MIST: 音楽用自動選択アイコンとその一覧表示に関する一提案

MIST: Music Icon Selector Technique and Its All-in-one Visualization Technique

小田 瑞穂 伊藤 貴之*

Summary. 近年の計算機の発達により、コンピュータに多数の楽曲が保存される機会が多々ある.しかし大量の楽曲の中から聴きたい曲を探し出すのは、困難なことが多い.そこで、楽曲の印象に沿うアイコン画像を自動選択して表示し、さらにそれらを一画面に表示することで、一目で聴きたい楽曲を探し出せる表示手法を提案する.楽曲の印象に沿う画像の選択には、まず楽曲と画像の特徴値を抽出し、関連性をニューラルネットワークで予め学習させ、楽曲から算出される理想的な画像の特徴値にもっとも近い画像を選択する.提案手法では階層型データ可視化手法「平安京ビュー」を用いて楽曲ファイルの階層構造を表現し、個々の楽曲をアイコン表示する.

1 概要

近年のコンピュータの記憶容量の増加により,個人の1台のコンピュータに大量の音楽が保存できるようになった.これに伴い,コンピュータで音楽を鑑賞する人が増えている.またプレイリストに代表されるように,楽曲の雰囲気や印象に応じて楽曲を選択できるインタフェースが提供されている.しかし,プレイリストで楽曲を提示するために表示する情報は,曲名などの文字列を列挙するだけである.このような表示方式では,大量の曲名の中から聴きたい雰囲気の曲を直感的に探し出すのは難しい.

そこで本報告では,ユーザが所有する楽曲の印象を表すアイコンを自動選出することを目的として,各々の楽曲に対して印象の近い画像を組み合わせる手法を提案する.本手法を用いることで,多くの楽曲をアイコンを用いて一覧表示した時に,ユーザが聴きたい印象の楽曲を自在に選択できると考えられる.

本手法ではまず,各々の楽曲と画像について,特徴量を自動算出する.続いてユーザにサンプル楽曲を聴かせ,この楽曲に印象の合う画像をサンプル画像の中から選択させる.この選択結果を収集した後,サンプル楽曲の特徴量と,このサンプル楽曲に対して選出されたサンプル画像の特徴量の相関性を学習する.我々の実装では,学習にはニューラルネットワーク (NN: Neural Network) を用いている.さらに,この学習結果に基づいて,個々の楽曲について,印象が最も合うと思われる画像の特徴値に近い画像を,アイコンとして選択する.

また本報告では,階層型データ可視化手法「平安京ビュー」[1] を用いて楽曲のアイコン画像選択結果を一覧表示する手法を提案する(図1参照).我々の実装では,楽曲はフォルダによって分類されてい

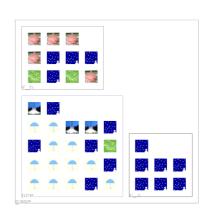


図 1. 平安京ビューを用いた楽曲表示例

ると仮定し,フォルダを枝ノード,各楽曲を葉ノードとみなして表示する.これにより,階層構造を構成する多量の楽曲を,一画面に一覧表示できる.

なお本手法は,大衆からのアンケート集計によって得られた平均的感性に合致するアイコン選択を目的とするのではなく,一人一人の感性に適応するアイコン選択を目的としている.また,タイトルMISTはMusic Icon Selector Technique の略であり,我々による既発表手法[2]の名前である.本報告は,その改良と拡張に関する報告と位置づけられる.

2 関連研究

楽曲の雰囲気に合致するアイコンを提示する研究は,我々以外にもいくつかのグループが着手している[3]が,その多くは曲名やタイトルなどのメタ情報から楽曲の印象を推定している.

我々の研究はそれに対して,メタ情報ではなく, 楽曲の特徴量に基づいてアイコン画像を選択するも のである.この考え方は我々自身が既に先行発表し ている[2]が,本報告ではその一部の改良を報告す

Copyright is held by the author(s).

^{*} Mizuho Oda and Takayuki Itoh, お茶の水女子大学大学院

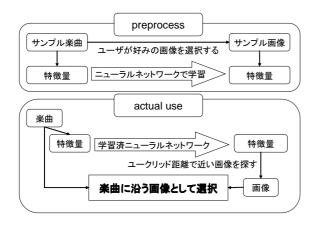


図 2. 処理手順

るとともに,楽曲群の一覧表示に関する一手法を提案するものである.

提案手法に適用する階層型データ可視化手法「平安京ビュー」[1] は,階層構造の最下位階層に属する大量の葉ノードを一覧表示する目的に向いている.我々は既に「平安京ビュー」を,静止画像や CG コンテンツなどの一覧表示に適用している.しかし,音楽コンテンツの一覧表示に関する報告は,本報告が初めてである.

3 提案内容

本手法の処理手順を図2に示す.本手法ではまず, 楽曲群と画像群を用意する.ただし楽曲,画像とも に後から入手したものを追加可能であるとする.

画像はデジタルカメラで撮影した写真や,コンピュータ上で描いた絵を,縦横48画素の正方領域に切り取り,BMP画像として用いる.画像の特徴量には,画素数が最も多い色,彩度と明度の合計値が最も大きい色,の2色の画素値を,RGB表色系からYCbCr表色系に変換して用いる.YCbCr表色系を用いる理由は,色情報から明度を分離し,色相や彩度に基づいて色の印象を学習させるためである.

楽曲にはクラシック音楽などを WAVE ファイル に変換して用いた.楽曲の特徴量には,平均音量,単位時間当たりの音数,最頻出の音高,調を用いている.なお本研究では楽曲の特徴量を重視するため,歌詞やアーティスト名などのメタ情報は参照しない.

続いて,楽曲と画像の特徴量の関連性を,コンピュータに学習させる.具体的には,まずユーザにサンプル楽曲を聴かせ,この楽曲に印象の合う画像をサンプル画像の中から選択させる.このサンプル楽曲の特徴量を入力とし,サンプル画像の特徴量を出力とし,その間にある関連性を学習させる.我々

の実装では,学習には教師付き学習手法である NN を用い,その収束には誤差逆伝播法を用いている.

続いて各々の楽曲について,最も印象が合うとされる理想的な画像の特徴量を算出する.そして多次元の特徴量空間において,既に用意されている各々の画像の中から,理想的な特徴量との距離が最小であるものを,楽曲のアイコン画像として選出する.我々の実装では,特徴量間の距離の算出にはユークリッド距離を用いている.

これらのアイコン画像の選択結果を一覧表示する 一手段として,提案手法では「平安京ビュー」を用 いる(図1参照).平安京ビューでは,フォルダの 階層構造を入れ子の長方形領域で表現し,隙間なく 画面に敷き詰める手法であり,大量の楽曲の一覧表 示に向いている.このような一覧表示は,聴きたい 曲の選曲を支援するだけでなく,例えば同じ印象の 楽曲を同じフォルダにまとめるなど,コンピュータ 内の楽曲ファイルの整理にも使えると考えられる.

4 実験

被験者には,楽曲各々に最短距離となる画像を組みあわせた結果を,図1のように表示し提示した.なお,フォルダと楽曲の関連は筆者が設定した.ここで,実験結果の満足具合を調べるために満足度を定義する.満足度は満足した表示アイコンの数を,すべての表示アイコンの数で割ったものとする.

5 まとめと今後の課題

本報告では,音楽に印象の合う画像をアイコンとして自動で提示し,その結果を一画面に表示する手法を提案した.今後はこれらの実験の満足度をあげるために,楽曲と画像の特徴量をどのように採取するかについて再検討したい.また現時点では,同一階層内の楽曲の表示位置は無作為であるため,表示結果に統一性が見られないことがある.今後は同じアイコン画像を選択した楽曲の配置を近づけるなど,一覧表示の統一性にも重点を置きたい.

参考文献

- T. Itoh, H. Takakura, A. Sawada, and K. Koyamada: "Hierarchical Visualization of Network Intrusion Detection Data in the IP Address Space", IEEE Computer Graphics and Applications, Vol. 26, No. 2, pp. 40-47 (2006).
- [2] M. Oda and T. Itoh: "MIST: A Music Icon Selection Technique Using Neural Network", NICO-GRAPH International 2007, (2007).
- [3] P. Kolhoff, J. Preub and J. Loviscach: "Music Icons: Procedural Glyphs for Audio Files", Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI'06), pp. 289-296, (2006).