

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知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知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知化情報環境の実現

(3)「中村」グループ

- ① 主たる共同研究者: 中村 聡史 (京都大学大学院 情報学研究科、特定准教授)
- ② 研究項目
 - ・Web インタラクションを中心とした類似度可知化情報環境の実現

§ 2. 研究実施内容

■研究のねらい

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する(図1)。そして、創作支援技術と鑑賞支援技術を研究開発することで、コンテンツの創作や鑑賞を誰もが能動的に楽しめる社会や、過去のコンテンツに敬意を払う文化、感動体験重視型のコンテンツ文化の実現を目指す。具体的には、メディアコンテンツとして代表的で重要性が高い音楽あるいは音楽連動動画(ミュージックビデオやダンス動画等)等を中心に、音楽自動理解技術を開発および応用しつつ、人間と情報環境の調和的な相互作用を引き出せるよう、ユーザを巻き込んだ研究開発に取り組む。さらに、人間とコンテンツとの関係、過去のコンテンツと未来のコンテンツとの関係が豊かで持続発展可能な共生関係となる「コンテンツ共生社会」を実現するために、類似度を可知化する情報環境を活用した「コンテンツ生態系」という新たな概念を提唱し、メディアコンテンツの創作と鑑賞を総合的に支援する「コンテンツ生態系技術」を開発する。この「コンテンツ生態系技術」では、過去のコンテンツが新たなコンテンツを育む「コンテンツ循環」をユーザと共に持続発展させることを狙う。そのために Web 上で公開されている音楽関連コンテンツを主な対象とした「コンテンツ生態系 Web サービス」を構築・公開して実証実験する。



図1. 「コンテンツ共生社会」のための「類似度を可知化する(知ることができる)情報環境」

■研究進捗状況と今後の見通し

平成 24 年度は本「OngaCREST Project(音画 CREST(おんがくれすと)プロジェクト)」の 2 年目として、基礎研究と応用研究にバランス良く取り組むことで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」の重要な機能である能動的音楽鑑賞 Web サービス「Songle (ソングル)」(<http://songle.jp>) を実現し、平成 24 年 8 月には一般公開をして実証実験を開始して、その研究成果を産総研と JST が共同でプレス発表した。Songle 上では楽曲の中身の可視化に加え、歌声の類似度に基づく関連楽曲表示や、コード進行検索も可能にした。その上で新たに、コンテンツ間の関係性に基づく音楽鑑賞支援 Web サービスの検討も進め、プロトタイプシステム実現に着手した。さらに、音楽連動動画に関しては、音楽的要素に対応した映像ターゲットの抽出・トラッキング手法の開発および鑑賞支援プロトタイプシステムの実現を進め、Web インタラクションに関しては、ソーシャルアノテーション

の収集・管理およびその分析に基づく音楽連動動画の印象分類手法の研究開発を進めた。平成 24 年度以降は、これまでの成果をさらに発展させつつ、具体的な連携を深めていく。そして、基礎研究の成果を蓄積しつつ、応用研究としてコンテンツ生態系 Web サービスのプロタイプシステムを研究開発して世の中にアピールしながら、研究を進めていく。

■研究成果

以下、平成 24 年度の代表的な研究成果を述べる。

・ 能動的音楽鑑賞サービス「Songle」の一般公開による実証実験の開始(プレス発表)

「コンテンツ生態系 Web サービス」の鑑賞支援技術の実現に最優先で取り組み、音楽音響信号理解技術によって音楽の聴き方をより豊かにするための能動的音楽鑑賞サービス「Songle (シングル)」(<http://songle.jp>) を実現して一般公開し、実証実験を開始した(図2)。これは、ウェブブラウザでアクセスするだけで、誰でも無償で利用できるサイトである。既に登録されている様々な楽曲を鑑賞するだけでなく、インターネット上で公開されている楽曲 MP3 ファイルの URL も新たに登録できる。Songle のユーザは、楽曲の中身(サビ、ビート、メロディー、コード)を自動解析して「音楽地図」を表示する可視化機能により、繰り返しなどの様々な観点に気づき、楽曲に対する理解を深めることができる。また、楽曲中で盛り上がるサビ区間などへ自在にジャンプできるサビ出し機能により、興味のある箇所を容易に見つけて楽しみながら鑑賞することができる。さらに、自動解析の誤りをユーザが自発的に訂正できるインターフェースも提供することで、ユーザの訂正協力によってより正確な「音楽地図」を共有して表示できる仕組みも導入した。本成果を、査読付き国際ワークショップ「CrowdSearch 2012」(WWW 2012 併設)で発表し⁶⁾、情報処理学会論文誌にも投稿して採択された²⁾。さらにその先進性から国際ワークショップで基調講演 2 件の招待を受けた。平成 24 年 8 月 29 日には産総研と JST が共同でプレス発表「インターネット上の楽曲の中身を自動解析する音楽鑑賞システム」をしてアピールし、Wired、日本経済新聞、ITmedia、インプレス Watch、マイナビニュース、週刊アスキーWeb、DTM MAGAZINE 等の 22 件以上の多数の報道 (Yahoo!、Infoseek、excite、BIGLOBE 等への同一記事転載を含めれば 43 件以上) がなされた。



図2. 能動的音楽鑑賞サービス「Songle (シングル)」の実現と一般公開

・ 音楽連動動画中での同一人物の登場シーンに基づく鑑賞支援プロトタイプシステムの構築

音楽連動動画における生成モデル構築等のために、音楽と同期した動画(特にライブ演奏動画等)の解析によって、注目すべき音楽的要素に対応した映像ターゲットの抽出・トラッキング手法の開発を行った。さらにこの手法を応用し、同一人物が登場するシーンをカテゴリ化して音楽的要素と共に一覧としてユーザに提示する鑑賞支援プロトタイプシステムの実現に取り組んだ。このシステムでは、シーンチェンジやパン・チルトなどのカメラ効果に加えて、演奏者が正面を向いていない場合でも安定してターゲットを追跡できる機能¹⁾を実現し、動画共有サイト上に存在する膨大な数の音楽連動動画コンテンツを逐一視聴しなくても、短時間に内容を把握できる動画鑑賞支援インタフェースを構築した。一方、人間の感じ方に基づく音楽要素および画像要素のクラスタリング手法にも着手し、その第一歩として入力された画像のシーンにふさわしい音楽コンテンツの推薦を行うプロトタイプシステムを実現した。

・ ソーシャルアノテーションの収集・管理と音楽連動動画の印象分類手法の研究開発

前年度に開発したソーシャルアノテーション収集・管理システムを改良し、各種のインデックスを付与することにより、ソーシャルアノテーションを柔軟に取り出せるようにした。また、ソーシャルアノテーションを活用した絞り込み検索のために、ソーシャルアノテーション分析と音楽分析をベースとした楽曲動画の印象分類手法^{4), 14)}を提案・実現した。さらに、この印象分類手法を利用することによって、音楽連動動画の主観的特徴量を抽出する手法を実現した。一方、音楽連動動画中にシーンチェンジやカメラ効果が付与された場合でも安定してターゲット(登場人物)を追跡することを支援するため、登場人物名や登場人物に関するキーワードが、その登場人物が登場しているシーンにどの程度現れているかなどの検討を行った。また、膨大な音楽連動動画の中から、ユーザが好みの動画と出会うことを容易にするため、ソーシャルアノテーション分析と音楽分析に基づくサムネイル動画(15秒の動画)の自動生成技術^{5), 13)}を開発した。

§3. 成果発表等

(3-1) 原著論文発表

● 論文詳細情報

1. 平井 辰典, 中野 倫靖, 後藤 真孝, 森島 繁生: "シーンの連続性と顔類似度に基づく動画コンテンツ中の同一人物登場シーンの同定", 映像情報メディア学会誌, Vol.66, No.7, pp.J251-J259, July 2012.
2. 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, April 2013. (in press)
3. 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, April 2013. (in press)
4. 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, June 2013. (accepted)
 5. 中村 聡史, 山本 岳洋, 後藤 真孝, 濱崎 雅弘: "視聴者反応と音楽的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, June 2013. (accepted)
 6. Masataka Goto, Jun Ogata, Kazuyoshi Yoshii, Hiromasa Fujihara, Matthias Mauch and Tomoyasu Nakano: PodCastle and Songle: Crowdsourcing-Based Web Services for Retrieval and Browsing of Speech and Music Content, Proceedings of the First International Workshop on Crowdsourcing Web Search (CrowdSearch 2012), pp.36-41, April 2012.
 7. Kazutoshi Umemoto, Takehiro Yamamoto, Satoshi Nakamura, Katsumi Tanaka: "Search Intent Estimation from User's Eye Movements for Supporting Information Seeking", Proceedings of the 11th International Working Conference on Advanced Visual Interfaces (ACM AVI 2012), pp.349-356, May 2012.
 8. 大矢 隼士, 森島 繁生: "動画サイトコンテンツ再利用による HMM に基づく音楽からの動画自動生成システム", Visual Computing / グラフィクスと CAD 合同シンポジウム 2012, June 2012.
 9. 平井 辰典, 大矢 隼士, 森島 繁生: "既存の音楽動画を用いて音楽に合った映像を自動生成するシステム", Visual Computing / グラフィクスと CAD 合同シンポジウム 2012, June 2012.
 10. Tatsunori Hirai, Hayato Ohya, Shigeo Morishima: Automatic Mash up Music Video Generation System by Perceptual Synchronization of Music and Video Features, ACM SIGGRAPH2012, Poster, Los Angeles, August 2012.
 11. Tomoyasu Nakano and Masataka Goto: A Spectral Envelope Estimation Method Based on F0-Adaptive Multi-Frame Integration Analysis, Proceedings of SAPA-SCALE Conference 2012, pp.11-16, September 2012.
 12. Kazuyoshi Yoshii and Masataka Goto: Infinite Composite Autoregressive Models for Music Signal Analysis, Proceedings of the 13th International Society for Music Information Retrieval Conference (ISMIR 2012), pp.79-84, October 2012.
 13. 中村 聡史, 山本 岳洋, 後藤 真孝, 濱崎 雅弘: "視聴者反応と音楽的特徴量を用いたサムネイル動画の自動生成", 第 5 回 Web とデータベースに関するフォーラム (WebDB Forum 2012) 論文集, November 2012.
 14. 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推定", 第 5 回 Web とデータベースに関するフォーラム (WebDB Forum 2012) 論文集, November 2012.

15. Hayato Ohya and Shigeo Morishima: Automatic music video creation system by reusing existing contents in video-sharing service based on HMM, Proc. of Image Electronics and Visual Computing Workshop 2012, November 2012.
16. Hironori Doi, Tomoki Toda, Tomoyasu Nakano, Masataka Goto, and Satoshi Nakamura: Singing Voice Conversion Method Based on Many-to-Many Eigenvoice Conversion and Training Data Generation Using a Singing-to-Singing Synthesis System, APSIPA (Asia-Pacific Signal and Information Processing Association) ASC (Annual Summit and Conference) 2012, December 2012.
17. 中野 倫靖, 後藤 真孝: "VocaRefiner: 歌を歌って歌い直して統合できる新しい歌声生成インタフェース", 日本ソフトウェア科学会 第20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 2012) 論文集, December 2012.
18. 濱崎 雅弘, 後藤 真孝: "Songrium: 関係性に基づいて音楽星図を渡り歩く音楽視聴支援サービス", 日本ソフトウェア科学会 第20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 2012) 論文集, December 2012.

(3-2) 知財出願

- ① 平成24年度特許出願件数(国内 2件)
- ② CREST研究期間累積件数(国内 2件)