九州大学学術情報リポジトリ
Kyushu University Institutional Repository

印象に基づくメディアデータベース検索およびメディア変換システム

野田, 寿彦
九州芸術工科大学大学院

趙, 東
九州芸術工科大学大学院

高木, 英行
九州芸術工科大学芸術情報設計学科

寺岡, 章人
(財)九州システム情報技術研究所

https://hdl.handle.net/2324/4481618

出版情報：pp.1-4, 1999-12-18. 日本知能情報ファジィ学会
バージョン：
権利関係：

KYUSHU UNIVERSITY

第1回日本ファジィ学会九州支部学術講演会(1999年12月18日，佐賀大学), pp. 1〜4

A101

印象に基づくメディアデータベース検索およびメディア変換システム

Media Database Retrieval and Conversion System based on Impression

野田 寿彦* 趙 東* 高木 英行** 寺岡章人***

Toshihiko NODA* Dong ZHAO * Hideyuki TAKAGI** Akito TERAOKA***

九州芸術工科大学 *大学院, **芸術情報設計学科

***(財) 九州システム情報技術研究所

Kyushu Institute of Design, *Graduate School, **Dept. of Art and Information Design

***Institute of Systems and Information Technologies/Kyushu

Abstract: We propose a multi-media database retrieval system based on KANSEI and a media converter that keeps same media impression. We first construct a universal psychological space which expresses impressions caused by physical features of several kinds of media. The psychological space is obtained from adjective-pair space by using component analysis. As its axes are orthogonal, it is expected to be able to search the space and retrieve media effectively. Retrieving media through the psychological space is conducted by neural network that nonlinearity maps one coordinate on physical feature space to corresponding coordinate on the psychological space and genetic algorithm that searches on the physical feature space.

***Key Words*: emotional information processing, database retrieval, genetic algorithm**

1. 背景・目的

マルチメディアの時代と呼ばれる現在、データベースはどんどん大きくなってきており、効率の良い検索方法が必要となってきている。しかし、画像や音楽などのマルチメディアを見たり聞いたりする対象は人であるので、“人がどのように感じるか” ということを考えたうえで検索を行うことは重要なことである。

このような、いろいろなメディアとそれが与える印象の関係を利用したメディア検索の方法も提案されている。対話型進化計算により、検索したい画像に表示された画像がどれだけ近いかをもとに検索する方法（[2], [3]）や、形容詞対を用いて検索する方法 （[1], [6]） などがある。しかし形容詞対空間を用いた場合は軸の直交性などの問題で冗長度が高く、検索効率を悪くしていることが考えられる。

そこで本研究では、主成分分析により形容詞空間を直交化した心理空間を構成することにより、図1のような、印象に基づく検索を効率良く行うシステムを提案する。

また、音楽と画像など、複数のメディアで共通の心理空間を構成することができれば、それを仲介にして相互に検索することにより、図2のような印象を共有するメディア変換を行うことができる。

この論文では、2節では検索システムで用いる心理空間の構成法、3節で心理空間を用いたメディアデータベース検索、4節で印象に基づくメディア変換についてを述べる。

2. 心理空間の構成

図1、図2のような感性モデルを利用したシステムは、人が画像を見たり音楽を聴いたりしたときに受ける印象を用いる。そのため、人が画像や音楽などのメディアに触れた時の印象がどのような因子であるかを調べ、それらの物理量との関係を調べる必要がある。この対応関係が分かると、印象に基づくデータベース検索・変換を行うことができる。この関係を調べるために、SD法、主成分分析を用い、その結果得られた人がメディアに触れたときに用いられている尺度を心理空間として表した。この心理空間がメディア検索・変換システムの中心となる。

心理空間は、画像を用いてSD法によって構成した。画像を提示し、被験者は画像から受けた印象を14組の形容詞対について7段階で答えてもらった。この結果から主成分分析によって心理空間を構成した。

2.1 形容詞空間の構成

心理空間の構成で用いられた形容詞は、次のようにして決められた。50枚の画像を3名の被験者に提示し、受けた印象を形容詞でできるだけ多く書き出してもらい151語の形容詞を得た。その形容詞の中から、人によって評価の基準が変わるもの、画像の内容に依存するもの、意味が似ているものを削除し、残った形容詞から対を作った。対を作る際には、画像と音楽とで形容詞が表す印象が同じになるようにした。こうすることで、画像と音楽との間で共有できる心理空間を構成する。この実験に用いた

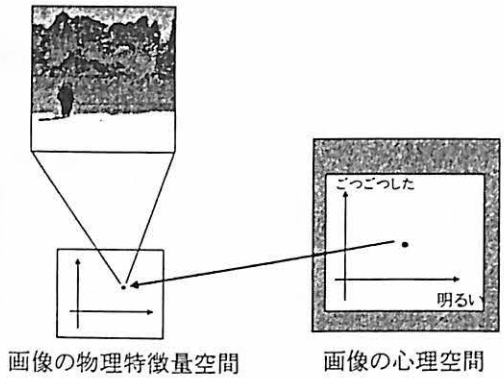

Fig. 1 検索システムの概略。明るく、ごつごつした画像を検索したい場合、そのような印象を与える物理量を持つ画像を検索する。

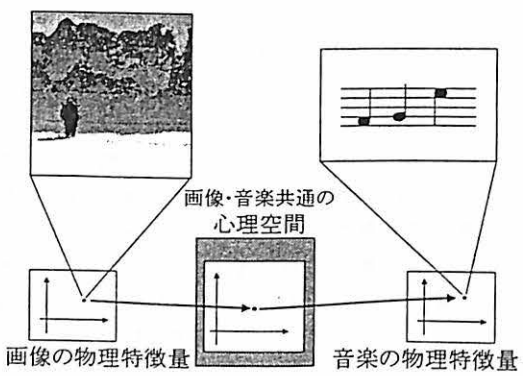

Fig. 2 メディア変換システムの概略。コンピュータに画像を入力すると、その物理特徴量が与える印象を予測し、その印象を与えるような音楽を検索する。

Table 1 心理空間構成で用いた14対の形容詞

透き通った	–	濁った	明るい	–	暗い
生き生きした	–	よどんだ	暖かい	–	冷たい
はっきりした	–	ぼやけた	単純な	–	複雑な
コミカルな	–	シリアスな	派手な	–	地味な
色っぽい	–	色気のない	軽快な	–	鈍重な
広がりのある	–	広がりのない	固い	–	柔らかい
激しい	–	おとなしい	乾いた	–	湿った

形容詞対を表1に示す。

また、心理空間の構成の際に用いられる画像は、予備実験を通して3000枚の画像から似ている画像を削除していくことにより、610枚を選んだ。

このようにして選ばれた形容詞対、画像を用いて、10名の被験者が1枚ずつ提示される画像から受ける印象を、図3のインターフェースを用いて14組の各形容詞対について7段階で評価してもらった。

14対の形容詞対に対する評価によって、各画像に対する形容詞空間の座標を調べることになるが、各形容詞対の間には相関があると思われるため、主成分分析により軸を直交化し、また14次元から低い次元の心理空間へ写像することにより検索・システム構築の両面で取り扱いや

Fig. 3 画像の心理空間構成実験のインターフェース

Table 2 本研究で用いる心理空間構成で用いる4軸。画像や、音楽などのメディア検索は今後この4軸によって行う

激しさ、軽快さ	などの 動的因子
暖かさ、色っぽさ	などの 情的因子
複雑さ	などの 複雑性因子
広がり感	などの 空間因子

すくする。この低次元化により、検索速度も速くなる。

2.2 形容詞空間から心理空間への変換

前節の形容詞空間を主成分分析により低次元化を行った。3〜7次元空間に形容詞空間を近似した結果、4次元の場合が検索、メディア追加時の評価実験の両方で操作がしやすい形容詞の組み合わせが得られたので、表2のような4つの軸によってあらわされる空間を本研究の心理空間とする。

この心理空間の構成は画像を用いて行ったが、音楽などの別のメディアを追加する場合は、表2の4軸を用いて評価し、心理空間上の座標を調べる。新しく追加されたメディアの心理空間の座標は、画像の心理空間の軸を用いて評価したため、共通の心理空間上の座標となる。つまり、4次元の心理空間上に、複数のメディアが混在しているため、距離を測ることで印象の近い異なるメディアを検索することができる。

このような方法で、メディアを追加したときにも心理空間を構成することができる。また、このような方法でなければ、異なるメディアを追加したい場合でも、因子を構成する形容詞軸が変わる可能性があり、追加されたすべてのメディアに対し心理空間をそれぞれ構成することは被験者の疲労の面などから難しいため、共有できる心理空間とすることができない。

2.3 心理空間の普遍性

本論文で用いる心理空間は、効率よく普遍的に人の印象を表すことができるため、心理空間を利用して感性に

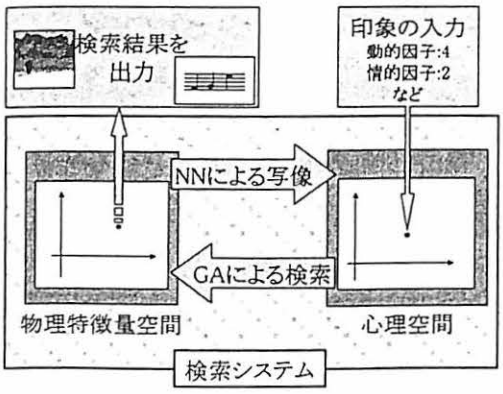

Fig. 4 データベース検索の概略。検索者が求めるメディアの印象を入力すると、GAが検索する。GAの評価値は心理空間上での距離なので、ニューラルネットによる写像が必要になる。

基づくシステムの評価に用いることができる。感性語を用いた検索システムを評価するとき、検索の性能が検索テーマによってばらつきがないかを調べなければならない。その評価をするために、いろいろな検索テーマを定め検索結果を調べることになるが、その検索テーマを定める際に心理空間を用いる。心理空間は主成分分析によって構成されるため、形容詞空間とは異なり各軸が直交している。そのためこの心理空間から検索テーマを定めることによって、評価のための検索テーマを偏りなく定めることができる[4],[5]。

3. 印象に基づくメディアデータベース検索

データベース検索に関する研究では、メディアが持つ物理量とそれが人に与える心理量との関係を用いたものがある。本論文で提案する方法は、人が受ける印象を表す心理空間を構成することによって検索をする方法である。

まず、SD法によって心理空間を構成する。多くのメディアを提示し、その印象をあらかじめ選んでおいた形容詞対で答えてもらい、主成分分析によって、形容詞空間から心理空間を構成する。その結果を主成分分析により形容詞空間の各軸を直交化することにより検索効率を高め、次元下げを行うことにより検索のしやすさを高める。このようにして心理空間を構成しこれをシステムの中心に据える。この心理空間は人がメディアに触れたとき、どのような尺度を使って印象を受けているかを表している。

次に、メディアの物理的な特徴量を計算する。例えば画像の場合であれば輝度や周波数の分布、音楽であれば音の高さの平均、分散など多くの特徴量が考えられる。

各メディアが持つ物理量と、それが与える印象との写像関係をニューラルネットなどの非線型な方法で構成することで、印象に基づく検索ができる(図4)。

3.1 画像データベース検索

画像検索システムを構成するために、2節で構成した心理空間を用い、物理量と心理量との関係を利用して図4のように検索を行う。

まず、あらかじめ物理量として、心理空間構成実験で用いた画像から6つの特徴量を計算し、この6つの物理量を入力として、4次元の心理空間の座標を出力するニューラルネットを学習により構成する。

検索をする際は、表2の心理空間の4軸に対し、探したい画像の印象を入力する。入力された印象をもとに、遺伝的アルゴリズムによって入力された印象に近い印象を与えるような物理量を持つ画像を検索する。ここでGAは、検索された画像が、検索者の意図する印象に近いかどうかをニューラルネットによって調べ、これを評価値として検索を行う。このシステムを用いることにより、「部屋に明るい絵を飾りたい」というような要求にこたえることができる。

3.2 音楽データベース検索

音楽データベース検索システムでも、画像検索と同じように心理空間を用い、物理量と心理量の写像関係から検索する。

2節の4次元の心理空間から、4軸を用いてそれぞれの軸に対し音楽を聴いた印象がどれだけ当てはまるかを5段階で評価してもらった。用いる音楽素材はMIDIファイルを8小節再生したもの250フレーズである。被験者は12名で、全員に同時に聞かせて評価してもらった。物理量として、8小節のフレーズで、2小節ごとに5つの物理量を計算する。5×4=20の物理量を入力するとそれが与える心理量を出力するようなニューラルネットを構成する。検索は印象を入力することにより3.1節と同じ方法で音楽を検索することができる。このシステムを用いることにより、曲名や作者などの音楽の知識が無い場合でも、知らない曲をデータベースから検索することができる。

4. 印象に基づくメディア変換

心理空間を仲介とすることで、メディア変換システムを構成することができる。前節の画像や音楽などのメディア検索システムは印象を入力することで検索をしているが、メディア自体を入力することで、メディア変換を行うことができる。メディア自体を入力することは、メディアの物理量を入力することであるから、入力された物理量からニューラルネットを通してその印象を予測することができる。その印象を入力として、入力されたメディアとは異なるメディアを検索することによって異なるメディアであるが同じような印象を与える素材を検索する、つまりメディア変換を行うことができる。

このシステムを用いることにより、ある画像をコンピュータに入力し、それを人が見たときに受ける印象と同じ印象を与えるような音楽を検索することができる。これを利用すると、情景にマッチする音楽の検索をしたり、音楽データベースの画像インデックス検索を行うこともできる。このようなシステムを構成するためには、音楽と画像など、複数のメディアで共有できる心理空間を構成する必要がある(図2)。

5. まとめ、今後の課題

本論文では、印象に基づく画像や音楽などのマルチメディアデータベース検索と、それを用いた印象を共有するメディア変換を行うシステムを提案した。そのシステムで用いる、画像や音楽などのメディアの物理的な特徴量が人に与える印象を普遍的に表すことのできる心理空間を印象評価実験を通して構成した。

今後は遺伝的アルゴリズムによる検索システムを構成することにより、異なるメディア間での相互検索を感性的に行うシステムを構成する。

参 考 文 献

[1] Hayashi, T. and Hagiwara, M., "An image retrieval system to estimate impression words from images using neural network," IEEE Int'l Conf. on Systems, Man and Cybernetics (SMC'97), Orlando, Florida, USA, pp.150–155 (Oct., 1997).

[2] Lee, J.-Y. and Cho, S.-B., "Interactive genetic algorithm for content-based image retrieval," Asian Fuzzy Systems Symposium (AFSS'98), Masan, Korea, pp.470–484, (June, 1998).

[3] Lee, J.-Y. and Cho, S.-B., "Interactive genetic algorithm with wavelet coefficients for emotional image retrieval," Int'l Conf. on Soft Computing and Information/Imtelligent Systems (IIZUKA'98), Iizuka, Fukuoka, Japan, pp. 829–832, (Oct., 1998).

[4] 野田 寿彦, 高木 英行, 張 明智, "画像検索のための心理因子空間構成と対話型GAベース画像検索システムの主観評価," 15th Fuzzy System Symposium, Osaka, Japan, pp. 683–684, (June, 1999).

[5] Takagi, H., Cho, S.-B., and Noda, T., "Evaluation of an iga-based image retrieval system using wavelet coefficients," IEEE Int'l Conf. on Fuzzy Systems (FUZZ-IEEE'99), Seoul, Korea, Vol.3, pp. 1775–1780, (Aug., 1999).

[6] Yoshida, K., Kato, T., and Yanaru, T., "Image retrieval system based on subjective interpretation," Int'l Conf. on Soft Computing and Information/Imtelligent Systems (IIZUKA'98), Iizuka, Fukuoka, Japan, pp. 247–250, (Oct., 1998).

[問い合わせ先]
〒815-8540 福岡市南区塩原 4-9-1
九州芸術工科大学
高木 英行 (takagi@kyushu-id.ac.jp)
Tel & Fax: 092-663-4555