

楽曲プレイリストに基づく観光写真スライドショー生成システムの提案

山崎 敦士[†] 奥 健太[†]

[†] 龍谷大学理工学部 〒520-2194 滋賀県大津市瀬田大江町横谷 1-5
E-mail: †t150549@mail.ryukoku.ac.jp, ††okukenta@rins.ryukoku.ac.jp

あらまし 入力された楽曲プレイリストに合った観光地の写真をスライドショー形式で提示するシステムを提案する。まず、事前処理として楽曲と観光資源に対して感情値を付与する。感情モデルとして Russell の感情モデルを用いる。楽曲と観光資源に与えられた感情値に基づき、楽曲と観光資源の対応付けを行う。楽曲プレイリストを再生したとき、再生中の楽曲に対応付けられた観光資源の写真を提示する。被験者実験の結果より、ランダム手法に比べて、提案手法では楽曲に合った観光写真スライドショーを生成できることを示した。

キーワード 楽曲プレイリスト, スライドショー, 観光情報推薦, 感情モデル

1. はじめに

2017 年の訪日外客数が 2,800 万人に達する^(注1)など観光を楽しむ人々は年々増大してきている。また、観光情報サイトの一つである TripAdvisor には、約 800 万件もの旅行情報（観光スポットやレストラン、ホテルなど）が登録されている^(注2)。観光者はこのような膨大な観光スポットの中から、訪れたい観光スポットを選択する必要がある。例えば、TripAdvisor では、行先や観光カテゴリ、キーワード等による検索機能が提供されている。ユーザはこれらの検索機能を用いて、訪れたい観光スポットを能動的に探すことができる。一方で、テレビや電車内広告で美しい風景を目にすることで、思いがけず興味のある観光情報に遭遇するという、受動的な体験もある。ここで、観光情報に関しては、写真から受ける印象は大きい。写真を目にすることがきっかけで、その観光地に訪れてみたいという印象をもつということもある。

大量の写真を受動的に閲覧する形態の一つとしてスライドショーが挙げられる。観光地の風景をスライドショーで提示することで、ユーザは旅行気分を味わいながら、興味のある観光情報に遭遇することが期待できる。また、写真スライドショーをより楽しむために、スライドショーに音楽を付加するという研究もある [1]。我々の先行研究 [2] では、入力された楽曲プレイリストに合った景観写真をスライドショーの形式で提示するシステムを提案した。ただし、[2] では、ドライブルート推薦を想定していることから、道路リンクの景観写真に限定したものであった。本研究では、対象を観光スポットの写真に拡張し、楽曲プレイリストに合った観光スポットの写真をスライドショーの形式で提示するシステムを提案する。さらに、提示された写真に関連する観光スポットの情報を添えることで、観光情報への興味を喚起する機能も付加する。

2. 関連研究

2.1 Russell の円環モデル

感情モデルの一つとして Russell の円環モデル [3] がある。Russell の円環モデルは、快適-不快 (pleasure), 覚醒-眠気 (arousal) の 2 次元で表現される平面上にすべての感情が配置されることを示したモデルである。ある対象について、Russell の円環モデルに基づく感情特徴を与えたい場合は、質問紙調査による方法がある [4]。質問紙調査により、対象について、覚醒-眠気 (A-S), 興奮-陰気 (E-G), 快適-不快 (P-U), 悲惨-休息 (D-R) の 4 軸に関するスコアを得る。得られた各スコアを基に次式を算出することで、対象の感情特徴を Russell の円環モデル上の座標空間 (x, y) に落とし込むことができる：

$$x = 1.00(P-U) + .707(E-G) - .707(D-R) \quad (1)$$

$$y = 1.00(A-S) + .707(E-G) + .707(D-R) \quad (2)$$

感情モデルとしてはさまざまなものが提案されているが、Russell の円環モデルはシンプルな構造で表現されており、かつすべての感情について網羅的に適用することができる。このような手頃さから、本研究では Russell の円環モデルを採用する。

2.2 スライドショー生成システム

Hua ら [1] は、付随音楽と写真の特徴に基づく写真スライドショーシステムを提案している。写真の遷移位置を決定するために、楽曲の立ち上がり（オンセット）位置を検出しておく。オンセット位置に基づき、楽曲を複数の部分楽曲に分割する。部分楽曲の切り替わり時点に合わせて写真を遷移させることでスライドショーを提示する。このシステムでは、楽曲の内容は写真を遷移させるタイミングを検出するために用いられており、楽曲の内容と写真の内容とのマッチングは行っていない。

Li ら [5] は、感情に基づく印象派絵画のスライドショーシステムを提案している。このシステムは感情をベースに伴奏音楽付きの絵画スライドショーを生成する。絵画の色特徴およびテクスチャ特徴とアノテーションされた感情の対を学習データとし、これら特徴と感情との相関ルールを抽出する。抽出された相関ルールに基づき絵画に対する感情を推定し、この感情に基

(注1) : <https://statistics.jnto.go.jp/>

(注2) : <https://tripadvisor.mediaroom.com/jp-about-us>

づき絵画をクラスタリングする。楽曲についても同様に楽曲の特徴（メロディ、リズム、テンポ）と感情との相関ルールを抽出し、感情に基づきクラスタリングする。絵画と楽曲それぞれに対応付けられた感情の類似度に基づき、絵画と楽曲とをマッチングさせ、スライドショーを生成している。

我々の先行研究である杉本らの研究 [2] では、入力された楽曲プレイリストに合った景観写真をスライドショーの形式で提示するシステムを提案した。ただし、[2] では、ドライブルート推薦を想定していることから、道路リンクの景観写真に限定したものであった。本研究では、対象を観光スポットの写真に拡張し、楽曲プレイリストに合った観光スポットの写真をスライドショーの形式で提示するシステムを提案する。

2.3 観光情報推薦システム

観光情報推薦システム [6] は、ユーザの要求に合った観光情報や娯楽施設を提供するシステムである。これまでも多くの観光情報推薦システムに関する研究が行われている [7]。

観光情報推薦システムは、Web インタフェースとして提供されることが多い。Lee ら [8] は、Web ページに Google Maps を埋め込み、地図上で旅行ルートを表示するインタフェースを提案している。台南市での旅行者向けに、文化遺産や郷土料理など個人の嗜好に合った旅行プランを提示する。EnoSigTur [9] では、ユーザが自身のデモグラフィック情報および嗜好を入力することで、観光に関する情報が推薦される。

さらに、スマートフォンなどのモバイルインタフェースとして観光情報推薦システムを適用しているものもある。Yu ら [10] は、PDA において個人向け旅行プランニングをサポートするための位置情報ベースの推薦システムを提供している。推薦情報は個人の嗜好や位置情報、時間に基づいて生成され、Google Maps 上に提示される。LiveCities [11] は、Android での通知サービスを用いて、ユーザコンテキストに合った情報をプッシュ提供する。

以上のように、観光情報推薦システムは Web インタフェースやモバイルインタフェースとして提供されていることが多い。本研究においても、Web インタフェースを想定した推薦システムを提案する。さらに、提示された写真に関連する観光スポットの情報を添えることで、観光情報への興味を喚起する機能も付加する。

3. 準備

定義 1 : 観光資源. 本稿では、清水寺や姫路城などの観光資源を $r_i \in R$ と表す。観光資源 r_i は、その資源の名称 $r_i.name$ をもつ。

定義 2 : 楽曲. 楽曲を $t_j \in T$ と表す。

定義 3 : 感情ベクトル. Russell の感情モデルに基づき、観光資源 r_i および楽曲 t_j の感情ベクトルを、それぞれ $\mathbf{r}_i, \mathbf{t}_j$ で表す。感情ベクトルは、快適-不快 (x) および覚醒-眠気 (y) の 2次元の特徴ベクトル ($-4 \leq x \leq +4, -4 \leq y \leq +4$) で表現される。

定義 4 : 楽曲に関連する観光資源集合. 楽曲 t_j に関連する観光資源集合を R^j と表記する。

定義 5 : 楽曲プレイリスト. 楽曲プレイリスト \mathcal{P} を、楽曲のシーケンスとみなし、 $\mathcal{P} = \langle t_1, t_2, \dots \rangle$ と表現する。これは、楽曲 t_1, t_2, \dots の順に時系列に楽曲が登録されていることを表す。

定義 6 : 観光資源シーケンス. 観光資源シーケンス \mathcal{L} を、観光資源のシーケンスとみなし、 $\mathcal{L} = \langle r_1, r_2, \dots \rangle$ と表現する。これは、観光資源 r_1, r_2, \dots の順に時系列に観光資源が登録されていることを表す。

4. システム

本研究は楽曲の聴取中に観光写真スライドショーを提示することで、観光地推薦の価値を生み出すセレンディピティ効果の発現を目的としている。提案システムでは、楽曲を聴取するユーザの感性に適した観光資源写真を推薦する。本研究では、楽曲と観光資源の関連付けには、それぞれから受ける感情に着目する。本章では楽曲プレイリストに基づく観光写真スライドショーシステムの概要と各処理の説明を行う。

4.1 システム概要

インタフェースは、楽曲プレイリストビューと観光資源スライドショービューから構成される。ユーザは、楽曲プレイリストビューの「プレイリスト作成」ボタンから楽曲を選択してプレイリストを作成する。作成したプレイリストを再生することで、楽曲の再生と同時にスライドショーが開始される。

楽曲プレイリストを再生すると、以下の流れにより、楽曲に合った観光資源写真のスライドショーが提示される：

- (1) 再生中の楽曲に対応する観光資源 ID リストを取得する。
- (2) 観光資源 ID に対応する観光資源情報を取得する。
- (3) 観光資源に対応する写真データを取得する。
- (4) 観光資源情報および観光資源写真を観光資源スライドショービューに提示する。

このようなシステムを実現するためには、楽曲と観光資源とのマッチングが課題となる。まず、4.2 節および 4.3 節で観光資源および楽曲への感情付与について述べたあと、4.4 節で感情に基づく楽曲と観光資源とのマッチング方法について述べる。そして、4.5 節で、楽曲プレイリストに合った観光資源シーケンスの生成方法について述べる。

4.2 観光資源への感情付与

観光資源 $r_i \in R$ に対し、感情を付与する。観光資源の感情モデルとして、2.1 節で説明した Russell の感情モデルを用いる。各観光資源 r_i について、感情ベクトル \mathbf{r}_i を付与する。

まず、写真共有サイトである Flickr^(注3) に対し、資源の名称 $r_i.name$ をクエリとして関連する写真を検索する。検索された写真リストを資源 r_i の写真リストとする。このとき、検索された写真数が 10 件未満の資源は以降の処理の対象から除外する。

残りの資源について、下記の手順により、クラウドソーシングを用いて資源 $r_i \in R$ に感情を付与する：

- (1) クラウドソーシングのワーカに、資源 r_i の写真リストを提示する。

(注3) : <https://www.flickr.com/>

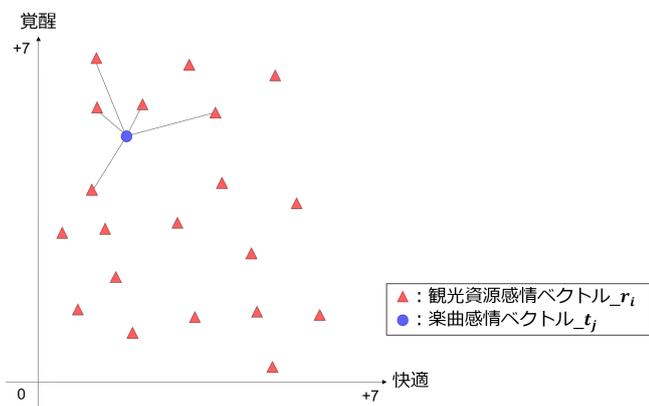


図1 楽曲に対する観光資源 ID の取得 ($n = 5$ のとき)

(2) ワーカは、提示された写真リストに相応しいと思う感情を数値で回答する。ここでは、覚醒-眠気 (A-S)，興奮-陰気 (E-G)，快適-不快 (P-U)，悲惨-休息 (D-R) の 4 軸について、それぞれ $\{-4, -3, -2, -1, 0, +1, +2, +3, +4\}$ のいずれかの値を選択する。システムの不具合等により写真が表示されない場合は「写真が表示されない」を選択する。

なお、資源 1 件あたりのワーカ数は 20 名とする。

資源 r_i について、与えられた 4 軸の値の平均値 (A-S, E-G, P-U, D-R) を算出する。これら 4 軸の平均値を用いて、式 (1) および式 (2) より、 x, y の値を算出する。これら x, y の値を要素とした感情ベクトル r_i を作成する。

4.3 楽曲への感情付与

4.2 節と同様に、楽曲 $t_j \in T$ に対し、感情を付与する。ここでも、楽曲の感情モデルとして、Russell の感情モデルを用いる。各楽曲 t_j について、感情ベクトル t_j を付与する。

楽曲 t_j には、楽曲ファイル (MP3 形式、WAV 形式など) が含まれる。事前に全楽曲 T について特定の区間を切り出し、評価用楽曲ファイルとしておく。以下の手順により、クラウドソーシングを用いて楽曲 $t_j \in T$ に感情を付与する：

(1) クラウドソーシングのワーカは楽曲 t_j の評価用楽曲ファイルを聴取する。

(2) ワーカは、聴取した楽曲に相応しいと思う感情を数値で回答する。4.2 節と同様に、覚醒-眠気 (A-S)，興奮-陰気 (E-G)，快適-不快 (P-U)，悲惨-休息 (D-R) の 4 軸について、それぞれ $\{-4, -3, -2, -1, 0, +1, +2, +3, +4\}$ のいずれかの値を選択する。システムの不具合等により写真が表示されない場合は「楽曲が再生されない」を選択する。

なお、楽曲 1 件あたりのワーカ数は 20 名とする。

4.2 節と同様に、4 軸の平均値を用いて、式 (1) および式 (2) より、 x, y の値を算出する。これら x, y の値を要素とした感情ベクトル t_j を作成する。

4.4 感情に基づく観光資源と楽曲とのマッチング

楽曲 t_j に関連する観光資源集合 $R^j \in R$ を取得する。図 1 のように、楽曲 t_j および観光資源集合 R を同一の感情特徴空間に配置する。このとき、楽曲の感情ベクトル t_j の n 件の最近傍の観光資源集合を R^j とする。ここで、楽曲と観光資源の距

離は、それぞれの感情ベクトル間のユークリッド距離に基づき算出する。

4.5 楽曲プレイリストに合った観光資源推薦

楽曲プレイリスト $\mathcal{P} = \langle t_1, t_2, \dots \rangle$ が与えられたとき、その楽曲プレイリスト \mathcal{P} に合う観光資源シーケンス \mathcal{L} を生成する。

観光資源シーケンス \mathcal{L} の取得は以下の手順で行う：

(1) 楽曲プレイリスト \mathcal{P} 内の各楽曲 t_k に関連する観光資源集合 R^k を取得する。

(2) 取得した観光資源集合 R^k を $\langle R^1, R^2, \dots \rangle$ のようにつなげ、これを観光資源シーケンス \mathcal{L} とする。

5. 評価

5.1 データセット

実験用データとして、国土数値情報の観光資源データ^(注4)から関西圏の観光資源 1,256 件を用いた。また、著作権フリーの楽曲を公開している『DOVA-SYNDROME』^(注5)から楽曲 750 曲をダウンロードした。

観光資源データおよび楽曲データについては、4.2 節および 4.3 節で述べたとおり、クラウドソーシングにより感情を付与した。楽曲については、評価用ファイルとして、楽曲長の中心から前後 30 秒間とした区間を用いた。ダウンロードした楽曲 750 曲のうち区間の切出しに成功した 706 曲を対象とした。この 706 曲に対し、感情を付与した。

5.2 実験方法

楽曲プレイリスト \mathcal{P} を $\mathcal{P} = \langle t_1, t_2, t_3, t_4 \rangle$ として与え、4.5 節のとおり、楽曲プレイリスト \mathcal{P} に合う観光資源シーケンス \mathcal{L} を生成した。

観光資源と楽曲の関連付けにおける優位性を調べるため、 $n = 5$ の提案手法と観光資源をランダムに取得したものと比較実験を行った。実験協力者として、20 代男性 7 名と 20 代女性 2 名の計 9 名を用意した。各被験者は、プレイリスト \mathcal{P} を再生する。プレイリスト \mathcal{P} が再生されている間、観光資源シーケンス \mathcal{L} 内の各観光資源の写真がスライドショーとして提示される。

スライドショーが終了後、被験者は、再生された楽曲と提示された観光資源が合っていたかという観点で評価する。評価尺度として、次の 4 段階評価を用いた：

- 4: 非常に合っていた、
- 3: 合っていた、
- 2: 合っていなかった、
- 1: まったく合っていなかった。

5.3 実験結果

図 2 は、被験者による評価の平均値を示したものである。図 2 のとおり、提案手法の平均評価値は約 2.19、ランダム手法の平均評価値は約 1.72 であった。この結果より、ランダムにスラ

(注4) : http://nlftp.mlit.go.jp/ksj/gml/datalist/KsjTmplt-P12-v2_2.html

(注5) : <https://dova-s.jp/>

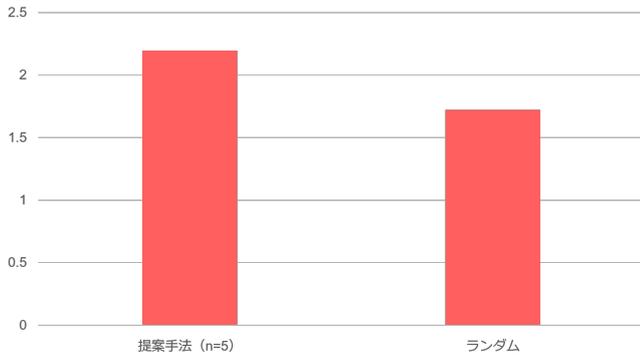


図2 提案手法とランダム手法に対する平均評価値。

イドショーを生成する方式に比べて、提案手法では楽曲に合った観光資源の写真が提示されていたことがわかる。

例として楽曲 t_3 の聴取中に流れたスライドショーについて考える。楽曲 t_3 は、楽曲の感情ベクトルと観光資源の感情ベクトルの値が近いものであった。実際に、楽曲 t_3 はアップテンポの曲で爽快感が感じられる曲であった。提示された観光資源の写真も道頓堀、鈴鹿サーキット、阪神甲子園球場といった、賑やかでありながら開放的である内容のものであった。

一方で、提案手法の平均評価値がランダム手法を上回ったとはいえ、提案手法の平均評価値は2.19であり、楽曲と合っていない観光資源写真も再生されることもあった。感情を介した楽曲と観光資源写真のマッチング方式について、さらなる検討が必要である。

6. おわりに

本稿では、楽曲プレイリストに合った観光写真スライドショー生成システムを提案した。Russellの感情モデルを用い、楽曲および観光資源それぞれを感情ベクトル化した。それぞれの感情ベクトルの類似度に基づき、楽曲と観光資源とのマッチング方式を提案した。

被験者実験の結果より、ランダム手法に比べて、提案手法では楽曲に合った観光写真スライドショーを生成できることを示した。しかしながら、平均評価値としては低く、楽曲に合っていない観光写真も提示されることがあった。今後は、感情を介した楽曲と観光資源写真のマッチング方式について、さらに検討する。

謝 辞

本研究はJSPS科研費JP15K12151,JP16H05932の助成を受けたものです。ここに記して謝意を表します。

文 献

- [1] Xian-Sheng Hua, Lie Lu, and Hong-Jiang Zhang. Content based photograph slide show with incidental music. In *Proceedings of the International Symposium on Circuits and Systems*, 2003.
- [2] 杉本蒼, 奥健太. playlist2scapeseq: 楽曲プレイリストに基づくシーケンス景観の生成. ARG 第11回Webインテリジェンスとインタラクション研究会, pp. 81–84, 2017.
- [3] James A. Russell. A Circumplex Model of Affect. *Journal of Personality and Social Psychology*, Vol. 39, No. 6, pp. 1161–1178, 1980.
- [4] James A. Russell and Geraldine Pratt. A description of the affective quality attributed to environments. *Journal of Personality and Social Psychology*, Vol. 38, No. 2, pp. 311–322, 1980.
- [5] Cheng-Te Li and Man-Kwan Shan. Emotion-based impressionism slideshow with automatic music accompaniment. In *Proceedings of the 15th ACM international conference on Multimedia*, pp. 1–5, 2007.
- [6] Francesco Ricci. Travel Recommender Systems. *IEEE Intelligent Systems*, Vol. 4, No. December, pp. 55–57, 2002.
- [7] Joan Borràs, Antonio Moreno, and Aida Valls. Intelligent tourism recommender systems: A survey. *Expert Systems with Applications*, Vol. 41, pp. 7370–7389, jun 2014.
- [8] Chang Shing Lee, Young Chung Chang, and Mei Hui Wang. Ontological recommendation multi-agent for Tainan City travel. *Expert Systems with Applications*, Vol. 36, No. 3 PART 2, pp. 6740–6753, 2009.
- [9] Joan Borràs, Joan de la Flor, Yolanda Pérez, Antonio Moreno, Aida Valls, David Isern, Alicia Orellana, Antonio Russo, and Salvador Anton-Clavé. SigTur/E-Destination: A System for the Management of Complex Tourist Regions. In R. Law, M. Fuchs, and F. Ricci, editors, *International Conference on Information and Communication Technologies in Tourism, ENTER 2011*, Vol. 3, pp. 39–50. Springer, 2011.
- [10] Chien Chih Yu and Hsiao Ping Chang. Personalized location-based recommendation services for tour planning in mobile tourism applications. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Vol. 5692 LNCS, pp. 38–49, 2009.
- [11] D Martin, a Alzua, and C Lamsfus. A contextual geofencing mobile tourism service. In R. Law, M. Fuchs, and F. Ricci, editors, *Information and communication technologies in tourism 2011*, pp. 191–202. Springer-Verlag, 2011.