九州大学学術情報リポジトリ
Kyushu University Institutional Repository

音声情報の能動的再構築処理に関する研究 : 騒音環境における音声認識を支援するシステム

勝瀬, 郁代

https://doi.org/10.15017/1398257

出版情報：九州芸術工科大学, 2001, 博士（工学）, 課程博士
バージョン：
権利関係：

Bibliography

[1] 白井克彦, 小林哲則, 工藤育男, "花開く音声処理技術," 情報処理学会誌, Vol.38, No.11, 971–975, 1997.

[2] 西村雅史, 伊東伸泰, "音声ワープロ －過去・現在・未来－," 情報処理学会誌, Vol.40, No.2, 164–167, 1999.

[3] 鹿野清宏, 伊藤克亘, 河原達也, 武田一哉, 山本幹雄編著, 音声認識システム, オーム社, 2001.

[4] Kodama, Y., Utsuro, T., Nishizaki, H. and Nakagawa, S., "Experimental Evaluation on Confidence of Agreement among Multiple Japanese LVCSR Models," Proc. of EUROSPEECH2001, 4, 2549–2550, 2001.

[5] 電子情報通信学会編, 電子情報通信ハンドブック第1分冊, オーム社, 1988.

[6] 金田豊, "マイクロホン系におけるディジタルフィルタの応用－不要な音を取り除く技術－," 日本音響学会誌, 45, 2, 125–128, 1989.

[7] Comon, P., "Independent component analysis, A new concept?," Signal Processing, 36, 287–314, 1994.

[8] Cardoso, J-F., "Statistical principles of source separation," proceedings of the SYSID '97, 11th IFAC symposium on system identification, 1837–1843, 1997.

[9] Atal, B. "Effectiveness of Linear Prediction Characteristics of the speech wave for automatic speaker identification and verification," J. Acoust. Soc. Am., 55, 1304–1312, 1974.

[10] Lockwood, P. and Boudy, J., "Experiments with a Non-linear spectral subtractor (NSS), hidden Markov models and the prediction, for robust speech recognition in cars," Proc. of Eurospeech1991, 79–82, 1991.

[11] Graf, J. T. and Hubing, N., "Dynamic time warping for the enhancement of speech degraded by white Gaussian noise," Proc. Int. Conf. Acoust. Speech Signal Process.(ICASSP), II,

339–342, 1993.

[12] Varga, A. P. and Moore, R. K., "Hidden Markov model decomposition of speech and noise," Proc. Int. Conf. Acoust. Speech Signal Process.(ICASSP), 845–848, 1990.

[13] Martin, F., Shikano, K., Minami, Y., "Recognition of noisy speech by composition of hidden Markov models," Proc. of EUROSPEECH1993, 1031–1034, 1993.

[14] Gales, M.J.F. and Young, S.J., "Robust Continuous Speech Recognition using Parallel Model Combination," IEEE Transactions on Speech and Audio Processing 4, 1996.

[15] Bregman, A. S., Auditory Scene Analysis, MIT Press, 1990.

[16] 河原英紀, "聴覚の工学的表現," 電子情報通信学会誌, 76, 11, 1197–1202, 1993.

[17] 河原英紀, "聴覚の情景分析," 日本聴覚医学会学術講演会, 1995.

[18] Warren, W. H. Jr. and Verbrugge, R. R., "壊れたのかバウンドしたのか？音で事象を判断することの心理物理学," ナチュラルコンピュテーション２, ホイットマンリチャーズ編, 石川正俊, 平原達也 訳, パーソナルメディア, I部６章, 82–95, 1994.

[19] 平原達也, "音声工学と聴覚科学," 日本音響学会誌, 53, 9, 714–719, 1997.

[20] 柏野牧夫, "音による環境理解," 第37回東北大通研シンポジウム「高次臨場感通信を目指して」, 157–171, 1997.

[21] Bregman, A. S., "Auditory scene analysis: hearing in complex environments," in Thinking in Sound, S. McAdams and E. Bigand, Eds. (Oxford Univ. Press, New York, 1993), Chap. 2., 1993.

[22] Lachman, R., Lachman, J. L., and Butterfield, E. C. 共著, 箱田祐司, 鈴木光太郎 共訳, 認知心理学と人間の情報処理II -意識と記憶-, 第6章意識と注意, サイエンス社, 1988.

[23] Lindsay, P. H. and Norman, D. A. 共著, 中溝幸夫, 箱田祐司, 近藤倫明 共訳, 情報処理心理学入門I 感覚と知覚 第２版, 第１章人間の知覚, サイエンス社, 1983.

[24] Lachman, R., Lachman, J. L., and Butterfield, E. C. 共著, 箱田祐司, 鈴木光太郎 共訳, 認知心理学と人間の情報処理III -言語と理解-, 第13章パターン認識, サイエンス社, 1988.

[25] Miller, G. A. and Licklider, J. C. R., "The Intelligibility of Interrupted Speech," J. Acoust. Soc. Am., 22(2), 167–173, 1950.

[26] Warren, R. M. "Perceptual restoration of missing speech sounds," Science, 167, 392–393, 1970.

[27] Warren, R. H., "Perceptual Restoration of Obliterated Sounds," Psychological Bulletin, 96, 371–383, 1984.

[28] Amano, S., "Phonemic restoration of a stop consonant in a vowel-consonant-vowel syllable in speeded and nonspeeded tasks," J. Acoust. Soc. Jpn. (E), 15, 4, 261–267, 1994.

[29] 川島尊之, 柏野牧夫, "日本語文章中の単語知覚における後続文脈の促進的効果," 日本音響学会秋季全国大会講演論文集, 1-3-4, 369–370, 1997.

[30] Repp, B. H., "Perceptual restoration of a "missing" speech sound: Auditory induction or illusion?," Perception & Psychophysics, 51, (1), 14–32, 1992.

[31] Samuel, A. G., "The Role of Bottom-Up Confirmation in the Phonemic Restoration Illusion," J. Experimental Psychology: Human Perception and Performance, 7(5), 1124–1131, 1981.

[32] Rubenstein, H. and Pollack, I., "Word Predictability and Intelligibility," J. Verbal Learning and Verbal Behavior, 2, 147–158, 1963.

[33] Kalikow, D. N., Stevens, K. N., and Elliott, L. L., "Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability," J. Acoust. Soc. Am., Vol. 61, 5, 1337–1351, 1977.

[34] Morton, J. and Long, J., "Effect of Word Transitional Probability on Phoneme Identification," Journal of Verbal Learning and Verbal Behavior, 15, 43–51, 1976.

[35] 東倉洋一 "聴覚機構と知覚モデル," 音声・聴覚と神経回路網モデル（甘利俊一監修, 東倉洋一, 鹿野清宏, 中川聖一共著, オーム社）第4章, 1990.

[36] 堂下修司, 音声による人間と機械の対話（堂下修司他編, オーム社）, 第1章 総説, 3–24, 1998.

[37] 岡田美智男, 鈴木紀子, 石井和夫, Altman, E., "雑談の構成的な理解に向けて," 音声言語情報処理, 17-7, 39–44, 1997.

[38] 勝瀬郁代, 津崎実, 相川清明, "連続母音の遷移部の知覚における後続雑音の効果," 日本音響学会誌,53(12), 922–932, 1997.

[39] Ikuyo Masuda-Katsuse, Hideki Kawahara, "Dynamic sound formation based on continuity of spectral change," Speech Communication, 27, 187–207, 1999.

[40] Ikuyo Masuda-Katsuse, "A New Method for Speech Recognition in the Presence of Non-stationary, Unpredictable and High-level Noise," Proc. of EUROSPEECH2001, Vol.2, 1119–1122, 2001.

[41] Masuda-Katsuse, I. and Sugano, Y., "Speech Estimation biased by phonemic expectation in the presence of non-stationary and unpredictable noise," Proc. of a workshop on Consistent and reliable acoustic cues for sound analysis, 2001.

[42] 勝瀬郁代, 高橋雅仁, 寺岡章人, 岸田哉生, 福田興一郎, 長迫挙志, "自律移動ロボットを遠隔制御するために音声対話インタフェースの開発," ヒューマンインタフェース学会論文誌, 3(2), 111-120, 2000.

[43] Klassner, F., Lesser, V. and Nawab, S. H., "The IPUS blackboard architecture as a framework for computational auditory scene analysis," in Computational Auditory Scene Analysis, D.F.Rosenthal and H. Okuno (Eds.), Mahwah, NJ, Lawrence Erlbaum, 105–114, 1998.

[44] 中谷智広, 後藤真孝, 川端豪, 奥乃博, "残差駆動型アーキテクチャの提案と音響ストリーム分離への応用," 人工知能学会誌, 12(1), 111–120, 1997.

[45] Cooke, M., "Modelling Auditory Processing and Organisation," Ph.D. Thesis, University of Sheffield, (1991), Published by Cambridge University Press, 1993.

[46] Brown, G. J., "Computational Auditory Scene Analysis: A Representational Approach," Ph. D. Thesis, Department of Computer Science, University of Sheffield, 1992.

[47] Brown, G. J., "Computational auditory scene analysis," Computer Speech and Language, 8, 297–336, 1994.

[48] Warren, R. M., Obusek, C. J., and Ackroff, J. M., "Auditory induction: Perceptual synthesis of absent sounds," Science, 176, 1149–1151, 1972.

[49] Kurakata, K., Matsui, M., and Nishimura, A., "Perceived trajectory of gliding tones though interrupting noise and the underlying frequency-extrapolation mechanism," Proceedings of the papers submitted to the Acoustical Society of Japan in ASA and ASJ third joint meeting, 695–700, 1995.

[50] Simons, M., "Time-swelling: An incomplete heterophonic induction in a non-repetitive design," Master's thesis of Leiden Univ., 1995.

[51] Kluender, K. R. and Jenison, R. L., "Effects of glide slope, noise intensity, and noise duration on the extrapolation of FM glides through noise," Perception & Psychophysics, 51 (3), 231–238. 1992.

[52] 相川清明, 津崎実, 河原英紀, 東倉洋一, "周波数変化音追跡の動特性," 日本音響学会誌, 52 (10), 741–751, 1996.

[53] 相川清明, 河原英紀, "複合音周波数変化音追跡神経演算モデル," 日本音響学会誌, 53(2), 95–102, 1997.

[54] 赤木正人, 古井貞熙, "音声知覚における母音ターゲット予測機構のモデル化," 信学論, J69-A, 1277–1285, 1986.

[55] Massaro, D. W. "Backward recognition masking," J. Acoust. Soc. Am., 58, 1059–1065, 1975.

[56] 岩澤秀紀, "母音知覚における聴覚逆向認知マスキング－音韻判断の場合－," 日本音響学会平成8年秋季研究発表会講演論文集第一分冊, 377–378, 1996.

[57] Klatt, D. H., "Software for a cascade / parallel formant synthesizer," J. Acoust. Soc. Am. 67 (3), 971–995, 1980.

[58] 田中良久, "数値尺度法," 心理学的測定法　第2版, 東京大学出版会, 東京, 117–118, 1977.

[59] 広中平祐, 現代数理科学辞典, 大阪書籍, 508–510, 1991.

[60] 森敏昭, 吉田俊夫, "2要因の分散分析, 5.2 要因とも対応がある場合," 心理学のためのデータ解析テクニカルブック　第3章　第2節, 北大路書房, 116–121, 1990.

[61] Darwin, C.J., "The relationship between speech perception and the perception of other sounds," In Modularity and the Motor Theory in Speech Perception (Matteingly, I.G. and Kennedy, M.S. Eds.), Lawrence Erlbaum, Hillsdale, NJ, 239–259. 1991.

[62] Akagi, M., Wieringen, A., and Pols, L., "Perception of central vowel with pre- and post-anchors," Proc. ICSLP 94, S10-7, 503–506, 1994.

[63] Brady, P.T., House, A.S., Stevens, K.N., "Perception of sounds characterized by a rapidly changing resonant frequency," J. Acoust. Soc. Am., 33(10), 1357–1362, 1961.

[64] Kawahara, H., "Speech representation and transformation using adaptive interpolation of weighted spectrum: VOCODER revised," Proc. of ICASSP 97, SPCH5P.7, Vol.2, 1303–1306, 1997.

[65] Kawahara,H., Masuda-Katsuse, I., and de Cheveigné, A., "Restructuring speech representations using a pitch-adaptive time-frequency smoothing and instantaneous-frequency based F0 extraction: possible role of a repetitive structure in sounds," Speech Communication, 27, 187–207, 1999.

[66] de Cheveigné, A., McAdams, S., Laroche, J., and Rosenberg, M., "Identification of concurrent harmonic and inharmonic vowels: A test of the theory of harmonic cancellation and enhancement," J. Acoust. Soc. Am., 97 (6), 3736–3748, 1995.

[67] de Cheveigné, A., "Concurrent vowel identification III. A neural model of harmonic interference cancellation," J. Acoust. Soc. Am., 101(5), 2857–2865, 1997.

[68] 平原達也, "母音知覚と聴覚系内スペクトル表現," 日本音響学会春季大会講演論文集, 2-7-13, 341–342, 1995.

[69] Divenyi, P. L. and Algazi, A. P., "Auditory segregation of simultaneous events based on residue pitch and spectrum," J. Acoust. Soc. Am., 100(4), Pt. 2, p.2682, 1996.

[70] 東山恵祐, 陸金林, 中村哲, 鹿野清宏, 河原英紀, "4kHz 帯域のSTRAIGHTの品質評価と情報圧縮について," 日本音響学会秋季大会予稿集1-2-6, 207–208, 1997.

[71] Cooke, M. et al., "Robust automatic speech recognition with missing and unreliable acoustic data," Speech Communication 34(3), 2001.

[72] de Cheveigné, A. and Kawahara, H., "Missing-data model of vowel identification," J. Acoust. Soc. Am., 105(6), 3497–3508, 1999.

[73] 後藤真孝, "音楽音響信号を対象としたメロディーとベースの音高推定," 信学論D-II, J84(1), 12–22, 2001.

[74] Goto, M., "A predominant-F0 estimation for CD recordings: MAP estimation using EM algorithm for adaptive tone models", Proc. of ICASSP2001, 2001.

[75] Kawahara, H. et al., "Fixed point analysis of frequency to instantaneous frequency mapping for accurate estimation of F0 and periodicity," Proc. of Eurospeech99, 2781–2784, 1999.

[76] 板倉秀一, "騒音データベースと日本語共通音声データDAT版," 日本音響学会誌47, 2, 951–953, 1991.

[77] Hain, T., Woodland, P.C., Evermann, G., and Povey, D., "The CU-HTK March 2000 HUB5E Transcription system," Proc. Speech Transcription Workshop 2000, 2000.

[78] Samuel, A.G., "Lexical uniqueness effects on phonemic restoration," J. Mem. Lang.26, 36–56, 1987.

[79] de Saussure, F., "Course in General Linguistics Ferdinand de Saussure," Chapter.3, 7–15, Bally, C. B. and Sechehaye A. (Eds), First McGraw-Hill Paperback Edition, 1966.

[80] 中川聖一, 傳田明弘, 伊藤敏彦, "マルチモーダル観光案内対話システム," 人工知能学会誌, 13(2), 241–251, 1998.

[81] 長田誠也, 土井伸一, 亀井真一郎, "自然言語を用いて家庭機器操作を行う対話システム," 信学技報SP98-71, 23–30, 1998.

[82] 竹林洋一, "音声自由対話システムTOSBURG II," 電子情報通信学会論文誌D-II, J77(8), 1417–1428, 1994.

[83] 西村靖史, 伊東伸泰, "音声認識技術の基礎と応用プログラム開発手法," インターフェース, 8, 100–108, 1998.

[84] 中川聖一, 堂下修司, "音声言語情報処理の現状と研究課題," 情報処理 36(11), 1995.

[85] 堂下修司, 新美康永, 白井克彦, 田中穂積, 溝口理一郎 共編, 音声による人間と機械の対話, オーム社, 1998.

[86] 長尾真 編, 自然言語処理, 岩波書店, 1996.

[87] 藤田義弘, "パーソナルロボットR100," 日本ロボット学会誌, 18(2), 198–199, 2000.

[88] 藤田雅博, "エンタテインメントロボット：AIBO," 映像情報メディア学会誌, 54(5), 657–661, 2000.

[89] Engelberger, J., "A day in the life of Isaac," Industrial Robot, 27(3), 176–180, 2000.

[90] Schraft, R. D., "Care-O-bot: The concept of a system for assisting elderly or disabled persons in home environments," Proceedings of the IEEE IECON, 4, 2476–2481, 1998.

[91] トマスJモーブレイ, ラファエルCマルボー著, "CORBA Design Pattern アーキテクチャからプログラミングまで", IDGコミュニケーションズ, 1998.

[92] 木室義彦, 高野茂, 徳島新一, 松本三千人, "顔認識と音声認識を組み合わせた移動ロボットナビゲーション," 日本機械学会 ロボティクス・メカトロニクス講演会(Robomec'00) 1A1-26-025, 2000.

[93] Fraser, N. M. et.al., "Simulating speech systems," Computer Speech and Language, 5(1), pp.81-99, 1991.

[94] 松本祐治 他, "日本語形態素解析システムJUMAN使用説明書ver.2," NAIST Technical Report, NAIST-IS-TR94025, 1994.

[95] 荒木雅弘, 堂下修司, "対話の自由度と発話の自由度に基づく音声対話システムの設計法," 人工知能学会研究資料 SIG-SLUD-9502-6, 1995.

[96] 西村雅史, 伊東伸泰, "音声認識の最新技術," bit 7, pp.8-13, 1998.

[97] 宮崎淳史, "音声認識プログラムVia Voiceの技術と今後," bit Vol.31, No.11, 1999.

[98] Saito, T. et.al., "ProTALKER: a Japanese Text-To-Speech System for personal Computers," IBM TRL Research Report, RT0110, June, 1995.

[99] 坂本憲治, 綿貫啓子, 外川文雄, "マルチモーダル対話解析," 人工知能学会研究会資料SIG-SLUD-9401, 39–46, 1994.

[100] 小林 豊, 中島大輔, 新美康永, "音声模擬対話における対話制御と快適性について," 情報処理学会研究報告, SLP 6-2, 1995.

[101] 古井貞熙, 前川喜久雄, 井佐原均, "科学技術振興調整費開放的融合研究推進制度 －大規模コーパスに基づく「話し言葉工学」の構築－," 日本音響学会誌, 56(11), 752–755, 2000.