5Q-9

複合特徴量での適合を考慮した音楽検索手法の提案と評価

二本 真† 高山 毅† 池田哲夫† 奥村一守†
岩手県立大学ソフトウェア情報学部†

1. はじめに

特徴量を用いたマルチメディア検索において，検索結果に対するユーザの満足度を高めるためのアプローチの一つとして，適合フィードバックがある[1]．しかしながら，適合性に関する情報は，特徴量ごとに別個に処理されており，複数の特徴量が融合して検索意図を満たす場合に対応できていない．本研究では，ある特徴量が単体で検索意図を満たす場合に加えて，複数の特徴量が融合して検索意図を満たす場合にも対応可能な検索手法を提案する．具体的には，適合/不適合情報から，単体の特徴量と，複数の特徴量の組み合わせ(以降「複合特徴量」と呼ぶ)のそれぞれに対して，以降の検索で与えるべき重みを決定する．これらの各重みを用いることで，より検索意図に見合う検索結果を得ることが可能になる．提案手法の有効性は，試作システムを用いた評価実験で確認できた．

2. 特徴量ごとに別個に処理する類似度評価方式[1]

2.1 各特徴量に与えるべき重みWの導出

データベースに格納されているオブジェクトOがn個の特徴量$O=\{o^1,o^2,...,o^n\}$を持つとする．文献[1]では検索ベクトルとして，オブジェクト

$$K=\{k^1,k^2,...,k^n\} \tag{1}$$

を指定し，検索の際，どの特徴量をどれだけ重視するかを指定する重み

$$W=\{w^1,w^2,...,w^n\} \tag{2}$$

を以下の(3)式によって決定している．

$$w^\ell=\frac{\sum_{i=1}^{p} f\left(k^\ell,o_{(i)}^\ell\right)}{p}-\frac{\sum_{j=1}^{q} f\left(k^\ell,\hat{o}_{(j)}^\ell\right)}{q} \tag{3}$$

ただし，

$$O_{(i)}=\{o_{(i)}^1,o_{(i)}^2,...,o_{(i)}^n\},(i=1,2,...,p) \tag{4}$$

は，Kに類似とユーザに評価されたp個のオブジェクト群であり，

$$\hat{O}_{(j)}=\{\hat{o}_{(j)}^1,\hat{o}_{(j)}^2,...,\hat{o}_{(j)}^n\},(j=1,2,...,q) \tag{5}$$

は，Kに非類似とユーザに評価されたq個のオブジェクト群である．また，$f\left(k^\ell,o^\ell\right)$は，近傍を示す充分小さな数$\varepsilon$に対して

$$\begin{aligned}&if(k^\ell-\varepsilon\le o^\ell\le k^\ell+\varepsilon)\quad return\ 1\\&else\quad return\ 0\end{aligned} \tag{6}$$

となる類似判定関数である．

2.2 オブジェクトと検索ベクトルの類似度評価

オブジェクトOと検索ベクトルKとの類似度d_{total}は，導出した重みWを用いて，以下の式で評価している．

$$d_{total}=\sum_{\ell=1}^{n} d^\ell \tag{7}$$

ただし，

$$d^\ell=w^\ell f\left(k^\ell,o^\ell\right) \tag{8}$$

3. 複合特徴量での適合を考慮した類似度評価方式の提案

3.1 複合特徴量ごとに見た類似判定

本稿では手始めに，二つの特徴量の組み合わせとして，${}_nC_2$個の複合特徴量

$$O_{comb}=\{o^{(1,2)},o^{(1,3)},..,o^{(1,n)},o^{(2,3)},..,o^{(2,n)},..,o^{(n-1,n)}\} \tag{9}$$

を考慮する．

また，複合特徴量に対する類似判定関数$f_{comb}\left(k^{(u,v)},o^{(u,v)}\right)$を，以下のように定義する．

$$\begin{aligned}&if(f\left(k^u,o^u\right)=f\left(k^v,o^v\right)=1)\quad return\ 1\\&else\quad return\ 0\end{aligned} \tag{10}$$

これはk^u，k^vに対し，同時にo^u，o^vがε近傍内に入っている場合のみ，$k^{(u,v)}$について$o^{(u,v)}$と類似であると評価する関数である．

3.2 各複合特徴量に与えるべき重みW_{comb}の導出

各複合特徴量がとるべき重み

$$W_{comb}=\{w^{(1,2)},w^{(1,3)},..,w^{(1,n)},w^{(2,3)},..,w^{(2,n)},..,w^{(n-1,n)}\} \tag{11}$$

を以下(12)式で評価する．

p，qを(4)，(5)式中と同一の含意として，

$$w^{(u,v)}=w_{relev}^{(u,v)}-w_{irrelev}^{(u,v)} \tag{12}$$

"Content-based Music Retrieval Method Considering Relevance in Multiple Features and its Evaluation"
† M.Futamoto, T.Takayama, T.Ikeda and K. Okumura
Faculty of Software and Information Science,
Iwate Prefectural University

$$w_{relev}^{(u,v)} = \frac{\left\{\sum_{i=1}^{p} f_{comb}\left(k^{(u,v)}, o_{(i)}^{(u,v)}\right)\right\}^2}{E_{relev}\, p} \tag{13}$$

$$w_{irrelev}^{(u,v)} = \frac{\left\{\sum_{j=1}^{q} f_{comb}\left(k^{(u,v)}, \hat{o}_{(j)}^{(u,v)}\right)\right\}^2}{E_{irrelev}\, q} \tag{14}$$

$$E_{relev} = \sum_{r=\max(more+fewer-p,0)}^{fewer} r\left\{\frac{{}_{more}C_r{}^{*}{}_{(p-more)}C_{(fewer-r)}}{{}_{p}C_{fewer}}\right\} \tag{15}$$

$$E_{irrelev} = \sum_{s=\max(more+fewer-q,0)}^{fewer} s\left\{\frac{{}_{more}C_s{}^{*}{}_{(q-more)}C_{(fewer-s)}}{{}_{q}C_{fewer}}\right\} \tag{16}$$

ただし,

$$\begin{aligned} &if\ \left(\sum_{i=1}^{p} f\left(k^{u}, o_{(i)}^{u}\right) \geq \sum_{i=1}^{p} f\left(k^{v}, o_{(i)}^{v}\right)\right) \\ &\qquad more = \sum_{i=1}^{p} f\left(k^{u}, o_{(i)}^{u}\right),\quad fewer = \sum_{i=1}^{p} f\left(k^{v}, o_{(i)}^{v}\right) \\ &else\quad more = \sum_{i=1}^{p} f\left(k^{v}, o_{(i)}^{v}\right),\quad fewer = \sum_{i=1}^{p} f\left(k^{u}, o_{(i)}^{u}\right) \end{aligned} \tag{17}$$

3.3 オブジェクトと検索ベクトルの類似度評価

オブジェクトOと検索ベクトルKとの類似度d_{total}は，特徴量単体での適合と，複合特徴量での適合の双方を考慮して，次式で評価する.

$$d_{total} = \sum_{\ell=1}^{n} d^{\ell} + \sum_{u=1}^{n-1}\sum_{v=u+1}^{n} d_{comb}^{(u,v)} \tag{18}$$

ただし,

$$d_{comb}^{(u,v)} = w^{(u,v)} f_{comb}\left(k^{(u,v)}, o^{(u,v)}\right) \tag{19}$$

4. 試作システムと評価

4.1 試作システム

類似度d_{total}について，評価式(7)を用いた場合(従来手法)と，評価式(18)を用いた場合(提案手法)の二つを実装した．オブジェクトは楽曲データ，具体的には 18 世紀から 20 世紀のピアノ曲 100 曲の SMF(Standard MIDI file)である．特徴量は文献[2][3]を参考に 15 個とした.

4.2 実験方法

被験者 50 名に対し以下の手順で実験を行った.

STEP1) 被験者は検索キーとする楽曲Mと，ほか 19 曲とを比較試聴する.

STEP2) 試聴した 19 曲をMに対する類似楽曲群，非類似楽曲群に分ける.

STEP3) M，類似楽曲群，非類似楽曲群を入力として，残りの 80 曲について従来手法でMと類似と算出される楽曲上位 4 件をもとめる.

STEP4) STEP3 と同じく，提案手法で楽曲上位 4 件をもとめる.

STEP5) 被験者はMと STEP3,4 で得た楽曲計 8 曲について比較試聴を行い，Mに対する適合度を 5 段階(1:類似していない - 5:類似している)で主観評価する.

4.3 実験結果

(i) 平均適合度は，従来手法が 3.215，提案手法が 3.550 で 10.42%の向上が見られた．t検定(両側)を用いて評価したところ，有意水準 1%で，提案手法は有効との結論を得た.

(ii) 検索結果における順位ごとの平均適合度を以下の図 1 に示す．順位 1 位から 4 位までの全ての平均適合度において，提案手法が従来手法を上回る結果となった.

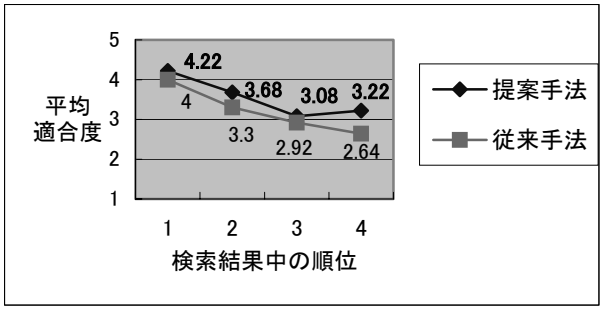

図 1：検索結果中の順位ごとの平均適合度

5. 結論と今後の展望

本稿では，複合特徴量が検索意図を満たす場合も考慮した音楽検索手法を提案した．SMF データを用いて，比較評価実験を行なった結果，平均適合度で 10.42%の向上が見られた．本提案の有効性は確認できたと言える.

今後の展望として，以下の二項目を検討している:(i) 三つ以上の特徴量による複合特徴量に対して与えるべき重みの導出，(ii) 提案手法を画像データに適用しての評価.

参考文献

[1] 寺本純司，紺谷精一，山室雅司，串間和彦：“視覚化と適合フィードバックを用いた画像検索インタフェース”，第 11 回データ工学ワークショップ DEWS2000,5B-1,2000.

[2] 宮下芳明，西本一志：“MIDI データからのクロマプロファイルの抽出と分析”，情処研報 2003-MUS-51,pp.97-101,2003.

[3] 黒瀬崇弘，梶川嘉延，野村康雄：“感性情報を用いた楽曲推薦システム”，第 14 回データ工学ワークショップ DEWS2003,8P-6,2003.