

MIST: 音楽アイコン自動選択とその一覧表示の一手法

小田 瑞穂 (指導教官: 伊藤 貴之)

1 概要

近年、計算機の記憶容量の増加の恩恵により、個人の1台の計算機に大量の音楽が取り込まれ、鑑賞する機会が増えている。しかし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。

そこで本論文では、ユーザが所有する楽曲の印象を表すアイコンを自動付与することを目的として、楽曲と楽曲に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、アイコンを用いて多くの楽曲を一覧表示した時に、ユーザが聴きたい印象の楽曲を選択しやすくなると考えられる。

なお本手法は、大衆からアンケートにより集計した平均的印象に合致するよう音楽と画像を組み合わせるのではなく、一人一人の印象に合致する組み合わせを提示することを目的としている。

2 関連研究

楽曲の雰囲気に沿った画像やアイコンを提示する研究の多くは、曲名やタイトルなどのメタ情報を利用して楽曲の印象を決定している。しかし本研究ではメタ情報ではなく、楽曲の音響的情報に着目してアイコンを提示する手法を提案する。また Kolhoff らの研究 [1] では、楽曲の信号処理を基にアイコンを提示しているが、あらかじめ規定された幾何学模様状のアイコン内から選ぶため、そのアイコン表現能力は非常に限られている。本研究では、写真画像などユーザが自由にアイコンとして採用することを前提としており、[1] よりも柔軟なアイコン表現を目指している。

3 提案内容

本研究で提案する MIST (Music Icon Selector Technique) は、楽曲に合う画像を自動付与し、階層型データ可視化手法「平安京ビュー」を用いて一覧表示するものである。図 1 に楽曲表示結果の一部を示す。この表示では、階層構造を二次元の長方形群の入れ子構造で表現し、各楽曲と印象の合うアイコン画像を葉ノードとして貼り付ける。

本手法の処理手順を図 2 に示す。本手法ではまず、組み合わせたい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。画像は縦横 48 画素の正方領域の BMP 画像を用い、楽曲には MP3 ファイルを WAVE ファイルに変換した楽曲を用いた。

次に本手法では、これらの特徴値をそれぞれ抽出する。本手法では、画像の特徴値には、「画素数が最多の色」と「彩度と明度の合計値が最多の色」の計 2 色を

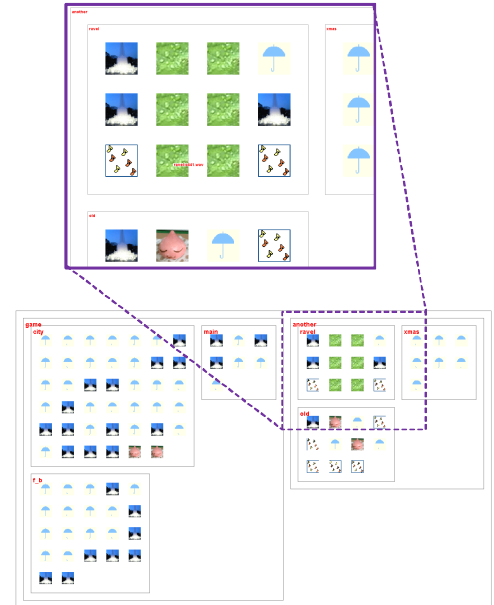


図 1: MIST による楽曲表示例

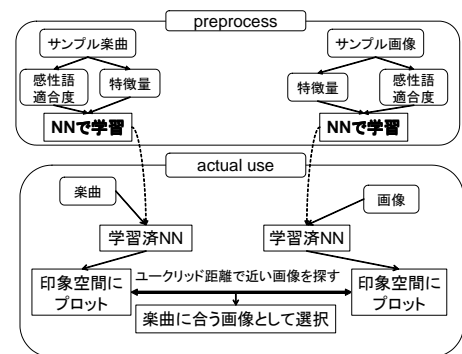


図 2: 処理手順

抽出し、RGB 表色系から YCbCr 表色系に変換した値の一部を利用する。楽曲の特徴値には、フーリエ変換などを用いた信号解析結果から、音楽の 3 要素に基づき、音量ピークの出現回数、メロディ音域の音量率、長和音と短和音の比率、を算出して利用した。

続いて、これらの特徴値から受ける印象を求めるために、ニューラルネットワークを利用して感性語適合度を算出する。感性語は表 1 にある通り、反対語の対からなる評価尺度を、8 対の感性語として利用した。これらの感性語は、池添らの研究 [2] にも用いられているものである。

続いて、感性語適合度を利用して、音楽に距離の近い画像を決定し組み合わせる。感性語は 8 項目あるので、8 次元空間に画像の感性語適合度の値をプロットし、同様に 8 次元空間に楽曲の感性語適合度の値をプ

表 1: 感性語一覧

感性語	
明るい	暗い
重い	軽い
硬い	柔らかい
安定	不安定
澄んだ	濁った
滑らか	歯切れ良
激しい	穏やか
厚い	薄い

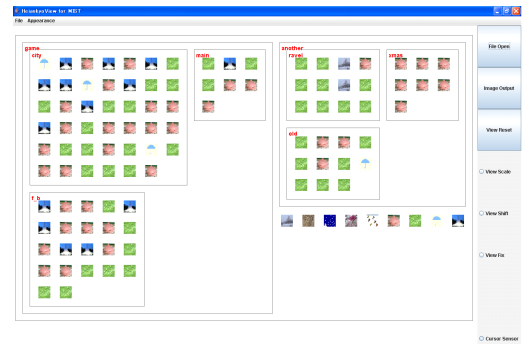


図 3: ある被験者によるアイコン付与結果

ロットする．続いて各々の楽曲に対し，全ての画像との距離をユークリッド距離で計算し，最短距離を算出した画像を組み合わせる．

最後に，各楽曲から最短距離を算出した画像をアイコンとして付与した結果を，階層型データ可視化手法のひとつである「平安京ビュー」を用いて表示する．このとき，ファイルシステムにおけるフォルダを枝ノードとして長方形の枠で表現し，各楽曲を葉ノードとして表現し，付与結果となるアイコン画像を表示する．各楽曲ファイルのファイルシステム上の位置は，利用者があらかじめ設定するものとする．

4 実験

9 枚の画像と 11 曲の楽曲を用いて学習を行い，9 枚の画像と 89 曲の楽曲を用いて実行し，結果を表示した例を，図 3 に示す．また，本実験において被験者には，各楽曲に合う画像をあらかじめリストアップしてもらい，その後本手法を実行した結果を示し，各楽曲に対してアイコン画像の付与結果が満足か否かを回答させた．

本実験結果に対する被験者の満足度を検証するための比較対象として，ランダムに画像を付与した結果も併せて作成した．これを基に，提案手法による学習結果がどれくらい向上したかについて，向上度を定義する．向上度は，提案手法による付与結果で満足できた楽曲数を，ランダムな付与結果で満足できた楽曲数で割ったものとする．

また本研究では，画像と楽曲の特徴を数値化しているため，楽曲と画像の組み合わせに感性語適合度を用いず，楽曲の特徴値から直接画像の特徴値を直接算出することが論理的に可能である．そこで比較実験として，感性語を使用しないで画像を付与した実験も行った．この手法でも同様に，ランダム付与結果との向上度を求める．

4 名の被験者による実験結果を表 2 に示す．向上度 (感性語無) と向上度 (感性語有) を比較した結果，向上度 (感性語有) の方が満足度が高いことがわかった．

表 2: 提案手法による実験結果

被験者	向上度	向上度
	(感性語無)	(感性語有)
A	1.11	0.92
B	1.18	2.18
C	0.70	0.80
D	0.57	1.07

また，楽曲にランダムにアイコン画像を付与するよりも，提案手法により特徴値に重点を置いたアイコン画像に満足することがわかった．

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自動で選択し，アイコンとして付与する手法を提案した．また付与結果の一覧表示を行った．満足度の検証として，ランダムに付与した結果よりも，感性語を用いた提案手法による付与結果は満足度が高いことがわかった．

今後はこれらの実験の満足度を上げるために，本手法のうち教師付き学習の部分を再検討し強化したい．また一覧表示の部分について，ユーザにとってより使いやすい GUI を構築できるように，検討を進めたい．

参考文献

- [1] 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).
- [2] 池添剛, 梶川嘉延, 野村康雄, 音楽感性空間を用いた感性語による音楽データベース検索システム, 情報処理学会論文誌, Vol. 42, No. 12, pp. 3201-3202 (2005).