

LK-010

仮想音環境のための頭部伝達関数コーパス

Head-Related Transfer Function Corpus for Virtual Auditory Environment

渡邊 貫治^{†‡}
Kanji Watanabe岩谷 幸雄[†]
Yukio Iwaya行場 次朗[§]
Jiro Gyoba鈴木 陽一[†]
Yôiti Suzuki

1. はじめに

聴覚ディスプレイシステムとは、音源の方向や距離などの空間的な情報を合成することにより、聴取者に音空間情報を提示するシステムである。音空間の合成には、頭部伝達関数 [1] を音源信号に畳み込む方法が一般的である。頭部伝達関数は、聴取者の頭部や胴体、耳介などでの反射・回折の影響を含むため、身体形状の方向依存性、個人性が強く反映される。そのため、高精度な音空間の実現のためには、聴取者に専用の頭部伝達関数を用いることが重要である。しかしながら、多数の聴取者の頭部伝達関数コーパスは世界的に少なく、個人性の原因となる身体形状を含んだものはさらに限定される [2, 3]。そこで、多数の聴取者に対し、頭部を囲む全方向の頭部伝達関数を測定し、身体寸法も含めたコーパスを構築した。本稿では、その測定法および内容について説明する。

2. 頭部伝達関数の測定

頭部伝達関数は、「頭の中心に相当する自由空間上の1点と、耳道内の1点との間の伝達関数」と定義される [1]。これにしたがい、各方向から外耳道入口までの伝達関数を求め、各伝達関数を聴取者がいない状態での音源から頭部中心相当の点までの伝達関数で割ることで得たものを頭部伝達関数として測定を行った。実際に測定するのは、各音源方向から外耳道入口および頭部中心相当位置までの系のインパルス応答である。

測定は、東北大学電気通信研究所の無響室で行われた。測定の音源信号には、時間引き延ばしパルス (TSP, 図 1)[4] を用い、聴取者の外耳道入口に設置したマイクロホン (knowles FG3329) で集録した。この信号は、周波数領域で定義され、その周波数特性 $H(k)$ は、

$$H(k) = \begin{cases} \exp(j4m\pi k^2/N^2), & 0 \leq k \leq N/2, \\ H^*(N-k), & N/2 < k < N, \end{cases} \quad (1)$$

と表される [4]。ここで、* は複素共役を示す。この信号は、単位インパルスに周波数の2乗に比例して位相が変化するフィルタを畳み込んだもので、同振幅のインパルスに比べて全体のパワーが大きいので、S/N を大きく取ることができる。また、インパルス応答を求めるための逆フィルタの収束が良いので、逆 DFT によって簡単にインパルス応答を求めることができる。

測定用スピーカ (FOSTEX FE83E) は、図 2, 3 に示すように、円形の枠に 10 度間隔で取り付けられており、枠は水平方向に 1 度間隔の精度で回転させることができる。聴取者の頭部中心からスピーカ前面までの距離は 1.5 m である。聴取者には頭部を動かさないよう指示し、専

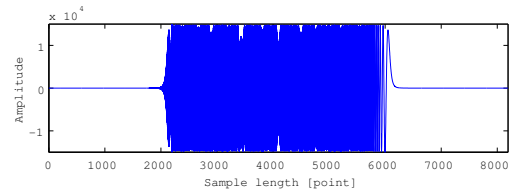


図 1: 測定に使用した TSP 信号の時間波形

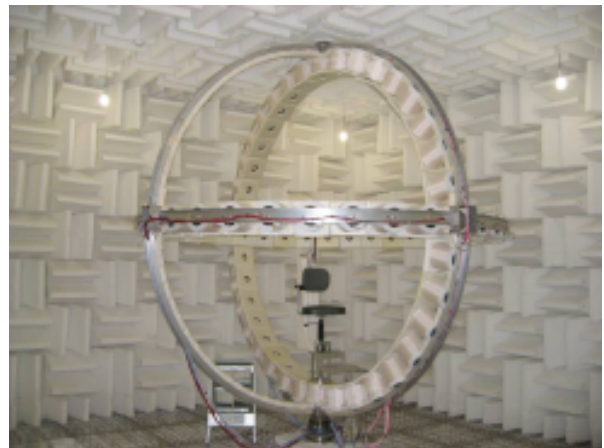


図 2: 頭部伝達関数の測定装置

用の治具を用いて頭部を強制的に固定した。聴取者の外耳道入口は、Møller[5] の提案する外耳道をふさぐ方法に従い、聴取者ごとにマイクロホンを埋め込んだ耳栓を作成して測定を行った。その結果得られた頭部伝達関数は外耳道の影響が除かれ、外見の形状のみで決まると考えられる。

TSP 信号はサンプリング周波数 48 kHz, 8192 ポイントとし、各方向について 4 回の同期加算を行った。測定角度間隔は真下を除く水平方向 5 度、仰角方向 10 度間隔の 1225 方向である。

3. 聴取者の身体特徴量の測定

頭部伝達関数の周波数特性は、耳に入るまでに人体での反射・回折により特徴づけされた結果と考えられる。耳からの距離を考えると、特に、頭部や耳介、肩の影響が大きいと思われる。それらの頭部伝達関数への寄与に関する分析も行われている [6] が、未解明の部分も多い。さらに、身体形状と空間知覚との関係も明らかではない。もし、それらの関係が明らかになれば、聴取者に最適な頭部伝達関数を測定を要せず推定することが可能と考えられ、聴取者への負担の減少につながる。したがって、

[†] 東北大学電気通信研究所/大学院情報科学研究科

[‡] 現在、山梨大学工学部コンピュータ・メディア工学科勤務

[§] 東北大学大学院文学研究科

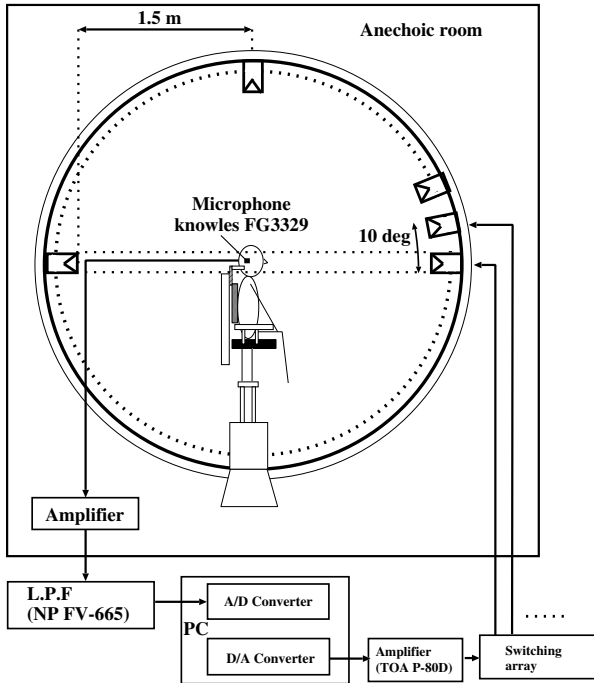


図 3: 頭部伝達関数の測定系

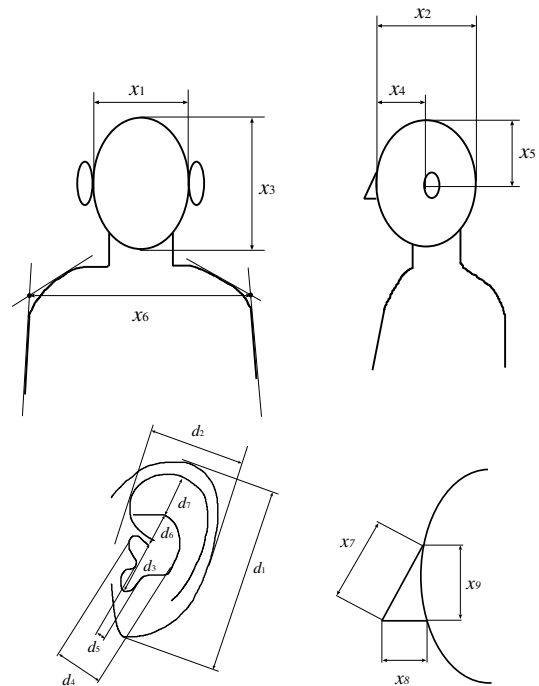


図 4: 測定した身体特徴量部位

頭部伝達関数コーパスに聴取者の身体寸法データが含まれていれば、個人性の研究のみならず音空間知覚の研究への貢献が大いに期待できる。

以上のことを踏まえ、頭部伝達関数を測定する際に3次元デジタルカメラ(VIVID910 KONICA MINOLTA)を用い、聴取者の肩より上の立体画像データを全方向について測定した。それを基に、図4に示す身体特徴量を得た。それぞれ、頭部の大きさ(x_1, x_2, x_3)、耳の位置(x_4, x_5)、肩幅(x_6)、鼻の大きさ(x_7, x_8, x_9)、耳介の大きさ・形状($d_1, d_2, d_3, d_4, d_5, d_6, d_7$)に関する特徴量に対応する。本報告では、それぞれ x_1 :頭幅、 x_2 :頭長、 x_3 :頭高、 x_4 :頭耳長、 x_5 :頭耳高、 x_6 :肩幅、 x_7 :鼻長、 x_8 :鼻深、 x_9 :鼻高、 d_1 :耳長、 d_2 :耳幅、 d_3 :耳甲介長、 d_4 :耳甲幅(長)、 d_5 :耳甲幅(短)、 d_6 :耳甲介艇、 d_7 :舟状窩とする[7]。

4. 構築したコーパスの内容および評価

測定は、59名の聴取者に対して行われた。年齢は20代から30代の成人であり、男性43名、女性15名、ダミーヘッド(高研, SAMRAI)である。

構築したコーパスの評価の一例として、頭部伝達関数から抽出される空間情報の重要な手がかりである、両耳間時間差および両耳間レベル差を図5, 6に示す。各図において、上に示してあるのは水平面上の方向に対する変化であり、横軸は音源方向、縦軸は右耳に対する左耳での音の到達時間差(図5)および音圧レベル差(図6)である。太線は全聴取者の平均値、細線は標準偏差を示している。また、各図において下に示してあるのは、90度における度数分布である。図の形状から、山を一つ持つ分布で、十分なばらつきをもっていると見なせる。ま

た、図5から、最大約0.1msの偏差が見て取れる。両耳間時間差の弁別限が数 μs であることを考えると、十分大きい値であるといえる。両耳間レベル差についても、同様のことがいえ、定位知覚として考えた場合でも個人性を分析するために、十分なばらつきをもったコーパスが構築されたと見なせる。

一方、図4で示した身体特徴量の名称および測定結果を表1に示す。これらの身体特徴量からでも、聴取者の身体形状の推測が可能である。例えば、頭幅(x_1)より頭長(x_2)が長いことから、頭部の形状は、円ではなく楕円に近いことが言え、平均値から扁平率を考えると、 $(199-156)/156 = 0.2161$ である。公開されている人体寸法・形状データベース(AIST人体寸法・形状データベース1991-92)[8]と比較すると、成人の頭長は平均185.1mm、頭幅(データベースでは耳珠間幅)は平均145.7mmであるので、同様に扁平率を計算すると0.2129であり、非常に近い値となった。また、頭耳長(x_4)は耳の前後位置に対応する特徴量である。頭長との比較から、平均値は頭長が頭耳長のほぼ2倍であり、耳がほぼ頭部の真横と言えそうであるが、最小値は頭耳長が79mmであるのに対し、頭長は180mmと倍以上の長さである。したがって、個人によっては必ずしも耳が真横とは限らず、前後にずれた位置に耳がついていると思われる。以上のことは、頭部をモデル化する際、楕円球状とし耳を真横からずらして考えた方が、個人性を考えたより良いモデルになりうることを示す。頭部伝達関数との関連が明らかになれば、個人性の研究に大きく寄与できると思われる。

以上のように構築したコーパスは、単に多数の頭部伝達関数を提供するだけでなく、個人性の研究など、応用範囲が広く、空間音響の分野において大きな貢献が期

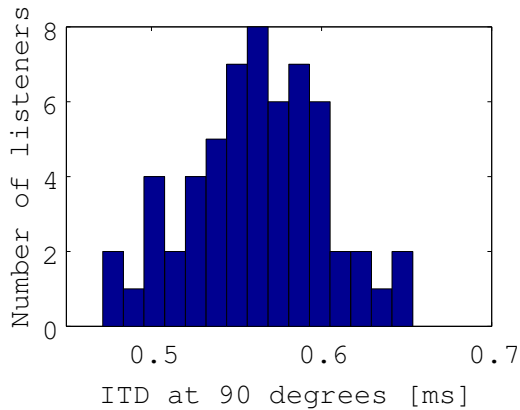
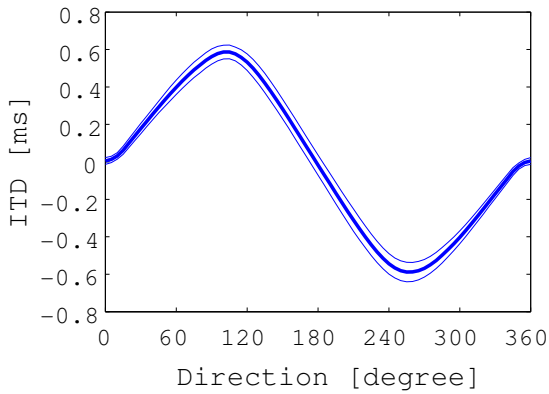


図 5: コーパスから得られる両耳間時間差 (上) および 90 度における両耳間時間差に対する聴取者の度数分布 (下)

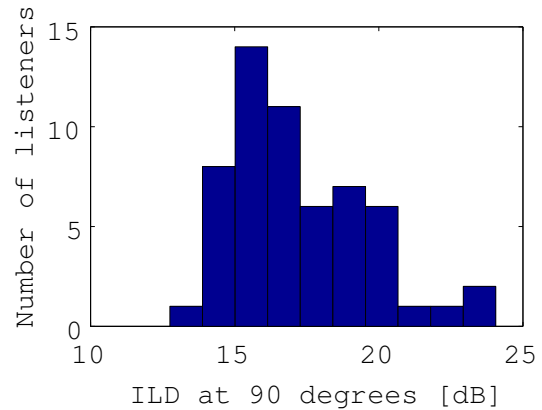
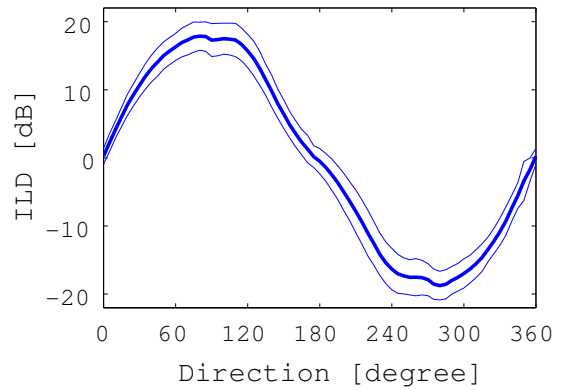


図 6: コーパスから得られる両耳間レベル差 (上) および 90 度における両耳間レベル差に対する聴取者の度数分布 (下)

待できる。

5. コーパスを利用した両耳間時間差と身体特徴量の分析

頭部伝達関数と身体特徴量の大规模な標本群があれば、その関係を統計的な方法を用いて分析が可能である。もし両者の関係が明らかになれば、容易に測定可能な身体特徴量から頭部伝達関数を推定することが可能と考えられ、聴覚ディスプレイなどの高精度化に寄与するところが多い。

そこでここでは、定位の基本的な手がかりとして重要な両耳間時間差と身体特徴量との関係を分析し、その結果を示す。両耳間時間差は、音源から左右耳までの到達時間差、すなわち経路差に対応し、経路差は頭部の幅など、身体特徴量が大きく影響すると考えられる。

5.1 分析方法

両耳間時間差は方向によって変化する多変数であり、一方、身体特徴量も多変数のパラメータである。本報告では、多変数同士の相関を分析する一方法である正準相関分析を用いて分析を行った。

p 個、 q 個の 2 つの変量群 x, y があり、それぞれの変量群について、各変数の線形結合により定義される合成

変量、

$$U = a_1x_1 + a_2x_2 + \dots + a_px_p \quad (2)$$

$$V = b_1y_1 + b_2y_2 + \dots + b_qy_q \quad (3)$$

を作ったとする。合成変量 U, V は、正準変量と呼ばれ、正準相関分析ではこれらの変量間の相関係数が最大になるような重み係数 a, b を求める。また、得られた U, V の相関係数を正準相関係数と呼ぶ。

本報告では、両耳間時間差については、水平面 30 度間隔の各値の 12 種類を変量とし、身体特徴量については図 4 に示したものを使用した。ただし、耳介の特徴量については、耳長 (d_1) および耳幅 (d_2) のみとし、全部で 11 種類の変量とした。これは、コーパスの聴取者数に対して変量を多くしすぎると分析の信頼性が低下するための制限であり、今後標本数を増やすことが、より詳細な分析のための課題である。

5.2 正準相関分析の結果および考察

正準相関分析の結果、得られた正準相関係数の値は 0.90 であり、非常に高い相関が得られた。このことは、両耳間時間差に関する第 1 正準変量の分散の 81% ($= 0.90^2$) が、身体特徴量に関する正準変量から説明できることを意味する。また、表 2 に、身体特徴量に関する構造係数を示す。ここで示されている構造係数は、第 1 正準変量

表 1: 身体特徴量の測定結果 (単位: mm)

測定部位	平均	最大	最小	標準偏差
x_1 : 頭幅	156	175	138	8
x_2 : 頭長	199	217	180	9
x_3 : 頭高	227	250	197	11
x_4 : 頭耳長	94	107	79	7
x_5 : 頭耳高	145	165	131	7
x_6 : 肩幅	408	481	340	33
x_7 : 鼻長	45	58	31	6
x_8 : 鼻深	16	20	11	2
x_9 : 鼻高	45	58	31	6
d_1 : 耳長	58	71	47	12
d_2 : 耳幅	28	33	24	5
d_3 : 耳甲介長	15	24	17	9
d_4 : 耳甲幅 (長)	15	26	14	9
d_5 : 耳甲幅 (短)	6	11	6	4
d_6 : 耳甲介挺	10	19	10	7
d_7 : 舟状窩	10	20	8	7

と各身体特徴量との相関を示したものである。表から、頭幅 (x_1) が非常に高い値を示しており、頭長 (x_2) や頭高 (x_3) も比較的高い値を示している。興味深い点として、頭耳長 (x_4) すなわち耳の前後位置や肩幅 (x_6) に対応する構造係数も高い値を示していることから、頭部の大きさだけでなく、耳の位置や肩幅が両耳間時間差を説明するのに重要な身体特徴量であることが考えられる。両耳間時間差の合成に関する研究として、Algazi らは、頭部を球で近似したモデルについて検討した [9]。まず、各聴取者の両耳間時間差を球のモデルに合わせ、その半径を求めた。次に、頭幅、頭長、頭高を用い、聴取者に最適な球の半径を推定し、モデルの半径と比較した結果、1 cm 以下の誤差であった。しかしながら、あくまで頭部を球として扱っているため、得られた両耳間時間差がどの程度実測に近いかは不明である。そこで、頭部の大きさに関する身体特徴量である、頭幅 (x_1)、頭長 (x_2)、頭高 (x_3) のみを用いた場合の正準相関分析を行った結果、得られた正準相関係数は 0.84 であり、同様に両耳間時間差に関する第 1 正準変量の分散の約 70% を説明できることが示された。

以上のことから、両耳間時間差に影響する身体特徴量として、頭部の大きさに関するものが最も重要であるが、より厳密に個人性を考慮する場合には、耳の位置や肩幅なども重要であるという、新しい知見が得られた。

6. まとめ

59 名の成人を聴取者とする頭部伝達関数およびその身体寸法データを含むコーパスを構築した。また、その応用例として、両耳間時間差と身体特徴量の関係を正準相関分析に基づき分析した。その結果、従来のモデルで考慮されていた頭部の大きさのほかに、耳の位置 (頭耳長) や肩幅も重要であることを明らかにした。この例からも、構築したコーパスを用いることで、頭部伝達関数

表 2: 正準相関分析の結果得られた身体特徴量に関する構造係数

身体特徴量	構造係数
x_1 : 頭幅	0.904
x_2 : 頭長	0.634
x_3 : 頭高	0.556
x_4 : 頭耳長	0.695
x_5 : 頭耳高	0.293
x_6 : 肩幅	0.623
x_7 : 鼻長	0.174
x_8 : 鼻深	0.021
x_9 : 鼻高	0.255
d_1 : 耳長	0.315
d_2 : 耳幅	0.263

について新たな知見が得られ、空間音響分野の深化に貢献しうることが大いに期待できる。

謝辞

本研究の一部は、経済産業省地域新生コンソーシアム研究開発事業の補助を受けて行われた。

参考文献

- [1] Jens Blauert, 森本政之, 後藤俊幸, 空間音響, 鹿島出版会, 1986.
- [2] <http://interface.cipic.ucdavis.edu/>
- [3] Nishino *et al.*, "Estimating the Head Related Transfer Function with the Multiple Regression Analysis," CREST workshop on Computational Models of Auditory Processing, 2002.
- [4] Y. Suzuki *et al.*, "An Optimum computer generated pulse signal suitable for the measurement of very long impulse responses," J. Acoust. Soc. Am. Vol. 97, pp. 1119-1123(1995).
- [5] H. Møller, "Fundamentals of Binaural Technology," Applied Acoustics, **36**, pp. 171-218(1992).
- [6] E. A. G. Shaw, "Acoustical Features of the Human External Ear," in Binaural and Spatial Hearing in Real and Virtual Environments, edited by R. H. Gilkey and T. R. Anderson, pp. 25-47, 1997.
- [7] 人類学講座編纂委員会 編, 人類学講座 (別巻 1) 人体計測法, 雄山閣出版 (1991).
- [8] <http://www.dh.aist.go.jp/AIST91DB/>
- [9] V. R. Algazi, C. Avendano and R. O. Duda, "Estimation of a Spherical-Head Model from Anthropometry," J. Audio Eng. Soc., Vol. 49, pp. 472-479(2001).