

印象に基づく映像と音楽の相互検索に関する一考察

宝珍 輝尚 井田 俊博 都司 達夫
福井大学 工学部 情報・メディア工学科
〒 910-8507 福井市文京 3 丁目 9 - 1
hochin@pear.fuis.fukui-u.ac.jp

本論文では、印象に基づいて動画クリップと音楽クリップの相互検索を可能とすることを目的として、動画クリップと音楽クリップを感性の主因子により対応付けを行った試みについて述べる。まず、動画クリップと音楽クリップに対する感性の主因子について概説する。次に、感性の主因子を用いた動画クリップと音楽クリップの対応付け法を示す。その後、動画クリップを基にして音楽クリップを検索した結果を示す。検索結果としては、動画クリップに合った音楽クリップが求まる場合と、音楽クリップを聞きながら動画クリップを見ると異なる感じを受けるような音楽クリップが求まる場合があった。

On the Mutual Access of Video and Music Clips Based on Impression

Teruhisa HOCHIN Toshihiro IDA Tatsuo TSUJI
Dept. of Information Science, Faculty of Eng., Fukui University
3-9-1, Bunkyo, Fukui-shi, Fukui 910-8507 Japan
hochin@pear.fuis.fukui-u.ac.jp

This paper describes an attempt of mutual adaptation of video clips and music ones by using the *Kansei* factors. The end of this paper is obtaining the video clips (music clips, respectively) having the similar impression of the specified music clip (video clip). First, the *Kansei* factors of video and music clips are briefly described. Next, the method of the mutual adaptation of video clips and music ones is presented. This method uses the *Kansei* factors of video and music clips. After that, music clips are tried to be obtained by specifying some video clips. In some cases, the music clips suitable to the specified video clip can be obtained. In other cases, the music clips that influence the feeling of the video clips are obtained.

1 はじめに

近年、インターネット上に、画像・動画・音といったマルチメディアデータが遍在するようになってきている。これらのメディアデータを内容に基づいて検索したいという要求は古くからあり、様々な研究が行われてきている。ここで、画像・動画・音といったマルチメディアデータは人間に対してある種の印象を与える。例えば、小川の写真が清涼感を与えるといったことである。従って、印象に基づいてマルチメディアデータを検索したいという要求も当然存在し、印象に基づいて画像・動画・音を検索する研究も盛んに行われている [1-10]。

著者らは、あるメディアデータにふさわしい他メディアのデータを取り出す試みを行っている [20, 21, 22]。画像、音クリップ、音楽クリップ、動画クリップに対して感性の主因子を求めたところ、メディアに共通の因子が存在するのではないかという結果が得られている [21, 22]。また、この結果を利用した、画像、音クリップ、動画クリップの相互検索システムを試作し、明らかに適合していないものを正解集合とすると、様々なメディアデータを感性の主因子により対応付ける方法は良い検索特性を持つことを明らかにしてきた [23, 24, 25]。

しかしながら、音楽クリップについては未サポートであった。特に、動画クリップに対しては、印象の合った音楽を検索したいという要求は強く、これに対する検討が必要である。

そこで本論文では、印象に基づく動画クリップと音楽クリップの相互検索を可能とすることを目的として、動画クリップと音楽クリップを感性の主因子により対応付けを行った試みについて述べる。

以下、2では感性の主因子について述べ、3で試作した検索システムについて述べる。4で考察を行い、最後に、5でまとめる。

2 感性の主因子

2.1 感性の主因子の導出法

心理学の研究では、印象語が表す感性的な性質はいくつかの基本的な因子（感性の主因子）によって表現されることが明らかとなっている。これは、Semantic Differential(SD)法という手法を用いて解析した結果得られるものである。SD法とは心理学者 C. E. Osgood が考案した手法で、複数の反対の

意味を持つ印象語の対を尺度とし、その間をいくつかの段階に分けてある対象物を被験者に評価させるものである [15, 16, 17]。図1は「父」についてどのような印象を受けるかを評価した例である。この例では、7段階で評価している。例えば、happy-sadでは、いくぶん happy であり、hard-soft では、かなり hard であるといった具合である。

```
happy  :_:_:X:_:_:_: sad
hard   :_:_:X:_:_:_: soft
slow   :_:_:_:_:X:_:_: fast
```

図 1: SD 法の例

これらの印象語対がそれぞれ単一次元の特徴を抽出すると仮定し、因子分析を適用すると、次元の少ない、より簡潔なものとすることができる。因子分析では、 n 個の観測対象に対する p 変量のデータの行列を X とすると、 X を下式のように表現する [19]。

$$X = FA' + E$$

ここで、 F は $n \times m$ の行列、 A' は $p \times m$ の行列 A の転置行列、 E は $n \times p$ の行列であり、 m をできるだけ小さくとり、かつ、 E を十分小さくなるように分解し、 F と A を求める。 F は因子得点行列、 A は因子負荷量行列、 E は残差行列と呼ばれる。変数の数 p よりもかなり小さい m を用いることで潜在的な因子を求める。

ここで、 E を十分小さくするにはいくつかの方法がある [19]。 E の分散・共分散行列を $U = E'E$ とすると、 U の各要素の 2 乗和を最小にする方法は主成分分析法と呼ばれる。また、 U の非対角要素の 2 乗和を最小にする方法は Minres 法と呼ばれる。

因子負荷行列 A は一意に決定できるものではなく、自由度がある。通常は、ある変数は絶対値が大きく、他の変数は絶対値が小さくなるような回転を施し、説明をしやすいのが一般的である。良く利用される方法にバリマックス法がある [19]。

また、因子得点行列 F も一意に決定できるものではない。良く知られている方法には以下の方法がある [19]。

$$F_1 = XR^{-1}A \tag{1}$$

$$F_2 = XWA(A'WA)^{-1} \tag{2}$$

ここで、 W は重み行列である [19]。この因子分析を SD 法によって求めた得点に適用すると、ものや概念に内在する因子が得られる [15]。これを感性の主因子と呼ぶ。

表 1: 音楽クリップに対する因子負荷量

形容詞対		第一因子	第二因子	第三因子	第四因子
美しい	- 醜い	-0.927864	0.187161	-0.139654	-0.137400
自然な	- 不自然な	-0.862504	0.197756	-0.133837	-0.394612
新鮮な	- 古くさい	0.143693	-0.366938	0.019719	0.069283
澄んだ	- 濁った	-0.858464	0.162458	-0.215991	0.030834
潤いのある	- 渴いた	-0.861612	-0.017717	-0.187395	-0.377424
かたい	- 柔らかい	0.522295	0.209327	0.229344	0.677070
暖かい	- 冷たい	-0.312847	-0.544928	-0.142103	-0.697552
明るい	- 暗い	0.145262	-0.954926	0.172095	-0.165238
単純な	- 複雑な	0.181876	-0.541903	-0.012666	-0.603547
大胆な	- 繊細な	0.642414	-0.389970	0.634508	-0.074972
動的な	- 静的な	0.569398	-0.565679	0.497459	0.311576
緊張した	- ゆったりした	0.416875	-0.041651	0.284217	0.848456
大きい	- 小さい	0.145097	0.035741	0.920146	0.160475
重い	- 軽い	-0.151100	0.811485	0.423871	0.259657
強い	- 弱い	0.323190	-0.054137	0.881561	0.212338
うれしい	- 悲しい	-0.027394	-0.934982	0.170345	-0.238042

2.2 音楽クリップと動画クリップに対する感性の主因子

画像, 音クリップ, 音楽クリップ, 動画クリップに対して, SD法により評価し因子分析してきている [21]. ここでは, 主に音楽クリップと動画クリップを対象にして, これまでに得られてきている結果を概説する.

ここでは, 被験者がSD法により, 2, 1, 0, -1, -2の5段階で各メディアデータを評価し, このデータを各画像, 各印象語対ごとに平均を求め, Minres法 [19] により因子分析を行って得ている. ここで, 使用した印象語対の数は16である. また, 因子は, 1より大きい固有値を持つ因子を採用することにして得ている. さらに, 因子負荷行列は, バリマックス回転を行って得ている.

対象の音楽クリップは, 約30秒程度の器楽による演奏20個で, おのおの, 約30秒である. これらは, 著作権フリーの素材集 [26, 27] から選択している. 被験者は男子大学生13名である.

対象の動画クリップは, 大道芸やストリートダンス等のストリートパフォーマンスのもの41個で, おのおの約30秒である. これらは著者らが撮影したものである. 被験者は男子大学生6名である.

なお, 対象のメディアデータの一覧を付録に示す. 音楽クリップに対しては, 4つの因子が得られている. 得られている因子負荷量を表1に示す.

動画クリップに対しては, 5つの因子が得られている. 得られている因子負荷量を表2に示す.

さらに, 感性の主因子の意味を説明する変数の決定方法を明確化し, 感性の主因子の説明を試みている [22]. この説明変数の決定方法は, 因子負荷量行列の因子負荷量の大きい印象語対がその因子を表しているという従来の考え方に従っているが, 他の因子と比較して特に重みの重い印象語対を選択するようにしている [22]. この方法により得られた印象語対をまとめて表3に示す.

明快性と名付けた因子は, 音楽クリップと動画クリップで非常に良く似た因子と考えられる. その他の因子は, 大体合っているという程度の因子であろうと考えられる. また, 因子の数が異なっており, 動画クリップにおける活動性の因子が音楽クリップでは現われていない.

3 検索システム

ここでは, 音楽クリップと動画クリップを対象とした相互検索について述べる.

表 2: 動画クリップに対する因子負荷量

形容詞対		第 1 因子	第 2 因子	第 3 因子	第 4 因子	第 5 因子
速い	- 遅い	0.538545	0.244507	0.283440	-0.576095	0.039596
規則的な	- 不規則な	0.011389	-0.039909	0.042983	-0.857792	-0.055398
動的な	- 静的な	0.432494	0.458802	0.089447	-0.630052	-0.194658
美しい	- 醜い	0.318047	0.274800	0.554906	-0.357700	0.097178
なめらかな	- なめらかでない	0.214270	0.826630	0.185568	-0.304658	0.125032
メリハリのある	- メリハリのない	0.781336	0.078301	0.295757	-0.183973	0.088458
直線的	- 曲線的	0.141999	-0.729723	0.027660	0.129440	0.268317
明るい	- 暗い	0.123178	-0.171992	0.921306	-0.217026	-0.216238
愉快的な	- 不愉快的な	0.097099	0.201310	0.770028	-0.010341	0.008387
複雑な	- 単純な	0.360257	0.552325	0.290819	-0.254394	-0.206072
自然な	- 人工的な	-0.024389	0.163046	0.070997	-0.086998	-0.612873
大胆な	- 繊細な	0.776792	0.142005	-0.034649	0.041973	-0.172107
緊張した	- ゆったりした	0.726090	-0.164509	0.013147	-0.147228	0.212057
迫力のある	- 迫力のない	0.879302	0.016049	0.210259	-0.084357	-0.025818
堅い	- 柔らかい	0.027794	-0.837093	-0.023316	-0.043819	0.102004
軽快な	- 重厚な	0.109458	0.411858	0.323982	-0.730394	-0.109490

表 3: 感性の主因子を説明する印象語対

主因子	音楽クリップ	動画クリップ
自然性	美しい - 醜い 自然な - 不自然な 潤いのある - 渴いた 澄んだ - 濁った	自然な - 人工的な
明快性	明るい - 暗い うれしい - 悲しい	明るい - 暗い 愉快的な - 不愉快的な
堅鋭性	緊張した - ゆったりした	なめらかな - なめらかでない 堅い - 柔らかい 直線的 - 曲線的
力量性	大きい - 小さい 強い - 弱い	迫力のある - 迫力のない メリハリのある - メリハリのない 大胆な - 繊細な 緊張した - ゆったりした
活動性	-	規則的な - 不規則な

3.1 対応付け法

試作している検索システムでは、メディアデータ

の因子得点をもとにする。ただし、動画クリップにおける活動性の因子が音楽クリップには存在しない。そこで、今回は、活動性の因子を使用しないで対応付けを行った。したがって、あるメディアデータが与えられたとき、その i 番目の因子の因子得点を tm_i とし、検索対象のデータの i 番目の因子の因子得点を m_i とすると、 $\sqrt{\sum_{i=1}^4 (tm_i - m_i)^2}$ により距離を求め、この距離の小さいもの（相違度が低いもの）を候補とする。ただし、因子得点は、2.1で示した式(2)を使用して求めている。

3.2 検索概要

試作した検索システムの検索法の概要を示す。

動画クリップをもとに検索を行う場合の検索画面の例を図 2 に示す。アイコンをクリックすると、動画が再生される。最上位の動画は、赤い T シャツを着た男性がストリートパフォーマンスを行っているものである。アイコンの下に、対応する画像、音クリップや音楽クリップを求めるボタンがある。ここで、対応する音楽クリップを求めるボタンを押すと、図 3 に示す画面が表示される。

図 3 において、検索結果を示すアンカーをクリックすると、音楽クリップが再生される。この例では、

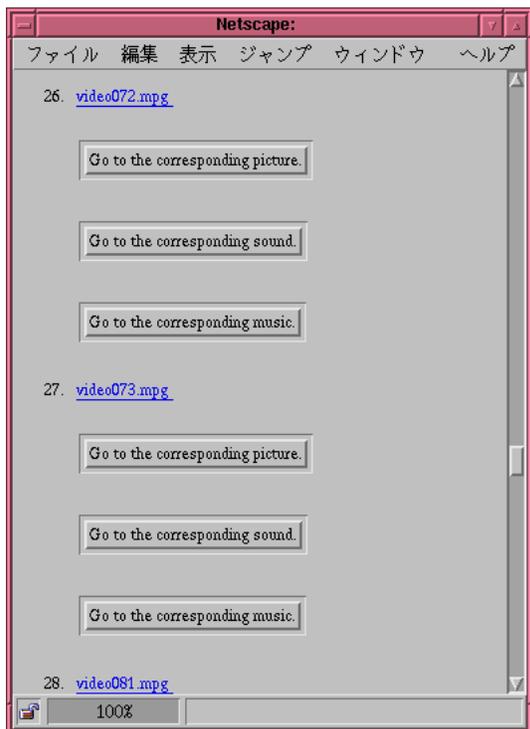


図 2: 動画クリップからの検索画面

ロックやラップ等が得られた。また、アンカーの下には、その音楽クリップに対応する画像、音クリップや動画クリップを求めるボタンがある。このように、多種のメディアデータをキーにして相互に検索することが可能である。

3.3 検索例

ここでは、いくつかの動画クリップから音楽クリップを求めた結果を示す。それぞれ、検索結果の上位5件を示す。また、結果に付した番号は、付録に示した対象メディアデータの番号である。

(1) 動画クリップ「5. ジャグリング, 輪, 1人, 男性」に対する検索結果を以下に示す。

1. ウィリアムテル序曲 (17)
2. Sound Track(20)
3. Rock(5)
4. Rap(12)
5. French Waltz(9)

これらの音楽クリップは動画クリップに良く合っているものであった。

(2) 動画クリップ「6. ジャグリング, こん棒, 1人, 3人, 白黒」に対する検索結果を以下に示す。

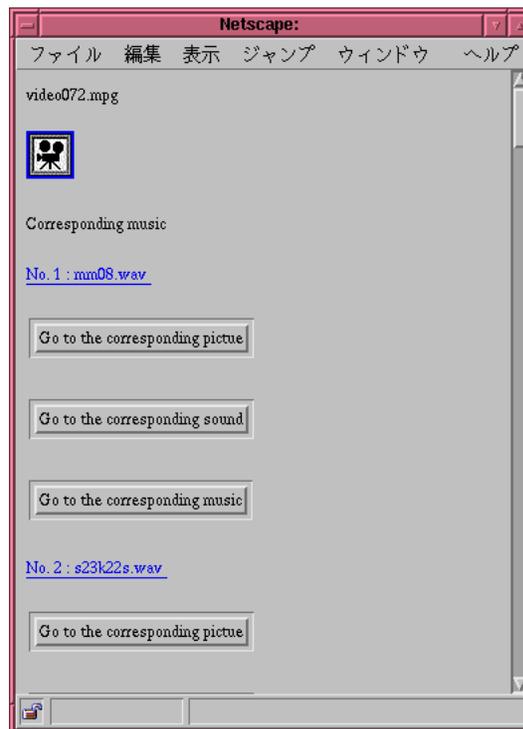


図 3: 動画クリップから音楽クリップの検索結果画面

男性」に対する検索結果を以下に示す。

1. 和 2(18)
2. Overture(3)
3. French Waltz(9)
4. オルゴール (Mozart)(13)
5. Slow Jazz(トランペット)(6)

「9. French Waltz」は非常に良く合っている。その他の4曲は静かな曲であり、合っていないことのないという程度のものであった。

(3) 動画クリップ「12. 大道芸, 呼び出し, おじさん」に対する検索結果を以下に示す。

1. Blue Grass(バンジョー)(10)
2. Rap(12)
3. Rock(5)
4. Bridge(7)
5. Flamenco(11)

「10. Blue Grass(バンジョー)」は非常に良く合っているが、その他は合っているとは言い難い。

(4) 動画クリップ「19. ストリートダンス, 女の子, 3人, 白黒」に対する検索結果を以下に示す。

1. Piano(4)
2. Bach Jazz(1)
3. Overture(3)
4. French Waltz(9)
5. Allegro(15)

「4. Piano」と「3. Overture」はあまり合っているとは言いがたく、その他は当たらずとも遠からずという程度である。

(5) 動画クリップ「26. ストリートパフォーマンス, 男, 1人, 赤, 人形の動き」に対する検索結果を以下に示す。

1. Rock(5)
2. Rap(12)
3. Bridge(7)
4. Reggae(8)
5. Flamenco(11)

これらの音楽クリップは動画クリップに良く合っているものであった。

(6) 動画クリップ「34. 女のピエロ, パントマイム, 赤」に対する検索結果を以下に示す。

1. Reggae(8)
2. コミカル 2(オルガン)(14)
3. Rock(5)
4. Rap(12)
5. Blue Grass(バンジョー)(10)

「8. Reggae」と「14. コミカル 2(オルガン)」は良く合っているが、その他は合っているとは言いがたい。

4 考察

動画クリップ「5. ジャグリング, 輪, 1人, 男性」, 「12. 大道芸, 呼び出し, おじさん」, 「26. ストリートパフォーマンス, 男, 1人, 赤, 人形の動き」, ならびに「34. 女のピエロ, パントマイム, 赤」については、上位の検索候補がそれぞれの動画クリップと良く合っている。12と34については、下位の検索候補はあまり合っていないが、これは音楽クリップが20個しかなく、ふさわしい音楽クリップがないためであろうと考えられる。また、今回使用した動画クリップはストリートパフォーマンスのものであるため、リズムカルな音楽クリップが選ばれる傾向がある。

動画クリップ「6. ジャグリング, こん棒, 1人, 男性」は、男性が1人でこん棒でジャグリングをしているクリップであるが、音楽クリップ「9. French Waltz」はこれに非常に良く合っている。一方、この動画に対しては、音楽クリップ「18. 和 2」や「6. Slow Jazz(トランペット)」も得られている。音楽クリップ「18. 和 2」はゆったりとした和風の曲であり、この曲を聞きながらこの動画クリップを見ると、昼下がりのけだるさを感じるような感じがする。また、音楽クリップ「6. Slow Jazz(トランペット)」は哀愁に満ちたトランペットの曲であるが、この曲を聞きながらこの動画クリップを見ると、ジャグリングをしている男性がかわいそうに感じてくる。音楽クリップ「18. 和 2」や「6. Slow Jazz(トランペット)」は、ともにこの動画クリップに合っていないとは言えないが、これらの音楽クリップを聞きながら動画クリップを見ると、別の雰囲気を感じる。これは、動画クリップ「19. ストリートダンス, 女の子, 3人, 白黒」についてもいえることである。以上のことは、動画クリップの許容範囲の広さを示すものであるとも考えられる。また、音楽クリップが別の雰囲気にさせてしまうともいえる。この意味では、音楽クリップと動画クリップが互いに作用し合ってそれぞれにはない雰囲気を創り出しているとも考えられる。この特性は、画像、音クリップ、ならびに、動画クリップの相互検索ではあまり見られなかった。動画クリップと音楽クリップに特に顕著に見られる特性であるとも考えられる。

以上より、動画クリップと音楽クリップの相互検索はかなり困難であると考えられる。得られたものが求める雰囲気と合致しているとは限らないからである。したがって、求める雰囲気を何らかの方法で得る必要があると考えられる。一方、これを積極的に利用すると、考えてもみなかった雰囲気の動画クリップと音楽クリップの組合せが求まる可能性もある。これらについての検討は今後の課題である。

5 おわりに

感性の主因子を用いた動画クリップと音楽クリップの相互検索を実現した。実際に検索を行ったところ、動画クリップに合った音楽クリップが求まった場合と、音楽クリップを聞きながら動画クリップを見ると異なる感じを受けるような音楽クリップが求まった場合があった。

今後は、動画クリップと音楽クリップの相互作用についての検討、各メディアデータの特徴量と因子得点の関係の明確化、感性の個人差への対応、大規模データベースへの適用などが課題である。

謝辞

本研究は、一部、財団法人 電気通信普及財団の助成による。ここに記して謝意を表す。

参考文献

- [1] Uemura, S., Arisawa, H., Arikawa, M., and Kiyoki, Y. : "Digital Media Information Base", IEICE Trans. Inf. & Syst., Vol. E82-D, No. 1, pp.22-33 (1999).
- [2] 栗田 多喜夫, 加藤 俊一, 福田 郁美, 板倉 あゆみ: "印象語による絵画データベースの検索", 情処論, Vol. 33, No. 11, pp.1373-1383 (1992).
- [3] 清木 康, 金子 昌史, 北川 高嗣: "意味の数学モデルによる画像データベース探索方式とその学習機構", 信学論 D-II, Vol. J79-D-II, No. 4, pp. 509-519 (1996).
- [4] 八村 広三郎, 英保 茂: "色彩分布と印象語に基づく絵画データの検索", 情処研報 CH-27, Vol. 95, No. 91, pp.37-44 (1995).
- [5] 丸山 竜也, 勝木 道哲, 柴田 義孝: "イメージ語を用いたデザイン画像データベース検索システム", 情処学第45回全国大会 F-9, pp.4-87 - 4-88 (1993).
- [6] 家出 太郎 他: "景観画像を対象にした感性語による画像検索システム", 平成12年度電気関係学会北陸支部連合大会 F-7, pg.308 (2000).
- [7] 佐々木 和也, 清水 裕子, 春日 正男, 庄 健二: "二色配列が視覚イメージに与える影響", 第2回日本感性工学会大会予稿集, pg.106 (2000).
- [8] 佐藤 聡, 菊地 幸平, 北上 始: "音楽データを対象としたイメージ検索のための感情価の自動生成", 情処研報, DBS 118-8 FI 54-8, pp. 57-64 (1999).
- [9] 吉野 太智, 高木 秀幸, 清木 康, 北川 高嗣: "楽曲データを対象としたメタデータ自動生成方式とその意味的連想検索への適用", 情処研報, DBS 116-41, pp. 109-116 (1998).
- [10] 辻 康博, 星 守, 大森 匡: "曲の局所パターン特徴量を用いた類似曲検索・感性語による検索", 信学技報, 音声研究会 SP96-124, pp. 17-24 (1997).
- [11] 鄭 載旭, 原田 昭: "二色配列が視覚イメージに与える影響", 第2回日本感性工学会大会予稿集, pg. 67(2000).
- [12] 神里 志穂子 他: "舞踏における手指軌道の運動特性と主観的印象との関係", 信学技報 HIP2000-14, pp. 47-51 (2000).
- [13] Chen, P., Hearst, M. Kupiec, J., Pedersen, J. and Wilcox, L. : "Metadata for Mixed-Media Access," SIGMOD RECORD, Vol. 23, No. 4, pp.64-71 (1994).
- [14] 柳沼 良知, 坂内 正夫: "DP マッチングを用いたドラマ映像・音声・シナリオ文書の対応付け手法の一提案", 信学論 D-II, Vol. J79-D-II, No. 5, pp.747-755 (1996).
- [15] Snider, J. G. and Osgood, C. E. : "Semantic Differential Technique - A Sourcebook," Aldine Publishing Company (1969).
- [16] 金子隆芳: "色彩の心理学", 岩波書店, 岩波新書 134 (1990).
- [17] 井口 征士 他: "感性情報処理", オーム社 (1994).
- [18] 太田 昭雄, 河原 英介: "色彩と配色", グラフィック社 (1976).
- [19] 奥野 忠一, 久米 均, 芳賀 敏郎, 吉澤 正: "多変量解析法", 日科技連 (1981).
- [20] Hochin, T., Yamada, K., and Tsuji, T.: "Multimedia Data Access Based on the Sensitivity Factors, Proc. of the 2000 International Database Engineering & Applications Symposium, pp. 319-326 (2000).
- [21] 宝珍 輝尚, 都司 達夫: 感性に基づくマルチメディアデータの関連付けに関する一考察, 電子情報通信学会技術報告, CS2000-113, 19-24 (2000).
- [22] 宝珍 輝尚, 高田 真介, 都司 達夫: 感性に基づくマルチメディアコーディネーションの一検討, 日本感性工学会 感性工房部会研究会 KF-4, 11-18(2001).
- [23] 宝珍 輝尚, 高田 真介, 都司 達夫: 感性に基づくマルチメディアデータの相互アクセスについて, 情報処理学会研究報告 DBS124-3 FI62-3, 2001, 44, 17-24 (2001).

- [24] T. Hochin, T. Tsuji: Mutual Multimedia Access using Kansei Factors, *Kansei Engineering International*, Vol. 2, No. 4, pp. 9–18 (2001).
- [25] 宝珍 輝尚, 都司 達夫: 感性に基づくマルチメディアデータの相互アクセス法, *情報処理学会論文誌*, Vol. 43, No. SIG 2(TOD 13), pp. 69–79 (2002).
- [26] (株) ウィネット: Music Wizard, (株) ウィネット.
- [27] (株) ウィネット: Music Wizard II, (株) ウィネット.

付録 対象メディアデータ

A.1 音楽クリップ

1. Bach Jazz
2. Capriccio Italiano
3. Overture
4. Piano
5. Rock
6. Slow Jazz(トランペット)
7. Bridge
8. Reggae
9. French Waltz
10. Blue Grass(バンジョー)
11. Flamenco
12. Rap
13. オルゴール (Mozart)
14. コミカル 2(オルガン)
15. Allegro
16. Invention(Bach)(チェンパロ)
17. ウィリアムテル序曲
18. 和 2
19. トッカータとフーガ
20. Sound Track

A.2 動画クリップ

1. ジャグリング, 玉, 1人, 男性
2. ジャグリング, 棒, 1人, 男性
3. ジャグリング, 玉, 2人, 男性
4. ジャグリング, 玉, 2人, 男性
5. ジャグリング, 輪, 1人, 男性
6. ジャグリング, こん棒, 1人, 男性
7. ジャグリング, こん棒, 2人, 男性
8. ジャグリング, 玉, 4人, 男性
9. 弾き語り, 一人, 男性
10. 弾き語り, 二人, 男性

11. 大道芸, 皿回し, おじさん
12. 大道芸, 呼び出し, おじさん
13. 大道芸, 国旗, 花火, 柳, おじさん
14. 大道芸, 大根切り, おじさんと女の子
15. ストリートダンス, 子供, 2人, カラフル
16. ストリートダンス, 子供, 3人, カラフル
17. ストリートダンス, 子供, 5人, カラフル
18. ストリートダンス, 女の子, 2人, 白黒
19. ストリートダンス, 女の子, 3人, 白黒
20. ストリートダンス, 女の子, 2人から多数, 白黒
21. ストリートダンス, 女の子, 多数と1人, 白黒と赤
22. 空手の組み手, 男, 5人
23. 側転と起き上がり, 男, 1人
24. 頭上かわら割り, 男, 1人
25. ストリートダンス, 男, 5人, 赤
26. ストリートパフォーマンス, 男, 1人, 赤, 人形の動き
27. ストリートダンス, 男, 1人, 赤, 回転
28. ストリートダンス, 女子, 4人, 白と赤, イェーイェー
29. ストリートダンス, 女子, 4人, 白と赤, 活発
30. ストリートダンス, 女子, 4人, 白と赤, グニャグニャ
31. ピエロ, 風船, 携帯電話
32. ピエロ, 風船, ハート
33. ピエロ, ジャグリング, 箱, 活発
34. 女のピエロ, パントマイム, 赤
35. 女のピエロ, 手品, 赤
36. 女のピエロと男性, 皿回し, 皿移し
37. 女のピエロと男性, 風船
38. 女のピエロとカップル, 告白劇
39. ストリートダンス, 女子, 4人, タンクトップとジャージ, グニャグニャ
40. ストリートダンス, 女子, 4人, タンクトップとジャージ, 活発
41. ストリートダンス, 女子, 4人, タンクトップとジャージ, 移行