

後藤 真孝

(独)産業技術総合研究所 情報技術研究部門
首席研究員

コンテンツ共生社会のための類似度を可視化する情報環境の実現

§ 1. 研究実施体制

(1)「後藤」グループ

① 研究代表者:後藤 真孝 (産業技術総合研究所 情報技術研究部門、首席研究員)

② 研究項目

・音楽を中心とした類似度可視化情報環境の実現と全体統括

(2)「森島」グループ

① 主たる共同研究者:森島 繁生 (早稲田大学 理工学術院、教授)

② 研究項目

・音楽連動動画を中心とした類似度可視化情報環境の実現

(3)「中村」グループ

① 主たる共同研究者:中村 聡史 (明治大学 総合数理学部、准教授)

② 研究項目

・Web インタラクションを中心とした類似度可視化情報環境の実現

§ 2. 研究実施の概要

本研究は、音楽や動画のようなメディアコンテンツを豊かで健全に創作・利用する「コンテンツ共生社会」の実現へ向けて、音楽を中心とした研究開発によって、膨大なメディアコンテンツ間の類似度を人々が知ることができる(可知化する)情報環境のための技術基盤を構築している。平成 25 年度は、より一層対外的に見える形で研究成果を創出しながら、本研究プロジェクトを進めた。

本研究が目指す「コンテンツ生態系 Web サービス」の関係性理解支援機能として、コンテンツ間の関係性を可視化する音楽鑑賞支援 Web サービス「Songrium (ソングリウム)」^{19), 27)}を実現し、平成 25 年 8 月には一般公開をして実証実験を開始した。その研究成果は産業技術総合研究所と JST が共同でプレス発表「ウェブ上の音楽コンテンツの関係性を可視化する音楽視聴支援システム」をした。Songrium では動画共有サービス上の音楽動画 60 万件の関係性を抽出し可視化することに成功し、派生関係を可視化する機能や、楽曲間の関係性をユーザが自由に追加できる機能、曲調や歌声の特徴に基づく楽曲群や派生作品群の可視化機能を実現した。

これは平成 24 年度の成果の能動的音楽鑑賞サービス「Songle (ソングル)」^{1), 31)}と連携して実現したが、Songle の方も拡張を続け、登録楽曲を 6000 曲から 60 万曲に大規模化させ、実世界デバイス制御にも取り組んだ。時系列的に変化する複雑な混合音を自動理解する技術においても基礎的かつ汎用的な研究成果を得た^{15), 16), 23), 24)}。例えば無限半正定値テンソル分解 PSDTF^{16), 23)}は、従来の非負値行列分解 NMF よりも優れた音源分離品質を持つ自然な拡張となっている。音楽において重要な歌声を扱う歌声情報処理に関する研究も幅広く進めた^{9), 17), 20), 21), 29), 30), 32)}。

さらに、音楽連動動画に関しては、鑑賞・創作支援技術確立に向けて、要素技術となるフォトリアルな顔画像の合成方式、ダンス動作制御方式、顔印象類似検索に関する検討を行った。発話に同期した唇形状のみならず、口内(歯や舌)の表現に関する検討を行い、小規模なデータベースのみで実写クオリティの発話アニメーション生成技術を実現した^{10), 22)}。またダンスにおいて動作間の類似性を定義する距離尺度を新たに提案し、表現に感情を付加し、類似するダンスと一部を置換可能な動作生成手法を提案した。さらに、人物の顔の印象類似度として頭髮の形状、色、毛並みのパラメータを導入して、印象類似人物検索システムを試作し、高い検索性能を実現した。

Web インタクションに関しては、主観類似度の判定のため、ある一定以上のコメントが付与されている 500 件の動画について、8 つの印象クラス、3 人以上が評価した主観的類似度判定データベースを構築し、そのデータベースによる主観的類似度に基づく検索システムを開発した。また、コンテンツ生態系を活性化させるため、鑑賞者が動画に装飾可能とする N 次装飾技術を提案及び実装し、ウェブブラウザ拡張として配布するとともに、動画装飾に関する事例を収集した¹⁴⁾。



図: 音楽鑑賞支援サービス「Songrium (ソングリウム)」の実現と一般公開

§ 3. 成果発表等

(3-1) 原著論文発表

論文詳細情報(国内)

1. 後藤 真孝, 吉井 和佳, 藤原 弘将, Matthias Mauch, 中野 倫靖: "Songle: 音楽音響信号理解技術とユーザによる誤り訂正に基づく能動的音楽鑑賞サービス", 情報処理学会論文誌, Vol.54, No.4, pp.1363-1372, April 2013.
2. 平井辰典, 大矢隼士, 森島繁生, "既存音楽動画の再利用による音楽に合った動画の自動生成システム", 情報処理学会論文誌, Vol.54, No.4, pp.1254-1262, April 2013.
3. 山本岳洋, 中村聡史: "視聴者の時刻同期コメントを用いた楽曲動画の印象分類", 情報処理学会論文誌(トランザクション)データベース(TOD58), Vol.6, No.3, pp.61-72, June 2013.
4. 中村聡史, 山本岳洋, 後藤真孝, 濱崎 雅弘: "視聴者反応と音響特徴量に基づくサムネイル動画の生成手法", 情報処理学会論文誌(トランザクション)データベース(TOD58), Vol.6, No.3, pp.148-158, June 2013.
5. 梅本和俊, 中村聡史, 山本岳洋, 田中克己: "Web 検索時の行動情報を用いたクエリ修正タイプの予測", 情報処理学会論文誌(トランザクション)データベース(TOD58), Vol.6, No.3, pp.132-147, June 2013.
6. 梅本和俊, 山本岳洋, 中村聡史, 田中克己: 視線情報からの注目語抽出に基づく検索意図のリアルタイム推定, 情報処理学会論文誌(トランザクション)データベース(TOD58), Vol.6, No.3, pp.120-131, June 2013.
7. 福里司, 平井辰典, 大矢隼士, 森島繁生, "アニメ作品におけるキーフレーム自動抽出に基づく映像要約手法の提案", 画像電子学会誌ビジュアルコンピューティング論文特集号, Vol.42, No.4, p.448-456, July 2013.
8. Hayato Ohya, and Shigeo Morishima, "Automatic Music Video Generation System by Reusing Posted Web Content with Hidden Markov Model", IIEEEJ Transactions on Image Electronics and Visual Computing, Vol.1, No.1, pp.65-73, December 2013.
9. Kazuhiro Kobayashi, Tomoki Toda, Hironori Doi, Tomoyasu Nakano, Masataka Goto, Graham Neubig, Sakriani Sakti, and Satoshi Nakamura: Voice Timbre Control Based on Perceived Age in Singing Voice Conversion, The IEICE Transactions on Information and Systems, June 2014. (in press)

[proceedings(査読審査の入るものに限る)]

10. Masahide Kawai, Tomoyori Iwao, Daisuke Mima, Akinobu Maejima, and Shigeo Morishima, "Photorealistic Inner Mouth Synthesis in Speech Animation", Proceedings of Visual Computing / グラフィクスと CAD 合同シンポジウム 2013, June 2013.
11. Shoto Sasaki, Tatsunori Hirai, Hayato Ohya, and Shigeo Morishima, "Affective Music Recommendation System based on Images", Proceedings of Visual Computing / グラフィクスと CAD 合同シンポジウム 2013, June 2013.
12. Tsukasa Fukusato, Tatsunori Hirai, Hayato Ohya, and Shigeo Morishima, "Automatic Comic-like Video Summarization by Key-frame Detection from Anime film", Proceedings of Visual Computing / グラフィクスと CAD 合同シンポジウム 2013, June 2013.
13. Takahiro Kosaka, Akinobu Maejima, Hiroyuki Kubo, and Shigeo Morishima, "Automatic Region-based Realistic Expression Cloning for Blendshapes", Proceedings of Visual Computing / グラフィクスと CAD 合同シンポジウム 2013, June 2013.
14. 中村聡史, 石川直樹, 渡邊恵太: "個人的な小さな幸せを実現するブラウザ上での動画編集・共有手法", 日本ソフトウェア科学会 第21回インタラクティブシステムとソフトウェアに関するワークショップ (WISS 2013) 論文集, December 2013.

論文詳細情報(国際)

[proceedings(査読審査の入るものに限る)]

15. Kazuyoshi Yoshii and Masataka Goto: Infinite Kernel Linear Prediction for Joint Estimation of Spectral Envelope and Fundamental Frequency, Proceedings of the 2013 IEEE International Conference on Acoustics, Speech, and Signal Processing (IEEE ICASSP 2013), pp.463-467, May 2013.
16. Kazuyoshi Yoshii, Ryota Tomioka, Daichi Mochihashi, and Masataka Goto: Infinite Positive Semidefinite Tensor Factorization for Source Separation of Mixture Signals, Proceedings of the 30th International Conference on Machine Learning (ICML 2013), pp.576-584, June 2013.
17. Tomoyasu Nakano and Masataka Goto: VocaRefiner: An Interactive Singing Recording System with Integration of Multiple Singing Recordings, Proceedings of the 10th Sound and Music Computing Conference (SMC 2013), pp.115-122, July 2013.
18. Takehiro Yamamoto and Satoshi Nakamura: "Leveraging Viewer Comments for Mood Classification of Music Video Clips", Proceedings of the 36th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (ACM SIGIR 2013), pp. 797-800, July 2013.

19. Masahiro Hamasaki and Masataka Goto: Songrium: A Music Browsing Assistance Service Based on Visualization of Massive Open Collaboration Within Music Content Creation Community, Proceedings of the 9th International Symposium on Open Collaboration (ACM WikiSym + OpenSym 2013), pp.1-10, August 2013.
20. Hironori Doi, Tomoki Toda, Tomoyasu Nakano, Masataka Goto, and Satoshi Nakamura: Evaluation of a Singing Voice Conversion Method Based on Many-to-Many Eigenvoice Conversion, Proceedings of the 14th Annual Conference of the International Speech Communication Association (Interspeech 2013), pp.1067-1071, August 2013.
21. Kazuhiro Kobayashi, Hironori Doi, Tomoki Toda, Tomoyasu Nakano, Masataka Goto, Graham Neubig, Sakriani Sakti, and Satoshi Nakamura: An Investigation of Acoustic Features for Singing Voice Conversion based on Perceptual Age, Proceedings of the 14th Annual Conference of the International Speech Communication Association (Interspeech 2013), pp.1057-1061, August 2013.
22. Masahide Kawai, Tomoyori Iwao, Akinobu Maejima, and Shigeo Morishima, "Video-Realistic Inner Mouth Reanimation", Proceedings of Pacific Graphics 2013, October 2013.
23. Kazuyoshi Yoshii, Ryota Tomioka, Daichi Mochihashi, and Masataka Goto: Beyond NMF: Time-Domain Audio Source Separation without Phase Reconstruction, Proceedings of the 14th International Society for Music Information Retrieval Conference (ISMIR 2013), pp.369-374, November 2013.
24. Philippe Hamel, Matthew E. P. Davies, Kazuyoshi Yoshii, and Masataka Goto: Transfer Learning In MIR: Sharing Learned Latent Representations for Music Audio Classification and Similarity, Proceedings of the 14th International Society for Music Information Retrieval Conference (ISMIR 2013), pp.9-14, November 2013.
25. Matthew E. P. Davies, Philippe Hamel, Kazuyoshi Yoshii, and Masataka Goto: AutoMashUpper: An Automatic Multi-Song Mashup System, Proceedings of the 14th International Society for Music Information Retrieval Conference (ISMIR 2013), pp.575-580, November 2013.
26. Satoru Fukayama, Kazuyoshi Yoshii, and Masataka Goto: Chord-Sequence-Factory: A Chord Arrangement System Modifying Factorized Chord Sequence Probabilities, Proceedings of the 14th International Society for Music Information Retrieval Conference (ISMIR 2013), pp.457-462, November 2013.
27. Masahiro Hamasaki, Masataka Goto, and Tomoyasu Nakano: Songrium: A Music Browsing Assistance Service with Interactive Visualization and Exploration of a Web of Music, Proceedings of the 23rd International World Wide Web Conference:

Web Science Track (WWW 2014), April 2014. (in press)

28. Tomohiko Nakamura, Hirokazu Kameoka, Kazuyoshi Yoshii, and Masataka Goto: Replacing Drum Timbres and Frequency Characteristics between Different Musical Pieces, Proceedings of the 2014 IEEE International Conference on Acoustics, Speech, and Signal Processing (IEEE ICASSP 2014), May 2014. (in press)
29. Kazuhiro Kobayashi, Tomoki Toda, Tomoyasu Nakano, Masataka Goto, Graham Neubig, Sakriani Sakti, and Satoshi Nakamura: Regression Approaches to Perceptual Age Control in Singing Voice Conversion, Proceedings of the 2014 IEEE International Conference on Acoustics, Speech, and Signal Processing (IEEE ICASSP 2014), May 2014. (in press)
30. Matt McVicar, Daniel P W Ellis, and Masataka Goto: Leveraging Repetition for Improved Automatic Lyric Transcription in Popular Music, Proceedings of the 2014 IEEE International Conference on Acoustics, Speech, and Signal Processing (IEEE ICASSP 2014), May 2014. (in press)
31. Kazuyoshi Yoshii, Hiromasa Fujihara, Tomoyasu Nakano, and Masataka Goto: Cultivating Vocal Activity Detection for Music Audio Signals in a Circulation-type Crowdsourcing Ecosystem, Proceedings of the 2014 IEEE International Conference on Acoustics, Speech, and Signal Processing (IEEE ICASSP 2014), May 2014. (in press)
32. Tomoyasu Nakano, Kazuyoshi Yoshii, and Masataka Goto: Vocal Timbre Analysis Using Latent Dirichlet Allocation and Cross-Gender Vocal Timbre Similarity, Proceedings of the 2014 IEEE International Conference on Acoustics, Speech, and Signal Processing (IEEE ICASSP 2014), May 2014. (in press)

(3-2) 知財出願

- ① 平成 25 年度特許出願件数(国内 1 件)
- ② CREST 研究期間累積件数(国内 3件)