_j^T (1 \times [p+1])$, adalah ukuran-ukuran antropometris yang digunakan dari suatu telinga seorang subyek dan $e_i(\theta, \phi) (n \times 1)$ adalah vektor kolom kesalahan estimasi MLR. Dalam penelitian ini, digunakan jumlah subyek $n=37$ dan jumlah ukuran antropometris $p=8$. Koefisien-koefisien regresi $\beta_i(\theta, \phi)$ dicari berdasarkan estimasi *least-square*. Hal ini diperoleh dengan memecahkan masalah optimasi seperti diberikan oleh persamaan (B.9). Kemudian estimasi-estimasi koefisien-koefisien regresi dapat dituliskan sebagai berikut:

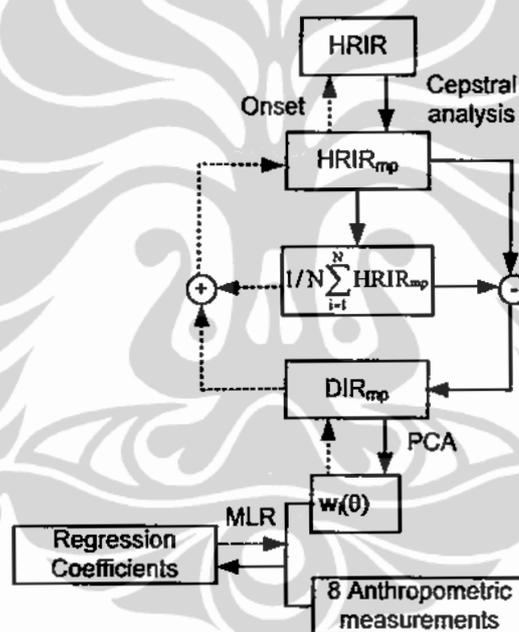
$$\hat{\beta}_i(\theta, \phi) = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T w_i(\theta, \phi) \quad (5.2)$$

Terlihat dari persamaan di atas bahwa untuk meningkatkan performansi dari model MLR, harus dipilih secara hati-hati baik variabel bebas, \mathbf{x}_j^T , maupun variabel tak bebas, $w_i(\theta, \phi)$. Dengan mengaplikasikan PCA pada HRIR ataupun HRTF, dimensi dari variabel-variabel tak bebas berkurang secara signifikan, demikian juga dengan kompleksitas model. Agar diperoleh suatu model yang lebih sederhana dan akurat, analisis korelasi digunakan untuk memilih variabel-variabel bebas berupa ukuran-ukuran antropometris subyek, seperti dijelaskan dalam Bab 4.

5.2 Metode Individualisasi Model HRIR Fasa Minimum yang Diusulkan

Salah satu kontribusi dari penelitian ini adalah membangun model-model MLR antara HRIR fasa minimum (HRIR_{mp}) dan 8 ukuran antropometris hasil

seleksi, sehingga seorang pendengar dapat mengindividualisasi model HRIR yang tersedia dengan ukuran-ukuran antropometris miliknya sendiri. Gambar 5.1 menunjukkan struktur algoritma dari metode individualisasi model HRIR fasa minimum yang diusulkan. Seperti dibahas dalam Bab 1, bahwa metode individualisasi seperti pada Gambar 5.1 merupakan metode baru yang dikembangkan dalam penelitian ini, yang belum pernah dilakukan oleh peneliti-peneliti dunia lain. Individualisasi model HRIR fasa minimum dilakukan untuk sekumpulan respon-respon $HRIR_{mp}$ di bidang horisontal [41] dan sekumpulan respon-respon $HRIR_{mp}$ di bidang median [42]. Jadi dalam individualisasi di sini digunakan dua kumpulan data $HRIR_{mp}$ yang dihitung dari respon-respon HRIR asli yang bersangkutan dari CIPIC HRTF Database.



Gambar 5.1. Metode Individualisasi Model HRIR Fasa Minimum yang Diusulkan

Selanjutnya dijelaskan individualisasi model HRIR fasa minimum untuk sumber-sumber bunyi di bidang horisontal. Dari Gambar 5.1, terlihat bahwa semua HRIR asli di bidang horisontal dari 37 subyek, mula-mula diproses dengan analisis *cepstral* untuk mengkonversinya menjadi respon-respon $HRIR_{mp}$ dengan 67 sampel yang bersangkutan dengan menggunakan persamaan (C.9). Berikutnya, respon rata-rata dari seluruh $HRIR_{mp}$, μ , dihitung menggunakan persamaan (3.22). Untuk mendapatkan data $HRIR_{mp}$ dengan *zero empirical mean*, yang disebut dengan data *direct impulse response* fasa minimum (DIR_{mp}), maka respon rata-

rata μ dikurangi dari setiap $HRIR_{mp}$. Data DIR_{mp} adalah masukan bagi PCA. Hasil dari pemodelan PCA adalah vektor-vektor basis ortonormal dan bobot-bobotnya. Kombinasi linier dari 10 vektor basis yang terbobot menghasilkan model-model DIR_{mp} . Algoritma yang digunakan untuk mengindividualisasi model DIR_{mp} adalah MLR, seperti dijelaskan dalam subbab 5.1. MLR menggunakan nilai-nilai dari setiap PCW dengan azimuth θ di bidang horisontal, $w_i(\theta)$, $i=1,2,\dots,10$, dan 8 ukuran antropometris yang terseleksi, agar dihasilkan estimasi dari koefisien-koefisien regresi $\hat{\beta}_i(\theta)$ yang bersangkutan.

Proses rekonstruksi untuk menghasilkan model HRIR individual yang diinginkan ditunjukkan oleh garis putus-putus pada Gambar 5.1. Estimasi dari koefisien-koefisien regresi $\hat{\beta}_i(\theta)$ dan 8 ukuran antropometris dari seorang pendengar baru digunakan oleh MLR untuk mengestimasi nilai-nilai dari 10 PCW. Estimasi dari nilai-nilai 10 PCW ini dikombinasikan dengan vektor-vektor basisnya untuk membangun model-model DIR_{mp} individual. Kemudian setiap model DIR_{mp} individual ditambahkan kembali dengan respon rata-rata μ , sehingga dihasilkan model-model $HRIR_{mp}$ individual. Waktu-waktu tunda yang berkaitan dengan jarak dari sumber bunyi ke masing-masing gendang telinga kiri dan kanan (*onset time delays*) disisipkan pada model-model $HRIR_{mp}$ individual untuk masing-masing telinga agar dihasilkan model-model HRIR individual untuk masing-masing telinga dari posisi bunyi dengan azimuth θ di bidang horisontal.

Penjelasan mengenai metode individualisasi model-model $HRIR_{mp}$ di bidang median sama seperti penjelasan di atas dengan mengganti data HRIR asli di bidang horisontal dengan data HRIR asli di bidang median.

5.3 Analisis Hasil Individualisasi Model HRIR Fasa Minimum

Beberapa ukuran antropometris dari 8 subyek tidak diberikan dalam basis data yang digunakan. Pemodelan HRIR fasa minimum dengan PCA dan individualisasi model tersebut dengan MLR hanya menggunakan data HRIR dari 37 subyek dari total 45 subyek. Hasil-hasil individualisasi model HRIR fasa minimum di bidang horisontal telah dipublikasikan dalam [41] dan di bidang median dalam [42]. Evaluasi kinerja metode individualisasi ini dilakukan dengan

melakukan percobaan simulasi obyektif dan tes pendengaran subyektif dalam menentukan posisi sumber bunyi antara model HRIR individual dan HRIR non-individual yang berasal dari subyek 021 basis data (manikin KEMAR dengan telinga besar).

5.3.1 Hasil Simulasi Obyektif

Sebelum dilakukan individualisasi model HRIR fasa minimum, MSE rata-rata dengan pemodelan PCA dari seluruh basis data dihitung. Pertama-tama digunakan data HRIR fasa minimum untuk 45 subyek dari semua posisi sumber bunyi, diperoleh MSE rata-rata di antara semua posisi dan semua subyek sebesar 7,26%, tetapi MSE rata-rata di sepanjang bidang horisontal adalah sebesar 8,88%. Kedua, digunakan data dari 37 subyek, dari semua posisi sumber bunyi. Hasil MSE rata-rata yang diperoleh dengan percobaan ini sebesar 7,59% serta MSE rata-rata yang diperoleh di bidang horisontal adalah 9,23%. Ketiga, digunakan data dari 45 subyek di bidang horisontal, dengan MSE rata-rata sebesar 8,19%. Terakhir dilakukan pemodelan PCA dengan menggunakan data dari 37 subyek di bidang horisontal, MSE rata-rata adalah 8,32%.

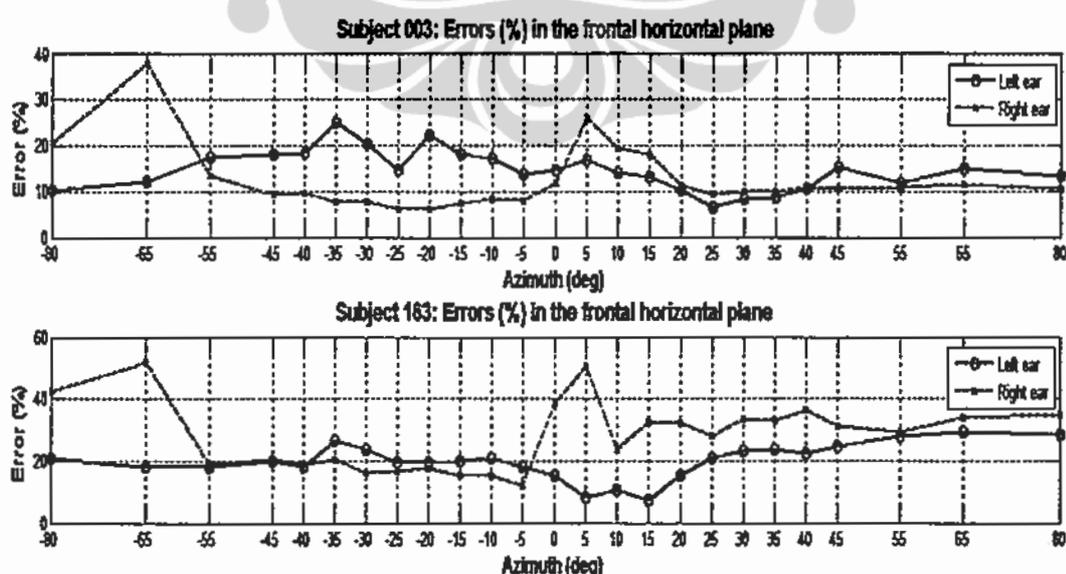
Dari hasil-hasil di atas dapat disimpulkan bahwa pertama, pemodelan dengan data dari subyek yang lebih banyak (secara umum lebih banyak data dilibatkan) menghasilkan MSE rata-rata yang lebih kecil. Kedua, penggunaan data yang lebih spesifik (hanya data pada bidang horisontal) juga menghasilkan MSE rata-rata yang lebih kecil. Fenomena ini juga terjadi pada pemodelan HRIR fasa minimum dengan PCA di bidang median. Dengan hanya menggunakan data HRIR dari kedua telinga 37 subyek, hanya pada posisi-posisi di bidang median, pemodelan HRIR fasa minimum memberikan MSE rata-rata sebesar 5,87%. Terlihat dari nilai-nilai MSE rata-rata bahwa kinerja pemodelan di bidang median lebih baik daripada pemodelan di bidang horisontal. Hal ini dapat dipahami karena semua HRIR fasa minimum di bidang median memiliki kemiripan satu sama lain baik dalam hal waktu-waktu tunda maupun struktur-strukturnya (dapat diamati dari Gambar 2.16).

Dengan menggunakan *setup* PCA yang sama seperti dalam [36], model-model HRIR fasa minimum di bidang median dalam penelitian ini, lebih baik

daripada model-model HRIR pulsa langsung di bidang median dari Hwang dan Park [36]. Dengan menggunakan 10 vektor basis, model-model mereka menghasilkan MSE rata-rata sebesar sekitar 6,67%, sedangkan model-model HRIR fasa minimum, juga dengan 10 vektor basis, memberikan MSE rata-rata hanya sebesar 5,31% [40]. Tetapi kemudian Hwang dan Park mengindividualisasi model-model mereka secara subyektif dengan cara mengatur tiga bobot yang memiliki deviasi standar terbesar pada setiap elevasi.

Kemudian dilakukan individualisasi model HRIR fasa minimum dengan MLR. Diamati bahwa MSE rata-rata berbeda dari subyek ke subyek dan MSE rata-rata untuk telinga kiri dan kanan seorang subyek tidaklah sama. Gambar 5.2 menunjukkan MSE telinga kiri dan MSE telinga kanan di sepanjang bidang horisontal depan dari subyek 003 dan subyek 163 basis data.

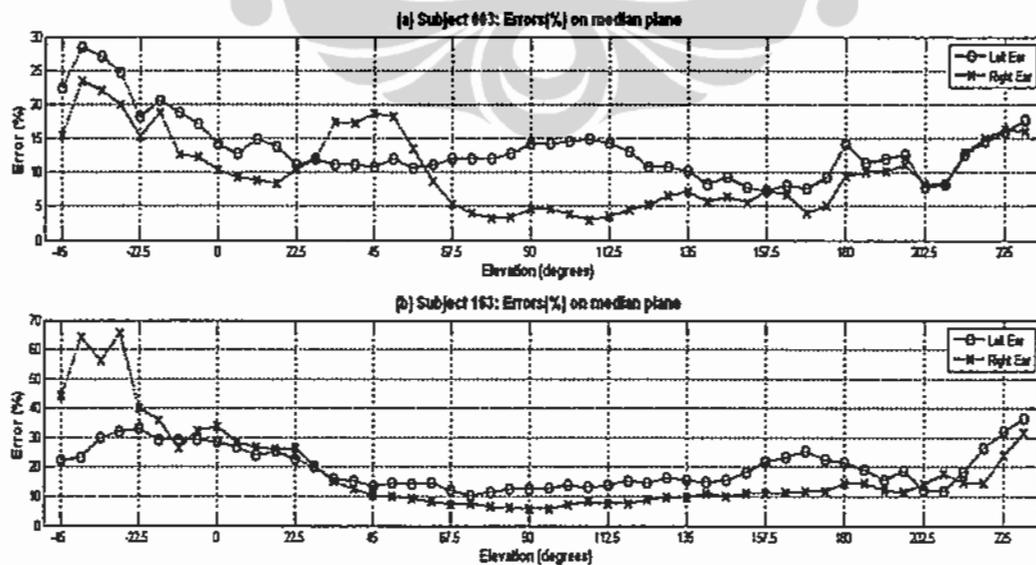
Nilai-nilai MSE telinga kiri dan MSE telinga kanan dari subyek 003 pada bidang horisontal depan pada umumnya kecil di bawah 20% kecuali pada 4 azimuth. Sedangkan nilai-nilai MSE dari subyek 163 dapat dikatakan kurang bagus pada sisi kanan bidang, yaitu di atas 20%. Tetapi pada sisi kiri bidang, diamati nilai MSE sebesar kira-kira 20% kecuali pada azimuth -80° and -65° . Dihitung dari nilai-nilai MSE pada Gambar 5.2, diperoleh MSE rata-rata di bidang horisontal depan untuk telinga kiri subyek 003 adalah 14,21% dan 11,26% untuk telinga kanannya.



Gambar 5.2. Kesalahan Model HRIR_{mp} Individual Subyek 003 dan Subyek 163 di Bidang Horisontal Depan

Tetapi untuk telinga kiri dan telinga kanan subyek 163, nilai-nilai MSE rata-rata di bidang horisontal depan adalah masing-masing 22,68% dan 28,71%. Jika dilihat dari nilai-nilai MSE rata-rata untuk kedua telinga, maka individualisasi model HRIR_{mp} di bidang horisontal depan dari subyek 003 adalah lebih baik daripada dari subyek 163.

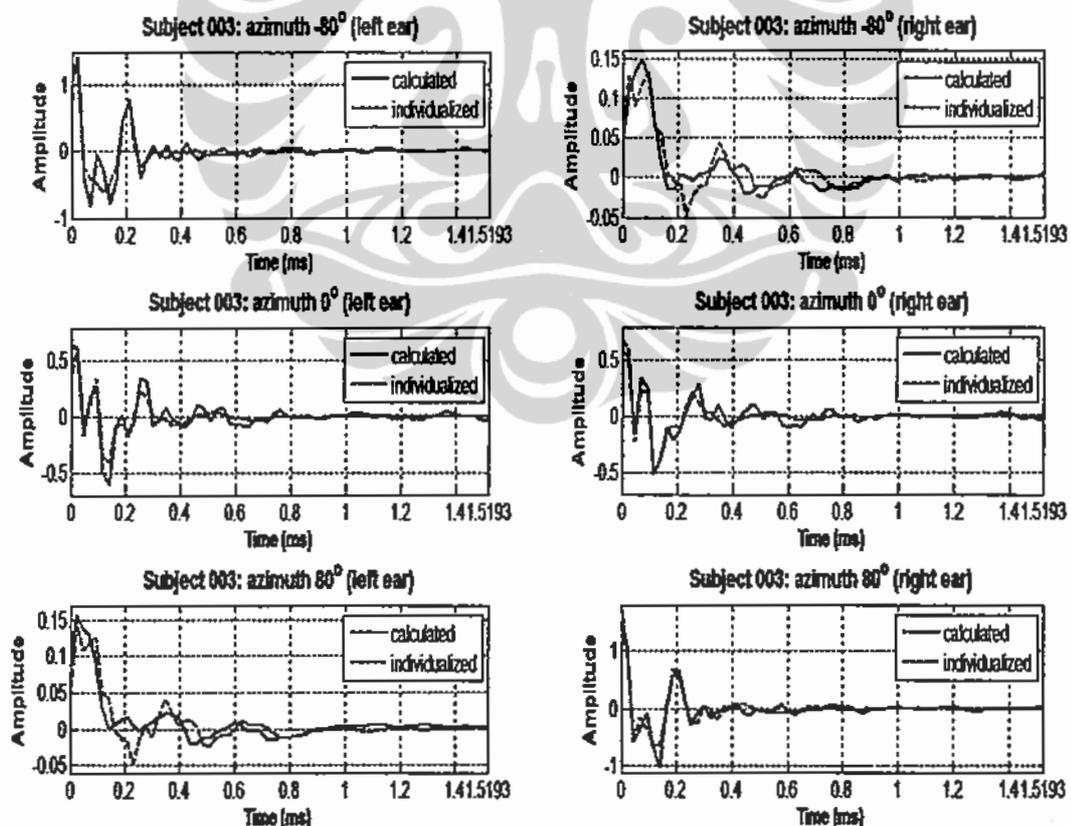
Gambar 5.3 menunjukkan kesalahan-kesalahan model HRIR_{mp} individual telinga kiri dan telinga kanan subyek 003 pada panel atas dan subyek 163 pada panel bawah; pada elevasi-elevasi di bidang median. Kesalahan-kesalahan model HRIR_{mp} individual telinga kiri dan telinga kanan dari subyek 003 pada umumnya kecil dengan MSE di bawah 20%, kecuali pada beberapa elevasi di depan bawah. MSE rata-rata di bidang median untuk telinga kiri adalah 13,45% dan 10,19% untuk telinga kanan. Individualisasi model HRIR_{mp} untuk telinga kanan subyek 163 menghasilkan kesalahan-kesalahan yang besar, dengan MSE lebih dari 40%, pada elevasi-elevasi antara -45° sampai $-22,5^\circ$. Pola kesalahan-kesalahan di sepanjang bidang median adalah serupa untuk kedua telinga subyek 163, kecuali pada elevasi-elevasi antara -45° sampai $-22,5^\circ$. Kinerja yang baik dari model-model individual untuk subyek 163 ditemui pada elevasi-elevasi antara $22,5^\circ$ dan $157,5^\circ$ (posisi-posisi di depan atas dan belakang atas kepala). Nilai-nilai MSE rata-rata untuk telinga kiri dan telinga kanan subyek 163 di bidang median masing-masing adalah 19,81% dan 18,76%.



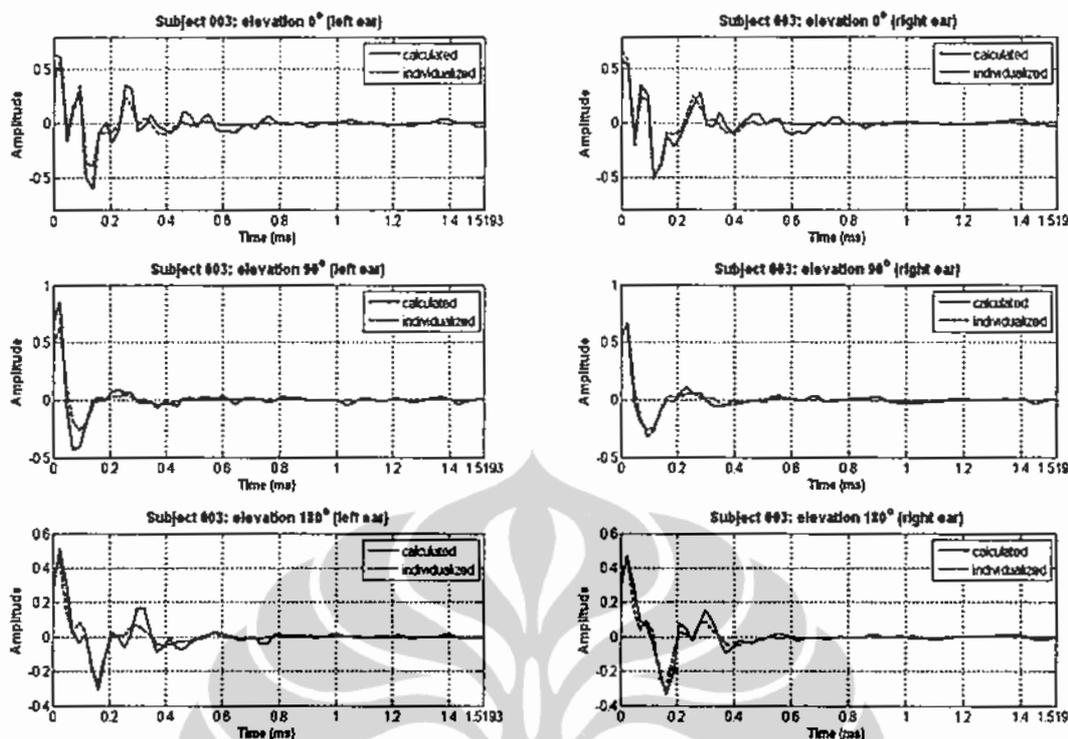
Gambar 5.3. Kesalahan Model HRIR_{mp} Individual Subyek 003 dan Subyek 163 di Bidang Median

Jika dilihat dari nilai-nilai MSE rata-rata untuk kedua telinga, maka individualisasi model-model $HRIR_{mp}$ di bidang median dari subyek 003 adalah lebih baik daripada individualisasi model-model $HRIR_{mp}$ dari subyek 163.

Model $HRIR_{mp}$ individual dari subyek 003 dapat mengaproksimasi dengan baik $HRIR_{mp}$ asli terutama pada beberapa sampel pertama, baik untuk bidang horisontal maupun bidang median. Gambar 5.4 menunjukkan model $HRIR_{mp}$ individual dan $HRIR_{mp}$ asli dari kedua telinga pada posisi-posisi ekstrim pada bidang horisontal depan, sedangkan Gambar 5.5 pada bidang median. Terlihat dari kedua gambar tersebut bahwa pola-pola dari model $HRIR_{mp}$ individual adalah sama dengan pola-pola dari $HRIR_{mp}$ asli yang relevan; di sini terdapat sedikit perbedaan dalam struktur-struktur detailnya. Gambar 5.4 bagian atas, tengah dan bawah berkaitan, secara berurutan, dengan azimuth -80° , 0° , and 80° . Sedangkan Gambar 5.5 bagian atas, tengah dan bawah secara berurutan berkaitan dengan elevasi 0° (tepat di depan), 90° (tepat di atas), and 180° (tepat di belakang).



Gambar 5.4. Model $HRIR_{mp}$ Individual dan $HRIR_{mp}$ Asli dari Subyek 003 di Bidang Horisontal Depan



Gambar 5.5. Model $HRIR_{mp}$ Individual dan $HRIR_{mp}$ Asli dari Subyek 003 di Bidang Median

Nilai MSE rata-rata dari 37 subyek di bidang horisontal adalah 22,5% sedangkan MSE rata-rata yang diperoleh dari PCA hanya 8,32%. Nilai MSE rata-rata dari individualisasi model $HRIR_{mp}$ di bidang median dari 37 subyek adalah 20,42%, tetapi pemodelan PCA sebelum individualisasi menghasilkan MSE rata-rata 5,87%. Kesalahan tambahan disebabkan oleh MLR karena di sini hanya digunakan regresi-regresi linier dalam mengestimasi PCWs. Sifat yang tidak sistematis dari PCWs di antara semua subyek dan semua posisi menyebabkan MLR cukup sulit mengestimasi.

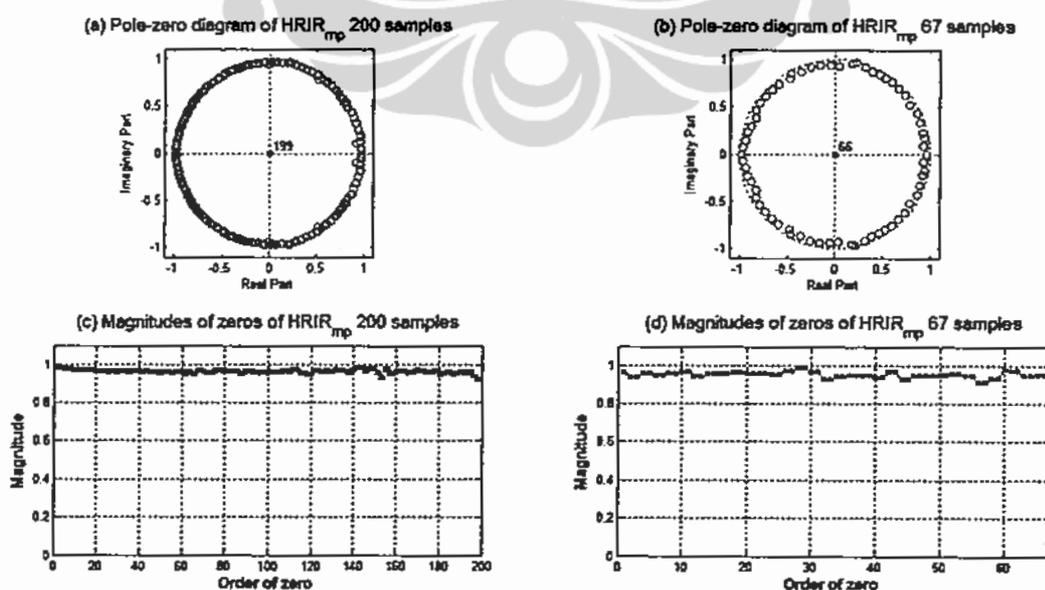
5.3.2 Diagram Pole-Zero Model HRIR Fasa Minimum Individual

Pada bagian ini, dibahas letak *pole* dan *zero* dari respon-respon $HRIR_{mp}$. Karena respon-respon $HRIR_{mp}$ adalah filter-filter FIR, maka pasti terdapat nilai-nilai *pole* pada titik pusat bidang-*z* sebanyak jumlah *zero* yang ada dalam setiap $HRIR_{mp}$. Individualisasi jenis data terbaik HRIR pada subbab 5.2 menggunakan data $HRIR_{mp}$ 67 sampel. Pertama-tama data HRIR asli 200 sampel dari basis data

dikonversi menjadi data $HRIR_{mp}$ 200 sampel. Kemudian hanya diambil 67 sampel dari setiap $HRIR_{mp}$ 200 sampel untuk individualisasi.

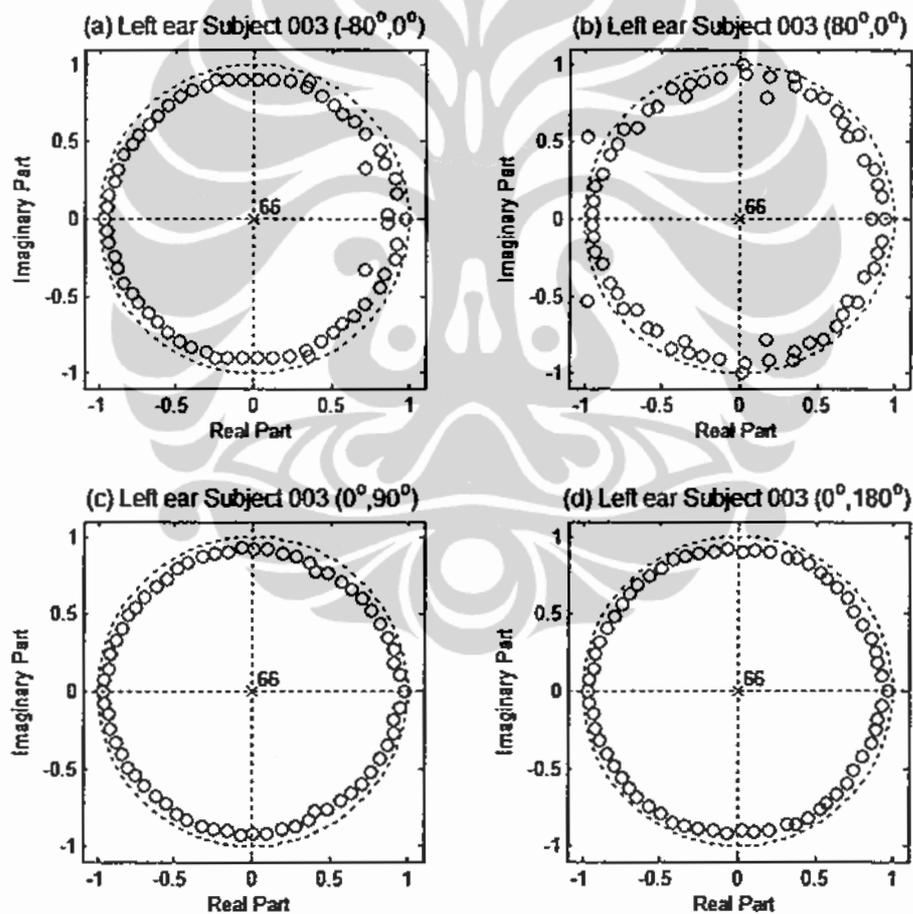
Letak semua *zero* dari setiap $HRIR_{mp}$ 200 sampel diselidiki, apakah semua *zero* terletak dalam lingkaran satuan dalam bidang-z, sesuai dengan syarat fasa minimum dari suatu respon impuls. Dari hasil pengamatan, diperoleh bahwa hampir semua $HRIR_{mp}$ 200 sampel memiliki nilai-nilai *zero* yang semuanya terletak di dalam lingkaran satuan. Seperti diketahui, terdapat seluruhnya 3700 $HRIR_{mp}$ untuk individualisasi di bidang horisontal dan 3700 $HRIR_{mp}$ untuk individualisasi di bidang median. Kemudian $HRIR_{mp}$ 200 sampel tersebut dipotong menjadi $HRIR_{mp}$ dengan 67 sampel pertama. Diamati bahwa hampir semua $HRIR_{mp}$ 67 sampel memiliki nilai-nilai *zero* yang semuanya terletak di dalam lingkaran satuan, sehingga dihasilkan respon-respon impuls fasa minimum.

Gambar 5.6a menunjukkan diagram *pole-zero* dari $HRIR_{mp}$ 200 sampel; Gambar 5.6b, diagram *pole-zero* dari $HRIR_{mp}$ 67 sampel; Gambar 5.6c, nilai-nilai *magnitude* dari setiap *zero* $HRIR_{mp}$ 200 sampel; Gambar 5.6d, nilai-nilai *magnitude* dari setiap *zero* $HRIR_{mp}$ 67 sampel; dimana masing-masing $HRIR_{mp}$ untuk posisi bunyi tepat di depan kepala ($0^{\circ}, 0^{\circ}$) dari telinga kiri subyek 003. Terlihat dari Gambar 5.6, bahwa semua *zero* dari $HRIR_{mp}$ 200 sampel dan semua *zero* dari $HRIR_{mp}$ 67 sampel terletak di dalam lingkaran satuan.



Gambar 5.6. Diagram *Pole-Zero* dari $HRIR_{mp}$ 200 Sampel dan $HRIR_{mp}$ 67 Sampel Beserta *Magnitude* dari Nilai-nilai *Zero* yang Bersangkutan

Berikutnya hampir semua model PCA dari $HRIR_{mp}$ 67 sampel juga memberikan semua *zero* terletak di dalam lingkaran satuan. Terakhir, hasil individualisasi model $HRIR_{mp}$ 67 sampel, seperti dijelaskan dalam bagian 5.3.1, menunjukkan semua *zero* masih terletak di dalam lingkaran satuan. Gambar 5.7 menunjukkan diagram-diagram *pole-zero* untuk model-model $HRIR_{mp}$ individual 67 sampel dari telinga kiri subyek 003, yang ditunjukkan pada panel-panel kiri dalam Gambar 5.4 dan Gambar 5.5, yakni untuk posisi-posisi bunyi $(-80^\circ, 0^\circ)$, $(80^\circ, 0^\circ)$, $(0^\circ, 90^\circ)$, $(0^\circ, 180^\circ)$. Terlihat dari Gambar 5.7b, bahwa terdapat dua *zero* dari model $HRIR_{mp}$ individual telinga kiri subyek 003 pada posisi bunyi $(80^\circ, 0^\circ)$ yang terletak di luar lingkaran satuan.



Gambar 5.7. Diagram *Pole-Zero* dari Model-Model $HRIR_{mp}$ Individual 67 Sampel dari Telinga Kiri Subyek 003

Dari masing-masing kumpulan data yang disebut di atas, yaitu $HRIR_{mp}$ 200 sampel, $HRIR_{mp}$ 67 sampel, model $HRIR_{mp}$ 67 sampel, model $HRIR_{mp}$ individual 67 sampel, diamati satu sampai tiga *zero* dari sebagian kecil data

HRIR_{mp} terletak di luar lingkaran satuan, tetapi dengan *magnitude* masing-masing *zero* yang tidak signifikan mendekati 1 (maksimum hanya 1,05). Nilai-nilai *zero* yang terletak di luar lingkaran satuan, dari model-model HRIR_{mp} individual 67 sampel, direfleksikan dan diambil nilai konjugasi kompleksnya, sehingga nilai-nilai *zero* tersebut masuk ke dalam lingkaran satuan. Kemudian dilakukan rekonstruksi ulang model-model HRIR_{mp} individual tersebut dengan menggunakan nilai-nilai *zero* yang telah direfleksikan. Hasil pengamatan mengenai percobaan ini adalah bahwa terdapat lebih banyak *zero* yang terletak di luar lingkaran satuan daripada sebelum direfleksikan, tetapi semua *zero* tersebut dengan *magnitude* yang lebih kecil. Dari percobaan ini ditarik kesimpulan bahwa terjadi *rounding errors* dalam perangkat lunak yang digunakan dalam menghitung data HRIR_{mp} awal dan dalam menghitung respon impuls dari nilai-nilai *zero*-nya.

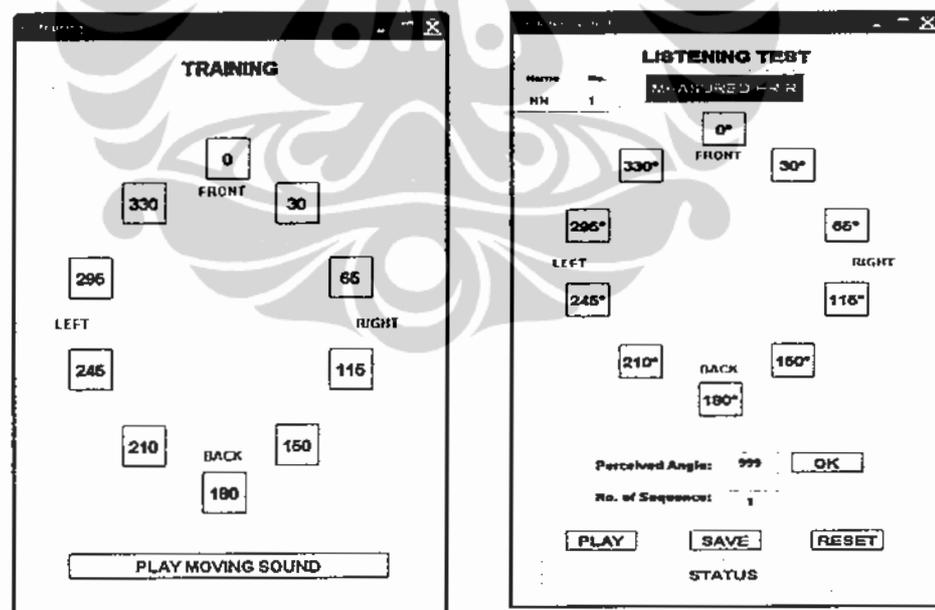
5.3.3 Hasil Penentuan Posisi Bunyi

Kinerja metode individualisasi dapat dievaluasi lebih jauh oleh percobaan penentuan posisi sumber bunyi dengan *headphones*. Tujuan dari percobaan ini adalah untuk membandingkan posisi bunyi yang didengar telinga pada bidang horisontal dari bunyi yang disaring dengan HRIR non-individual dan model HRIR individual yang berasal dari sintesis model HRIR_{mp} individual menjadi model HRIR individual. Percobaan dilakukan di ruangan kantor 3 x 3 m dengan latar belakang derau yang kecil. Dua mahasiswa laki-laki, yang berumur 19 dan 20 tahun, yang mempunyai sistem pendengaran normal, berpartisipasi dalam percobaan penentuan posisi bunyi. Kedua mahasiswa yang terlibat tadi disebut dengan subyek HK dan subyek TD. Mereka sebelumnya belum pernah mengikuti percobaan serupa. Ukuran-ukuran antropometris mereka yang diperlukan diukur melalui foto-foto.

Percobaan ini dibagi dalam 2 sesi yaitu sesi latihan dan sesi tes. Kedua subyek diminta untuk menjalani sesi latihan agar mereka terbiasa mendengar bunyi yang berasal dari posisi-posisi yang dipilih di bidang horisontal dan kemudian menentukan secara benar posisi yang sebenarnya. Sesi latihan berlangsung selama kira-kira satu jam. Sepuluh azimuth target dipilih di sepanjang bidang horisontal yaitu 0°, 30°, 65°, 115°, 150°, 180°, 210°, 245°, 295° dan 330°,

seperti terlihat dari Gambar 5.8, dimana 0° dan 180° adalah posisi depan dan belakang dari pusat kepala. Sumber bunyi yang digunakan berupa stimulus selebar 250 ms. Stimulus ini merupakan derau putih (*white noise*) yang memiliki pita frekuensi lebar. Stimulus diulang 3 kali untuk satu posisi bunyi dengan interval *silent* sebesar 150 ms di antara dua bunyi yang berurutan. Dalam sesi latihan, kedua subyek juga dapat mendengar bunyi-bunyi bergerak mengelilingi bidang horisontal sesuai dengan azimuth target.

Sesi tes dibagi lagi menjadi dua sesi yaitu sesi penentuan posisi bunyi dari HRIR non-individual dan dari model HRIR individual milik mereka. Setiap sesi terdiri dari 100 stimulus yang berisi 10 stimulus untuk setiap azimuth target yang berjumlah 10. Sepuluh stimulus untuk setiap azimuth target dapat merepresentasikan jumlah sampel untuk azimuth tersebut. Semua stimulus tadi diputar dalam urutan yang acak. Pada setiap urutan, kedua subyek diminta memutar satu stimulus dengan menekan tombol PLAY dan kemudian memutuskan di mana bunyi terdengar. Jumlah pengulangan stimulus sebelum menentukan pilihan tidak dibatasi.



Gambar 5.8. Panel untuk Sesi Latihan dan Sesi Tes dari Percobaan Penentuan Posisi Bunyi

Hasil dari sesi tes adalah sebagai berikut. Faktor kebenaran (*correct rate*) dari subyek HK untuk HRIR non-individual adalah 13% dan 19% untuk model HRIR individual. Subyek TD memberikan faktor kebenaran sebesar 14% untuk

HRIR non-individual dan 19% untuk model HRIR individual. Walaupun faktor kebenaran yang diperoleh rendah untuk kedua jenis HRIR, tetapi faktor kebenaran dari kedua subyek untuk model HRIR individual lebih baik dari pada faktor kebenaran untuk HRIR non-individual. Faktor keterbalikan depan dan belakang juga dihitung di sini. Faktor tersebut dari subyek HK adalah 7% untuk model HRIR individual dan 12% untuk HRIR non-individual serta 9% dan 15% untuk model HRIR individual dan HRIR non-individual dari subyek TD. Diperoleh keterbalikan yang lebih kecil dari kedua subyek untuk model HRIR individual.

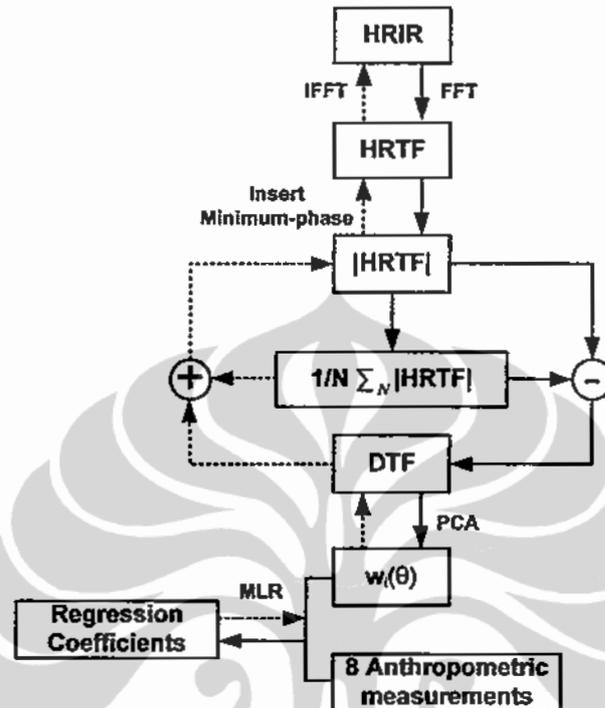
Hasil keseluruhan menunjukkan bahwa penentuan posisi bunyi lebih baik dengan menggunakan model HRIR individual daripada HRIR non-individual. Faktor keterbalikan depan dan belakang menunjukkan bahwa subyek tidak berpengalaman dalam percobaan penentuan posisi menggunakan *headphones* atau skenario percobaan tes kurang tepat. Gambar 5.8 menunjukkan panel-panel yang digunakan untuk percobaan penentuan posisi bunyi.

5.4 Metode Individualisasi Model HRTF *Magnitude* yang Diusulkan

Pada bagian berikut ini dijelaskan metode individualisasi model HRTF *magnitude* di sepanjang bidang horisontal. Metode ini mengindividualisasi model-model HRTF *magnitude* menjadi model-model HRIR individual yang sesuai untuk seorang pendengar, dengan menggunakan 8 ukuran antropometris miliknya. Model HRIR individual sangat diperlukan ketika pendengar tadi menggunakan aplikasi audio spasial. Diagram skematik dari metode individualisasi model HRTF *magnitude* yang diusulkan, ditunjukkan oleh Gambar 5.9.

Pertama-tama, dikumpulkan semua HRIR asli pada bidang horisontal (bidang dengan elevasi 0°) dari 37 subyek, yang terdiri dari 50 HRIR untuk setiap telinga dari seorang subyek. Pemodelan dan individualisasi HRIR di sini menggunakan seluruhnya 3700 HRIR. Setiap HRIR kemudian ditransformasi ke dalam kawasan frekuensi dengan DFT 256 titik, menjadi HRTF kompleks. Sebagai obyek dari pemodelan HRTF, hanya digunakan 128 komponen frekuensi dari *magnitude* HRTF kompleks. Pada langkah ini, fasa dari HRTF kompleks diabaikan. Kemudian, dihitung rata-rata dari semua HRTF *magnitude*. Rata-rata ini dikurangi dari setiap HRTF *magnitude* untuk mendapatkan DTF *magnitude*.

Pengurangan ini perlu dilakukan dalam PCA agar diperoleh data terpusat dari HRTF *magnitude* sehingga vektor basis pertama tidak mendekati rata-rata yang dihitung tadi.



Gambar 5.9. Metode Individualisasi Model HRTF *Magnitude* yang Diusulkan

Untuk keperluan pemodelan HRTF, semua DTF *magnitude* dimasukkan ke dalam PCA. PCA menghasilkan 128 vektor basis yang berurutan (disebut juga vektor eigen atau *principal component / PC*) dan bobot-bobot yang bersangkutan (PCWs). Seluruh 128 vektor eigen diurutkan dari vektor eigen yang memiliki nilai eigen terbesar sampai dengan vektor eigen dengan nilai eigen terkecil. Setiap nilai eigen merepresentasikan persentase varians dari semua data DTF *magnitude*. Vektor eigen pertama yang berhubungan dengan nilai eigen terbesar merepresentasikan persentase varians terbesar dari semua DTF *magnitude*. Untuk mendapatkan spektrum-spektrum HRTF yang terindividualisasi untuk seorang pendengar baru, maka MLR dilakukan antara PCWs dan 8 ukuran antropometris yang terpilih dari 37 subyek basis data. Proses seleksi 8 ukuran antropometris dari total 27 ukuran dijelaskan dalam Bab 4.

Metode MLR memberikan aproksimasi koefisien-koefisien regresi, $\hat{\beta}_i(\theta, \phi)$, yang berkorelasi dengan PCWs dan 8 ukuran antropometris dari 37

subyek. Aproksimasi dari koefisien-koefisien regresi ini digunakan pada 8 ukuran antropometris dari seorang pendengar baru untuk mendapatkan estimasi-estimasi PCWs untuk pendengar tersebut. Kombinasi linier dari vektor-vektor basis yang terbobot, menggunakan estimasi-estimasi PCWs, menghasilkan suatu model DTF *magnitude* individual.

Model-model HRIR individual dari seorang subyek diperoleh dengan menggunakan proses rekonstruksi yang ditunjukkan oleh garis putus-putus dalam Gambar 5.9. Setiap model DTF *magnitude* individual yang diperoleh dari gabungan metode PCA dan MLR, ditambahkan kembali dengan rata-rata dari semua DTFs *magnitude* yang dihitung sebelumnya, untuk menghasilkan model HRTF *magnitude* individual. Pada akhirnya, fasa minimum disisipkan pada model HRTF *magnitude* individual untuk memberikan suatu model HRTF kompleks individual. Kulkarni [22] menyatakan bahwa fasa dari HRTF dapat diaproksimasi oleh fasa minimum. Fasa minimum dari model HRTF *magnitude* tertentu dihitung dengan menggunakan transformasi Hilbert dari logaritma natural model HRTF *magnitude* tersebut. Fasa minimum, φ_{mp} , dari suatu model HRTF *magnitude*, $|\hat{h}_i|$, (kolom ke- i dari matriks \hat{H} pada persamaan (3.27)), diberikan oleh:

$$\varphi_{mp} = \text{Imag}\{ \mathcal{H}\{-\ln|\hat{h}_i|\} \}, \quad (5.1)$$

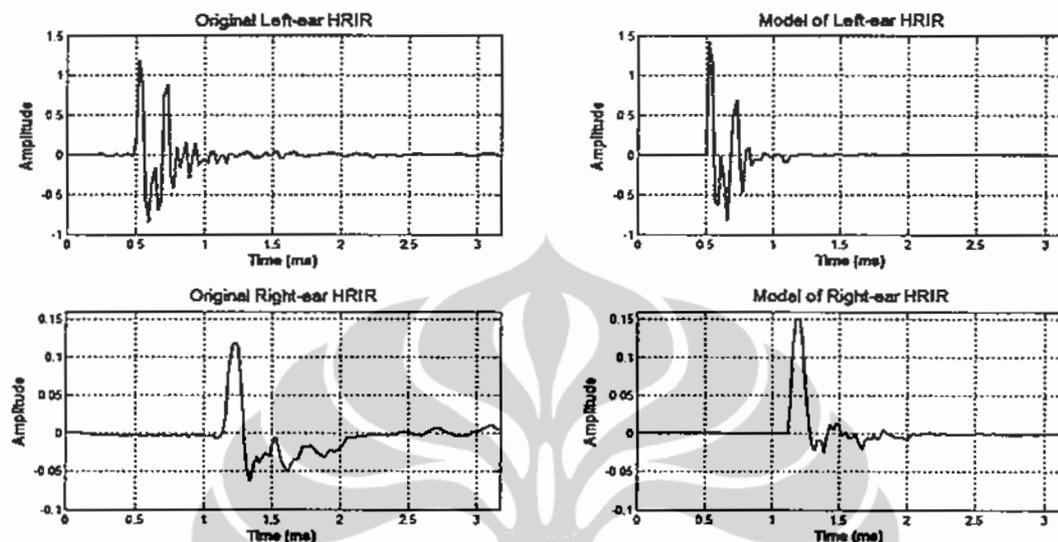
dimana $\text{Imag}\{\}$ menyatakan bagian imajiner dari suatu bilangan kompleks; \mathcal{H} adalah transformasi Hilbert dan \ln adalah logaritma natural. Jadi, model HRTF kompleks dengan fasa minimum, \hat{h}_{ci} , dapat dihitung sebagai berikut:

$$\hat{h}_{ci} = |\hat{h}_i| \cdot \exp(j \cdot \varphi_{mp}), \quad (5.2)$$

dimana $\exp()$ melambangkan fungsi eksponensial.

Invers DFT akhirnya diimplementasikan pada model HRTF kompleks individual sehingga didapatkan model HRIR individual. Waktu-waktu tunda awal dari telinga kiri dan telinga kanan, yang berkaitan dengan jarak antara sumber bunyi pada posisi tertentu dengan gendang telinga, disisipkan masing-masing pada HRIR fasa minimum telinga kiri dan HRIR fasa minimum telinga kanan. Waktu-waktu tunda tersebut diperoleh dari rata-rata waktu-waktu tunda dari posisi yang bersangkutan dari seluruh 37 subyek dalam basis data. Gambar 5.10 menunjukkan pada panel kiri HRIR asli untuk telinga kiri dan telinga kanan dari subyek 003 dari basis data pada posisi dengan azimuth -80° dan elevasi 0° .

Sedangkan pada panel kanan terlihat model-model HRIR yang bersangkutan. Tetapi, model-model HRIR pada Gambar 5.10 tidak dilakukan individualisasi. Terlihat kemiripan antara HRIR asli dan modelnya jika dilihat dari pola-pola, waktu-waktu tunda dan struktur detail dari HRIR.



Gambar 5.10 Perbandingan HRIR Asli dan Model HRIR dari Rekonstruksi Model-model HRTF *Magnitude*

5.5 Analisis Hasil Individualisasi Model HRTF *Magnitude*

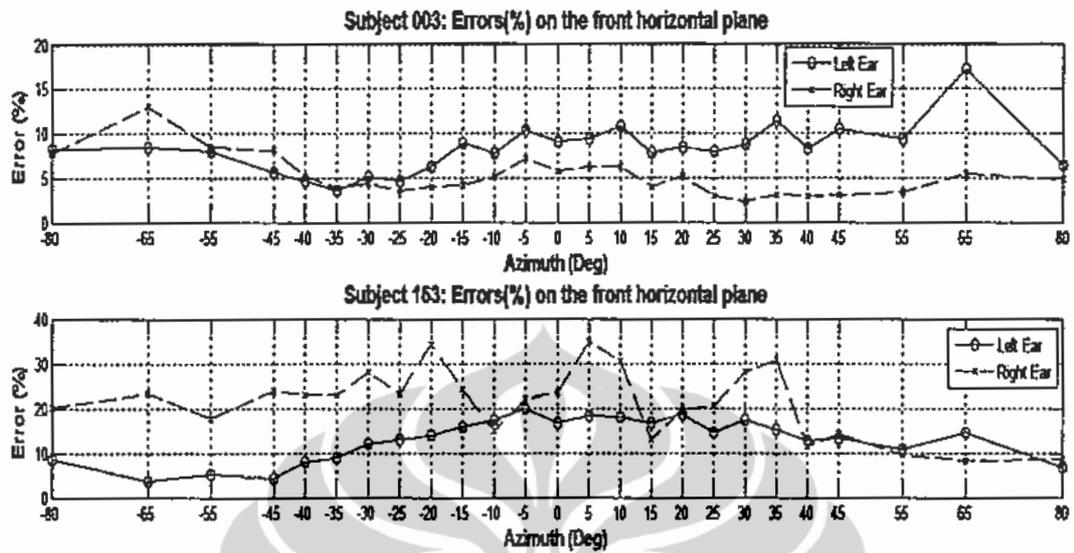
Sebelum model-model HRTF *magnitude* diindividualisasi menggunakan MLR, MSE rata-rata dari pemodelan HRTF *magnitude* dengan PCA dihitung dari semua data HRTF *magnitude*. Mula-mula pemodelan dilakukan untuk semua data HRIR dari semua posisi sumber bunyi dari 45 subyek. Percobaan ini menghasilkan MSE rata-rata 3,30% di antara semua posisi dan semua subyek, tetapi MSE rata-rata di sepanjang bidang horisontal adalah 3,65%. Kedua, pemodelan dilakukan hanya dengan data 37 subyek, yang menghasilkan MSE rata-rata 3,32% di antara semua posisi, sedangkan MSE rata-rata di antara posisi-posisi di bidang horisontal adalah 3,68%. Dari kedua jenis percobaan di atas, diamati bahwa kesalahan rata-rata yang terjadi adalah hampir sama. Ketiga, digunakan data dari 45 subyek pada posisi-posisi hanya di bidang horisontal, yang menghasilkan MSE rata-rata 3,67%. Akhirnya, digunakan data dari 37 subyek pada posisi-posisi di bidang horisontal. MSE rata-rata yang dihasilkan dari percobaan terakhir adalah 3,68%. Teramati kembali bahwa MSE rata-rata yang diperoleh hampir sama untuk kedua jenis percobaan terakhir.

Dari hasil-hasil ini, dapat ditarik kesimpulan bahwa MSE rata-rata di sepanjang posisi-posisi yang bersangkutan, hampir sama jika digunakan data dari 45 subyek maupun 37 subyek. Data spesifik HRIR yang digunakan (hanya HRIR pada bidang horisontal) memberikan MSE rata-rata yang hampir sama seperti jika digunakan data umum HRIR pada semua posisi. Hasil pengamatan pada pemodelan di bidang horisontal ini juga terjadi pada pemodelan di bidang median [43]. Pemodelan HRTF *magnitude* di bidang median dengan data dari 37 subyek hanya pada posisi-posisi bidang median menghasilkan MSE rata-rata sebesar 2,52%. Kesalahan-kesalahan rata-rata yang terjadi adalah kurang dari separuh dari kesalahan-kesalahan rata-rata yang berkaitan dari pemodelan HRIR fasa minimum.

Dalam individualisasi model HRTF *magnitude*, juga digunakan, seperti pada pemodelan HRIR fasa minimum, data HRIR dari kedua telinga hanya dari 37 subyek pada posisi-posisi di bidang horisontal dan pada posisi-posisi di bidang median [43]. Dengan demikian di sini digunakan hasil-hasil dari percobaan keempat seperti dijelaskan di atas, untuk individualisasi. MSE rata-rata dari model HRTF *magnitude* di sepanjang bidang horisontal (3,68%), jauh lebih kecil dibandingkan MSE rata-rata dari model HRIR fasa minimum di sepanjang bidang horisontal (8,32%).

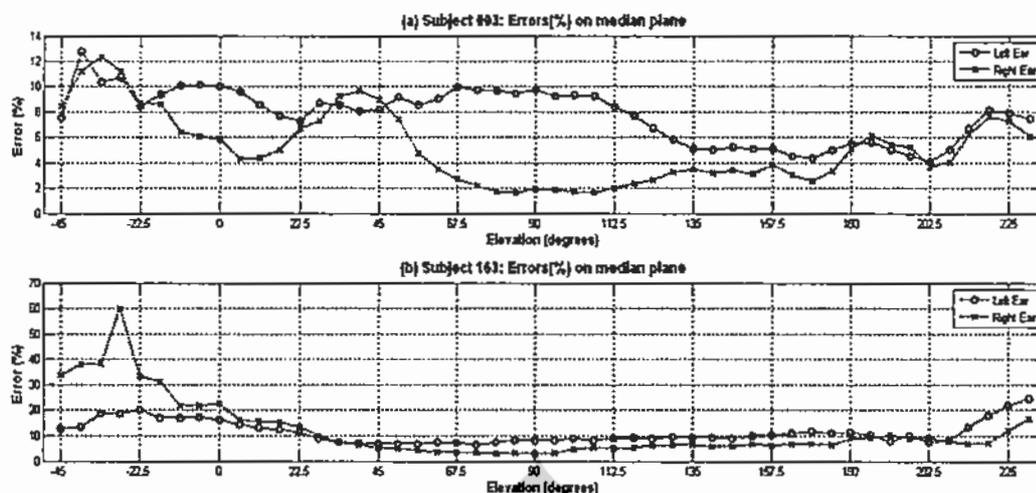
Pada saatnya, model-model PCA dari HRTF *magnitude* diindividualisasi menggunakan MLR dan 8 ukuran antropometris yang terpilih. Metode individualisasi model HRTF *magnitude* telah dijelaskan pada subbab 5.4 di atas. Diamati dari percobaan-percobaan, bahwa MSE rata-rata adalah berbeda untuk subyek yang berbeda dan kinerja individualisasi model untuk telinga kiri juga berbeda dengan kinerja individualisasi model untuk telinga kanan seorang subyek. Nilai MSE rata-rata keseluruhan dari individualisasi model-model HRTF *magnitude* di bidang horisontal adalah 12,17% dan di bidang median adalah 11,21%; dimana hasil-hasil ini lebih baik daripada nilai MSE rata-rata individualisasi model-model HRIR fasa minimum di bidang horisontal (22,5%) dan di bidang median (20,42%). Hal ini disebabkan karena spektrum-spektrum HRTF *magnitude* lebih halus daripada respon-respon HRIR fasa minimum.

PCA lebih baik memodelkan kurva-kurva yang lebih halus, demikian juga MLR lebih baik mengindividualisasi kurva-kurva yang lebih halus.



Gambar 5.11. Kesalahan Model HRTF *Magnitude* Individual Subyek 003 dan Subyek 163 di Bidang Horizontal Depan

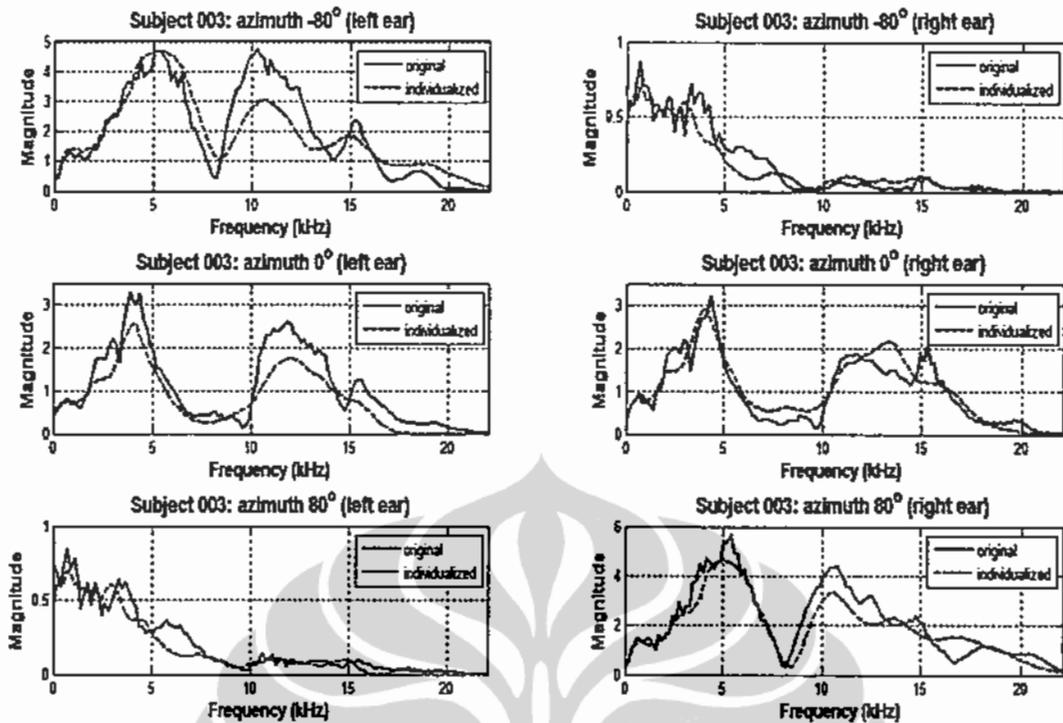
Gambar 5.11 menunjukkan kesalahan-kesalahan dari individualisasi model-model HRTF *magnitude* untuk telinga kiri dan telinga kanan dari subyek 003 (panel atas) dan dari subyek 163 (panel bawah), dengan azimuth-azimuth di sepanjang bidang horizontal depan. Secara umum, MSE rata-rata dari kedua telinga subyek 163 (12,92% untuk telinga kiri dan 21,20% untuk telinga kanan) adalah lebih buruk daripada untuk kedua telinga subyek 003 (8,27% untuk telinga kiri dan 5,18% untuk telinga kanan). MSE untuk telinga kiri subyek 003 di bidang horizontal depan secara umum sekitar 10%, kecuali pada azimuth 65° dan -65° ; dimana azimuth positif berkaitan dengan posisi sumber bunyi di sebelah kanan pusat kepala. Tetapi kesalahan-kesalahan untuk telinga kanan subyek 003 umumnya sangat baik sekitar 5% pada hampir semua azimuth, kecuali pada azimuth-azimuth -80° sampai -55° . Kesalahan-kesalahan dari telinga kiri subyek 163 terlihat lebih kecil daripada kesalahan-kesalahan dari telinga kanannya.



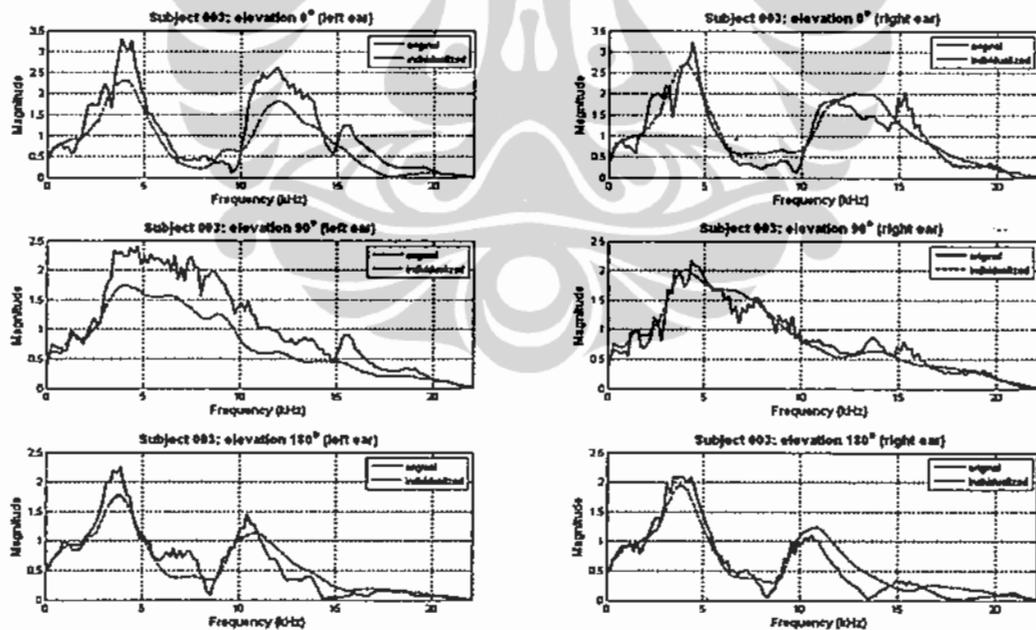
Gambar 5.12. Kesalahan Model HRTF *Magnitude Individual* Subyek 003 dan Subyek 163 di Bidang Median

Panel atas dari Gambar 5.12 menunjukkan kesalahan-kesalahan sebagai fungsi elevasi di bidang median dari subyek 003, sedangkan pada panel bawah ditunjukkan kesalahan-kesalahan untuk subyek 163 setelah proses individualisasi. MSE rata-rata dari telinga kiri subyek 003 adalah 7,75% dan dari telinga kanan subyek 003 adalah 5,18%. Kedua MSE rata-rata ini lebih baik daripada untuk subyek 163. MSE rata-rata untuk telinga kiri dan telinga kanan subyek 163 masing-masing adalah 11,38% dan 12,09%. Kesalahan-kesalahan untuk telinga kiri subyek 003 di bidang median sekitar kurang dari 10%, kecuali pada beberapa elevasi di depan bawah. Kesalahan-kesalahan untuk telinga kanan pada umumnya lebih kecil daripada untuk telinga kiri. Sedangkan kesalahan-kesalahan untuk telinga kiri subyek 163 diamati hampir sama seperti untuk telinga kanannya, kecuali pada beberapa elevasi dari -45° sampai $-16,875^\circ$ (di depan bawah). Kinerja yang pada umumnya buruk pada elevasi-elevasi di depan bawah disebabkan oleh refleksi-refleksi gelombang bunyi dari lantai.

Di dalam individualisasi model-model HRTF *magnitude* juga terdapat kesalahan-kesalahan tambahan yang disebabkan oleh metode MLR. Sifat yang tidak sistematis dari PCWs antar subyek dan antar posisi menyebabkan MLR kesulitan mengestimasi koefisien-koefisien regresi yang cukup akurat. Demikian juga di sini hanya digunakan regresi-regresi linier dari 8 ukuran antropometris untuk mengestimasi PCWs. Regresi-regresi dengan orde yang lebih tinggi dan kombinasi beberapa orde mungkin memberikan estimasi PCWs yang lebih baik.



Gambar 5.13. Model HRTF *Magnitude* Individual dan HRTF *Magnitude* Asli dari Subyek 003 di Bidang Horizontal Depan



Gambar 5.14. Model HRTF *Magnitude* Individual dan HRTF *Magnitude* Asli dari Subyek 003 di Bidang Median

Model-model HRTF *magnitude* individual dari subyek 003 dapat mengaproksimasi dengan baik HRTF *magnitude* asli yang bersangkutan, terutama pada frekuensi-frekuensi di bawah 8 kHz. Gambar 5.13 menunjukkan model-

model HRTF *magnitude* individual dan HRTF *magnitude* asli untuk kedua telinga dalam posisi-posisi ekstrim di bidang horisontal depan (azimut -80° , 0° , dan 80°); sedangkan Gambar 5.14 di bidang median (elevasi 0° , 90° dan 180°). Terlihat dari kedua gambar tersebut bahwa walaupun terdapat perbedaan yang cukup besar dalam struktur-struktur detail dari kedua jenis respon *magnitude*, tetapi pola-pola dari model HRTF *magnitude* individual adalah sama dengan pola-pola dari HRTF *magnitude* asli yang relevan.



Bab 6

KESIMPULAN

Penelitian ini menghasilkan metode baru dalam individualisasi model parametrik baik dari jenis data terbaik di kawasan waktu, yaitu HRIR fasa minimum dengan 67 sampel, maupun dari jenis data terbaik di kawasan frekuensi, yaitu HRTF *magnitude* dengan 256 komponen frekuensi, dengan menggunakan PCA dan MLR, berdasarkan 8 ukuran antropometris subyek, yaitu lebar kepala, tebal kepala, lebar leher, lebar pundak, tinggi *cavum concha*, lebar *cavum concha*, lebar pinna dan tinggi pinna.

Beberapa kesimpulan dapat ditarik dari penelitian ini, yaitu:

1. Dari analisis pemodelan PCA dari berbagai jenis data HRIR di kawasan waktu, diperoleh bahwa HRIR fasa minimum memberikan MSE rata-rata yang terkecil sebesar 7,26%.
2. Dari analisis pemodelan PCA dari berbagai jenis data HRTF di kawasan frekuensi, diperoleh bahwa HRTF *magnitude* memberikan MSE rata-rata yang terkecil sebesar 3,30%.
3. Analisis korelasi antara 27 ukuran antropometris baik terhadap ILD, ITD dan frekuensi lembah pinna maupun terhadap bobot-bobot vektor basis sangat efektif dan tepat dalam menyeleksi 8 ukuran antropometris yang dipakai dalam individualisasi.
4. Individualisasi model HRIR fasa minimum menggunakan MLR menghasilkan MSE rata-rata yang cukup memuaskan sebesar 22,5% dari di bidang horisontal dan 20,42% di bidang median.
5. Model-model HRIR fasa minimum individual yang dihasilkan masih memiliki sifat fasa minimum, yaitu semua *zero* terletak di dalam lingkaran satuan dalam bidang-z;

6. Di kawasan frekuensi, dihasilkan MSE rata-rata lebih kecil untuk individualisasi model HRTF *magnitude* menggunakan MLR, yaitu 12,17% di bidang horisontal dan 11,21% di bidang median.
7. Walaupun terdapat perbedaan-perbedaan dalam struktur-struktur detail antara respon-respon HRIR asli dan HRTF asli terhadap model-model HRIR dan HRTF individual yang bersangkutan, tetapi pola-pola dari model-model HRIR dan HRTF individual adalah sama dengan pola-pola dari respon-respon HRIR asli dan respon-respon HRTF asli yang relevan. Model-model HRIR dan HRTF individual dapat mengaproksimasi dengan baik respon-respon HRIR asli dan spektrum-spektrum HRTF asli yang bersangkutan.

Pengembangan lanjutan dari penelitian ini mencakup:

1. Pengujian model HRIR individual dengan menggunakan percobaan pendengaran yang mencakup posisi-posisi bunyi di bidang horisontal dan di bidang median; jumlah subyek yang lebih banyak; dan di dalam ruang kedap suara.
2. Penggunaan MLR dengan orde yang lebih tinggi atau kombinasi dari berbagai orde dalam mengestimasi bobot-bobot vektor basis PCA dengan ukuran-ukuran antropometris subyek.
3. Penggunaan metode individualisasi lain daripada MLR, misalnya algoritma genetika di dalam mencari hubungan antara bobot-bobot vektor basis PCA dengan ukuran-ukuran antropometris subyek.

DAFTAR ACUAN

- [1] Algazi, V. R., R. O. Duda, D. P. Thompson, and C. Avendano, "The CIPIC HRTF Database", *Proc. IEEE WASPAA01*, New Paltz, NY, pp. 99-102, 2001.
- [2] CIPIC HRTF Database Files, Release 1.2, September 23, 2004, tersedia di situs http://interface.cipic.ucdavis.edu/CIL_html/CIL_HRTF_database.htm, diakses 25 Juli 2005.
- [3] Carlile, S., "Virtual auditory space: Generation and applications", *R. G. Landes Company*, Austin, TX, 1996.
- [4] Duda, R. O., "Modeling head related transfer functions", *Proc. 27th Asilomar conf. on Signal, Systems and Computers*, Asilomar, CA, pp. 457-461, 1993.
- [5] Genuit, K., „Ein Modell zur Beschreibung von Aussenohruebertragungseigenschaften“, *Ph.D. dissertation*, Rheinisch-Westfaelischen Technischen Hochschule Aachen, Germany, 1984.
- [6] Blommer, M. A. and G. H. Wakefield, "Pole-zero approximations for head-related transfer functions using a logarithmic error criterion", *IEEE Trans. Speech Audio Processing*, vol. 5, pp. 278-287, 1997.
- [7] Durant, E. A. and G. H. Wakefield, "Efficient model fitting using a genetic algorithm: pole-zero approximations of HRTFs", *IEEE Trans. on Speech and Audio Processing*, vol. 10, No. 1, pp. 18 – 27, 2002.
- [8] Kahana, Y., P. A. Nelson, M. Petyt and S. Choi, "Numerical Modelling of the Transfer Functions of a Dummy-Head and of the External Ear", *Proc. AES 16th Int. Conf. Spatial Sound Reproduction*, Rovaniemi, Finland, 1999.
- [9] Katz, B. F. G., "Measurement and Calculation of Individual Head-Related Transfer Functions Using a Boundary Element Model Including the Measurement and Effect of Skin and Hair Impedance", *Ph.D. dissertation*, Graduate Program in Acoustics, Pennsylvania State University, 1998.
- [10] Speyer, G., "A Boundary Element Model for Predicting the Head-Related Transfer Function", *M.S. thesis*, Department of Electrical Engineering-Systems, Tel Aviv University, Israel, 1999.
- [11] Chen, J., B. D. Van Veen and K. E. Hecox, "External ear transfer function modeling: A beamforming approach", *J. Acoust. Soc. Amer.*, vol. 92, pp. 1933-1945, 1992.

- [12] Hasegawa, H., M. Kasuga, S. Matsumoto, and A. Koike, "Simply realization of sound localization using HRTF approximated by IIR filter", *IEICE Trans. Fundamentals*, vol. E83-A, No. 6, pp. 973-978, 2000.
- [13] Batteau, D.W., "The role of the pinna in human localization", *Proc. Royal Society London*, vol. 168 (series B), pp. 158-180, 1967.
- [14] Shaw, E. A. G., "Transformation of sound pressure level from the free-field to the eardrum in the horizontal plane", *J. Acoust. Soc. Am.*, vol. 56, pp. 1848-1861, 1974.
- [15] Algazi, V. R., C. Avendano, and D. P. Thompson, "Dependence of Subject and Measurement Position in Binaural Signal Acquisition", *J. Audio Eng. Soc.*, vol. 47, pp. 937-947, 1999.
- [16] Algazi, V.R., R. O. Duda, R. P. Morrison and D. M. Thompson, "Structural composition and decomposition of HRTF", *Proc. IEEE WASPAA01*, New Paltz, NY, pp. 103-106, 2001.
- [17] Algazi, V. R., R. O. Duda, R. Duraiswami, N. A. Gumerov and Z. Tang, "Approximating the head-related transfer function using simple geometric models of the head and torso", *J. Acoust. Soc. Am.*, vol. 112, pp. 2053-2064, 2002.
- [18] Brown, C. P. and R. O. Duda, "A structural model for binaural sound synthesis", *IEEE Trans. on Speech and Audio Processing*, vol. 6, no. 5, pp. 476-488, 1998.
- [19] Satarzadeh, P., V. R. Algazi and R. O. Duda, "Physical and filter pinna models based on antropometry", *122nd Convention of the Audio Engineering Society*, Vienna, Austria, paper 7098, 2007.
- [20] Cheung, Ngai-Man and S. Trautman, "Genetic Algorithm Approach to Head-Related Transfer Functions Modeling in 3-D Sound System", *Proc. IEEE First Workshop on Multimedia Signal Processing*, Princeton, NJ, USA, pp. 83 - 88, 1997.
- [21] Kistler, D. J. and F. L. Wightman, "A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction", *J. Acoust. Soc. Am.*, vol. 91 no. 3, pp. 1637-1647, 1992.
- [22] Kulkarni, A., S. K. Isabelle and H. S. Colburn, "On the minimum-phase approximation of head-related transfer functions", *Proc. IEEE WASPAA95*, pp. 84-87, 1995.
- [23] Middlebrooks, J. C. and D. M. Green, "Observations on a principal components analysis of head-related transfer functions", *J. Acoust. Soc. Am.*, vol. 92, no. 1, pp. 597-599, 1992.

- [24] Chen, J., B. D. Van Veen and K. E. Hecox, "A spatial feature extraction and regularization model for the head-related transfer function", *J. Acoust. Soc. Am.*, vol. 97, no. 1, pp. 439-452, 1995.
- [25] Sodnik, J. and S. Tomazic, "Directional information in head related transfer functions", *Proceedings of TENCON 2004 – 2004 IEEE Region 10 Conf.*, Chiang Mai, Thailand, vol. 2, pp.100-103, 2004.
- [26] Sodnik, J., A. Umek, R. Susnik, G. Bobojevic, and S. Tomazik, "Representation of head-related transfer functions with principal component analysis", *Proc. Acoustics 2004 – 2004 Conference of the Australian Acoustical Society*, Gold Coast, Australia, pp. 603 – 607, 2004.
- [27] Wu, Z., F. H. Y. Chan and F. K. Lam, "A time domain binaural model based on spatial feature extraction for the head-related transfer function", *J. Acoust. Soc. Am.*, vol. 102, no. 4, pp. 2211-2218, 1997.
- [28] Leung, J., and S. Carlile, "PCA compression of HRTFs and localization performance", *Int. Workshop on The Principles and Applications of Spatial Hearing*, Miyagi, Japan, pp. 1 – 3, 2009.
- [29] Pec, M., and P. Strumitto, "Principal component analysis for head-related transfer function low-order filter design", *Proc. Acoustics 2009 - 34th Int. Acoustical Conf. – EAA Symposium*, Adelaide, Australia, paper no. 26, 2009.
- [30] Shimada, S., M. Hayashi and S. Hayashi, "A clustering method for sound localization transfer functions", *J. Audio Eng. Soc.*, vol. 42, no. 7/8, pp. 577-584, 1994.
- [31] Jin, J., P. Leong, J. Leung, A. Corderoy and S. Carlile, "Enabling individualized virtual auditory space using morphological measurements", *Proc. First IEEE Pacific-Rim Conf. on Multimedia*, pp. 235-238, 2000.
- [32] Middlebrooks, J. C., "Individual differences in external-ear transfer functions reduced by scaling in frequency", *J. Acoust. Soc. Am.*, vol. 106, no.3, pp. 1480-1492, 1999.
- [33] Middlebrooks, J. C., "Virtual localization improved by scaling non-individualized external-ear transfer functions in frequency", *J. Acoust. Soc. Am.*, vol. 106, no.3, pp. 1493-1510, 1999.
- [34] Zotkin, D. N., J. Hwang, R. Duraiswami, L. S. Davis, "HRTF personalization using antropometric measurements", *Proc. IEEE WASPAA03*, New Paltz, NY, pp. 157-160, 2003.

- [35] Shin, K. and Y. Park, "Enhanced Vertical Perception through Head-Related Impulse Response Customization Based on Pinna Response Tuning in the Median Plane", *IEICE Trans. Fundamentals*, vol. E91-A, no. 1, pp. 345-356, 2008.
- [36] Hwang, S. and Y. Park, "HRIR customization in the median plane via principal components analysis", *Proc. AES 31st Int. Conf. New Directions in High Resolution Audio*, London, UK, 2007.
- [37] Hwang, S., Y. Park, and Y. Park, "Modeling and customization of head-related transfer functions using principal component analysis", *Proc. 2008 IEEE Int. Conf. on Control, Automation and Systems*, Seoul, Korea, pp. 227 - 231, 2008.
- [38] Hu, H., L. Zhou, J. Zhang, H. Ma, and Z. Wu, "Head Related Transfer Function Personalization Based on Multiple Regression Analysis," *Proc. IEEE Int. Conf. on Computational Intelligence and Security*, Guangzhou, China, vol. 2, pp.1829 - 1832, 2006.
- [39] Hu, H., L. Chen, and Z. Wu, "The estimation of personalized HRTFs in individual VAS", *Proc. IEEE 4th Int. Conf. on Natural Computation*, Jinan, China, pp. 203 - 207, 2008.
- [40] Hugeng, W. Wahab, and D. Gunawan, "Effective Preprocessing in Modeling Head-Related Impulse Responses Based on Principal Components Analysis", *Signal Processing: An International Journal (SPIJ)*, vol. 4, no. 4, pp. 201-212, 2010.
- [41] Hugeng, W. Wahab, and D. Gunawan, "Enhanced Individualization of Head-Related Impulse Response Model in Horizontal Plane Based on Multiple Regression Analysis", *Proc. IEEE 2010 2nd Int. Conf. on Computer Engineering and Applications (ICCEA 2010)*, pp. 226-230, 2010.
- [42] Hugeng, W. Wahab, and D. Gunawan, "A Novel Individualization of Head-Related Impulse Responses on Median Plane Using Listener's Anthropometries Based on Multiple Regression Analysis", *Jurnal Penelitian dan Pengembangan Telekomunikasi Institut Teknologi Telkom*, vol. 15 no. 1, in press.
- [43] Hugeng, W. Wahab, and D. Gunawan, "The Effectiveness of Chosen Partial Anthropometric Measurements in Individualizing Head-Related Transfer Functions on Median Plane", *ITB Journal of ICT*, submitted.
- [44] Duda, R. O, "3-D Audio for Human-Computer Interface", *Human-Computer Interface Design Course*, Department of Electrical Engineering San Jose State University, 1996.

- [45] Rayleigh, L., "On our perception of sound direction", *Philosophical magazine*, vol. 3, pp. 214 – 232, 1907.
- [46] Han, H. L., "Measuring a dummy head in search of pinna cues", *J. Audio Eng. Soc.*, vol. 42, no. 1/2, pp 15 – 37, 1994.
- [47] Berg, R. E., and D. G. Stork, "The physics of sound", Prentice-Hall, Englewood Cliffs, NJ, 1982.
- [48] Blauert, J., "Spatial hearing: The psychophysics of human sound localization", MIT Press, Cambridge, MA, 1997.
- [49] Gardner, W. G., and K. D. Martin, "HRTF measurements of a KEMAR", *J. Acoust. Soc. Am.*, vol. 97, no. 6, pp. 3907-3908, 1995.
- [50] Begault, D. R., „3D sound for virtual reality and multimedia“, *National Aeronautics and Space Administration*, Ames Research Center, Moffett Field, California, USA, 2000.
- [51] Jolliffe, I. T., "Principal component analysis", Second edition, *Springer-Verlag*, New York, 2002.
- [52] Inoue, N., T. Kimura, T. Nishino, K. Itou and K. Takeda. "Evaluation of HRTFs estimated using physical features", *Acoust. Sci. & Tech.*, vol. 26, no. 5, pp. 453-455, 2005.
- [53] Xu, S., Z. Li, and G. Salvendy. "Improved method to individualize head-related transfer function using anthropometric measurements", *Acoust. Sci. & Tech.*, vol. 29, no. 6, pp. 388-390, 2008.
- [54] Weisberg, S., "Applied linear regression", Third edition, John Wiley and Sons, Hoboken, New Jersey, pp. 21 - 65, 2005.
- [55] Oppenheim, A. V., R. W. Schaffer, and J. R. Buck, "Discrete-time signal processing", Second edition, Prentice Hall, Inc., New Jersey, pp. 280 – 291, 1999.
- [56] Mehrgardt, S., and V. Mellert, "Transformation characteristics of the external human ear", *J. Acoust. Soc. Am.*, vol. 61, no. 6, pp. 3907-3908, June 1977.
- [57] Nam, J., A. K. Miriam, and J. S. Abel, "On the minimum-phase nature of head-related transfer functions", *Proc. 125th AES Convention*, San Francisco, USA, paper no. 7546, October 2008.

PRINCIPAL COMPONENTS ANALYSIS

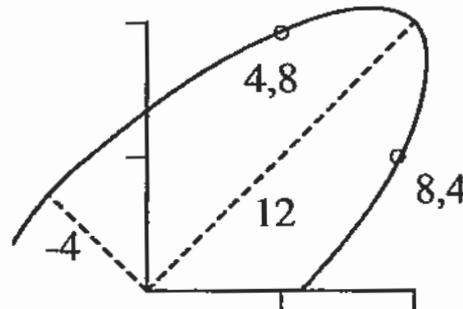
Principal Components Analysis (PCA) adalah suatu teknik statistikal yang mentransformasi secara ortogonal suatu kumpulan variabel asli menjadi suatu kumpulan lebih kecil dari variabel-variabel yang tidak berkorelasi yang merepresentasikan sebagian besar informasi dalam kumpulan variabel-variabel asli. Ide awal PCA pertama kali diungkapkan oleh Pearson pada tahun 1901 dan dikembangkan secara independen oleh Hotelling pada tahun 1933 [51]. PCA digunakan secara luas dalam pemrosesan sinyal, statistik, dan komputasi neural. Tergantung dari bidang aplikasinya, PCA disebut juga dengan Transformasi Diskrit Karhunen-Loève, Transformasi Hotelling, atau *proper orthogonal decomposition*. Transformasi ortogonal ini telah digunakan secara luas dalam analisis data dan kompresi data. Tujuan PCA adalah untuk mereduksi dimensi dari kumpulan data asli, tetapi tetap mempertahankan sebanyak mungkin variasi yang terdapat dalam data asli. Suatu kumpulan kecil variabel-variabel yang tidak berkorelasi jauh lebih mudah dimengerti dan digunakan dalam analisis- analisis selanjutnya daripada sekumpulan besar variabel-variabel yang berkorelasi.

PCA adalah suatu metode *multivariate*, dimana operasi matriks memegang peranan penting. Operasi ini melibatkan perhitungan nilai-nilai eigen (atau dikenal juga sebagai akar-akar *latent*) dan vektor-vektor eigen. Perhatikan matriks A berikut ini dimana baris-barisnya adalah koordinat-koordinat dari sepasang titik dalam ruang 2 dimensi.

$$A = \begin{bmatrix} 4 & 8 \\ 8 & 4 \end{bmatrix}$$

Sebuah elips dapat dibangun, dengan menggunakan koordinat (0,0) sebagai pusatnya, sedemikian rupa sehingga kedua titik terletak pada perimeter elips tersebut, seperti ditunjukkan oleh Gambar A.1. Karena matriks A terdiri dari dua baris, maka dapat dihitung dua nilai eigen dari matriks A , yaitu -4 dan 12 , yang terlihat dalam Gambar A.1 sebagai panjang sumbu minor dan panjang sumbu mayor dari elips. Demikian juga terdapat satu vektor eigen untuk setiap nilai eigen

dari matriks A . Perlu diperhatikan bahwa, nilai-nilai eigen dan vektor-vektor eigen hanya dapat dihitung dari matriks-matriks simetris seperti matriks A .



Gambar A.1. Deskripsi Grafikal dari Nilai-nilai Eigen

Sumbu-sumbu dari elips dapat digambarkan jika diberikan informasi mengenai panjang-panjang dan orientasi-orientasi sumbu mayor dan minor (koordinat-koordinatnya). Vektor-vektor eigen merupakan koordinat-koordinat yang mendefinisikan orientasi dari sumbu-sumbu koordinat yang memiliki panjang sesuai dengan nilai-nilai eigen yang bersangkutan. Tetapi, vektor-vektor eigen yang berpusat pada $(0,0)$ tidak memiliki nilai-nilai yang unik. Setiap vektor eigen memiliki tak berhingga banyaknya kemungkinan nilai-nilai. Hal ini disebabkan karena setiap basis pada sumbu akan sesuai.

PCA didasarkan pada representasi statistik dari suatu variabel acak. Misalkan terdapat suatu vektor acak populasi \mathbf{x} , dimana

$$\mathbf{x} = (x_1, x_2, \dots, x_M)^T, \quad (\text{A.1})$$

dan suatu rata-rata dari populasi tersebut diberikan oleh:

$$\boldsymbol{\mu} = E\{\mathbf{x}\}, \quad (\text{A.2})$$

dan matriks kovarians dari himpunan data yang sama dapat dituliskan sebagai:

$$\mathbf{S} = E\{(\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^T\}. \quad (\text{A.3})$$

Komponen-komponen dari \mathbf{S} , yang dilambangkan dengan s_{ij} , merepresentasikan kovarians-kovarians antara komponen variabel acak x_i dan komponen variabel acak x_j . Sedangkan komponen s_{ii} merupakan varians dari komponen variabel acak x_i . Varians sebuah komponen menunjukkan simpangan nilai-nilai komponen di sekitar nilai rata-ratanya. Jika dua komponen x_i dan x_j dari data tidak berkorelasi,

maka kovarians antara dua komponen tersebut adalah nol ($s_{ij} = s_{ji} = 0$). Karena definisi dalam persamaan (A.3), maka matriks kovarians selalu simetris.

Jika diambil vektor-vektor acak sampel x_1, x_2, \dots, x_N , maka dapat dihitung rata-rata sampel dan matriks kovarians sampel sebagai estimasi-estimasi dari rata-rata dan matriks kovarians populasi. Dari suatu matriks simetris seperti matriks kovarians, dapat dihitung suatu basis ortogonal dengan cara menemukan nilai-nilai eigen dan vektor-vektor eigen dari matriks simetris tersebut. Vektor-vektor eigen, v_i , and nilai-nilai eigen yang bersangkutan, λ_i , merupakan solusi dari persamaan eigen berikut ini:

$$S v_i = \lambda_i v_i, \text{ dimana } i=1,2,\dots,M. \quad (\text{A.4})$$

Agar menjadi lebih sederhana, dianggap nilai-nilai λ_i berbeda satu sama lain. Nilai-nilai ini dapat dihitung dengan menemukan solusi-solusi dari persamaan karakteristik berikut:

$$|S - \lambda I| = 0, \quad (\text{A.5})$$

dimana I adalah matriks identitas yang memiliki orde yang sama dengan S dan $|\cdot|$ melambangkan determinan matriks. Jika vektor data memiliki M komponen, persamaan karakteristik memiliki orde M . Persamaan ini mudah dipecahkan hanya jika nilai M kecil. Perhitungan nilai-nilai eigen dan vektor-vektor eigen adalah suatu pekerjaan penting dan terdapat banyak metode untuk keperluan tersebut.

Salah satu metode untuk memecahkan masalah nilai eigen adalah dengan menggunakan solusi neural terhadap masalah tersebut. Data diumpangkan sebagai masukan dan jaringan (*network*) konvergen menuju solusi yang diinginkan. Dengan mengurutkan vektor-vektor eigen dalam urutan berdasarkan nilai-nilai eigen yang semakin mengecil (nilai eigen yang terbesar pada urutan pertama), suatu basis ortogonal yang berurutan dapat diciptakan, dimana vektor eigen pertama memiliki arah pada varians data terbesar. Dengan cara ini, dapat ditemukan arah-arrah dimana himpunan data memiliki jumlah-jumlah energi yang sangat signifikan.

Vektor-vektor eigen dapat juga didefinisikan secara intuitif dengan menggunakan suatu formulasi rekursif. Misalkan vektor eigen pertama, v_1 , adalah:

$$\mathbf{v}_1 = \arg \max_{\|\mathbf{v}\|=1} E\{(\mathbf{v}^T \mathbf{d})^2\} \quad (\text{A.6})$$

dimana \mathbf{d} adalah vektor data terpusat (vektor data asli \mathbf{x} dikurangi dengan rata-rata μ), yang diberikan oleh:

$$\mathbf{d} = \mathbf{x} - \mu, \quad (\text{A.7})$$

dan $\arg \max_x f(x)$ adalah sekumpulan nilai-nilai x dimana $f(x)$ memiliki nilai maksimum ($\arg \max$ singkatan dari *argument of the maximum*). Vektor eigen pertama, \mathbf{v}_1 , memiliki dimensi M yang sama seperti vektor data terpusat \mathbf{d} . Semua vektor dalam subbab ini adalah vektor-vektor kolom. Jadi vektor eigen pertama adalah proyeksi pada arah dimana varians dari proyeksi tersebut maksimal.

Pengurangan rata-rata, μ , (*mean subtraction* atau *mean centering*) dari vektor data asli, \mathbf{x} , diperlukan dalam mengimplementasikan PCA untuk memastikan bahwa vektor eigen pertama mendeskripsikan arah varians maksimum. Jika pengurangan rata-rata tidak dilakukan, vektor eigen pertama hampir mirip dengan *mean* data. Suatu data dengan *zero empirical mean* diperlukan untuk menemukan suatu basis yang meminimalisasi *mean-square error* dari aproksimasi data.

Jika sebanyak $m-1$ vektor eigen telah ditentukan, vektor eigen yang ke- m dapat dihitung sebagai vektor eigen dari residual:

$$\mathbf{v}_m = \arg \max_{\|\mathbf{v}\|=1} E\{[\mathbf{v}^T (\mathbf{d} - \sum_{i=1}^{m-1} \mathbf{v}_i \mathbf{v}_i^T \mathbf{d})]^2\} \quad (\text{A.8})$$

Misalkan terdapat suatu himpunan data dengan rata-rata sampel dan matriks kovariansnya telah dihitung. Matriks \mathbf{V} adalah suatu matriks yang terdiri dari vektor-vektor eigen, \mathbf{v}_i , dari matriks kovarians sebagai vektor-vektor baris. Dengan mentransformasi suatu vektor data \mathbf{x} , diperoleh:

$$\mathbf{w} = \mathbf{V} (\mathbf{x} - \mu) \quad (\text{A.9})$$

yang merupakan satu titik pada sistem koordinat orthogonal yang didefinisikan oleh vektor-vektor eigen. Komponen-komponen dari \mathbf{w} dapat dilihat sebagai koordinat-koordinat dalam basis orthogonal. Vektor data asli \mathbf{x} dapat direkonstruksi dari \mathbf{w} , dengan cara berikut:

$$\hat{\mathbf{x}} = \mathbf{V}^T \mathbf{w} + \mu, \quad (\text{A.10})$$

di sini digunakan sifat dari suatu matriks ortogonal, yaitu $V^{-1} = V^T$. Matriks V^T adalah transpos dari matriks V . Vektor asli x diproyeksikan pada sumbu-sumbu koordinat yang didefinisikan oleh basis ortogonal. Vektor asli lalu direkonstruksi oleh suatu kombinasi linier dari vektor-vektor basis ortogonal.

Daripada menggunakan semua vektor eigen dari matriks kovarians, data dapat direpresentasikan oleh hanya beberapa vektor basis dari basis ortogonal. Jika matriks V_L adalah matriks V yang hanya terdiri dari L vektor eigen yang pertama, transformasi yang serupa dengan persamaan (A.9) dapat dilakukan, yaitu:

$$w = V_L (x - \mu), \quad (\text{A.11})$$

dan rekonstruksi vektor acak asli menjadi:

$$\hat{x} = V_L^T w + \mu. \quad (\text{A.12})$$

Hal ini berarti bahwa vektor data asli diproyeksikan pada sumbu-sumbu koordinat yang memiliki dimensi L dan mentransformasikan vektor tersebut kembali sebagai suatu kombinasi linier dari vektor-vektor basis. Cara ini meminimalisasi *mean-square error* antara data dan representasi berdasarkan jumlah vektor eigen yang diberikan. Jika data terkonsentrasi pada suatu sub ruang linier, PCA memberikan suatu cara untuk mengompres data tanpa kehilangan banyak informasi dan menyederhanakan representasi. Dengan mengambil vektor-vektor eigen yang berkaitan dengan nilai-nilai eigen terbesar, informasi yang hilang adalah minimal dari segi *mean-square*.

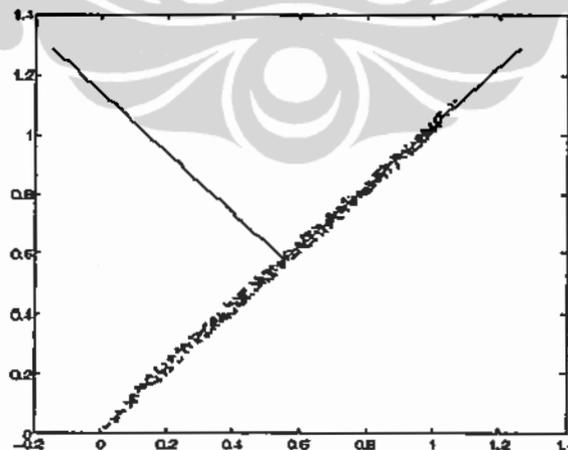
Jumlah vektor eigen beserta nilai-nilai eigennya dapat dipilih tetap sehingga diperoleh suatu representasi atau abstraksi data yang konsisten. Hal ini menyebabkan jumlah energi yang bervariasi dari data asli. Alternatif lain, jumlah energi yang hampir sama dapat dipilih dan jumlah vektor eigen yang bervariasi. Hal ini memberikan jumlah energi (informasi) yang hampir konsisten, tetapi didapat representasi-representasi yang bervariasi sesuai dengan jumlah vektor eigen yang digunakan.

Terlihat dua tujuan yang berlawanan: pada satu sisi, masalah yang dihadapi harus disederhanakan dengan mengurangi dimensi representasi; tetapi pada sisi lain isi informasi asli harus dipertahankan sebanyak mungkin. PCA menawarkan suatu cara yang baik dalam mengendalikan *trade-off* antara

kehilangan informasi dan penyederhaan masalah. Model-model linier sebagian dapat dihasilkan dengan membagi data masukan ke dalam daerah-daerah lebih kecil dan melakukan *fitting* data secara lokal.

Berikut ini dijelaskan suatu contoh kecil yang menunjukkan karakteristik-karakteristik dari vektor-vektor eigen. Beberapa data artifisial telah dibangkitkan, seperti terlihat dari Gambar A.2. Titik-titik kecil adalah titik-titik dalam kumpulan data. Rata-rata sampel dan matriks kovarians sampel dapat dihitung dengan mudah dari data. Kemudian vektor-vektor eigen dan nilai-nilai eigen dapat dihitung dari matriks kovarians. Arah-arrah dari setiap vektor eigen ditunjukkan sebagai garis-garis dalam Gambar A.2.

Vektor eigen pertama yang memiliki nilai eigen terbesar menunjuk pada arah varians terbesar (ke kanan atas), sementara itu, vektor eigen kedua ortogonal pada vektor eigen pertama (mengarah ke kiri atas). Dalam contoh ini, nilai eigen pertama yang berkaitan dengan vektor eigen pertama adalah $\lambda_1 = 0,1737$ sedangkan nilai eigen lain adalah $\lambda_2 = 0,0001$. Dengan membandingkan besar nilai-nilai eigen terhadap jumlah total besar nilai eigen, dapat diketahui seberapa banyak energi yang terkonsentrasi sepanjang vektor eigen tertentu. Dalam hal ini, vektor eigen pertama mengandung hampir semua energi. Data dapat diaproksimasi dengan baik dalam suatu representasi satu dimensi.



Gambar A.2. Vektor-vektor Eigen dari Data Artifisial

Sekarang dapat dijelaskan algoritma dari metode PCA dari sekumpulan vektor-vektor data, x_1, x_2, \dots, x_N , sebagai berikut:

1. Susun kumpulan data

Misalkan terdapat data yang terdiri dari sekumpulan observasi dengan M variabel. Diinginkan untuk mereduksi data sehingga setiap observasi dapat dideskripsikan dengan hanya L variabel, $L < M$. Misalkan juga data disusun sebagai suatu kumpulan N vektor data, $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$ dimana setiap \mathbf{x}_n merepresentasikan suatu observasi terkelompok tunggal dari M variabel.

- Tulis vektor-vektor $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$ sebagai vektor kolom, setiap vektor memiliki M baris;
- Tempatkan vektor-vektor kolom tersebut ke dalam satu matriks tunggal \mathbf{X} yang berukuran $M \times N$.

2. Hitung rata-rata empiris

- Hitung rata-rata empiris dari setiap dimensi $m = 1, 2, \dots, M$.
- Tempatkan nilai-nilai rata-rata yang sudah dihitung ke dalam vektor rata-rata empiris $\boldsymbol{\mu}$ yang berukuran $M \times 1$.

$$\mu[m] = \frac{1}{N} \sum_{n=1}^N X[m, n] \quad (\text{A.13})$$

3. Hitung simpangan setiap vektor observasi dari rata-rata

Pengurangan rata-rata adalah suatu bagian integral dari solusi menuju penentuan suatu vektor basis yang meminimalisasi *mean square error* dari aproksimasi data. Jadi proses *centering data* adalah sebagai berikut:

- Kurangkan vektor rata-rata empiris $\boldsymbol{\mu}$ dari setiap kolom matriks data \mathbf{X} ;
- Simpan data yang telah dikurangi dengan rata-rata ke dalam matriks \mathbf{D} yang berukuran $M \times N$;

$$\mathbf{D} = \mathbf{X} - \boldsymbol{\mu} \mathbf{y}, \quad (\text{A.14})$$

dimana \mathbf{y} adalah suatu vektor baris $1 \times N$ yang berisi semuanya 1; $y[n] = 1$ untuk $n=1, 2, \dots, N$.

4. Tentukan matriks kovarians

Tentukan matriks kovarians empiris, \mathbf{S} [$M \times M$], dari *outer product* matriks \mathbf{D} dengan dirinya sendiri:

$$\mathbf{S} = E[\mathbf{D} \otimes \mathbf{D}] = E[\mathbf{D} \cdot \mathbf{D}^*] = (\mathbf{D} \cdot \mathbf{D}^*) / (N-1) \quad (\text{A.15})$$

dimana

$E[\]$ adalah operator nilai ekspektasi;

⊗ adalah operator *outer product*; dan

* adalah operator *conjugate transpose*, jika D terdiri seluruhnya dari bilangan-bilangan riil, yang terjadi pada sebagian besar aplikasi, maka *conjugate transpose* sama seperti transpos biasa.

5. Tentukan nilai-nilai eigen dan vektor-vektor eigen dari matriks kovarians

- Hitung matriks V yang berisi vektor-vektor eigen yang membuat matriks kovarians menjadi terdiagonalisasi, dengan menggunakan persamaan berikut:

$$V^{-1}SV = \Lambda \quad (\text{A.16})$$

dimana Λ adalah matriks diagonal yang terdiri dari semua nilai eigen dari S pada diagonalnya. Langkah ini biasanya melibatkan penggunaan algoritma yang berbasis komputer untuk menghitung vektor-vektor dan nilai-nilai eigen. Algoritma-algoritma ini sudah tersedia sebagai sub-komponen dari kebanyakan sistem aljabar linier dalam banyak perangkat lunak matematika terkemuka.

- Matriks Λ berbentuk matriks diagonal $M \times M$, dimana:

$$\Lambda[p,q] = \lambda_m \text{ untuk } p=q=m \quad (\text{A.17})$$

adalah nilai eigen ke- m dari matriks kovarians S dan,

$$\Lambda[p,q] = 0 \text{ untuk } p \neq q. \quad (\text{A.18})$$

Matriks V [$M \times M$] mengandung M vektor kolom [$M \times 1$], yang merepresentasikan M vektor-vektor eigen dari matriks kovarians S .

Nilai-nilai dan vektor-vektor eigen diurutkan dan disusun berpasangan. Nilai eigen ke- m berhubungan dengan vektor eigen ke- m .

6. Susun ulang nilai-nilai eigen dan vektor-vektor eigen

- Sortir kolom-kolom dari matriks vektor eigen V dan matriks nilai eigen Λ dalam urutan nilai eigen yang berkurang.
- Pastikan untuk mempertahankan pasangan yang benar antara kolom-kolom dalam setiap matriks.

7. Hitung konten energi kumulatif untuk setiap vektor eigen

Nilai-nilai eigen merepresentasikan distribusi dari energi data sumber di antara vektor-vektor eigen, dimana vektor-vektor eigen membentuk suatu basis untuk

data tersebut. Konten energi kumulatif g untuk vektor eigen ke- m adalah konten energi sepanjang semua nilai eigen dari 1 sampai m :

$$g[m] = \sum_{q=1}^m \Lambda [q,q] \text{ untuk } m = 1,2,\dots,M. \quad (\text{A.19})$$

8. Pilih subset dari vektor-vektor eigen sebagai vektor-vektor basis

Simpan L kolom pertama dari V sebagai matriks $V_L [M \times L]$, seperti ditulis berikut ini:

$$V_L[p,q] = V[p,q] \text{ untuk } p=1,2,\dots,M; q=1,2,\dots,L; 1 \leq L \leq M. \quad (\text{A.20})$$

Gunakan vektor g sebagai suatu petunjuk dalam memilih nilai untuk L . Tujuannya adalah memilih suatu nilai L sekecil mungkin dengan mendapatkan nilai yang tinggi dari g dalam persentase. Misalnya nilai L dipilih sehingga energi kumulatif g di atas suatu batas, seperti 90%. Dalam hal ini, pilih nilai terkecil dari L sehingga:

$$(g[m=L]) / \left(\sum_{q=1}^M \Lambda[q,q] \right) \geq 90\%. \quad (\text{A.21})$$

9. Proyeksikan matriks data D pada basis yang baru

Vektor-vektor yang telah terproyeksi adalah kolom-kolom dari matriks:

$$W = V_L^* \cdot D = \text{KLT}\{X\} \quad (\text{A.22})$$

dimana V_L^* adalah *conjugate transpose* dari matriks eigen. Kolom-kolom dari matriks W merepresentasikan Transformasi Karhunen-Loève (KLT) dari vektor-vektor data dalam kolom-kolom matriks X .

10. Hitung matriks data terpusat dengan dimensi yang telah tereduksi, \hat{D}

$$\hat{D} = V_L W. \quad (\text{A.23})$$

11. Hitung matriks data asli dengan dimensi yang telah tereduksi, \hat{X}

$$\hat{X} = \hat{D} + \mu y = V_L W + \mu y. \quad (\text{A.24})$$

MULTIPLE LINEAR REGRESSION

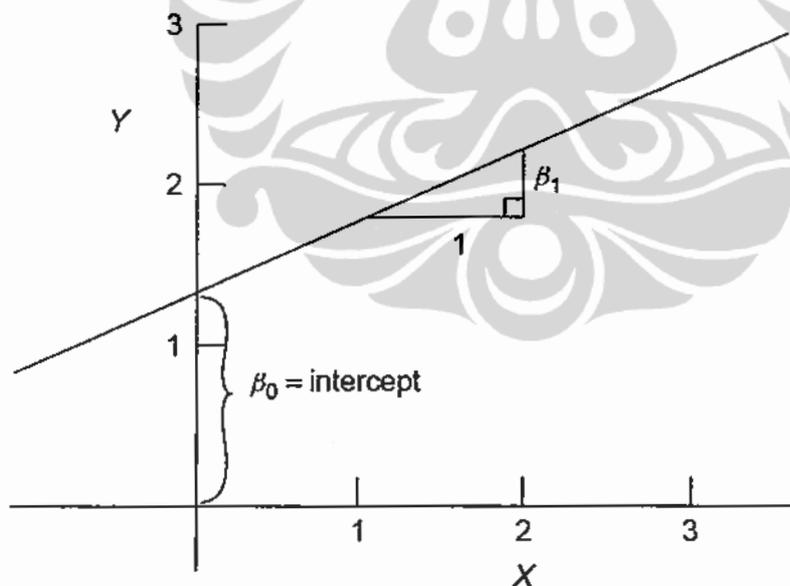
B.1 Multiple Linear Regression

Multiple linear regression (MLR) merupakan generalisasi dari model regresi linier sederhana dengan membolehkan banyak suku dalam suatu fungsi *mean* daripada hanya satu *intercept* dan satu *slope*. Model regresi linier sederhana terdiri dari fungsi *mean* dan fungsi *variance* sebagai berikut:

$$E(Y | X_1 = x_1) = \beta_0 + \beta_1 x_1, \quad (\text{B.1})$$

$$\text{Var}(Y | X_1 = x_1) = \sigma^2. \quad (\text{B.2})$$

Parameter-parameter dalam fungsi *mean* adalah *intercept* β_0 , nilai $E(Y|X_1=x_1)$ ketika $x_1=0$, dan *slope* β_1 , laju perubahan $E(Y|X_1=x_1)$ untuk setiap satuan perubahan dalam X ; seperti ditunjukkan Gambar B.1 berikut ini. Dengan mengubah-ubah kedua parameter tersebut, seluruh kemungkinan garis lurus dapat diperoleh. Dalam sebagian besar aplikasi, parameter-parameter tidak diketahui dan harus diestimasi menggunakan data. Fungsi *variance* dalam persamaan (B.2) dianggap konstan sebesar nilai positif σ^2 , yang biasanya tidak diketahui.



Gambar B.1. Persamaan Garis Lurus $E(Y | X = x) = \beta_0 + \beta_1 x$ [54]

Misalkan sekarang terdapat variabel kedua X_2 untuk memprediksi respon yang dikehendaki. Dengan penambahan X_2 ke dalam permasalahan, fungsi *mean* menjadi tergantung dari nilai X_1 dan nilai X_2 .

$$E(Y | X_1 = x_1, X_2 = x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \quad (\text{B.3})$$

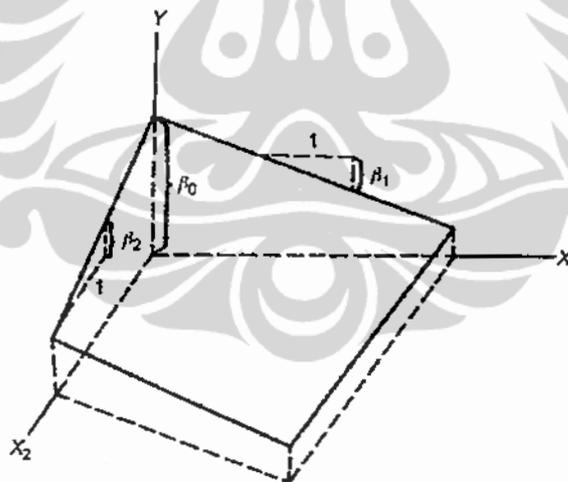
Ide dasar dari penambahan X_2 adalah untuk menjelaskan bagian dari Y yang belum dijelaskan oleh X_1 . Bentuk umum dari model MLR dengan respon Y dan suku-suku X_1, X_2, \dots, X_p dapat ditulis sebagai berikut:

$$E(Y|X) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p. \quad (\text{B.4})$$

Simbol X dalam $E(Y|X)$ berarti bahwa pusat perhatian terletak pada semua suku di sebelah kanan tanda sama dengan. Demikian pula ketika diperhatikan nilai-nilai tertentu untuk prediktor-prediktor x_1, x_2, \dots, x_p , yang disebut secara bersama \mathbf{x} , maka dapat ditulis persamaan berikut:

$$E(Y|X=\mathbf{x}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p. \quad (\text{B.5})$$

Nilai-nilai $\beta_i, i = 0, 1, \dots, p$, adalah parameter-parameter yang tidak diketahui dan perlu diestimasi. Persamaan (B.4) merupakan suatu fungsi linier dari parameter-parameter tersebut. Jika $p = 1$, X hanya memiliki satu elemen sehingga disebut regresi linier sederhana. Tetapi ketika $p = 2$, maka fungsi rata-rata dari persamaan (B.4) berkaitan dengan suatu bidang dimensi-3, seperti ditunjukkan oleh Gambar B.2 berikut ini.



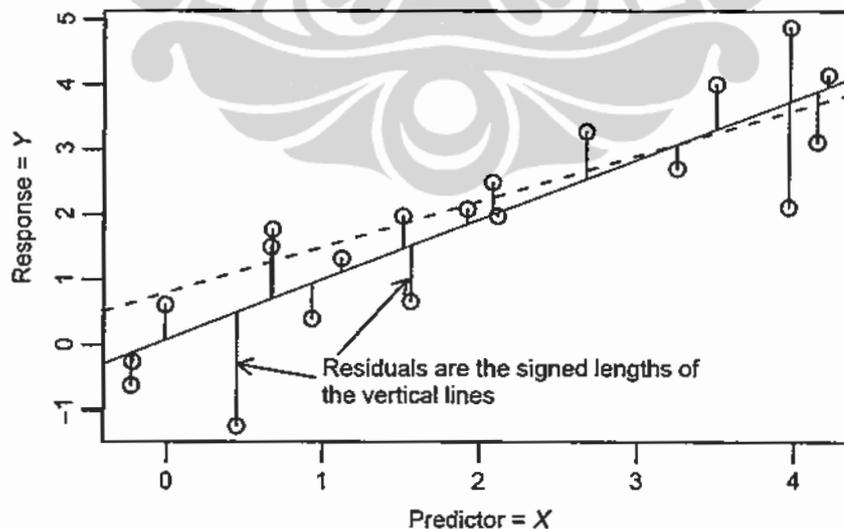
Gambar B.2. Suatu Permukaan Regresi Linier dengan $p = 2$ Prediktor [54]

Dalam hal $p > 2$, maka fungsi rata-rata yang dihasilkan adalah suatu *hyperplane*, suatu generalisasi dari bidang dimensi- p dalam ruang spatial dimensi- $(p+1)$. Suatu bidang dimensi- p tidak dapat digambarkan dalam dunia yang berdimensi-tiga.

B.2 Estimasi *Ordinary Least Squares*

Banyak metode telah diusulkan untuk mendapatkan estimasi-estimasi dari parameter-parameter dalam suatu model. Metode yang dibahas di sini disebut dengan *Ordinary Least Squares* (OLS), dimana estimasi-estimasi parameter dipilih untuk meminimalisasi suatu kuantitas yang disebut *residual sum of squares* (RSS). Parameter-parameter adalah kuantitas-kuantitas yang tidak diketahui yang memberikan ciri kepada suatu model. Estimasi-estimasi dari parameter-parameter adalah fungsi-fungsi data yang dapat dihitung dan karena itu bersifat statistik. Agar tidak terjadi kebingungan, maka parameter-parameter disimbolkan dengan huruf-huruf Yunani, seperti α , β , γ , dan σ , dan estimasi-estimasi dari parameter-parameter dilambangkan dengan menambahkan suatu 'topi' di atas huruf Yunani yang bersangkutan, misalnya $\hat{\beta}_1$ adalah estimator dari β_1 dan $\hat{\sigma}^2$ adalah estimator dari σ^2 .

Fungsi kriteria untuk mendapatkan estimator-estimator didasarkan pada residual-residual, yang secara geometris adalah jarak-jarak vertikal antara *fitted line* dan nilai-nilai y yang sebenarnya, seperti diilustrasikan oleh Gambar B.3. Residual-residual mencerminkan asimetri inheren antara peranan respon dan prediktor dalam masalah-masalah regresi.



Gambar B.3. Suatu Plot Skematik untuk *OLS Fitting* [54]

Dari kelompok awal dari prediktor-prediktor potensial, telah dihitung sekumpulan $p+1$ suku, termasuk satu *intercept*, $X = (X_0, X_1, \dots, X_p)$. Fungsi *mean*

untuk MLR diberikan oleh persamaan (B.4), dimana parameter-parameter β adalah parameter-parameter yang tidak diketahui dan perlu diestimasi. Misalkan terdapat data untuk n kasus atau satuan, yang berarti suatu nilai dari Y dan semua suku untuk setiap n kasus tersedia. Simbol-simbol untuk respon dan suku-suku dapat dituliskan dalam bentuk matriks-matriks dan vektor-vektor. Didefinisikan bahwa:

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ \vdots \\ y_n \end{pmatrix}; \quad X = \begin{pmatrix} 1 & x_{11} & \dots & x_{1p} \\ 1 & x_{21} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & \dots & x_{np} \end{pmatrix}$$

jadi Y adalah suatu vektor $n \times 1$ dan X adalah suatu matriks $n \times (p+1)$. Didefinisikan juga β adalah suatu vektor $(p+1) \times 1$ yang terdiri dari koefisien-koefisien regresi dan e adalah suatu vektor $n \times 1$ yang terdiri dari kesalahan-kesalahan statistik,

$$\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \vdots \\ \beta_p \end{pmatrix}; \quad e = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ \vdots \\ e_n \end{pmatrix}$$

Matriks X memberikan semua nilai-nilai observasi dari suku-suku. Baris yang ke- i dari X didefinisikan oleh simbol \mathbf{x}_i^T , yang adalah suatu vektor $(p+1) \times 1$ untuk fungsi-fungsi *mean* termasuk *intercept*. Walaupun \mathbf{x}_i adalah suatu baris dari X , tetapi di subbab ini digunakan konvensi bahwa semua vektor adalah vektor-vektor kolom dan karena itu perlu dituliskan \mathbf{x}_i^T untuk merepresentasikan suatu vektor baris. Persamaan untuk fungsi *mean* yang dievaluasi pada \mathbf{x}_i adalah:

$$E(Y|X=\mathbf{x}_i) = \mathbf{x}_i^T \beta = \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip}. \quad (\text{B.6})$$

Dalam notasi matrix, model MLR dapat dituliskan sebagai berikut:

$$Y = X \beta + e \quad (\text{B.7})$$

dimana baris ke- i dari persamaan (B.7) adalah:

$$y_i = \mathbf{x}_i^T \beta + e_i. \quad (\text{B.8})$$

Estimasi $\hat{\beta}$ dari β dipilih agar fungsi *residual sum of squares* berikut ini minimal:

$$\text{RSS}(\beta) = \sum (y_i - x_i^T \beta)^2 = (\mathbf{Y} - \mathbf{X} \beta)^T (\mathbf{Y} - \mathbf{X} \beta). \quad (\text{B.9})$$

Estimasi-estimasi OLS dapat dihitung dari persamaan (B.9) dengan diferensiasi dalam suatu matriks. Estimasi OLS diberikan oleh persamaan berikut ini:

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y} \quad (\text{B.10})$$

Dengan syarat terdapat invers $(\mathbf{X}^T \mathbf{X})^{-1}$. Estimator $\hat{\beta}$ hanya tergantung dari $\mathbf{X}^T \mathbf{X}$ dan $\mathbf{X}^T \mathbf{Y}$, yaitu matriks-matriks *sum of squares* dan *cross-products*.



PERHITUNGAN HRIR FASA MINIMUM

HRTF merupakan keluaran dari suatu sistem LTI, yakni difraksi dan refleksi-refleksi dari kepala manusia, daun telinga, dan *torso*, sebagai respon terhadap suatu stimulus medan bebas (*free-field stimulus*) yang terjadi pada suatu titik dalam kanal telinga. Jadi, secara teoritis, HRIRs dapat langsung direpresentasikan sebagai filter-filter *finite impulse response* (FIR) yang non rekursif.

Respon frekuensi dari suatu HRTF dapat diuraikan menjadi respon *magnitude* dan respon fasa sebagai berikut:

$$H(e^{j\omega}) = |H(e^{j\omega})| e^{j \arg\{H(e^{j\omega})\}} \quad (C.1)$$

dimana $|H(e^{j\omega})|$ adalah respon *magnitude* HRTF dan $\arg\{H(e^{j\omega})\}$ adalah respon fasa HRTF. Kemudian, suatu fungsi sistem LTI dapat lebih jauh dinyatakan sebagai suatu kaskade dari suatu sistem fasa minimum dan suatu sistem fasa *allpass* sebagai berikut [55]:

$$H(e^{j\omega}) = H_{mp}(e^{j\omega}) H_{ap}(e^{j\omega}), \quad (C.2)$$

$$H(e^{j\omega}) = |H_{mp}(e^{j\omega})| e^{j(\arg\{H_{mp}(e^{j\omega})\})} |H_{ap}(e^{j\omega})| e^{j(\arg\{H_{ap}(e^{j\omega})\})}, \quad (C.3)$$

karena $|H_{mp}(e^{j\omega})| = |H(e^{j\omega})|$ dan $|H_{ap}(e^{j\omega})| = 1$ sepanjang spectrum frekuensi, maka:

$$H(e^{j\omega}) = |H(e^{j\omega})| e^{j(\arg\{H_{ap}(e^{j\omega})\} + \arg\{H_{mp}(e^{j\omega})\})} \quad (C.4)$$

dimana $H_{ap}(e^{j\omega})$ adalah HRTF fasa *allpass* dan $H_{mp}(e^{j\omega})$ adalah HRTF fasa minimum.

Setiap HRIR di dalam basis data HRTF CIPIC dari jarak satu meter dari sumber bunyi menuju pusat kepala subyek. Jika dilihat dari grafik HRIR, terdapat waktu tunda yang berkaitan dengan jarak satu meter tadi yang diperlukan oleh sinyal bunyi untuk merambat dari sumbernya ke gendang telinga subyek sebelum terjadi amplitudo maksimum dari HRIR. Untuk menghilangkan waktu tunda ini, HRIR dapat direkonstruksi menjadi suatu deret yang berfasa minimum dengan menggunakan transformasi Hilbert.

Suatu filter di dalam kawasan frekuensi linier dan tidak berubah terhadap waktu (*LTI*), $H(z) = B(z)/A(z)$, dikatakan memiliki fasa minimum jika semua kutub (*pole*) dan nol (*zero*) terletak di dalam lingkaran satuan $|z|=1$ pada bidang- z .

Demikian juga dikatakan suatu filter $H(z)$ memiliki fasa minimum jika filter tersebut dan inversnya $1/H(z)$ stabil. Suatu filter dengan fasa minimum juga bersifat kausal karena suku-suku non kausal dalam respon impuls terkait dengan kutub-kutub pada tak berhingga.

Contoh yang paling sederhana dari filter fasa minimum adalah filter $H(z)=z$, yang terdiri dari sebuah nol pada $z = 0$ dan suatu kutub pada $z = \infty$. Filter yang berfasa minimum memiliki pembilang dan penyebut dari fungsi alihnya yang merupakan polinomial fasa minimum dalam z^{-1} . Suatu polinomial dalam bentuk:

$$\begin{aligned} B(z) &= b_0 + b_1 z^{-1} + b_2 z^{-2} + \dots + b_M z^{-M} \\ &= b_0 (1 - \zeta_1 z^{-1}) (1 - \zeta_2 z^{-1}) \dots (1 - \zeta_M z^{-1}) \end{aligned} \quad (C.5)$$

dikatakan berfasa minimum jika semua akarnya ζ_i , $i=1,2,\dots,M$, terletak di dalam lingkaran satuan, yaitu $|\zeta_i| < 1$. Suatu sifat umum dari deret-deret yang berfasa minimum adalah bahwa di antara semua respon impuls $h_i(n)$ yang mempunyai spektrum absolut yang identik, respon impuls yang berfasa minimum mengalami peluruhan yang tercepat berdasarkan persamaan berikut:

$$\sum_{n=0}^K |h_{mp}(n)|^2 \geq \sum_{n=0}^K |h_i(n)|^2, \quad n=0, 1, 2, \dots, K. \quad (C.6)$$

Persamaan di atas menyatakan bahwa energi dalam $K+1$ sampel yang pertama dari respon impuls fasa minimum paling sedikit sama besar dengan suatu respon impuls kausal yang memiliki spektrum absolut yang sama. Jadi respon impuls fasa minimum terkonsentrasi secara maksimal di sekitar waktu nol di antara ruang respon impuls kausal untuk suatu spektrum absolut yang diberikan. Oleh karena sifat ini, maka respon impuls fasa minimum kadang kala disebut respon impuls dengan waktu tunda minimum.

Dalam filter fasa minimum $H(z) = e^{a(z)} e^{i b(z)}$ juga berlaku hubungan $b(z) = -\mathcal{H}\{a(z)\}$ dan $a(z) = -\mathcal{H}\{b(z)\}$, dimana $\mathcal{H}\{\}$ adalah transformasi Hilbert dari suatu filter. Perubahan algoritmis dari hubungan ini untuk sistem-sistem waktu diskrit diperoleh terutama melalui perhitungan *real cepstrum* seperti dijelaskan berikut ini.

Dari suatu respon impuls, $h(n)$, dengan fasa sembarang dapat diperoleh suatu respon impuls fasa minimum, $h_{mp}(n)$ tetapi memiliki spektrum *magnitude*

yang sama. Pertama-tama dihitung suatu deret kawasan waktu $c(n)$ sebagai berikut:

$$c(n) = \text{Re}\{F_D^{-1}[\ln|F_D\{x(n)\}|]\} \quad (\text{C.7})$$

dimana \ln adalah logaritmik natural, $F_D\{\}$ dan $F_D^{-1}\{\}$ adalah DFT dan inversnya.

Deret ini lalu diboboti dengan fungsi jendela $\text{win}(n)$ berikut:

$$\text{win}(n) = \begin{cases} 0 & \text{jika } n < 0 \\ 1 & \text{jika } n = 0 \\ 2 & \text{jika } n > 0 \end{cases} \quad (\text{C.8})$$

Jika ditinjau dari segi transformasi-z rasional dari $h(n)$, $H(z)$, fungsi jendela ini dapat dilihat sebagai inversi kompleks konjugasi dari posisi-posisi nol (*zeros*) yang terletak di luar lingkaran satuan, yang menghasilkan respon impuls fasa minimum. Dengan demikian respon impuls fasa minimum, $h_{mp}(n)$, yang diinginkan dapat dihasilkan melalui persamaan berikut:

$$h_{mp}(n) = F_D^{-1}\{\exp(F_D[\text{win}(n).c(n)])\} \quad (\text{C.9})$$

MODEL-MODEL HRIR FASA MINIMUM INDIVIDUAL SUBYEK 003 DI BIDANG HORIZONTAL

Dalam lampiran ini, ditampilkan 25 model HRIR fasa minimum ($HRIR_{mp}$) individual dari telinga kiri dan telinga kanan subyek 003, yang dihasilkan dari individualisasi model HRIR fasa minimum seperti dijelaskan dalam subbab 5.2, untuk sumber-sumber bunyi di bidang horisontal. Untuk suatu telinga dari seorang subyek, dihasilkan keseluruhan 50 model HRIR fasa minimum individual di bidang horisontal sebagai fungsi dari azimut.

Dalam satu koordinat, $HRIR_{mp}$ asli ditunjukkan oleh garis (*solid*) dan model $HRIR_{mp}$ individual ditunjukkan oleh garis putus-putus (*dashed*) pada Gambar D1 sampai Gambar D5. Respon-respon impuls telinga kiri ditunjukkan pada panel-panel kiri dan respon-respon impuls telinga kanan ditunjukkan pada panel-panel kanan, dalam semua gambar tersebut.

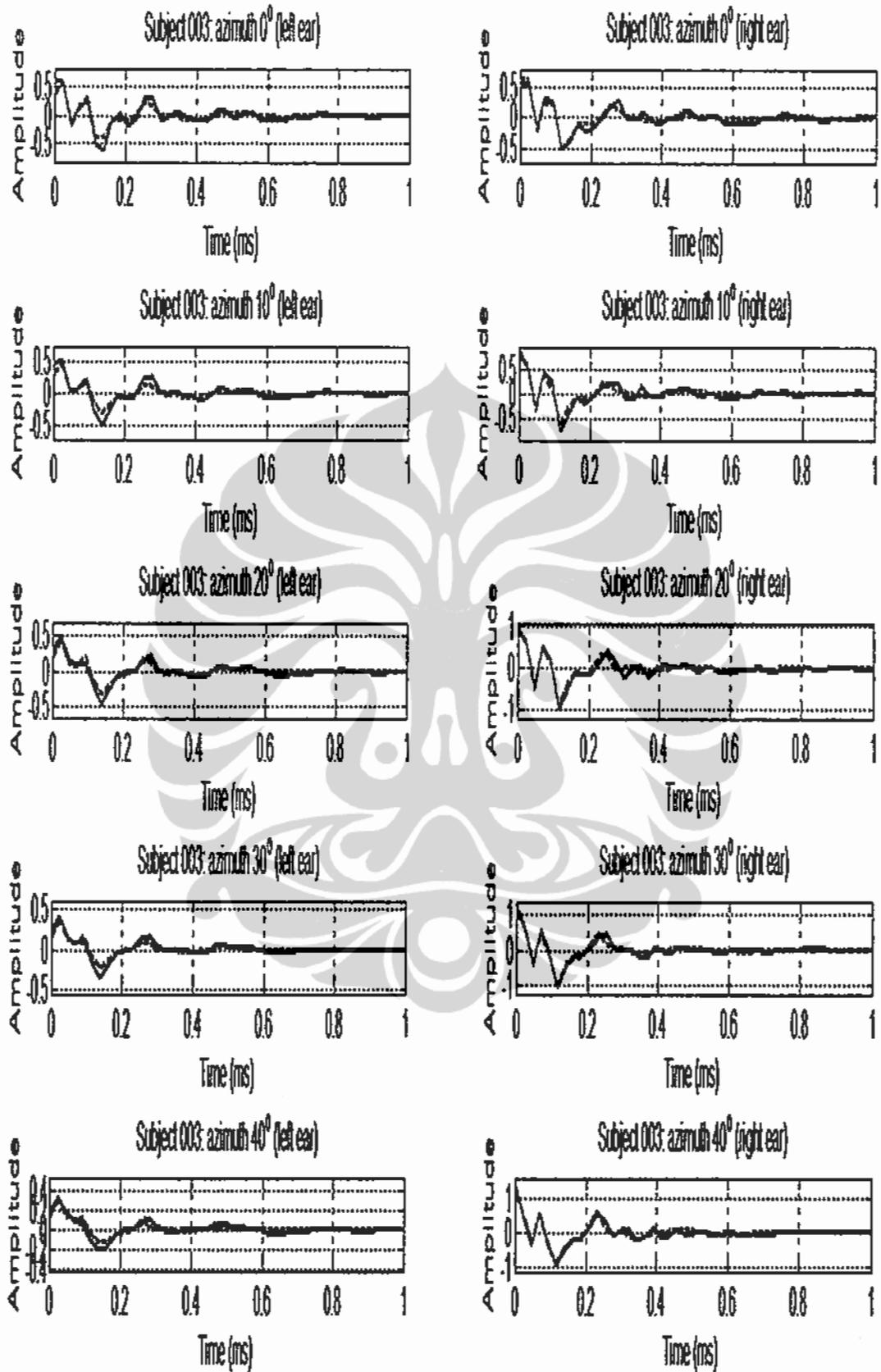
Gambar D1 menunjukkan $HRIR_{mp}$ asli dan model $HRIR_{mp}$ individual dari sumber bunyi di bidang horisontal dengan azimut 0° , 10° , 20° , 30° , 40° .

Gambar D2 menunjukkan $HRIR_{mp}$ asli dan model $HRIR_{mp}$ individual dari sumber bunyi di bidang horisontal dengan azimut 55° , 80° , 100° , 125° , 140° .

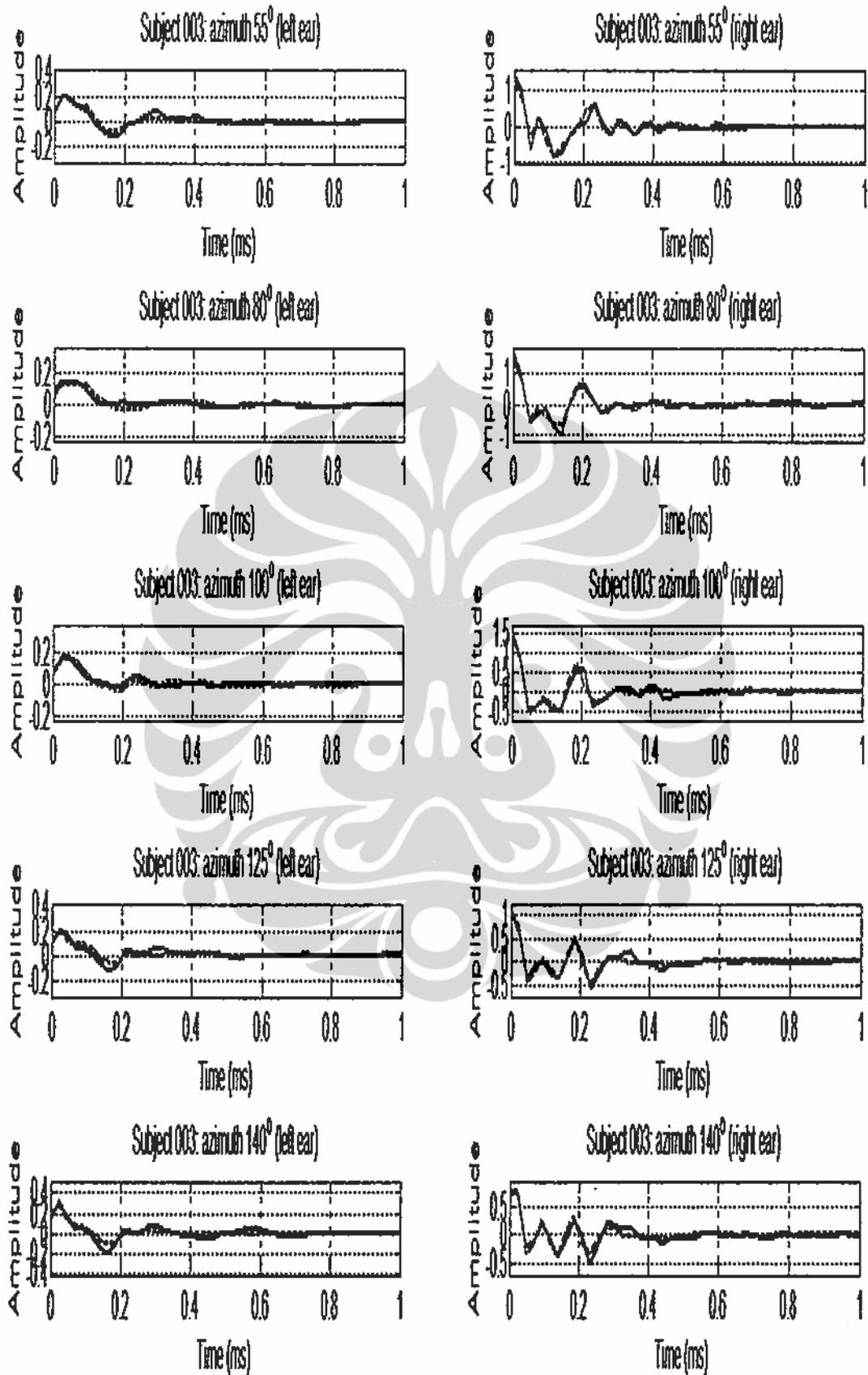
Gambar D3 menunjukkan $HRIR_{mp}$ asli dan model $HRIR_{mp}$ individual dari sumber bunyi di bidang horisontal dengan azimut 150° , 160° , 170° , 180° , 190° .

Gambar D4 menunjukkan $HRIR_{mp}$ asli dan model $HRIR_{mp}$ individual dari sumber bunyi di bidang horisontal dengan azimut 200° , 210° , 220° , 235° , 260° .

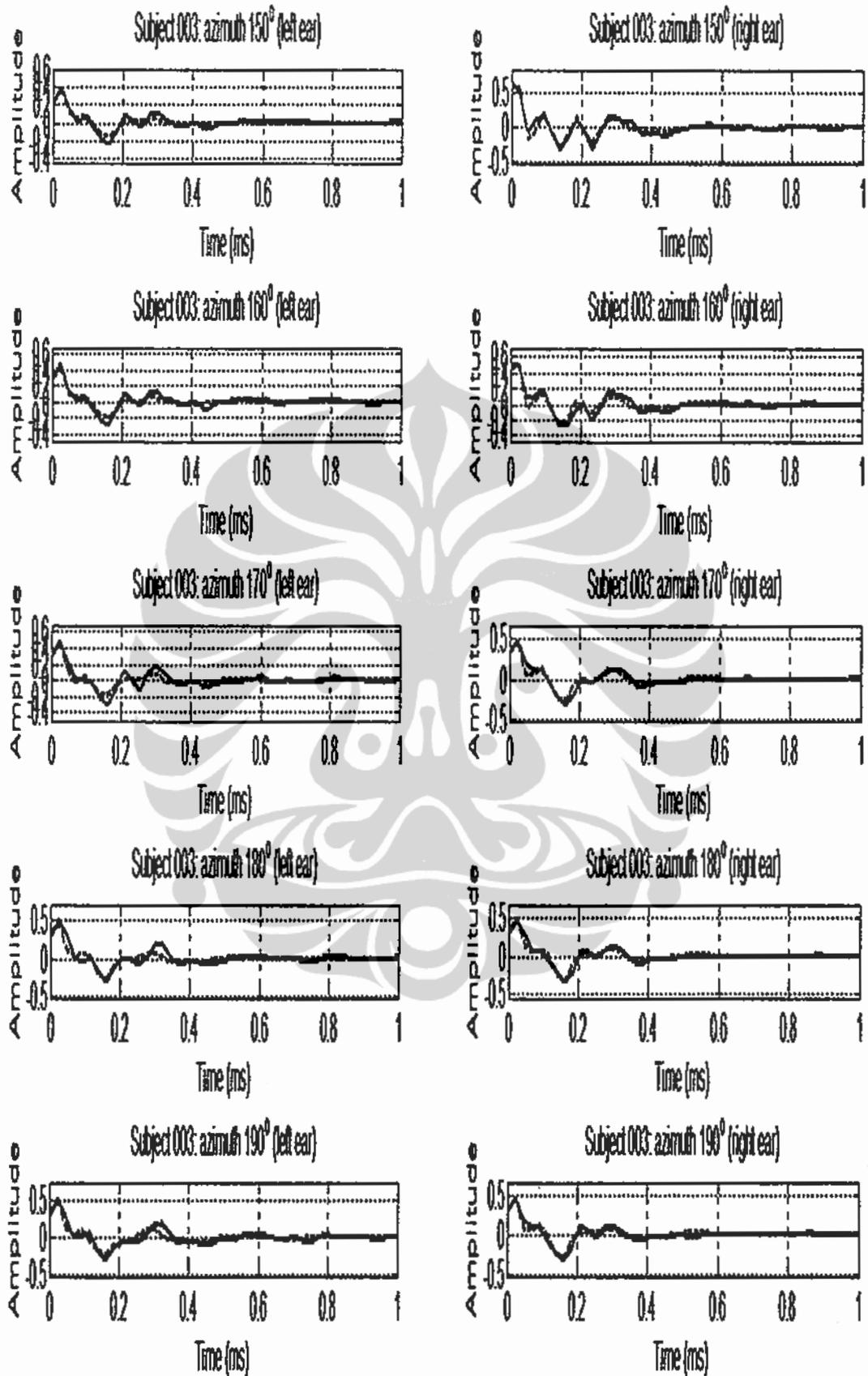
Gambar D5 menunjukkan $HRIR_{mp}$ asli dan model $HRIR_{mp}$ individual dari sumber bunyi di bidang horisontal dengan azimut 280° , 305° , 320° , 330° , 340° .



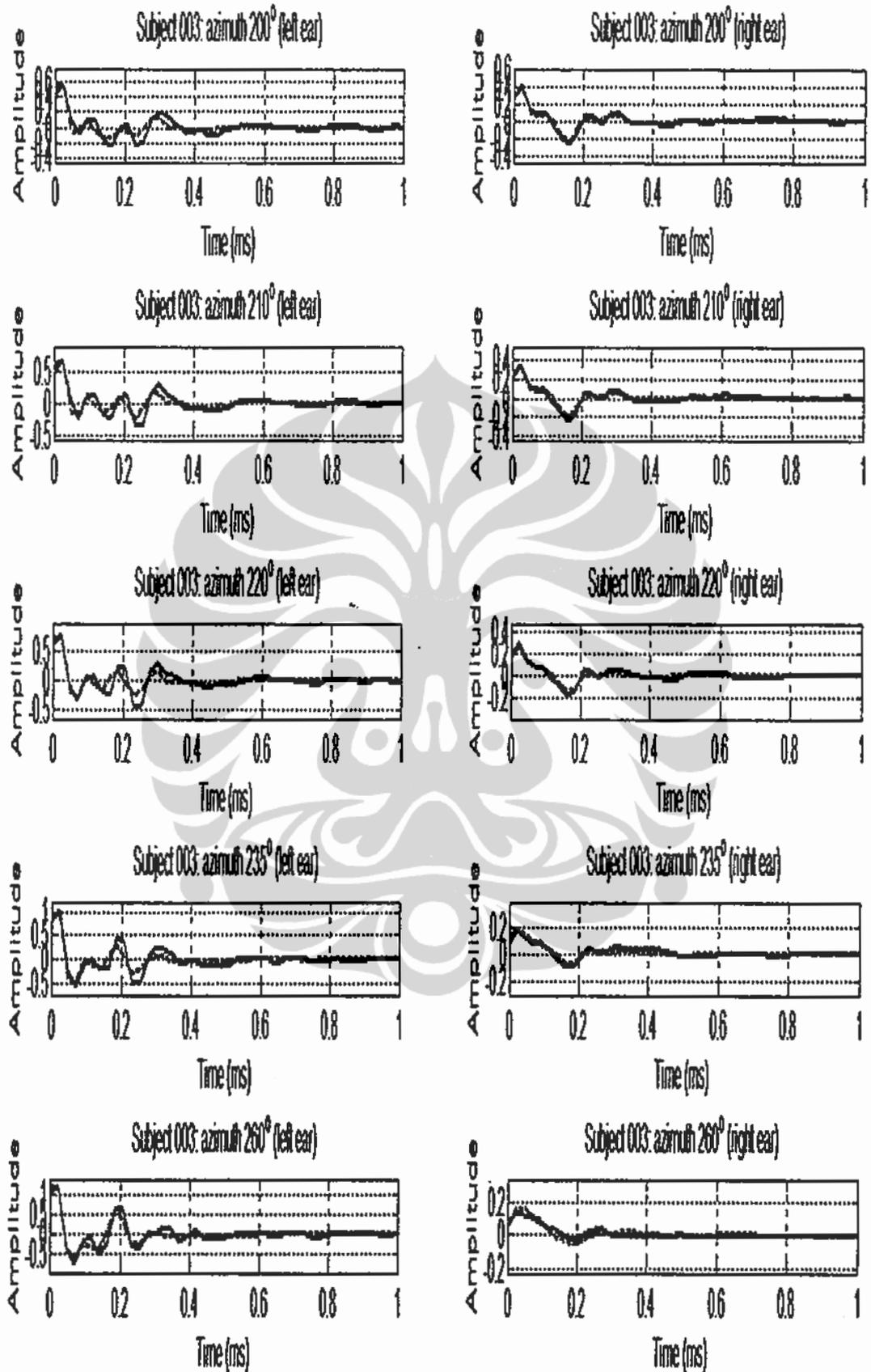
Gambar D.1. Model $HRIR_{mp}$ Individual (*dashed*) dan $HRIR_{mp}$ Asli (*solid*) dengan Azimut 0° , 10° , 20° , 30° , 40° di Bidang Horizontal



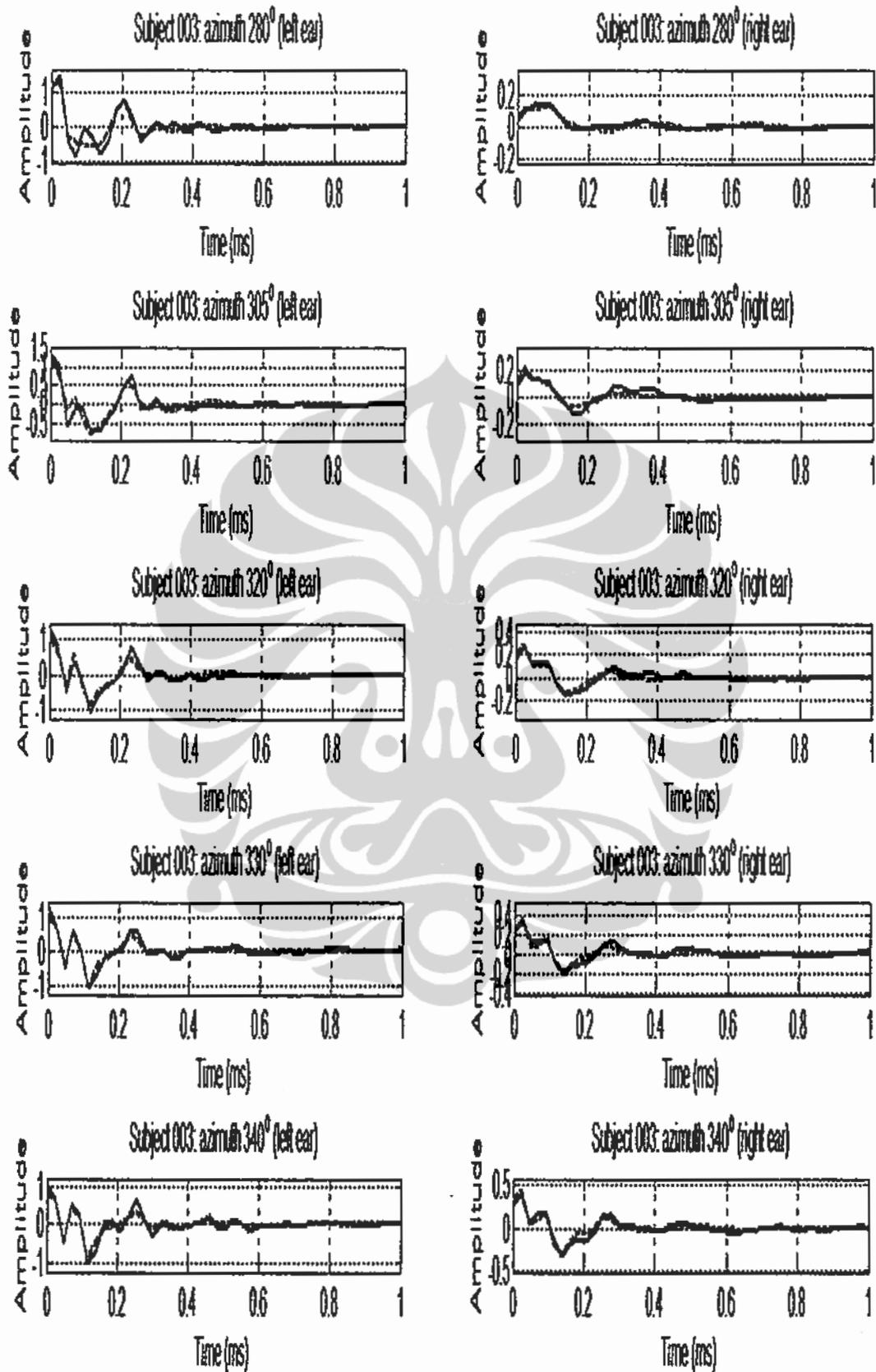
Gambar D.2. Model $HRIR_{mp}$ Individual (*dashed*) dan $HRIR_{mp}$ Asli (*solid*) dengan Azimut 55° , 80° , 100° , 125° , 140° di Bidang Horizontal



Gambar D.3. Model $HRIR_{mp}$ Individual (*dashed*) dan $HRIR_{mp}$ Asli (*solid*) dengan Azimuth 150° , 160° , 170° , 180° , 190° di Bidang Horizontal



Gambar D.4. Model $HRIR_{mp}$ Individual (*dashed*) dan $HRIR_{mp}$ Asli (*solid*) dengan Azimut 200° , 210° , 220° , 235° , 260° di Bidang Horizontal



Gambar D.5. Model $HRIR_{mp}$ Individual (*dashed*) dan $HRIR_{mp}$ Asli (*solid*) dengan Azimut 280°, 305°, 320°, 330°, 340° di Bidang Horizontal

MODEL-MODEL HRIR FASA MINIMUM INDIVIDUAL SUBYEK 003 DI BIDANG MEDIAN

Dalam lampiran ini, ditampilkan 25 model HRIR fasa minimum ($HRIR_{mp}$) individual dari telinga kiri dan telinga kanan subyek 003, yang dihasilkan dari individualisasi model HRIR fasa minimum seperti dijelaskan dalam subbab 5.2, untuk sumber-sumber bunyi di bidang median. Untuk suatu telinga dari seorang subyek, dihasilkan keseluruhan 50 model HRIR fasa minimum individual di bidang median sebagai fungsi dari elevasi.

Dalam satu koordinat, $HRIR_{mp}$ asli ditunjukkan oleh garis (*solid*) dan model $HRIR_{mp}$ individual ditunjukkan oleh garis putus-putus (*dashed*) pada Gambar E1 sampai Gambar E5. Respon-respon impuls telinga kiri ditunjukkan pada panel-panel kiri dan respon-respon impuls telinga kanan ditunjukkan pada panel-panel kanan, dalam semua gambar tersebut.

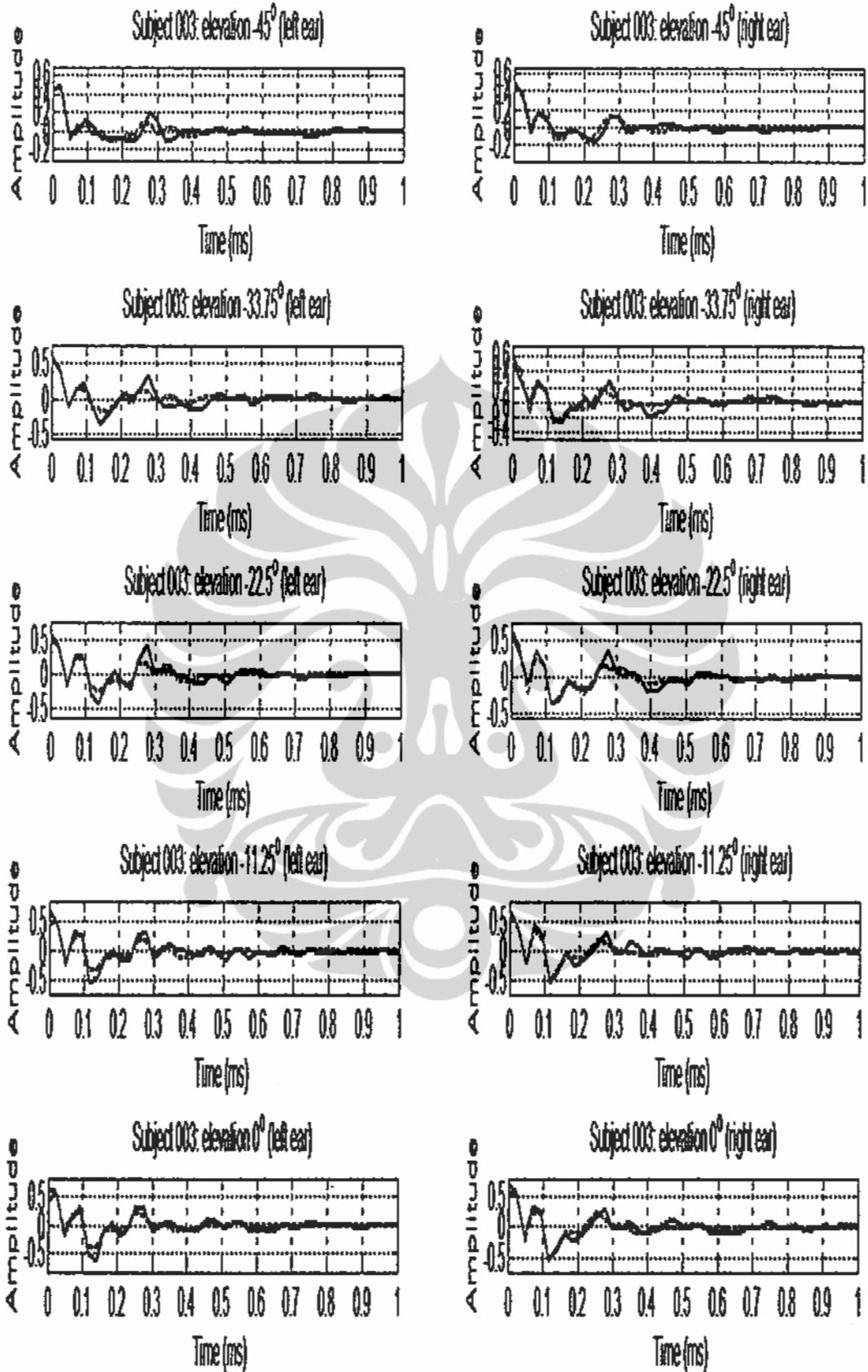
Gambar E1 menunjukkan $HRIR_{mp}$ asli dan model $HRIR_{mp}$ individual dari sumber bunyi di bidang median dengan elevasi -45° , $-33,75^\circ$, $-22,5^\circ$, $-11,25^\circ$, 0° .

Gambar E2 menunjukkan $HRIR_{mp}$ asli dan model $HRIR_{mp}$ individual dari sumber bunyi di bidang median dengan elevasi $11,25^\circ$, $22,5^\circ$, $33,75^\circ$, 45° , $56,25^\circ$.

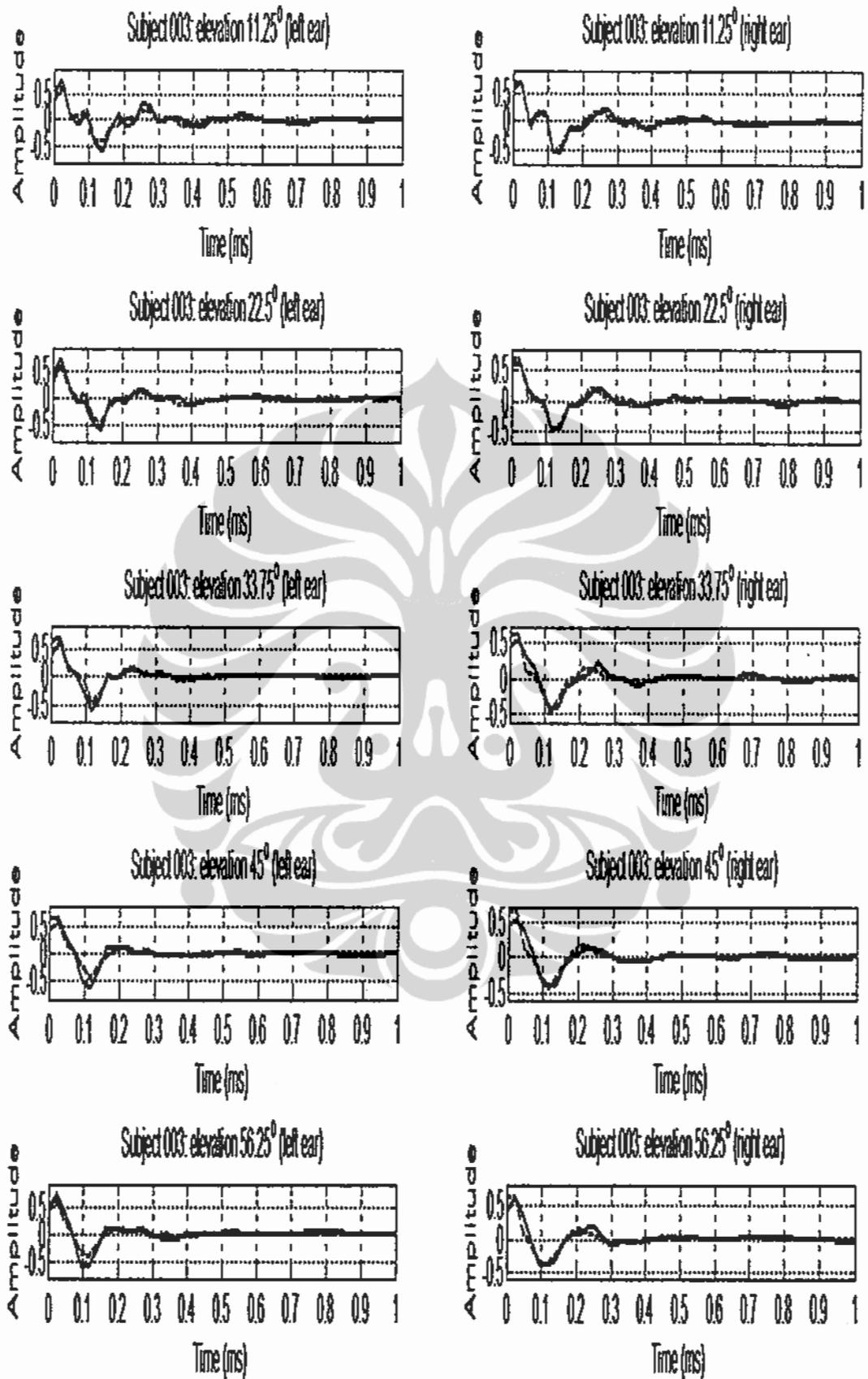
Gambar E3 menunjukkan $HRIR_{mp}$ asli dan model $HRIR_{mp}$ individual dari sumber bunyi di bidang median dengan elevasi $67,5^\circ$, $78,75^\circ$, 90° , $101,25^\circ$, $112,5^\circ$.

Gambar E4 menunjukkan $HRIR_{mp}$ asli dan model $HRIR_{mp}$ individual dari sumber bunyi di bidang median dengan elevasi $123,75^\circ$, 135° , $146,25^\circ$, $157,5^\circ$, $168,75^\circ$.

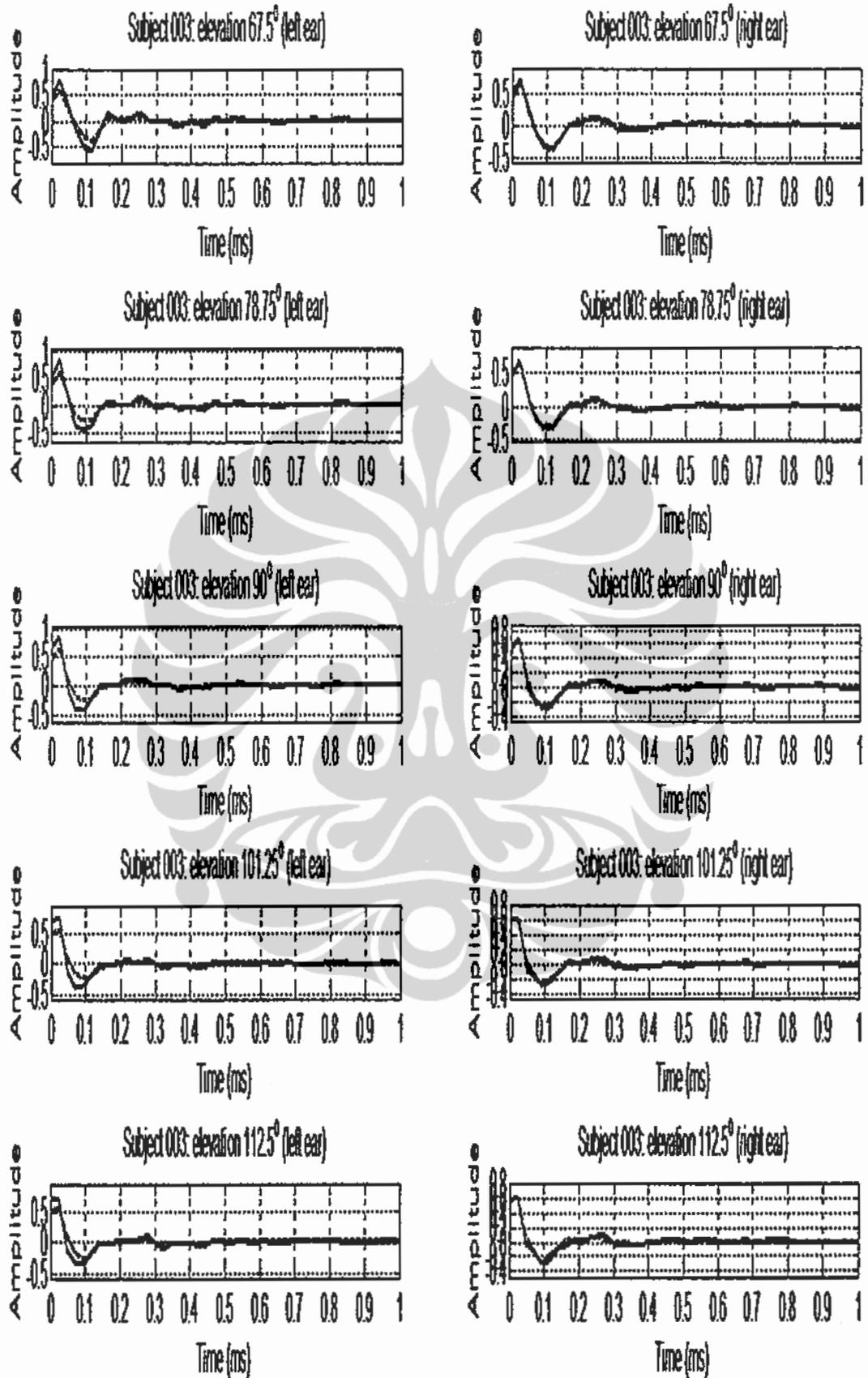
Gambar E5 menunjukkan $HRIR_{mp}$ asli dan model $HRIR_{mp}$ individual dari sumber bunyi di bidang median dengan elevasi 280° , 305° , 320° , 330° , 340° .



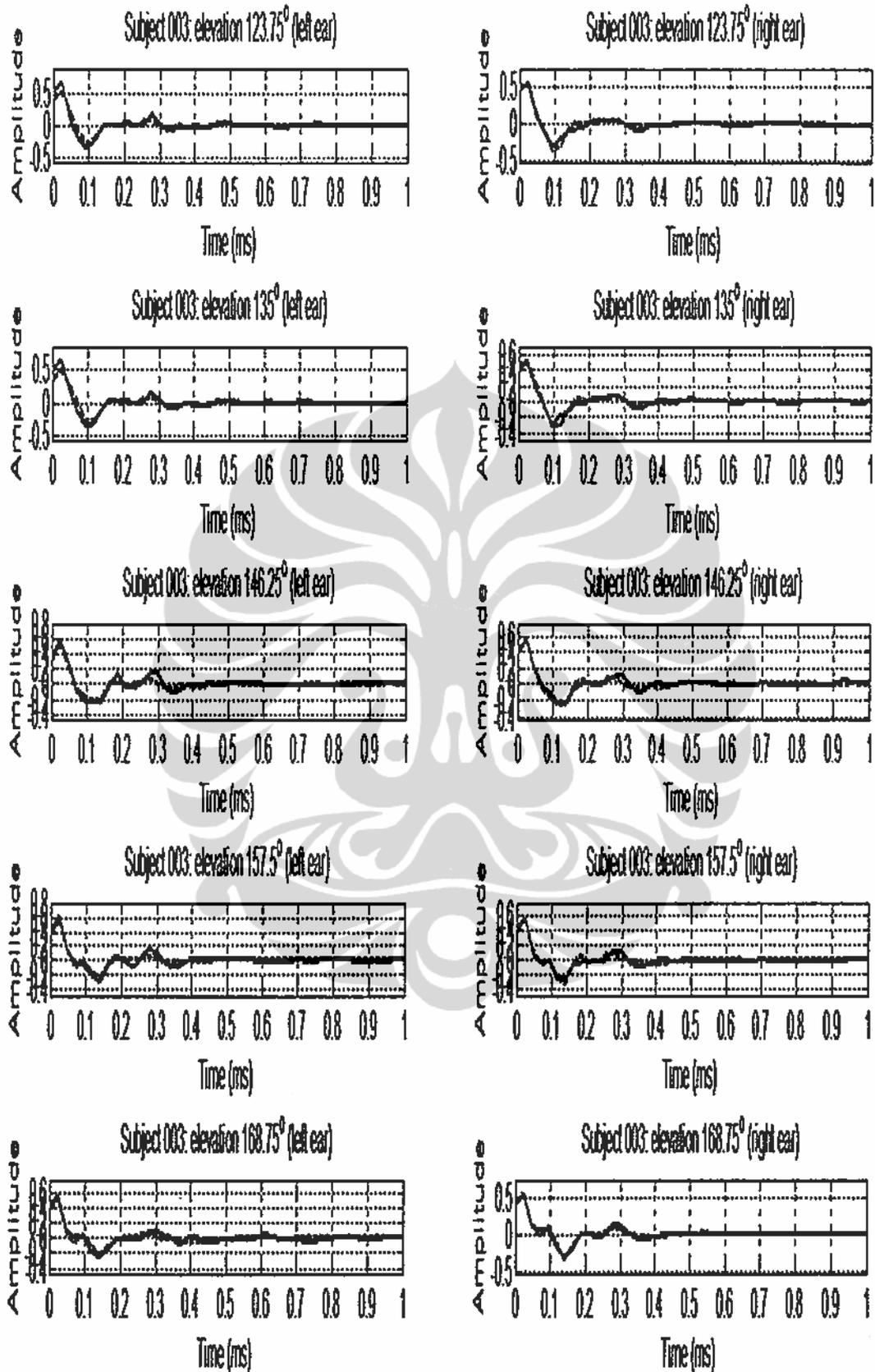
Gambar E.1. Model $HRIR_{mp}$ Individual (dashed) dan $HRIR_{mp}$ Asli (solid) dengan Elevasi -45° , $-33,75^\circ$, $-22,5^\circ$, $-11,25^\circ$, 0° di Bidang Median



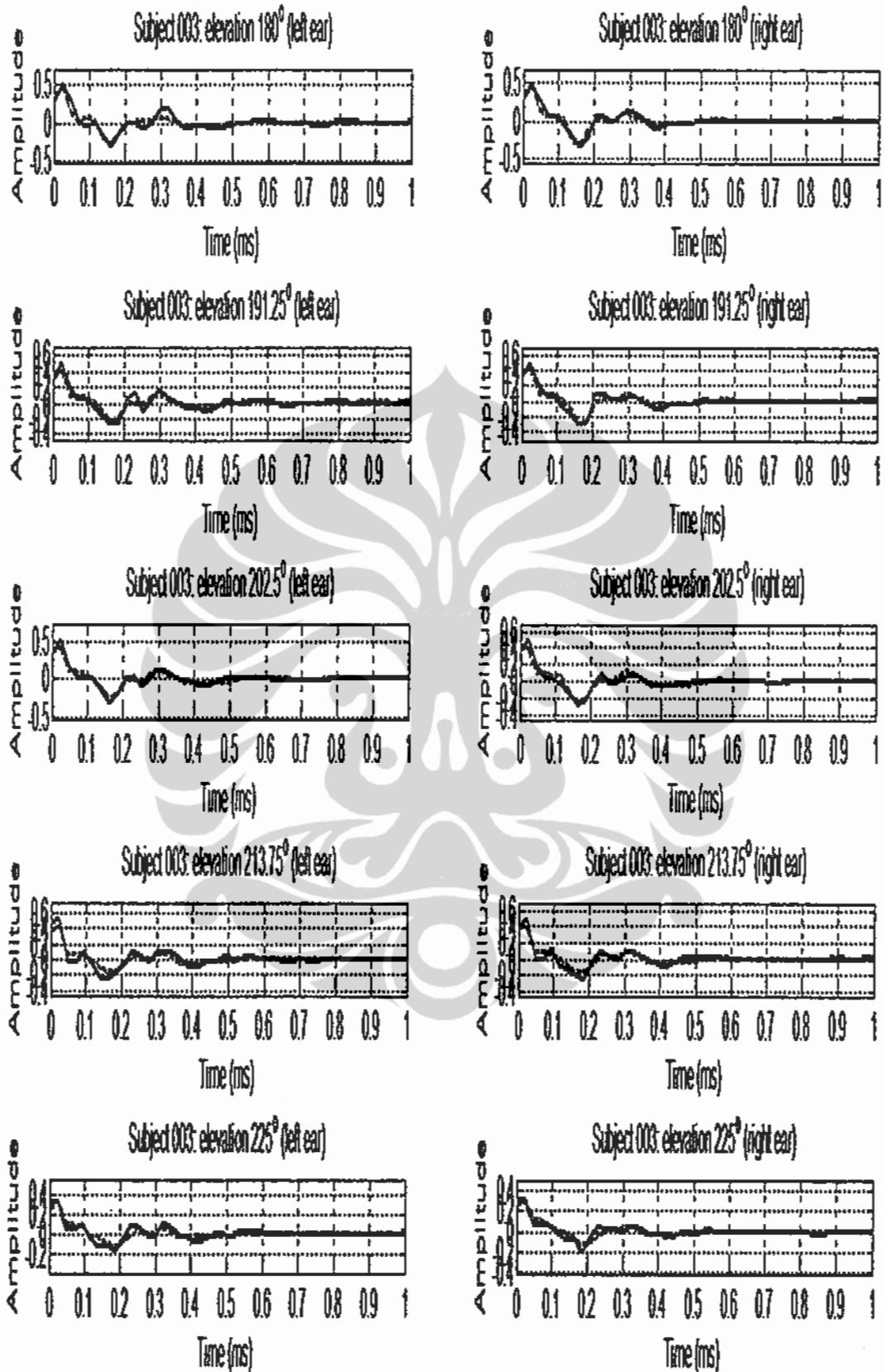
Gambar E.2. Model $HRIR_{mp}$ Individual (dashed) dan $HRIR_{mp}$ Asli (solid) dengan Elevasi $11,25^\circ$, $22,5^\circ$, $33,75^\circ$, 45° , $56,25^\circ$ di Bidang Median



Gambar E.3. Model $HRIR_{mp}$ Individual (*dashed*) dan $HRIR_{mp}$ Asli (*solid*) dengan Elevasi 67.5° , 78.75° , 90° , 101.25° , 112.5° di Bidang Median



Gambar E.4. Model $HRIR_{mp}$ Individual (*dashed*) dan $HRIR_{mp}$ Asli (*solid*) dengan Elevasi $123,75^\circ$, 135° , $146,25^\circ$, $157,5^\circ$, $168,75^\circ$ di Bidang Median



Gambar E.5. Model $HRIR_{mp}$ Individual (*dashed*) dan $HRIR_{mp}$ Asli (*solid*) dengan Elevasi 180°, 191,25°, 202,5°, 213,75°, 225° di Bidang Median

MODEL-MODEL HRTF *MAGNITUDE* INDIVIDUAL SUBYEK 003 DI BIDANG HORIZONTAL

Dalam lampiran ini, ditampilkan 25 model HRTF *magnitude* individual dari telinga kiri dan telinga kanan subyek 003, yang dihasilkan dari individualisasi model HRTF *magnitude* seperti dijelaskan dalam subbab 5.4, untuk sumber-sumber bunyi di bidang horisontal. Untuk suatu telinga dari seorang subyek, dihasilkan keseluruhan 50 model HRTF *magnitude* individual di bidang horisontal sebagai fungsi dari azimut.

Dalam satu koordinat, HRTF *magnitude* asli ditunjukkan oleh garis (*solid*) dan model HRTF *magnitude* individual ditunjukkan oleh garis putus-putus (*dashed*) pada Gambar F1 sampai Gambar F5. Spektrum-spektrum *magnitude* telinga kiri ditunjukkan pada panel-panel kiri dan spektrum-spektrum *magnitude* telinga kanan ditunjukkan pada panel-panel kanan, dalam semua gambar tersebut.

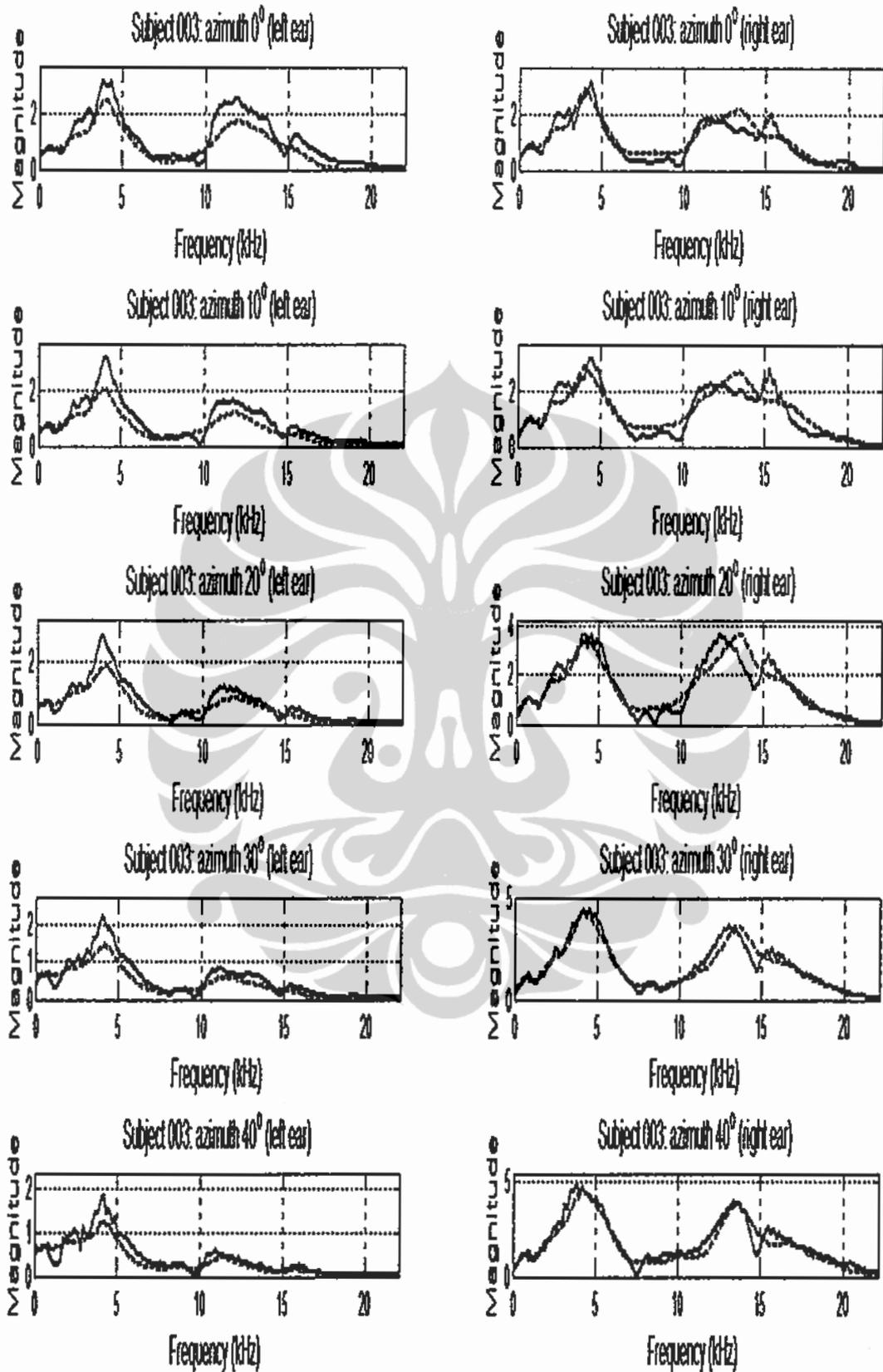
Gambar F1 menunjukkan HRTF *magnitude* asli dan model HRTF *magnitude* individual dari sumber bunyi di bidang horisontal dengan azimut 0° , 10° , 20° , 30° , 40° .

Gambar F2 menunjukkan HRTF *magnitude* asli dan model HRTF *magnitude* individual dari sumber bunyi di bidang horisontal dengan azimut 55° , 80° , 100° , 125° , 140° .

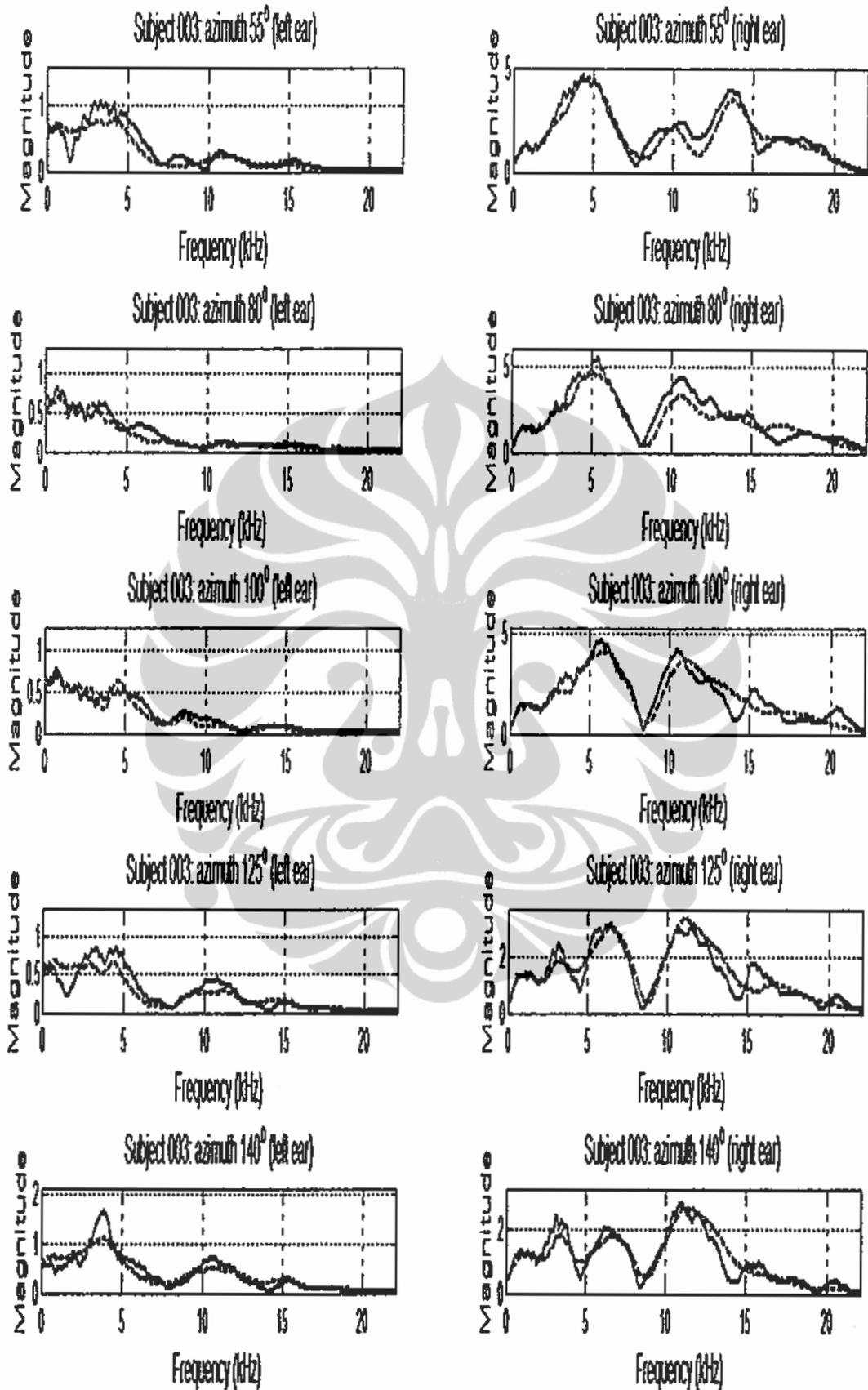
Gambar F3 menunjukkan HRTF *magnitude* asli dan model HRTF *magnitude* individual dari sumber bunyi di bidang horisontal dengan azimut 150° , 160° , 170° , 180° , 190° .

Gambar F4 menunjukkan HRTF *magnitude* asli dan model HRTF *magnitude* individual dari sumber bunyi di bidang horisontal dengan azimut 200° , 210° , 220° , 235° , 260° .

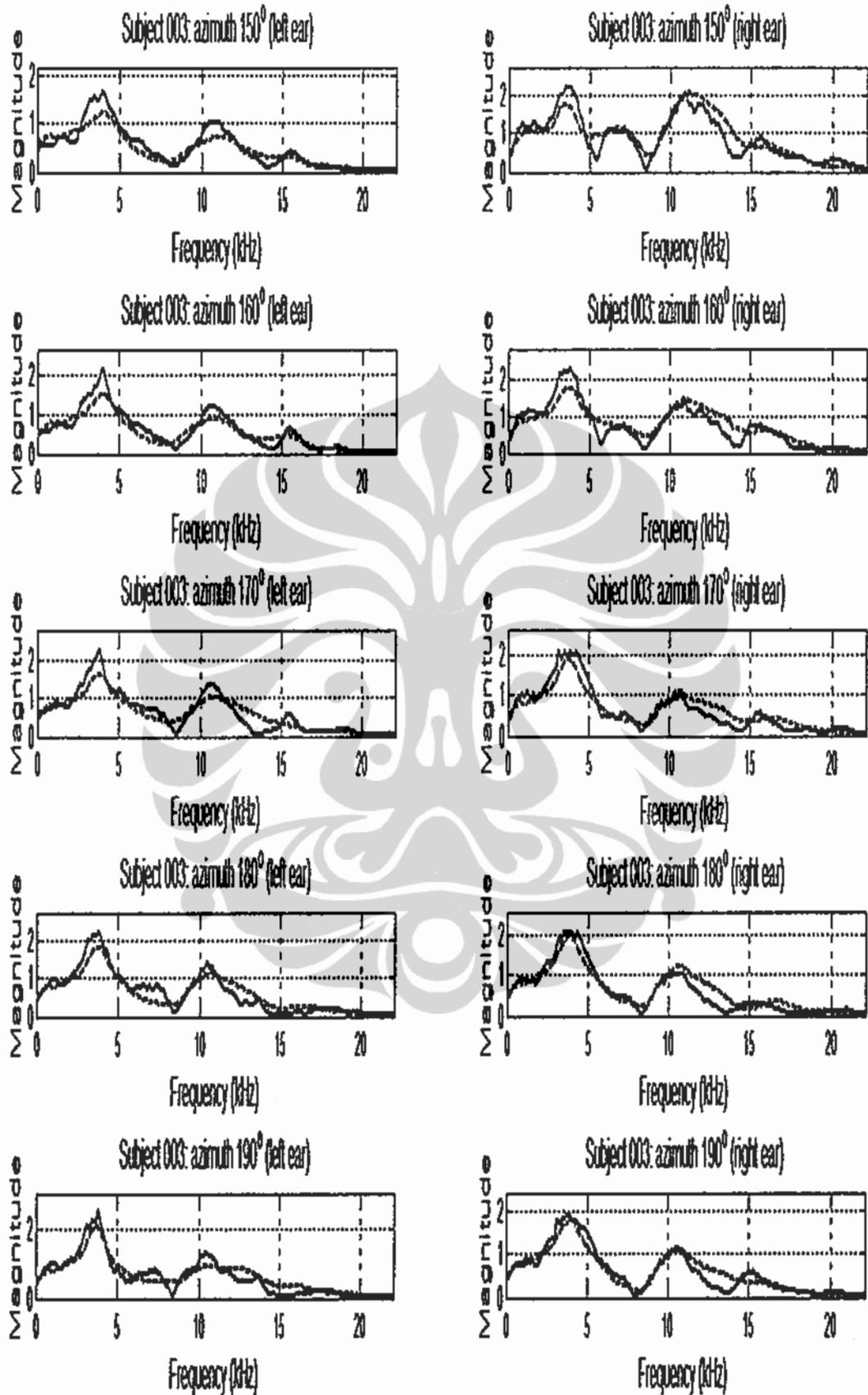
Gambar F5 menunjukkan HRTF *magnitude* asli dan model HRTF *magnitude* individual dari sumber bunyi di bidang horisontal dengan azimut 280° , 305° , 320° , 330° , 340° .



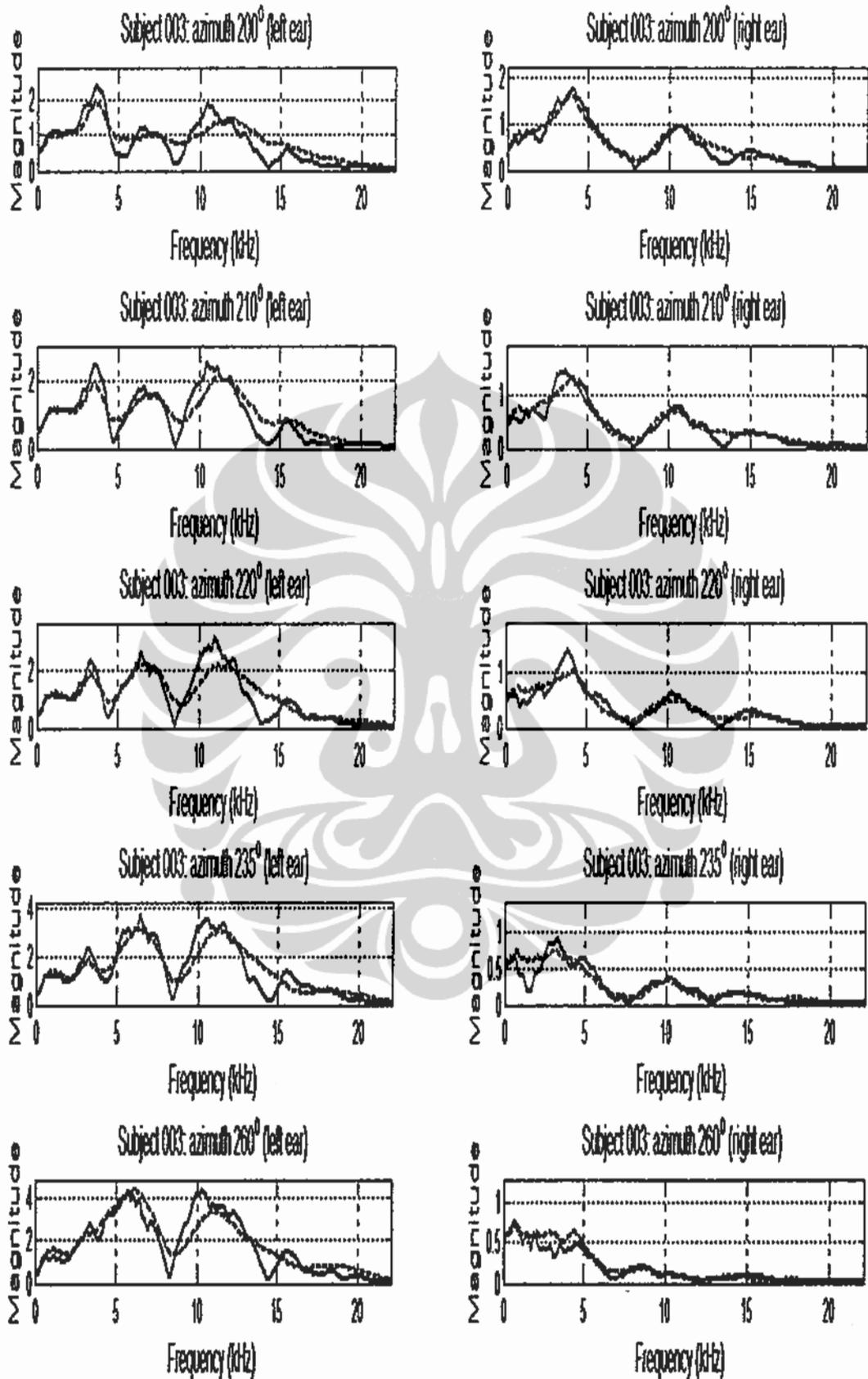
Gambar F.1. Model HRTF *Magnitude* Individual (dashed) dan HRTF *Magnitude* Asli (solid) dengan Azimuth 0°, 10°, 20°, 30°, 40° di Bidang Horizontal



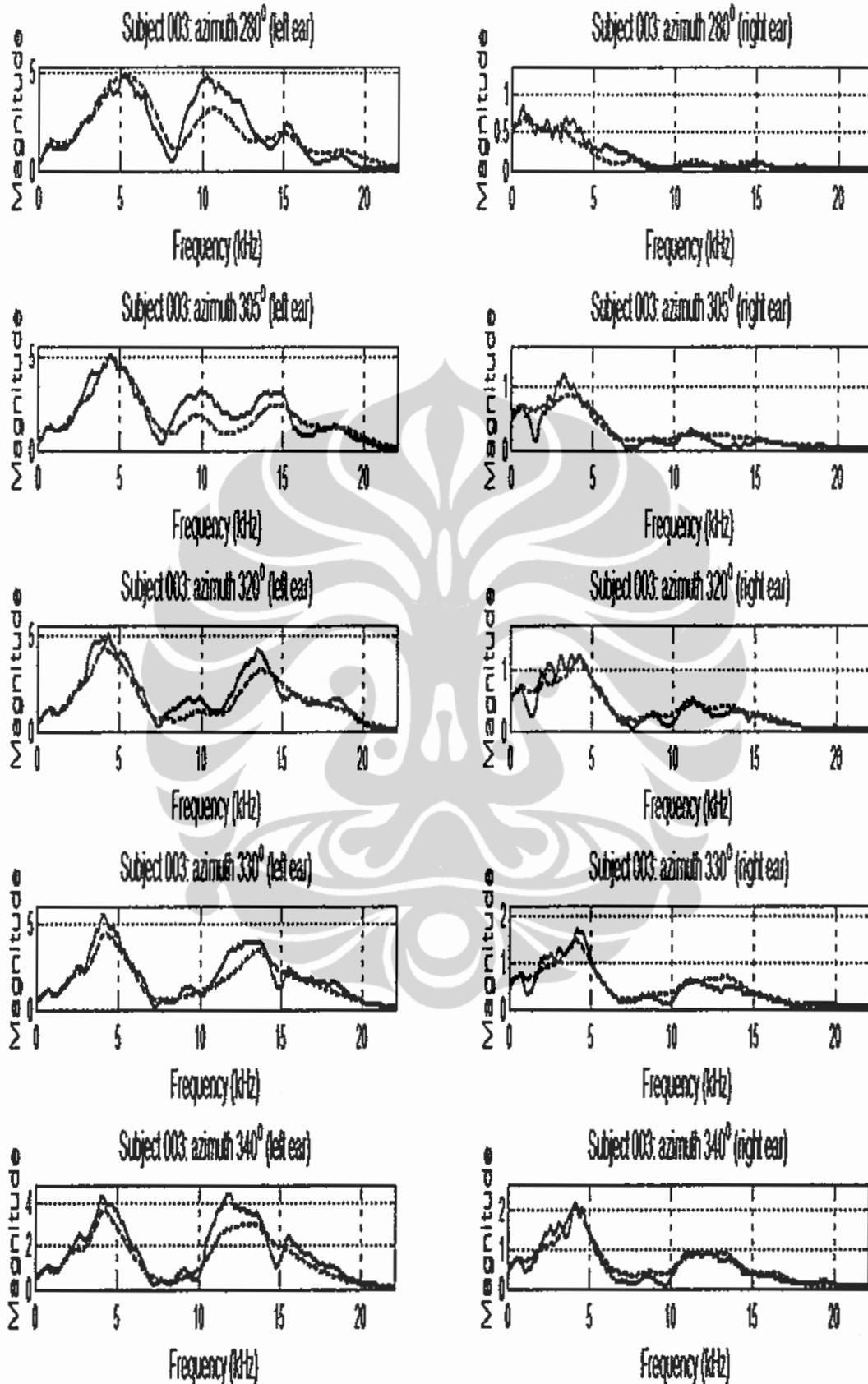
Gambar F.2. Model HRTF Magnitude Individual (*dashed*) dan HRTF Magnitude Asli (*solid*) dengan Azimut 55°, 80°, 100°, 125°, 140° di Bidang Horizontal



Gambar F.3. Model HRTF Magnitude Individual (*dashed*) dan HRTF Magnitude Asli (*solid*) dengan Azimut 150°, 160°, 170°, 180°, 190° di Bidang Horisontal



Gambar F.4. Model HRTF Magnitude Individual (dashed) dan HRTF Magnitude Asli (solid) dengan Azimut 200°, 210°, 220°, 235°, 260° di Bidang Horizontal



Gambar F.5. Model HRTF *Magnitude Individual (dashed)* dan HRTF *Magnitude Asli (solid)* dengan Azimut 280°, 305°, 320°, 330°, 340° di Bidang Horizontal

MODEL-MODEL HRTF *MAGNITUDE* INDIVIDUAL SUBYEK 003 DI BIDANG MEDIAN

Dalam lampiran ini, ditampilkan 25 model HRTF *magnitude* individual dari telinga kiri dan telinga kanan subyek 003, yang dihasilkan dari individualisasi model HRTF *magnitude* seperti dijelaskan dalam subbab 5.4, untuk sumber-sumber bunyi di bidang median. Untuk suatu telinga dari seorang subyek, dihasilkan keseluruhan 50 model HRTF *magnitude* individual di bidang median sebagai fungsi dari elevasi.

Dalam satu koordinat, HRTF *magnitude* asli ditunjukkan oleh garis (*solid*) dan model HRTF *magnitude* individual ditunjukkan oleh garis putus-putus (*dashed*) pada Gambar G1 sampai Gambar G5. Spektrum-spektrum *magnitude* telinga kiri ditunjukkan pada panel-panel kiri dan spektrum-spektrum *magnitude* telinga kanan ditunjukkan pada panel-panel kanan, dalam semua gambar tersebut.

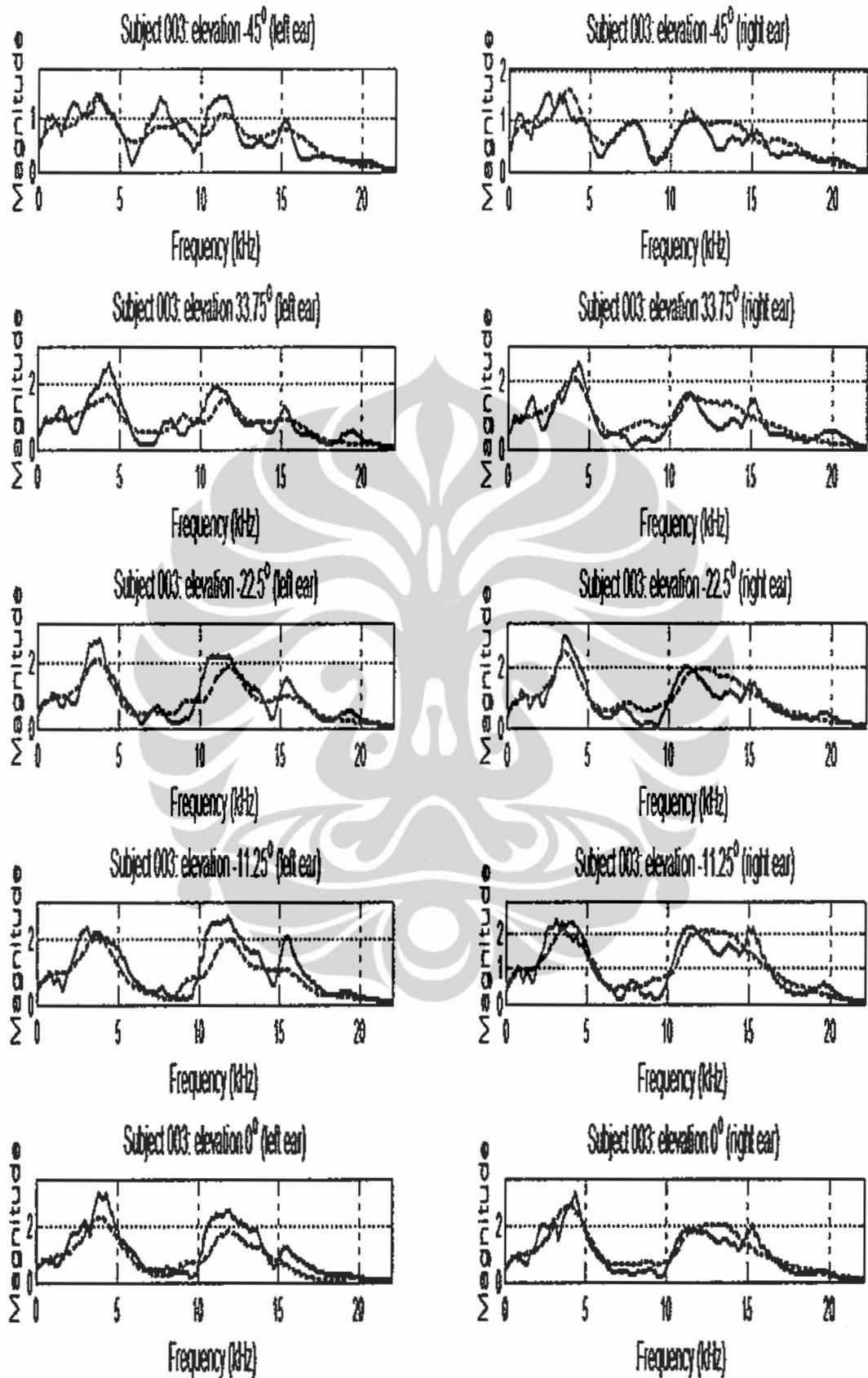
Gambar G1 menunjukkan HRTF *magnitude* asli dan model HRTF *magnitude* individual dari sumber bunyi di bidang median dengan elevasi -45° , $-33,75^\circ$, $-22,5^\circ$, $-11,25^\circ$, 0° .

Gambar G2 menunjukkan HRTF *magnitude* asli dan model HRTF *magnitude* individual dari sumber bunyi di bidang median dengan elevasi $11,25^\circ$, $22,5^\circ$, $33,75^\circ$, 45° , $56,25^\circ$.

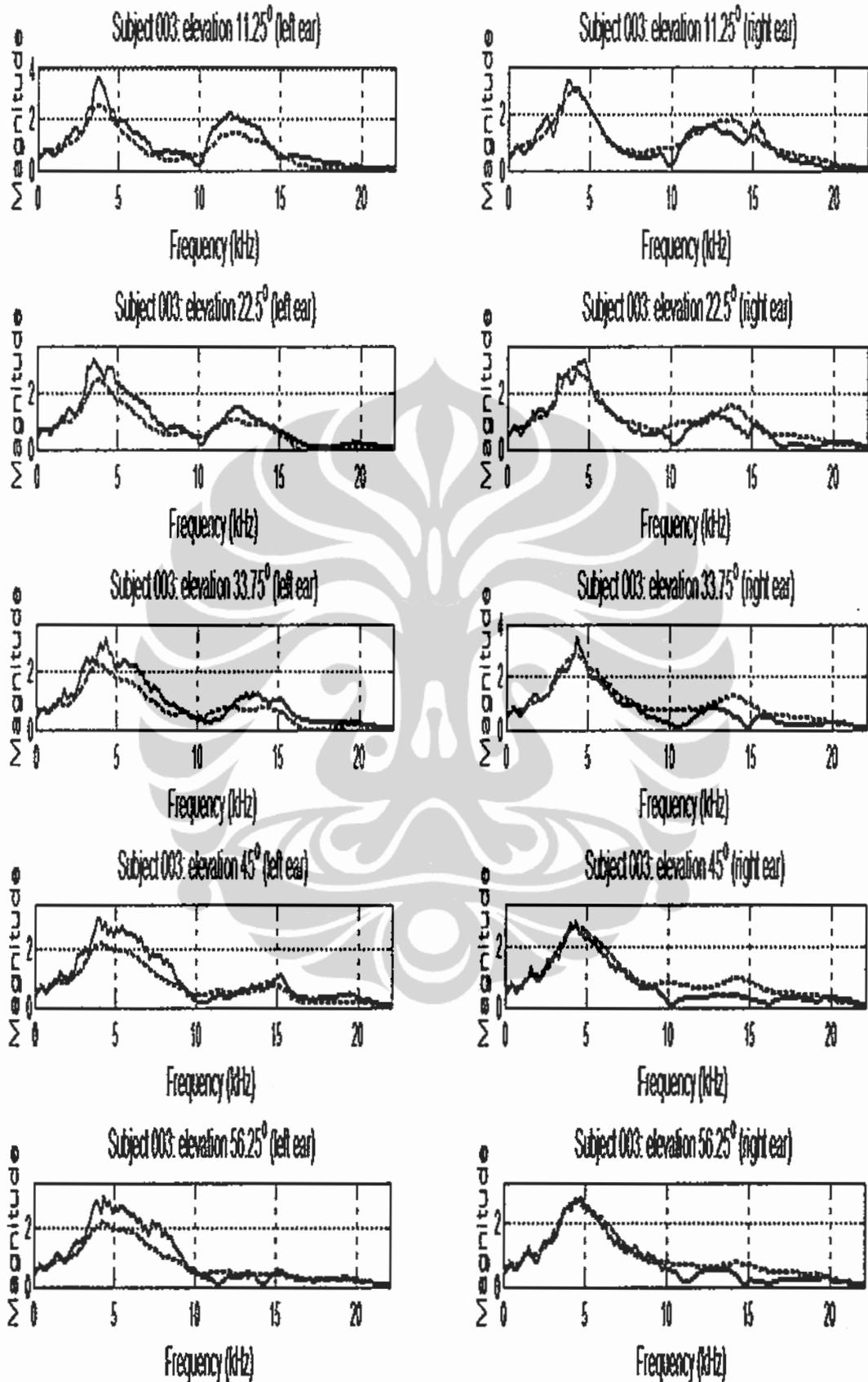
Gambar G3 menunjukkan HRTF *magnitude* asli dan model HRTF *magnitude* individual dari sumber bunyi di bidang median dengan elevasi $67,5^\circ$, $78,75^\circ$, 90° , $101,25^\circ$, $112,5^\circ$.

Gambar G4 menunjukkan HRTF *magnitude* asli dan model HRTF *magnitude* individual dari sumber bunyi di bidang median dengan elevasi $123,75^\circ$, 135° , $146,25^\circ$, $157,5^\circ$, $168,75^\circ$.

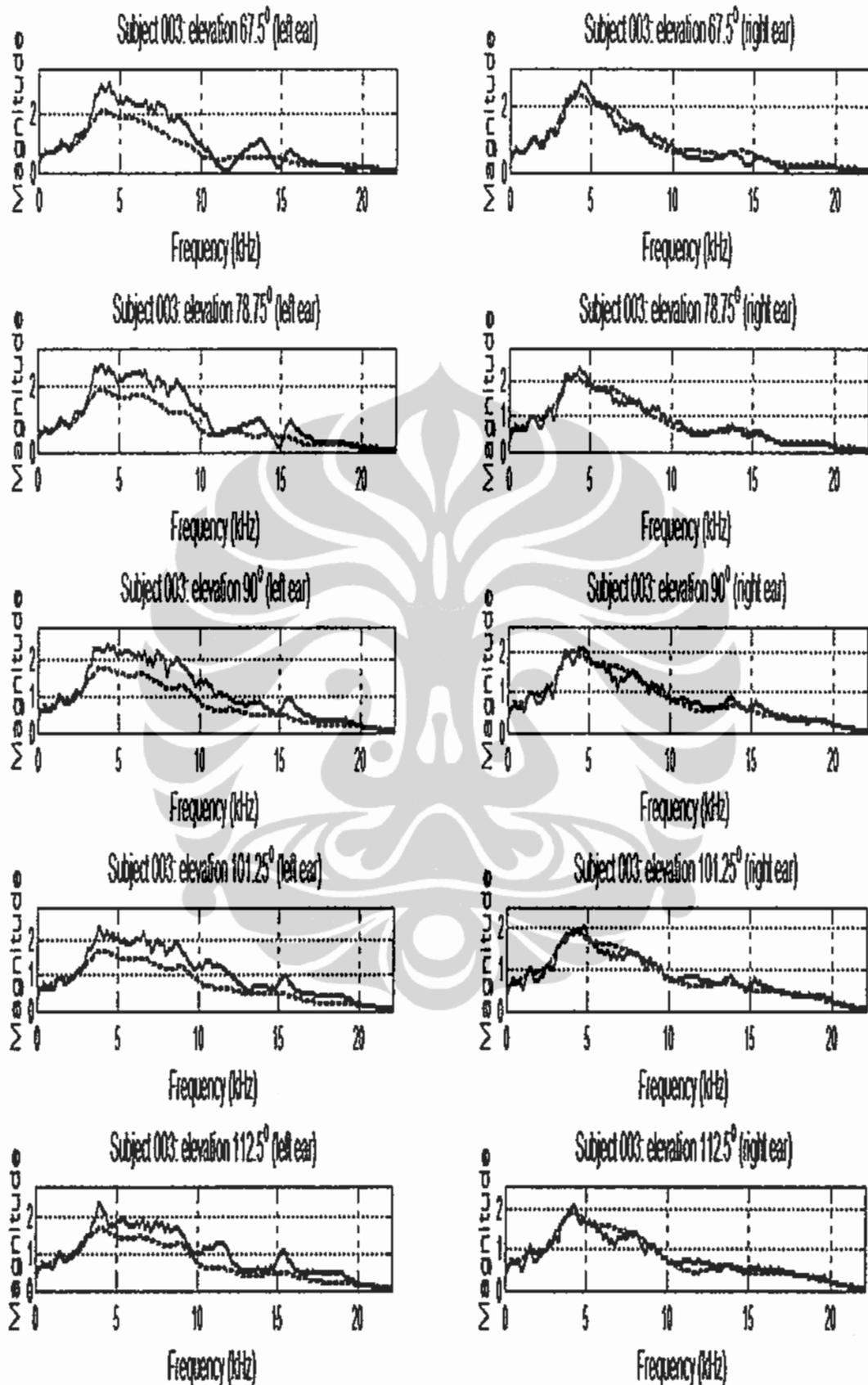
Gambar G5 menunjukkan HRTF *magnitude* asli dan model HRTF *magnitude* individual dari sumber bunyi di bidang median dengan elevasi 280° , 305° , 320° , 330° , 340° .



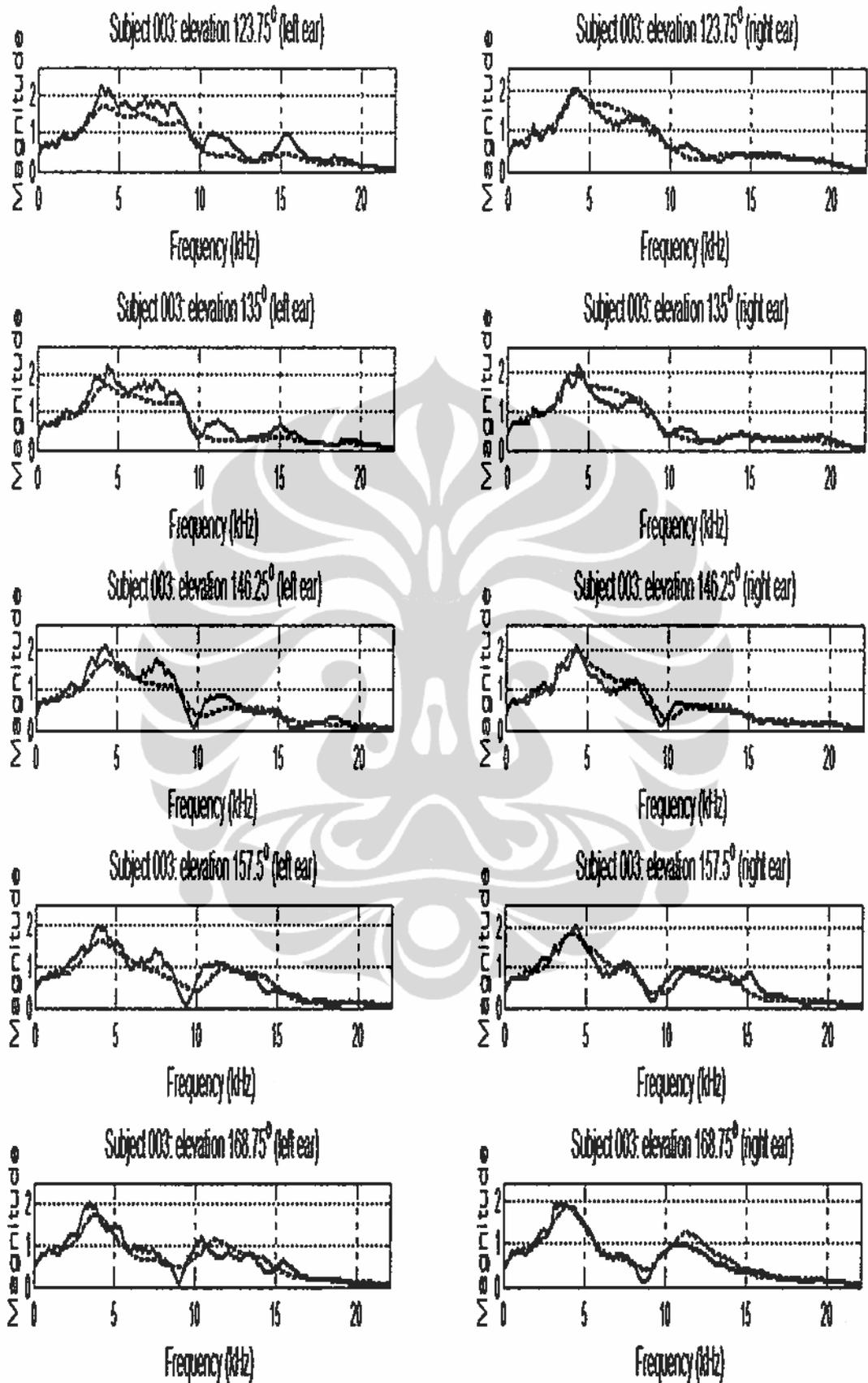
Gambar G.1. Model HRTF Magnitude (dashed) dan HRTF Magnitude Asli (solid) dengan Elevasi -45° , -33.75° , -22.5° , -11.25° , 0° di Bidang Median



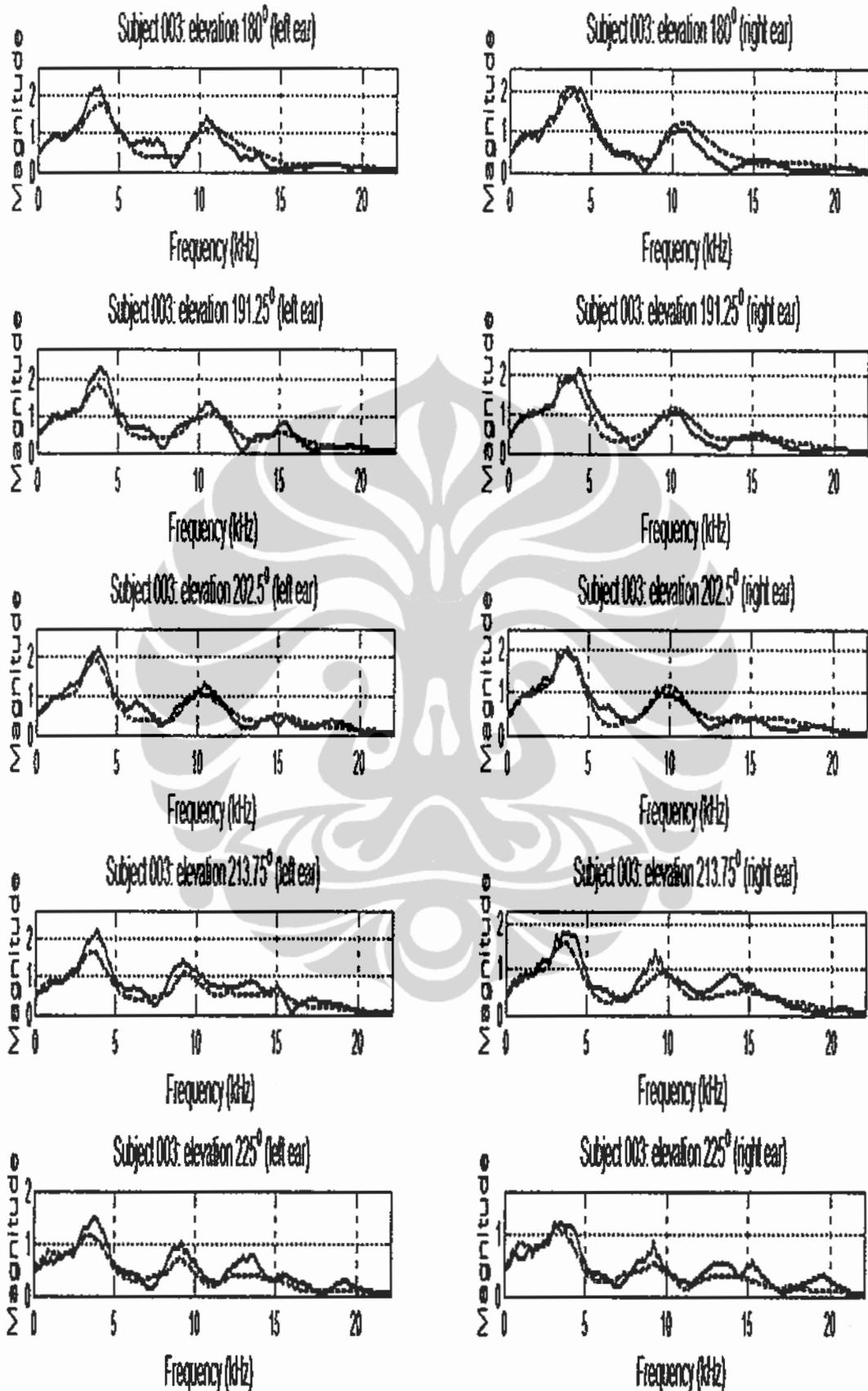
Gambar G.2. Model HRTF Magnitude Individual (*dashed*) dan HRTF Magnitude Asli (*solid*) dengan Elevasi 11,25°, 22,5°, 33,75°, 45°, 56,25° di Bidang Median



Gambar G.3. Model HRTF Magnitude Individual (*dashed*) dan HRTF Magnitude Asli (*solid*) dengan Elevasi $67,5^\circ$, $78,75^\circ$, 90° , $101,25^\circ$, $112,5^\circ$ di Bidang Median



Gambar G.4. Model HRTF Magnitude Individual (dashed) dan HRTF Magnitude Asli (solid) dengan Elevasi 123,75°, 135°, 146,25°, 157,5°, 168,75° di Bidang Median



Gambar G.5. Model HRTF Magnitude Individual (dashed) dan HRTF Magnitude Asli (solid) dengan Elevasi 180° , $191,25^\circ$, $202,5^\circ$, $213,75^\circ$, 225° di Bidang Median