

平成 30 年 6 月 18 日現在

機関番号：17102

研究種目：基盤研究(A) (一般)

研究期間：2013～2017

課題番号：25242002

研究課題名(和文) 公共空間における音響放送の改善：知覚的相互作用を考慮した音デザイン

研究課題名(英文) Improvement of speech signals in public spaces: Acoustic design from a viewpoint to optimize perceptual interactions

研究代表者

中島 祥好 (Nakajima, Yoshitaka)

九州大学・芸術工学研究院・教授

研究者番号：90127267

交付決定額(研究期間全体)：(直接経費) 31,000,000円

研究成果の概要(和文)：音声伝達環境に関する従来の技術は、個々の告知放送や環境に対する粗い印象に基づく場合が多い。しかし、告知放送においては内容が正しく伝わるのが何よりも重要であり、騒音や残響に対して頑強な音声信号を作らなければならない。ここでは、聴覚心理学、音声科学、神経生理学の観点から、基礎研究と応用研究の両方を行った。音声のスペクトル変化を多変量解析によってただか4個の因子によって表すことに成功し、因子得点の時間変化のみに基づいて、充分明瞭な雑音駆動音声再合成することができた。600-1700 Hz の近傍に関係する因子は、音韻論において「鳴音性 sonority」と呼ばれる音素の性質に対応している。

研究成果の概要(英文)：Conventional technologies related to speech transmission are often based on crude impressions of speech announcements or environments. However, the most important thing about speech announcements is whether the contents are conveyed correctly, and it is desired to develop speech signals resilient against noise and reverberation. We performed basic and applied studies from the viewpoints of auditory psychology, speech science, and neurophysiology. Our multivariate analyses revealed that four factors are sufficient to express essential spectral fluctuations of speech: Noise-vocoded speech resynthesized from the four factors was highly intelligible. One of the factors related to a frequency range around 600-1700 Hz corresponded to what is called 'sonority' in phonology.

研究分野：聴覚心理学

キーワード：音声強調 音声情報圧縮 モザイク音声 時間分解能 多変量解析 音韻論 脳波 鳴音性

1. 研究開始当初の背景

列車の駅、空港、大型店舗、遊戯施設などにおいて、さまざまな告知放送(アナウンス)がスピーカーから流されているが、騒音や残響のために内容を聞き取りにくいことが多い。また、他の告知放送、背景音楽、さまざまな告知音(サイン音)が同時に鳴らされるために、マスキングが生じ、聞き取りが一層困難になっていることが多い。このような場合、告知放送の音量が上げられ、それに対抗して他の放送音の音量も上げざるをえない。その結果、歪みによる音質の劣化と、騒音の増大とが同時に起こり、音環境が悪化すると考えられる。また、防災放送を想定した音響設備は音質が悪くとも、高周波成分に対する感度が低下している高齢者にも、告知内容が伝わるような工夫が必要である。さらには、災害時などの緊急を要する状況において、補聴器装用ができない条件でも聞き取りが可能な音声でなければならない。これらの点に関して体系的な研究はなされていない。

音環境に関する従来の技術は個々の告知放送、告知音、環境全体に対する主観的印象を中心とする場合が多いが、告知放送においては、告知内容が正しく伝わるのが重要である。そのため、騒音、残響に対して、頑強な音声信号を作ることが必要である。これについて、本研究グループが聴覚心理学・音声学・神経生理学の研究に基づいて開発した独自技術をもって、解決可能であると考え。

2. 研究の目的

公共空間においては、さまざまな告知放送がなされているが、騒音や残響、さらには他の告知放送、背景音楽、告知音が同時に聴こえることで、知覚的な相互作用が働き、内容を聞き取ることが困難であることが多い。これによって音響放送の音量の増加を引き起こし、ますます音環境を悪化させている。高齢者にも若齢者にも快適で、必要な情報を聞き取ることができるような音環境を作るために、人間の聴覚の仕組みを熟知することによって、「音量をあげる」という安易な方法に頼らずに、聞き取りやすさを改善する方法を見つける必要がある。そこで、これまでに開発してきた聴覚心理学の知見に基づく技術と従来の技術とを組み合わせることで、上述の問題を解決することを目指した。

3. 研究の方法

聴覚の仕組みについての基礎研究と、応用研究とを分けて行った。基礎研究では、音理解に必要な聴覚の時間分解精度と周波数分解精度とを定量的に測定するために、音声の時間軸方向と周波数軸方向に対して、モザイク処理を施し、これを使って聴取実験を行った。さらに、残響・騒音を付加した状況下での実験に発展させているところである。この実験結果を、音声の聞き取りを改善するために必要な信号加工技術に関連付ける。応用

研究では、独自技術を用いて加工した音声を使い、音声知覚の実験を行った。また、多重録音再生装置を用い、同時に複数の信号を鳴らし、その状況下での聞き取りやすさを調べた。今後、実験室での実験を経て、実環境での実験までを試みる。時間構造を有する聴覚刺激に対してどのような活動が行われるか、神経生理学の観点からも解析を行った。純正基礎研究・応用研究を統合し、音響放送が最も聞き取りやすくなる信号処理法について考察した。

4. 研究成果

騒音、残響、他の音声信号との混合に対して、音声信号を知覚のうえで頑強にする技術を開発した。この技術を利用して、音声信号を1秒あたり数キロビット以下の情報量に圧縮しても十分な明瞭度を得ることが可能になった(特許出願)。また、これまでに開発した子音強調、帯域圧縮、帯域パワー増分強調の手法を組み合わせ、最適の条件を探るプログラムをほぼ完成させた。音声信号に含まれる最重要な情報を明らかにするために、毎秒3000ビット未満の伝送率で十分明瞭な音声を伝える技術を完成させた。これは、20個程度の臨界帯域に分かれるパワー変化の情報を数個の因子の時間変化に縮約し、これに基本周波数の有無・高低の情報を付加するものである。

このような、情報圧縮の基礎研究として、音声知覚に必要な時間分解精度、周波数分解精度を測定するために、時間、周波数の座標上でパワーをモザイク化した音声、すなわち「モザイク音声」を作成し、時間分解精度に焦点を当てて、聴取実験を行った(図1)。

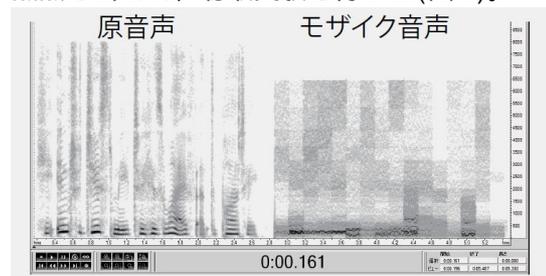


図1 原音声(左)とモザイク音声(右)

その結果、日本語音声を聞き取るために必要な時間分解精度は、40 ms程度であることが判った。時間方向に20 ms程度の精度が得られ、周波数方向に適切な10帯域が得られれば、日本語のモーラ正答率にして95%以上が達成されるが、時間分解精度、周波数分解精度のいずれが落ちてても音声の明瞭度が低下する。明瞭度がゼロに近付くまでは、その傾向が続く。英語音声についても処理時間幅が40 ms程度であれば十分な明瞭度の得られる場合が多いが、これについては参加者が英語母語話者かどうかによって多少の変動のあることが示された。

音声信号のスペクトル変化の様子を多変

量解析によって捉える研究を、従来行っていたものからさらに発展させた。ケプストラム分析によるスペクトルの平滑化、因子分析を音響パワーがゼロになる状態を起点として行うことなどの改良を加え、常に安定した4因子が得られるようになった。

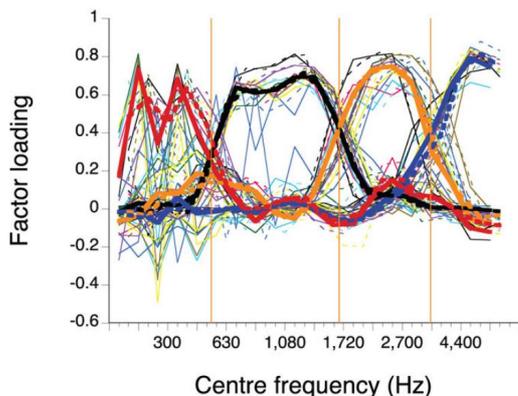


図2 因子分析によって4因子を取り出した結果である。中程度の帯域が「鳴音性 sonority」と関連がある。

このことにより、分析結果を音韻論の立場から解釈することが容易になった。たとえば、子音の鳴音性（後述）が知覚されるためのスペクトル変化の条件を特定することができた。今回は因子分析の結果から雑音駆動音声を再合成し、その明瞭度を求めた。この際、因子分析の方法を改良し、無音の箇所は必ず無音として再合成されるようにすることで、聴取実験の妥当性を高めた。因子の数が2から3に増加するときに明瞭度が飛躍的に増加し約70%になることが判った。因子負荷量のパターンを、以前の分析結果と合わせて見ることにより、3-4個の因子が音声コミュニケーションの基本をなすことが結論付けられた。この因子分析法をさらに改良し、因子負荷量、因子得点のいずれも非負の値とし、因子負荷量のベクトルが直交することは崩さないようにしたところ、上記の傾向が一層はっきりと現れた。これらの因子のうち、540-1700 Hzの周波数帯域に当たるものが、音韻論において「鳴音性 sonority」と呼ばれる音素の性質に対応することが確かめられた(図2)。鳴音性は音節の産出、知覚に密接に関係していると考えられ、そのことを考慮して音声強調のパラメーターを調整することが可能となった。

さらに、ささやき声と通常発声による音声との定量的な比較ができるようになった。加えて、因子分析によって得られた因子得点から、雑音駆動音声を合成することが可能になった。

音声のパワー変化から導かれる因子の因子得点の時間変化が、音声知覚とどのように関連付けられるのか検討した結果、日本語朗読文の場合、音声信号を2つの因子の時間変

化に縮約したときにはほとんど言語内容が把握できないのに対して、3つの因子の場合には、言語内容の約80%を聞き取ることができた。さらに、調音運動について、先進的な研究環境を有するトロント大学との共同研究を行い、調音運動のデータと音声データとを同時に得て比較することができた。

音声強調、音声情報圧縮に関して新しい技術を開発することができた。これは、音声情報を音源情報とフィルター情報とに分ける古典的な考え方を、徹底して単純化したことで可能となった。局部時間反転音声の知覚についての研究、および局部時間反転音声を用いた無関連音効果（視覚呈示を用いた短期記憶による系列再生課題に、無関連音が及ぼす妨害効果）についての研究を進めた。

音声コミュニケーションに関わる脳活動（脳波、脳磁図）の計測を行い、吃音症患者において、左右の聴覚野の機能連関が健常者に比べて強いことなどを示した。また、自閉症児における音声に対する脳反応の特徴について調べた。健常人を対象として、視覚・聴覚・体性感覚モダリティーにおける心理課題の負荷とその遂行中の脳活動を計測した。主な研究成果として、痛みの共感時には、運動感覚野および痛み情動関連域で脳の神経振動が周波数特異的に変化すること、聴覚の時間縮小錯覚課題を用いる状況において、脳の中の時計が右半球に認められ、側頭頂接合部、下側頭回が重要であること、聴覚の半球側方性は、刺激に対して、時間・周波数依存的に変化すること、などが挙げられる。

現在は、これまでに得た知見と、過去の知見とを総合し、協力企業とともに実際の雑音・残響環境で音声を放送する実証実験を手がかりにして、産学連携による起業の準備を進めている。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計71件)

Nakajima, Y., Ueda, K., Fujimaru, S., Motomura, H., & Ohsaka, Y., (2017).

“English phonology and an acoustic language universal,” *Scientific Reports*, 7, 46049. 10.1038/srep46049 (査読あり)

Ueda, K., & Nakajima, Y. (2017).

“An acoustic key to eight languages/dialects: Factor analyses of critical-band-filtered speech,” *Scientific Reports*, 7:42468. 10.1038/srep42468. (査読あり)

Komiya, R., Saitoh, T., Fuyuno, M., Yamashita, Y., & Nakajima, Y. (2017).

“Head pose estimation and motion analysis of public speaking videos,” *International Journal of Software*

- Innovation (IJSI), 4, 57-71.
 10.4018/IJSI.2017010105 (査読あり)
Remijn, G. B., Kikuchi, M., Shitamichi, K., Ueno, S., Yoshimura, Y., Tsubokawa, T., Kojima, H., Higashida, H., & Minabe, Y. (2017). "A NIRS study on cortical hemodynamic responses to normal and whispered speech in 3- to 7-year-old children.," *Journal of Speech, Language and Hearing Research*, 60, 465-470. (査読あり)
- Chatani, H., Hagiwara, K., Hironaga, N., Ogata, K., Shigeto, H., Morioka, T., Sakata, A., Hashiguchi, K., Murakami, N., Uehara, T., Kira, J.-I., & Tobimatsu, S. (2016). "Neuromagnetic evidence for hippocampal modulation of auditory processing.," *Neuro Image*, 1:124, 256-266. 10.1016/j.neuroimage.2015.09.006 (査読あり)
- Kishida, T., Nakajima, Y., Ueda, K., & Remijn, G. B. (2016). "Three factors are critical in order to synthesize intelligible noise-vocoded Japanese speech.," *Frontiers in Psychology*, 7. 10.3389/fpsyg.2016.00517 (査読あり)
- Morimoto, S., Remijn, G. B., & Nakajima, Y. (2016). "Computational-model-based analysis of context effects on harmonic expectancy.," *PLOS ONE*, 11:3, 1-17. 10.1371/journal.pone.0151374. (査読あり)
- Hasuo, E., Nakajima, Y., Wakasugi, M., Fujioka, T. (2016). "Effects of sound marker durations on the perception of inter-onset time intervals: a study with instrumental sounds," *基礎心理学研究 (Japanese Journal of Psychonomic Science)*, 34:1, 2-16. <http://doi.org/10.14947/psychono.34.2>. (査読あり)
- Ellermeier, W., Kattner, F., Ueda, K., Doumoto, K. Nakajima, Y. (2015). "Memory disruption by irrelevant noise-vocoded speech: Effects of native language and the number of frequency bands.," *Journal of the Acoustical Society of America*, 138, 1561-1569. <http://dx.doi.org/10.1121/1.4928954>. (査読あり)
- Mitsudo, T., Nakajima, Y., Takeichi, H., & Tobimatsu, S. (2014). "Perceptual inequality between two neighboring time intervals defined by sound markers: correspondence between neurophysiological and psychological data.," *Frontiers in Psychology*, 5;937, 1-12. 10.3389/fpsyg.2014.00937. (査読あり)
- Nakajima, Y., Hasuo, E., Yamashita, M. & Haraguchi, Y. (2014). "Overestimation of the second time interval replaces time-shrinking when the difference between two adjacent time intervals increases.," *Frontiers in Human Neuroscience*, 8:281, 1-12. 10.3389/fnhum.2014.00281. (査読あり)
- Nakajima, Y., Sasaki, T., Ueda, K., & Remijn, G. B. (2014). "Auditory Grammar.," *Acoustics Australia*, 42: 2, 97-101. (査読あり)
- 飛松省三 (2014). "脳波リズムの発現機序臨床神経生理学," 42, 358-363. (査読なし)
- Hasuo, E., Nakajima, Y., Tomimatsu, E., Grondin, S., & Ueda, K. (2014). "The occurrence of the filled duration illusion: A comparison of the method of adjustment with the method of Magnitude estimation," *Acta Psychologica*, 147, 111-121. 10.1016/j.actpsy.2013.10.003. Epub 2013 Nov 5. (査読あり)
- Yoshimura, Y., Kikuchi, M., Shitamichi, K., Ueno, S., Munesue, T., Ono, Y., Tsubokawa, T., Haruta, Y., Oi, M., Niida, Y., Remijn, G. B., Takahashi, T., Suzuki, M., Higashida, H., & Minabe, Y. (2013). "Atypical brain lateralisation in the auditory cortex and language performance in 3- to 7-year-old children with high-functioning autism spectrum disorder: a child-customised magnetoencephalography (MEG) study.," *Molecular Autism*, 4, 38. 10.1186/2040-2392-4-38 (査読あり)
- Yoshimura, Y., Kikuchi, M., Ueno, S., Okumura, E., Hiraishi, H., Hasegawa, & C., Remijn, G. B., Shitamichi, K., Munesue, T., Tsubokawa, T., Higashida, H., & Minabe, Y. (2013). "The brain's response to the human voice depends on the incidence of autistic traits in the general population.," *PLOS ONE*, 8, e80126. 10.1371/journal.pone.0080126 (査読あり)
- Yamashita, Y., Nakajima, Y., Ueda, K., Shimada, Y., Hirsh, D., Seno, T., & Smith, B. A., (2013). "Acoustic analyses of speech sounds and rhythms in Japanese- and English-learning infants," *Frontiers in Psychology*, 4:57. 10.3389/fpsyg.2013.00057 (査読あり)
- 飛松省三 (2013). "医学と心理学をつなぐ

脳生理学のアプローチ,” 基礎心理学研究 32, 88-93.(査読なし)

[学会発表](計 128 件)

Nakajima, Y., Ueda, K., Remijn, G. B., Yamashita, Y., & Kishida, T.(2017). “Phonology and psychophysics: Is sonority real?,” The 33rd Annual Meeting of the International Society for Psychophysics.

Nakajima, Y., Matsuda M., Ueda, K., & Remijn, G.B.(2017). “Temporal resolution needed for auditory communication: Measurement with mosaic speech,” The 33rd Annual Meeting of the International Society for Psychophysics.

Tobimatsu, S.(2017). “Functional significance of neural oscillations in humans.,” 6th International Conference on Cognitive Neurodynamics.

岸田拓也, 中島祥好, 上田和夫, Gerard B. Remijn, 中尾 貴志(2017). “臨界帯域パワー変動因子を用いた雑音駆動音声の合成: 因子の除去が明瞭度に与える効果,” 日本音響学会聴覚研究会.

Ueda, K., Nakajima, Y., Ellermeier, W., & Kattner, F.(2017). “Intelligibility of locally time-reversed speech: A multilingual comparison with normalizing speech rates,” 日本音響学会春季研究発表会.

仲田愛子, 上田和夫, Gerard B. Remijn, 中島祥好(2016). “モザイク処理した刺激における音声と雑音の分離,” 日本音響学会聴覚研究会.

Nakajima, Y.(2016). “Perceptual interactions between adjacent time intervals marked by sound bursts.,” 5th Joint Meeting of the Acoustical Society of America and Acoustical Society of Japan (招待講演).

Fujihira, H., Shiraishi, K., & Remijn, G.B.(2016). “Auditory brainstem responses to anechoic and reverberant two-syllable speech sounds in elderly listeners.,” 5th joint meeting of the Acoustical Society of American and Acoustical Society of Japan.

Nakajima, Y., Matsuda, M., Tomimatsu, E., & Hasuo, E.(2016). “Perceptual contrast between two short adjacent time intervals marked by clicks,” 31st International Congress of Psychology

Nakao, K., Kishida, T., & Nakajima, Y.(2016). “Perceptual validity and analytical advantages of non-negative

bases extracted from Factor Analyses of Japanese Speech,” 31st International Congress of Psychology.

Kishida, T., Nakajima, Y., Ueda, K., & Remijn, G. B.(2016). “Effects of factor elimination on intelligibility of noise-vocoded Japanese speech.,” 31st International Congress of Psychology.

Ueda, K. Nakajima, Y., Tamura, S., Ellermeier, W., Kattner, F., Daebler S., & Neo, N. N.(2016). “Intelligibility of locally time-reversed speech in Chinese, English, German, and Japanese,” 31st International Congress of Psychology.

Kojima, K., & Nakajima, Y.(2016). “Influence of the temporal unit duration on the intelligibility of English mosaic speech,” 31st International Congress of Psychology.

Fuyuno, M. Yamashita, Y., Komiya, R., Saitoh, T., & Nakajima, Y.(2016). “Multimodal corpus approach to speech psychology: the effect of public speakers’ facial movements on audience reception,” 31st International Congress of Psychology.

Kishida, T., Nakao, K., & Nakajima, Y.(2016). “Origin-shifted principal component analysis: A method suitable for reconstructing non-negative data,” 31st International Congress of Psychology.

中島祥好(2016). “音楽聴取における知覚体制化,” 第 18 回日本ヒト脳機能マッピング学会 (招待講演).

Kishida, T., Nakajima, Y., Ueda, K., & Remijn, G. B.(2015). “A critical number of power-fluctuation factors needed for Japanese noise-vocoded speech perception.,” Proceedings of the 31st Annual Meeting of the International Society for Psychophysics.

Ueda, K., Nakajima, Y., Tamura, S., Shichida, A., Ellermeier, W., Kattner, F., & Daebler, S.(2015). “The effect of segment duration on the intelligibility of locally time-reversed speech: A multilingual comparison.,” Proceedings of the 31st Annual Meeting of the International Society for Psychophysics.

Ueda, K., Nakajima, Y., Tamura, S., Ellermeier, W., Kattner, F., Daebler, S., & Neo, N. N.(2015).

- “Intelligibility of locally time-reversed speech in four different languages.” Proceedings of the 2nd Annual Meeting of the Society for Bioacoustics.
- Remijn, G. B. (2015). “Introduction to near-infrared spectroscopy: Hemodynamic blood flow patterns in the human brain in response to sound.” the Japan Audiology Society, 10th ERA/OAE Research meeting (招待講演).
- 21 松尾行雄 (2015). “音響学におけるバイオミメティクス,”バイオミメティクス研究会 (招待講演).
- 22 Kishida, T., Nakajima, Y., Ueda, K., & Remijn, G. B. (2015). “Demonstrations of the perceptual roles of power-fluctuation factors in Japanese and English speech,” The International Five-Sense Symposium.
- 23 Remijn, G. B. (2015). “About time processing and processing time in the auditory and visual modality,” The International Five-sense symposium.
- 24 Kishida, T., Nakajima, Y., Ueda, K., & Remijn, G. B. (2014). “Perceptual Roles of Power-Fluctuation Factors in Speech Perception: A New Method of Factor Analysis,” 日本音響学会聴覚研究会.
- 25 中島祥好, 佐々木隆之, 上田和夫, Remijn, G. B. (2014). “聴覚の文法: 聴覚体制化と聴覚コミュニケーション,” 日本心理学会.
- 26 上田和夫, 中島祥好, 佐々木隆之, G. B. Remijn (2014). “聴覚の文法: 音声知覚への応用,” 日本音響学会春季研究発表会.
- 27 Nakajima, Y., Sasaki, T., Ueda, K., & Remijn, G. B. (2014). “Auditory Grammar in Music,” ICMP13 - APSCOM 15.
- 28 Nakajima, Y., Matsuda, M., Remijn, G. B., & Ueda, K. (2014). “Temporal resolution needed to hear out Japanese morae in mosaic speech,” 日本音響学会聴覚研究会.
- 29 中島祥好, 佐々木隆之, 上田和夫, G. B. Remijn (2014). “聴覚の文法,” 日本音響学会春季研究発表会.
- 30 中島祥好 (2013). “音声信号の騒音、残響に対する耐性を増す: 聴覚の特性を考慮したアルゴリズム Simple algorithms to enhance speech signals against noise and reverberation: a proposal based on basic auditory research,” 九州大学新技術説明会.
- 31 Nakajima, Y., Takeichi, H., Mitsudo, T., & Tobimatsu, S. (2013). “Perceptual processing of pairs of acoustically marked time intervals:

Correspondence between psychophysical and electrophysiological data,” 29th Annual Meeting of the International Society for Psychophysics.

- 32 Nakajima, Y., Ueda, K., Fujimaru, S., & Ohsaka, Y. (2013). “Sonority in British English,” 21st International Congress on Acoustics, 165th Meeting of the Acoustical Society of America, 52nd Meeting of the Canadian Acoustical Association.

〔図書〕(計 12 件)

Tobimatsu, S. & Kakigi, R. (2016). “Clinical Applications of Magnetoencephalography,” Springer, 330pp.

中島祥好, 佐々木隆之, 上田和夫, ジェラード B. レメイン (2014). 『聴覚の文法(音響サイエンスシリーズ 8)』コロナ社, 169pp.

〔産業財産権〕

出願状況 (計 1 件)

名称: 音声のコーティング・デコーディング方法
 発明者: 松尾行雄、中島祥好、上田和夫、岸田拓也
 権利者: 東北学院大学
 種類: 特願
 番号: 2017-038988
 出願年月日: 2017 年 3 月 20 日
 国内外の別: 国内

6. 研究組織

(1) 研究代表者

中島 祥好 (NAKAJIMA, YOSHITAKA)
 九州大学・芸術工学研究院・教授
 研究者番号: 90127267

(2) 研究分担者

飛松 省三 (TOBIMATSU, SHOZO)
 九州大学・医学研究院・教授
 研究者番号: 40164008

松尾 行雄 (MATSUO, IKUO)
 東北学院大学・教養学部・教授
 研究者番号: 40323117

レメイン ジェラード (REMIJN, GERARD B.)
 九州大学・芸術工学研究院・準教授
 研究者番号: 40467098

上田 和夫 (UEDA, KAZUO)
 九州大学・芸術工学研究院・準教授
 研究者番号: 80254316