

聴覚における寸法知覚と最適末梢系

入野 俊夫^{†1,a)}

概要：

本発表では、聴覚末梢系の最適性を追求することから始まり、音源の寸法-形状知覚の特性測定や応用に至る研究の背景と経緯を紹介する。聴覚末梢系のインパルス応答は、時間-スケール表現における最小不確定性を満足する関数として解析的に求められることがわかった。得られた「ガンマチャープ関数」とその後の発展版は、生理実験データや心理物理実験データを良く説明できる関数として広く用いられている。また、この健聴者のみならず難聴者の分析特性を模擬できるばかりでなく、模擬難聴の音を合成することも可能となった。この最適性の背景となる時間-スケール表現の導出アルゴリズムとして、「安定化ウェーブレット-メルン変換」を提案した。この表現は、音響系のスケールすなわち音源の寸法や形状に関して、人間の知覚系で分離抽出処理が行われているという仮説に基づいている。日常生活では、大人と子供の同一発話を、同一内容として認識できると同時に発話者のおおよその寸法もわかるので、直感的には妥当である。そこで知覚処理の精度を測るため、様々な音声における寸法の弁別閾や認識率を、聴覚心理実験を通して明らかにしてきた。さらにこのモデルの工学的な優位性を示すため、話者の声道長の推定や大人子供判別に適用し、従来法に比べて性能が向上することも示してきた。世界の先陣を切って突入した高度高齢化社会における音支援を考える上で、最も重要な聴覚的信号処理の基盤となる技術であると考えている。

1. 文献紹介

研究背景に関しては、紙幅の制限上十分解説することができない。むしろ、過去の解説や論文を直接参照していただいた方がよいと思われる。そこでここでは、カテゴリごとに整理して文献番号を示すにとどめた。主に原著の英語論文や解説記事を引用するが、直近の研究会や大会等における日本語の発表もある。それらの情報を含め、現在の研究概要等は、Web page[1]に掲載されているので参照していただきたい。

1.1 ガンマチャープ聴覚フィルタ

- 最小不確定性による導出とデータ適合：[2]
- IIR フィルタによる分析合成系の実装：[3] [4]
- 圧縮特性導入と健聴者の聴覚フィルタ特性：[5] [6] [7]
- 動的な特性導入とソフトウェアパッケージ化：[8]
- 圧縮特性の推定と模擬難聴：[9]
- 聴覚フィルタの初心者向け解説：[10] [11] [12] [13]

1.2 安定化ウェーブレット-メルン変換

- 計算理論と背景：[14] [15] [16] [17]

- 聴覚イメージモデルと聴覚非対称性：[18] [19]
- 聴覚イメージモデルによる音源分離：[20]

1.3 音源寸法-形状知覚

- 音声からの寸法-形状知覚：[21] [22]
- 寸法-形状知覚の時間特性：[23] [24] [25]
- 寸法手掛かりによる音脈分凝/聴覚情景分析：[26] [27]
- ささやき声からの寸法-形状知覚：[28]

1.4 ガンマチャープの工学応用

- 声道長推定：[29] [30]
- 大人子供判別：[31]

謝辞 本発表で紹介した内容は、多くの研究者との共同研究によって発展させることができた。Cambridge 大 生理発達神経科学科 CNBH Patterson 博士、和歌山大学 河原英紀教授、京都市芸術大学 津崎実教授他、多くの方々に心より感謝する。

参考文献

- [1] <http://www.wakayama-u.ac.jp/~irino/> (2013.04.12)
- [2] Toshio Irino and Roy D. Patterson "A time-domain, level-dependent auditory filter: the gammachirp," J. Acoust. Soc. Am., 101 (1), pp.412-419, January 1997.

^{†1} 現在, 和歌山大学
Presently with Wakayama University

^{a)} irino@sys.wakayama-u.ac.jp

- [doi:10.1121/1.417975]
- [3] Toshio Irino and Masashi Unoki, "An analysis/synthesis auditory filterbank based on an IIR implementation of the gammachirp," *J. Acoust. Soc. Japan (E)*, 20 (6), pp. 397-406, Nov. 1999.
- [4] Masashi Unoki, Toshio Irino, and Roy D. Patterson, "Improvement of an IIR asymmetric compensation gammachirp filter," *Acoust. Sci. & Tech.* (ed. by the Acoustical Society of Japan), 22 (6), pp. 426-430, Nov. 2001. [doi:10.1250/ast.22.426]
- [5] Toshio Irino and Roy D. Patterson, "A compressive gammachirp auditory filter for both physiological and psychophysical data," *J. Acoust. Soc. Am.*, 109 (5), pp.2008-2022, May 2001. [doi:10.1121/1.1367253]
- [6] Roy D. Patterson, Masashi Unoki, and Toshio Irino "Extending the domain of center frequencies for the compressive gammachirp auditory filter," *J. Acoust. Soc. Am.*, 114 (3), pp.1529-1542, Sept 2003. [doi:10.1121/1.1600720]
- [7] Masashi Unoki, Toshio Irino, Brian Glasberg, Brian C. J. Moore, and Roy D. Patterson, "Comparison of the roex and gammachirp filters as representations of the auditory filter," *J. Acoust. Soc. Am.*, 120(3), pp.1474-1492, Sept., 2006. [doi:10.1121/1.2228539]
- [8] Toshio Irino and Roy D. Patterson, "A dynamic compressive gammachirp auditory filterbank" *IEEE Trans. Audio, Speech, and Language Process.*, 14(6), pp.2222-2232, Nov. 2006. [doi:10.1109/TASL.2006.874669]
- [9] Toshio Irino, Tomofumi Fukawatase, Makoto Sakaguchi, Ryuichi Nisimura, Hideki Kawahara, and Roy D. Patterson, "Accurate estimation of compression in simultaneous masking enables the simulation of hearing impairment for normal hearing listeners," 16th International Symposium on Hearing (ISH2012), St John's College, Cambridge UK, 23-27 July, 2012
- [10] 入野俊夫, "聴覚フィルタの測定と定式化について," 日本音響学会聴覚研究会資料, H-2009-73, Vol. 39, No. 6, pp.413 - 418, [聴覚研究会、レクチャー招待講演], 加太国民休暇村, 和歌山, 2009年10月9~10日
- [11] 入野俊夫, "はじめての聴覚フィルター - 心理物理実験デモで学ぶ聴覚フィルタ特性 -," 日本音響学会: 秋季研究発表会講演論文集, pp.1347 - 1348, 関西大学, 大阪, 2010年9月14日~16日
- [12] 入野俊夫, "はじめての聴覚フィルタ," 日本音響学会誌, 66巻10号, pp. 506-512, Oct. 2010.
- [13] 森周司, 香田徹 (編), 香田徹, 日比野浩, 任書晃, 倉智嘉久, 入野俊夫, 鶴木祐史, 鈴木陽一, 牧勝弘, 津崎実, (共著), "聴覚モデル," コロナ社, 東京, 233p., ISBN 978-4-339-01323-8.
- [14] Toshio Irino and Roy D. Patterson, "Segregating information about the size and shape of the vocal tract using a time-domain auditory model: The Stabilised Wavelet Mellin Transform," *Speech Communication*, 36 (3-4), pp.181-203, March 2002. [doi: 10.1016/S0167-6393(00)00085-6]
- [15] 入野俊夫, "初期聴覚系におけるスケール理論" 日本音響学会春季研究会講演論文集, I, pp.511-514, 早稲田大学, 2003年3月18日
- [16] 入野俊夫, 津崎実, "聴覚初期過程の機能モデル," in 新編感覚知覚心理学ハンドブック Part 2 (大山正・今井省吾・和氣典二・菊池正 編), 第III部聴覚, pp.349-358, 誠信書房, 東京, ISBN 978-4-414-30504-3, 615p, 2007年9月.
- [17] 入野俊夫, 河原英紀, "聴覚認知過程の数理," in 数理科学辞典 (第2版) (広中平祐 他 編), pp. 1058-1061, 丸善, 東京, ISBN 978-4-621-08125-9, 1450p.
- [18] Toshio Irino and Roy D. Patterson, "Temporal asymmetry in the auditory system," *J. Acoust. Soc. Am.*, 99 (4), pp.2316-2331, April 1996. [doi:10.1121/1.415419]
- [19] Roy D. Patterson and Toshio Irino, "Modeling temporal asymmetry in the auditory system," *J. Acoust. Soc. Am.*, 104 (5), pp.2967-2979, November 1998. [doi:10.1121/1.423879]
- [20] Toshio Irino, Roy D. Patterson, and Hideki Kawahara "Speech segregation using an auditory vocoder with event-synchronous enhancements," *IEEE Trans. Audio, Speech, and Language Process.*, 14(6), pp.2212-2221, Nov. 2006. [doi:10.1109/TASL.2006.872611]
- [21] David R. Smith, Roy D. Patterson, Richard Turner, Hideki Kawahara, and Toshio Irino, "The processing and perception of size information in speech sounds," *J. Acoust. Soc. Am.*, 117(1), pp. 305-318, Jan., 2005. [doi:10.1121/1.1828637]
- [22] 津崎実, 入野俊夫, "音色に潜む寸法と形状情報 - 混沌から紡ぎだす秩序 -," 日本音響学会: 秋季研究発表会講演論文集, pp.619-622, 金沢大学, 石川, 2006年9月13日-15日
- [23] 津崎実, 入野俊夫, "スケール変調音声に対する聴覚的追従性と聴覚的情景," 日本音響学会: 秋季研究発表会講演論文集, I, pp.521-524, 琉球大学, 沖縄, 2004年9月28日-30日
- [24] Chihiro Takeshima, Minoru Tsuzaki, and Toshio Irino, "Detection of temporal modulation of size in vowel sequences," *Acoust. Sci. & Tech.* (ed. by the Acoustical Society of Japan), 28 (5), pp. 349-351, Sep., 2007. (Letter) [doi:10.1250/ast.28.349]
- [25] Chihiro Takeshima, Minoru Tsuzaki, and Toshio Irino, "Perception of vowel sequence with varying speaker size," *Acoust. Sci. & Tech.* (ed. by the Acoustical Society of Japan), Vol. 31, No. 2, pp.156-164, Mar., 2010. [doi:10.1250/ast.31.156] *
- [26] Minoru Tsuzaki, Chihiro Takeshima, and Toshio Irino, "Perception of size modulated vowel sequence: Can we normalize the size of continuously changing vocal tract?," *Acoust. Sci. & Tech.* (ed. by the Acoustical Society of Japan), 30 (2), pp. 83-88, Mar., 2009. [doi:10.1250/ast.30.83]
- [27] 津崎実, 入野俊夫, 竹島千尋, 松井淑恵, "寸法知覚を中心とした聴覚情景分析 - 物理世界と心理世界をつなぐ聴覚 -," 日本音響学会: 秋季研究発表会講演論文集, pp.1437-1440, 島根大, 島根, 2011年9月20日~22日, (発表日9月21日)
- [28] Toshio Irino, Yoshie Aoki, Hideki Kawahara, and Roy D. Patterson, "Comparison of performance with voiced and whispered speech in word recognition and mean-formant-frequency discrimination," *Speech Commun.*, 54 (9), pp.998-1013, 2012. [doi:10.1016/j.specom.2012.04.002]
- [29] Erika Okamoto, Toshio Irino, Ryuichi Nisimura and Hideki Kawahara, "Evaluation of voice morphing using vocal tract length normalization based on auditory filterbank," *J. Signal Processing (信号処理)*, Vol.15, No. 4, pp.283-286, July, 2011.
- [30] 入野俊夫, 河原英紀, "安定な声道長推定のための聴覚フィルタバンクとその理論," 日本音響学会: 秋季研究発表会講演論文集, pp.505-508, 島根大, 島根, 2011年9月20日~22日.
- [31] Ryuichi Nisimura, Shoko Miyamori, Erika Okamoto, Hideki Kawahara, and Toshio Irino "Detecting child speaker based on auditory feature vectors for VTL estimation," in *Proc. APSIPA 2012 #117*, Hollywood, California, 3-6 Dec., 2012.