

■7群 (コンピュータソフトウェア) - 6編 (情報検索とデータマイニング)

4章 マルチメディア情報検索

【本章の構成】

本章では以下について解説する.

- 4-1 画像検索
- 4-2 音声情報検索
- 4-3 動画検索
- 4-4 他情報の検索：感性検索，その他

■7群-6編-4章

4-1 画像検索

(執筆者：田邊勝義) [2010年10月 受領]

画像検索 (Image Retrieval) は、キーとなる情報 (キーワード、キー画像など) を入力して、関連する画像をデータベースから検索し、あらかじめ定められた基準でランキングした結果を検索者に提示する技術である。

インターネットの普及に伴い、多くの種類の画像データにアクセスできるようになり、画像を検索する機会が増えてきている^{1),2)}。

検索結果には、画像だけではなく、テキスト、動画、音声などのマルチメディア情報を含む場合もある。本節では画像検索について、本章ではテキスト検索以外のマルチメディア情報検索について概説している。

本節ではまず、キーワードによる画像検索について述べる。画像検索を実現する実用システムでは、テキスト検索の延長としてキーワードを用いて画像を検索するシステムが多い。この場合、画像内容を表すテキストをメタデータとして付与することが行われている。しかしながら、画像内容を的確に表現するキーワードを自動的に付与することは、現時点では難いため、あらかじめデータベースに蓄積される大量の画像に人手でメタデータとしてキーワードを付与が行われており、大変な労力が必要となっている。

また、少ないテキストだけでは、画像の内容を十分表現することは難しい。画像は、言語情報と比較して、はるかに多義性や曖昧性が高いため、キーワードを用いた検索だけでは必ずしも満足した結果が得られないことがあるからである。

そこで、画像を検索キーとして用いて蓄積された画像から類似した画像を検索する方法についても研究開発が行われている。本節ではこの画像を検索キーとした画像内容に基づく検索 (CBIR: Content-Based Image Retrieval) についても述べる。更に、音声による画像検索についても紹介する。

4-1-1 キーワードによる画像検索

画像を蓄積する際には画像の属性情報として、タイトル、カテゴリー、内容を表すキーワード、作者や日時などのメタデータやタグと呼ばれるテキストをインデックス情報として画像とともに格納する。キーワードによる画像検索では、このメタデータ (タグ) 情報とキーワード (テキスト) とのマッチングにより、キーワードに関連する画像を検索する。

現在、実用に供されているシステムでは、テキスト検索の延長として、言語情報であるキーワードを検索キーとして用いているものが大部分である^{3),4)}。最近の Google⁵⁾、Yahoo!⁶⁾ や goo⁷⁾ などの Web (WWW: World Wide Web) 上の画像検索においてもキーワードから画像を検索する方式が主流である。YouTube などの動画共有サイトや Flickr などのフォトアルバムサイトでは、登録の際にメタデータ (タグ) 情報を登録者に登録させている。

キーワードによる画像検索では、画像内容そのものではなく、このようなメタデータ情報、例えば、画像のタイトル、カテゴリー、作者や日時などの関連ワード (属性情報) を、画像に付与し、付与されたメタデータとキーワードとのマッチングにより、画像検索が行われる。

メタデータ付与者の主観に左右されることもあるが、人間の感覚にあったメタデータが付与されるため、検索結果に納得性があり、違和感は少ない。画像の内容と付与されたメタデータに相違がない場合や、メタデータが複数の意味・評価を持たない場合には、有効な方法である。メタデータ付与で労力軽減のため、自動でメタデータを付与することも検討されているが容易ではなく修正が必要なため、現状ではメタデータ付与には人手による修正作業など多くの労力が必要である。

なお、画像に付与されるメタデータとしては、上記の属性情報のほかに、「派手」、「地味」のような感性情報を用いたものもあり⁸⁾⁹⁾、これについては、後述の4-4節の感性検索で述べる。

4-1-2 キー画像による画像検索

画像を検索キーとして、画像データベースから類似した画像を検索する方法^{10)~12)}について説明する。画像検索ではキー画像と類似しているか否かを判断する技術が必要である。この場合、画像内容を意識せずに、画像全体から特徴量を習得し、そのデータを基に類似度を算出し、画像比較を行う方法と、画像の中に写っている領域を抽出、あるいは対象を認識し、その領域あるいは対象から特徴量を習得し、そのデータを基に類似度を算出して画像比較を行う方法（CBIR：Content-Based Image Retrieval）がある。前者の場合は、画像全体同士の比較を行うのに対し、後者は、画像内に特定の対象物が含まれているか否かを判定する方法である。そのためには画像中に写っている対象物が何かを認識することが望まれる。システムは、画像を蓄積する際に画像の中に写っている対象物を認識し、対象物に関する情報をインデックスとして画像とともに格納する。この場合には、画像中から対象物の対象候補領域を特定し切り出すこと（画像領域抽出）を行わなければならない。

つまり、画像同士の類似性を判断するには、画像全体、もしくは、比較対象となる類似領域を特定する必要がある。システムはあらかじめ比較対象領域を特定し切り出すために必要な情報を記述しておく。この対象領域の記述情報に基づいて、画像中から対象領域に類似しているとみなされる類似画像領域の抽出技術を確認することが課題となる。

次に、画像から抽出された類似領域同士がどの程度類似しているのかを客観的に測る類似評価尺度をあらかじめ構築しておく必要がある。そのためには、各類似領域の状態・性質を（明確に）表現でき、区別できるような特徴量の選定と検索キー画像などの基準となる画像からの類似度の距離や各画像間の類似距離を表現できる特徴空間を構築する必要がある。

この特徴量に基づき構成される特徴空間の構成方法には多くの方法があるが、代表的には、線形に分類する方法と非線形に分類する方法がある。統計学習により線形な分類を行う方法では、判別分析や重回帰分析により2つの特徴空間（多次元心理空間と物理特徴空間）との対応付けを行う方法がある。

また、非線形の例では以下の方法がある。SVM（Support Vector Machine）は、教師あり学習で用いる識別方法で、カーネルトリックにより元の入力データを高次元特徴空間へ写像することにより、元の空間では非線形な分類を行う方法である。bag of features では、画像特徴を局所的な特徴量のヒストグラムで表現する。

また、扱う画像 DB によっても、検索の仕方は異なる。すなわち、広い範囲（the broad domain）の画像 DB の中から意味的に同じ内容を持つ同じカテゴリーに分類される画像を検索する検索方法と、狭い範囲（the narrow domain）の画像 DB の中から細かい違いを捉えて、類似した画像

を検索する検索方法がある。前者については、検索対象となる画像の種類は広範囲なため、検索の主眼は、検索キー画像と同じ種類(カテゴリ)の画像を探すこととなり、画像を荒く分類・検索することが求められる。一般的な画像が検索対象であり、例えばインターネット上の広い範囲の一般的な画像データベース群を横断的に検索する場合が該当する。world-image-translation model と呼ばれる領域分割した領域と単語を統計的な確率モデルで推定し単語同士のマッチングで検索を行う方法¹³⁾や、bag-of-keypoints という局所パターンの特徴量を位置に非依存に出現頻度をヒストグラムで表して比較し、類似する画像を検索する方法¹⁴⁾がある。また後者は、カテゴリとしては同一の対象画像であり、検索キー画像と類似している形、模様、色の配置具合など、詳細な画像間の違いを区別でき、画像を細かく分類することが求められる。例えば、花などの植物データベースや医用画像データベースのような特定の利用目的を持った専門データベースを対象に検索キー画像とのより詳細な類似度合に応じた検索を行う場合が該当する。あらかじめ画像から抽出しておいたオブジェクトをアイコン化しておき、検索キー画像からオブジェクトを複数抽出し、オブジェクトにより検索を行う方法¹⁵⁾や人間の感覚を多次元尺度法により反映した心理空間と画像特徴量とを重回帰分析で対応付けることで人間の感覚を反映した類似画像検索を行う方法¹⁶⁾など、多くの研究が行われている。

また、概略図をキー画像として類似画像検索を行う方法がある。求めたい画像のスケッチを用いる方法¹⁷⁾や幾何形状や閉曲線で塗りつぶした概略図(簡易画)を用いる方法¹⁸⁾がある。検索に相応しい概略図を検索者が作成するのは大変であるが、自分が欲しいイメージを簡易に表現して、検索者の意図にあった検索を実現しようとする方法である。

4-1-3 キー音声による画像検索

新しいサービスとして、キーワードを音声で入力して画像検索を行うものがある。Google は、テキスト検索について、キーワードを音声で入力する音声検索のサービスを開始している¹⁹⁾。携帯のマイクに音声でキーワードを話すと、サーバで音声認識を行い、テキストに変換して検索を行う方式である。Google では、Web サイトの検索だけでなく、写真、動画の検索も可能であるとしている。現在実現されているサービスは、Android と iPhone 向けに限られている。

4-1-4 将来展望

これまでに実現されている画像検索サービスは、画像中に付与されたメタデータを用いて検索するテキスト検索の延長線上によるサービスであり、画像特徴を用いた類似画像検索はまだ研究の段階にある。しかしながら、インターネットが普及し、世界中にアクセスできるようになり、言語依存性の無い画像の内容を直接ハンドリングして検索を行う画像検索技術が今後ますます重要になっていくと考えられる。

また、検索対象とする画像 DB (あるいは画像群) を限定するために、テキストから抽出した拘束条件を利用して、その条件を踏まえた画像処理を行うことにより、計算コストが軽減され、効率的な処理結果を得られる可能性がある。

そして、前処理の技術も重要である。画像処理は、入力された画像の状態にその後の処理が大きく左右される。画質の悪い画像に対しては、高度のアルゴリズムを考案・適用することが必要であるが、画質条件の良い入力画像に対してはその技術は不要となる。しかしながら、必ずしも条件の良い画像だけが入手できるわけではないため、特徴抽出処理の前処理で画像の品

質を上げておくことが重要となる。したがって、キー画像による画像検索では、前処理を軽視することなく、特徴抽出処理と両輪で技術検討を進めていくことが重要である。

■参考文献

- 1) 総務省：平成 21 年版情報通信白書，ぎょうせい，2009.
- 2) 財団法人データベース振興センター：データベース白書 2008，財団法人データベース振興センター，2009.
- 3) 遠藤隆也，谷口道夫，堀口真寿，若菜 忠：“プライベートデジタルビデオテックス方式，” NTT 電気通信研究所研究実用化報告，vol.36，no.3，pp.395-402，1987.
- 4) 藤本 眞：“画像検索技術の動向，” テレビジョン学会誌，vol.47，no.1，pp.59-66，1993.
- 5) <http://www.google.co.jp/>
- 6) <http://www.yahoo.co.jp/>
- 7) <http://www.goo.ne.jp/>
- 8) 栗田多喜夫，加藤俊一，福田郁美，坂倉あゆみ：“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.
- 9) 木本晴夫：“感性語による画像検索とその精度評価，” 情報処理学会論文誌，vol.40，no.3，pp.886-898，1999.
- 10) 小早川倫広，星 守：“画像の内容に基づいた画像検索システム，” bit，vol.31，no.10，pp.23-34，1999.
- 11) 柳井啓司：“一般物体認識の現状と今後，” 情処論 CVIM 19，vol.48，no.SIG16，pp.1-24，2007.
- 12) Arnold W.M. Smeulders, Marcel Worring, Simone Santini, Amarnath Gupta, and Ramesh Jain：“Content-Based Image Retrieval at the End of the Early Years，” IEEE Transactions on Pattern Analysis and Machine Intelligence，vol.22，no.12，pp.1349-1380，2000.
- 13) K. Barnard, P. Duygulu, N.d. Freitas, D. Forsyth, D. Blei, and M. Jordan：“Matching Words and Pictures，” Journal of Machine Learning Research，vol.3，pp. 1107-1135，2003.
- 14) G. Csurka, C. Bray, C. Dance, and L. Fan：“Visual categorization with bags of keypoints，” Proc. ECCV Workshop on Statistical Learning in Computer Vision，pp.1-22，2004.
- 15) 串間和彦，赤間浩樹，紺谷精一，木本晴夫，山室雅司：“オブジェクトに基づく高速画像検索システム：ExSight，” 情処論，vol.40，no.2，pp.732-741，1999.
- 16) 田邊勝義，大谷 淳，石井健一郎：“多次元心理空間を用いる類似画像検索法，” 信学論 (D-II)，vol.J75-D-II，no.11，pp.1856-1865，1992.
- 17) 長谷川純一，岡田憲武，鳥脇純一郎：“スケッチを利用した胸部 X 線像データベースの知的検索，” 信学論 (D)，vol.J71-D，no.3，pp.549-560，1988.
- 18) 松崎公紀，櫻村雅章，小沢慎治：“概略図を入力とした特徴グラフに基づく絵画画像検索，” 信学論 (D-II)，vol.J87-D-II，no.2，pp.521-533，2004.
- 19) <http://www.google.co.jp/mobile/default/onsei.html>

■7 群-6 編-4 章

4-2 音声情報検索

(※準備中)

■7群-6編-4章

4-3 動画検索

(執筆著者：谷口行信) [2009年8月 受領]

本節では、動画検索の代表的なアプローチと要素技術について説明する。

4-3-1 メタデータ検索

ハードディスクビデオレコーダ、オンデマンド型の動画配信サービスでは、テレビ番組表に相当する番組メタデータ（タイトル、制作者、俳優名、ジャンルなど）を手がかりにユーザが番組選択を行うのが一般的である。いわゆる「メタデータ」検索であり、この番組メタデータは放送局などの番組提供側で制作され、放送波または通信回線を通じて提供される。

用途に応じて様々な「メタデータ」が標準化されている¹⁾。例を挙げると、放送系では、放送と通信が連携した蓄積型映像サービスに向けた TV-Anytime Forum の規格、収録素材や番組の交換のための標準映像フォーマット MXF (Material Exchange Format) などがある。放送系に限定せずマルチメディアコンテンツ一般の内容記述のための国際規格 (ISO/IEC) MPEG-7 は、テキストベースの高レベル記述だけでなく特徴レベルの記述（画像であれば、色・テクスチャ情報）を可能とする枠組みを提供している。インターネット上のコンテンツフィードで使われる RSS もメタデータの一つといえる。

しかし、「メタデータ」だけですべての検索要求に対応できるわけではない。メタデータ検索の問題点として (1) メタデータ付与のコスト、(2) 動画の多義性、(3) 検索要求の曖昧性、が挙げられる。

メタデータ付与コストは大きな問題であり、放送番組のように再利用価値の高い動画に対しては、コストをかけて人手によるメタデータ付与が行われている。しかし、それでも付与対象のメタデータ項目を利用頻度が高いものに絞るなど、作業量を減らして対応しているのが現状である。メタデータ付与の経済化を目的として、音声認識やテロップ認識などのメディア解析技術を応用するメタデータ生成システムが開発されている²⁾。

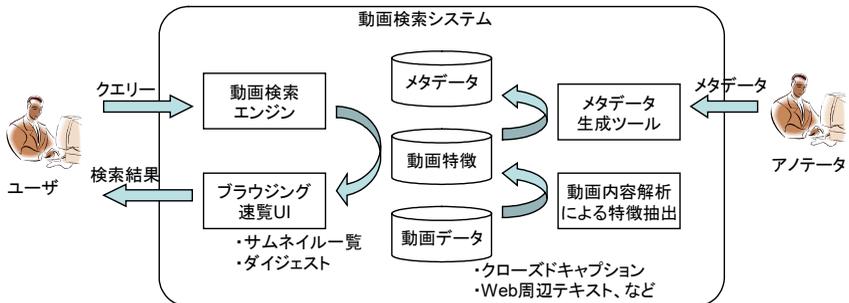


図 3・1 動画検索システムの構成

上述したメタデータ検索を含む「動画検索システム」の全体像を図 3・1 に示す。ユーザが入力したクエリーに対して、検索結果を返すだけでなく、検索結果をユーザがブラウジング・

速覧しながらクエリーを改良するというプロセスを繰り返すインタラクティブなシステムと捉えることができる。また、メタデータ付与コストと動画の多義性の問題を解決するために、画像・音声・テキスト解析技術を応用した動画特徴の自動抽出技術についても研究が盛んに行われている。

4-3-2 動画内容解析による特徴抽出

動画内容解析に基づく特徴抽出については 1990 年代から盛んに研究が行われ、主に低レベル特徴—ショット境界検出、カメラワーク検出、顔検出、音響分類（音声、音楽、拍手など）、テロップ文字認識、音声認識など—を自動抽出する技術が開発された。これらは、主に視覚的ブラウジングの手がかりとして用いられてきた。

インターネット上の動画検索では、動画からではなく、周辺テキストや投稿者が付与したタグ、コメントを利用して検索インデックスを構築している。

動画からの高レベル特徴（意味特徴、コンセプトとも呼ばれる）の自動抽出は、ハードルが高い課題ではあるが、近年、国際的な競争型ワークショップ TRECVID（TREC Video Retrieval Evaluation）³⁾の影響もあり研究が進展している。正解データ付きの大規模コーパスを整備したうえで、学習データから抽出された複数種の低レベル特徴量に対して、統計的手法や機械学習を駆使して分類を行う手法が盛んに研究されている^{4,5)}。2008 年の TRECVID の高レベル特徴抽出タスクでは 20 種類の特徴（classrooms, bridge, emergency vehicle, dog, kitchen など）を認識する問題が課せられた。TRECVID では参加者に対して、学習用データとして、ショット分割された 200 本以上もの実放送番組の動画ファイル（3 万以上のショット）が、ショットの代表フレーム、音声認識結果データ、正解データとともに提供されている点も特筆すべき点である。

4-3-3 動画速覧インタフェース

高レベル特徴が仮にすべて正確に抽出できたととしても、それだけでユーザにとって使いやすい動画検索が実現するわけではない。ユーザが、明確な検索要求を持っていない場合や、検索クエリーを明に表現できない場合も多いからである。そのため、検索結果から、ユーザにとって見る価値があるものを取捨選択する「動画速覧インタフェース」が必要である。

「動画速覧インタフェース」は、動画を空間的に展開する「静的速覧」と、時間軸方向に展開することにより動きや臨場感・雰囲気表現する「動的速覧」に大別される⁶⁾。

「静的速覧」の代表的な方法は、代表フレームを一覧表示したサムネイル一覧である。Informedia プロジェクト⁷⁾では、主にニュースを対象に、時間軸プロット、塗り分け地図、動的問合せプレビューヒストグラム、固有表現関連図などを組み合わせて視覚的検索の効率化を図っている。

「動的速覧」の代表的な手法は、動画の内容を解析し何らかの基準で重要なシーン（ハイライト）を検出し、重要シーンをつなぎ合わせたダイジェストを生成するものである。ハイライト検出の手がかりとして、動画中のイベント（ショット切替えの頻度、顔の出現など）に着目するもの、音声の韻律特徴に着目して音声強調や笑いを含むシーンを抽出するものなどがある⁶⁾。また、2007 年から TRECVID では、BBC の番組素材を対象としたラッシュ要約（Rush Summarization）のタスクが追加され、ハイライト検出だけでなく、重複シーン削除、不要シーン除去などのアルゴリズムを駆使した総合的な要約手法が研究されている。

4-3-4 性能評価 (TRECVID)

TRECVID は、2001 年に始まった動画検索に関する国際的な競争型ワークショップであり、大規模な評価用データと統一的な性能評価手順を提供している。2008 年のタスクは、高レベル特徴抽出 (high-level feature extraction)、検索 (search)、コピー検出 (content-based copy detection)、ラッシュ要約 (rush summarization)、監視映像イベント検出 (surveillance event detection) であった。詳細については参考文献 3) を参照されたい。検索タスクは、インタラクティブな動画検索システムの総合的な評価を目指したものである。

■参考文献

- 1) 石川清彦(編)：“最新メタデータ事情～規格・制作・利用サービス,” 映情学誌, vol.61, no.2, pp.127-161, 2007.
- 2) 桑野秀豪, 他：“ライブ番組向けダイジェスト視聴サービスのためのリアルタイムメタデータ生成技術,” NTT 技術ジャーナル, pp.14-17, 2005.
- 3) TREC Video Retrieval Evaluation, <http://www-nlpir.nist.gov/projects/trecvid/>
- 4) 佐藤真一：“映像内容解析における TRECVID の取り組み,” 信学誌, vol.91, no.1, pp.55-59, 2008.
- 5) 柳井啓司：“一般画像認識の現状と今後,” 情処学論 コンピュータビジョンとイメージメディア, vol.48, no.SIG 16 (CVIM 19), pp.1-24, 2007.
- 6) 日高浩太, 佐藤 隆：“映像の速覧技術,” 映情学誌, vol.63, no.1, pp.36-41, 2009.
- 7) M.G. Christel：“Supporting Video Library Exploratory Search: When Storyboards are not Enough,” Proc. of the 2008 Intl Conf. on Image and Video Retrieval (CIVR'08), pp.447-456, 2008.

■7群-6編-4章

4-4 他情報の検索：感性検索，その他

(執筆著者：木本晴夫) [2009年1月受領]

4-4-1 感性検索とは

感性検索とは、感性に基づく検索である。一般に検索においては、検索要求と検索対象があるが、感性検索の説明をするためには、まず、感性とは何かを議論し、どのような感性を扱うのかを明らかにした後、検索要求や検索対象において、それらの感性をどのように表現するかを述べる必要がある。この表現のことを感性表現と呼ぶ。検索要求と検索対象を対応付けるために検索照合を行う。また、検索過程においてインタラクティブに検索を進めるやり方や、学習をするやり方もある。また、感性は個人ごとに異なるので、個人に適合した感性検索も重要であり、これを個人適合と呼ぶ。以上について、以下の各項で説明する。

4-4-2 感性検索の現状

感性検索の実用システムは、以下に述べるように種々あるが、研究も活発に行われている。研究では、より人間の感性に合った検索を実現する試みが行われている。本節では、実用システムの解説も行うが、活発に行われている研究事例について、4-4-3節以降で解説する。

実用化されているシステムは、画像検索、音楽検索を含めて、Google 画像、YAHOO! JAPAN 画像検索、Baidu 百度画像検索、Live Search、goo 画像・動画・音楽検索、So-net 画像検索、Flickr などがある。実用システムの感性表現方法は、検索要求として感性語（印象語とも呼ばれる）を使っている。YAHOO! JAPAN 画像検索では検索要求として、「うきうき」を使った場合の検索結果は、画像タイトルが「デニム素材のうきうきエプロン・ベストタイプ」の画像である。また、検索要求として、画像を使って類似画像検索をするものもあるが、それは、goo 画像・動画・音楽検索だけである。

4-4-3 感性と感性表現

現状の実用システムや研究では、「賑やか」－「寂しい」、「派手」－「地味」などのような、意味的に相反する感性語の対によって表されるものを感性と呼んでいる。しかし、「感性とは何か」についての議論が十分でなく、今後は感性についての議論を深める必要がある。感性そのものについては、大きく分けて一般人の感性、芸術家の感性、性別の感性、年齢別の感性、国・地域・文化・歴史・時代別の感性などがある。現在、感性検索で扱われている感性は上記のうち、通常人レベルの感性である。

感性表現方法は、言語による表現や、絵による表現、詩や歌による表現、踊りによる表現、そして、物の色、形、素材による表現などがある。現在、感性検索での感性の表現は言語によるものが大部分である。絵（オブジェクト表現を含む）で表現した感性を扱う研究もある¹⁾。絵による表現では、線、繰り返し、配置・構成、向き、模様などが表現できて、これらは画像解析が可能な特徴であり、感性表現の幅を広げるもので、今後の研究開発が望まれる。

4-4-4 検索システム内部での感性表現

(1) 感性検索システムの構成

感性検索システムは、大きく分けて、検索要求、検索対象、及び検索照合をするための検索空間から構成される。以下におのおのについて述べる。

(2) 検索要求とその表現

検索要求の表現で、一般的に使われている表現方法は次のようである。

- ① 感性語による表現
- ② 感性語集合の各感性語について SD 法によって評価値を与えたものによる表現
- ③ 因子分析の共通因子による表現
- ④ 画像による表現
- ⑤ その他

①の感性語による表現では、検索のために一つまたは複数個の感性語を用いる。

②の感性語集合による表現では、感性語集合とは、「賑やか」－「寂しい」、「派手」－「地味」のような意味的に相反する感性語対の集まりのことで、どの程度「賑やか」か「寂しい」か、などを各感性語対において、5段階などで評価し、それを各感性語対の評価値としたものである。このような感性語集合による表現を感性工学では SD 法と呼んでいる。感性語集合は検索のための印象表現空間を構成し、各感性語はその空間の要素ベクトルで、検索要求は印象表現空間上での一つの点となる。

③の因子分析の共通因子による表現では、因子分析の共通因子によって構成される空間において検索要求を表現する。因子分析を使って、SD 法で用いた感性語対の評価を変量として、それらの変量間の共通因子を求める。共通因子は、多数の評価項目（ここでは、感性語対）の評価値をもとにして抽出された、評価者の回答に潜む少数の共通的な説明要因である。共通因子によって検索要求や検索対象を表現すると印象の表現が簡潔に行える。因子分析を用いた研究の例として宝珍ら²⁾の研究がある。

(3) 検索対象とその表現

検索対象は、一般的には画像、音楽などであるが、検索対象は、物理特徴量によって表現する方法と、印象に基づく感性語の集合によって表現する方法とがある。後者の方法は、検索対象を、感性語を使って SD 法によって印象評価して、その結果で表現する。

画像の物理特徴量としては、面積、エントロピー、周囲長、回転モーメント、角の数、画像中のエレメント数、色パラメータ¹⁾、偏平度、輪郭線の平均直線長の割合、複雑度、重心の高さ、凸閉包の円形度³⁾、カラー画像の 0 次自己相関特徴、1 次自己相関特徴⁴⁾、HSV ヒストグラム、フーリエパワースペクトル、画像コントラスト⁵⁾、Lab 局所コントラスト、3 次元オブジェクトの頂点の配置状況と頂点の明度の分布状況⁶⁾、などがある。

音楽の物理特徴量としては、音高、音の強さ、音の長さ、無音時間、音数、平均音高からの差、音高の最大と最小の差、音の調和度、リズムの平均値、調性、拍子、テンポの初期値設定、低音などの音数の割合、低音などの時間の割合などがある⁷⁾。

(4) 検索照合と検索空間

検索要求は一般的に、複数の感性語を使って検索要求感性空間を構成し、検索対象は複数の物理特徴量を使って検索対象感性空間を構成する。検索要求感性空間と検索対象感性空間とを対応付けして、検索照合を行うことが検索において重要である。この検索照合を行う空間を検

索空間と呼ぶ。

(5) 検索空間の構成法と対応付け

検索空間は検索照合を行う空間である。ここでは、検索空間の構成法、及び検索空間と検索要求感性空間及び検索対象感性空間との対応付けをする方法について述べる。この構成法と対応付けには以下のような方法があり、おのおの研究報告がされている。

- (1) 因子分析を用いる方法は、因子分析の結果から得られる共通因子によって検索空間としての因子空間を構成するものである^{2),3),7),8)}。検索要求感性空間及び検索対象感性空間から因子空間への対応付けは、ニューラルネットワーク^{7)~9)}や重回帰分析³⁾によってなされる。また、検索要求感性空間を因子分析の共通因子によって構成する場合は、そのまま対応付けがなされている。
- (2) 正準相関分析を用いる方法は、正準相関分析における正準変量によって検索空間を構成する方法である^{4),6)}。検索要求感性空間などとの対応付けは、主成分分析⁶⁾及び正準相関分析^{4),6)}によってなされる。
- (3) 対義語対を用いる方法は、対義語対によって検索空間を構成する方法である。検索空間と検索要求感性空間などとの対応付けは、検索要求に対応する教師用画像群の物理特徴量に基づいて、検索要求に対応する対義語空間上の対義語ベクトルを計算することなど、及び検索対象の画像の評価値計算などによってなされる⁵⁾。
- (4) 感性ベクトルを用いる方法は、基準となる評価者の感性ベクトルを用いて検索空間としての感性空間を構成する方法⁹⁾である。検索要求感性空間などとの対応付けは、ニューラルネットワークやSD法によってなされる。

(6) 検索空間への対応付けのための学習

検索空間への対応付けをするために、学習用のサンプルを用いて、対応付け情報を抽出している。この対応付け情報を用いて、実際の検索時にユーザから出された検索要求や検索対象を検索空間へ対応付けして、検索空間で照合を行う。

(7) 検索照合

一般的に、検索要求と検索対象は検索空間上の点として表現されて、点と点のユークリッド距離を計算して、距離の小さいものから順に検索要求に適合しているとされる。

4-4-5 検索過程（検索のためのインタラクション）と学習

適合フィードバック（Relevance Feedback）という方法がある。この方法は、検索の過程で、得られた結果についての適合性判定情報をユーザがシステムに与えて、それに基づいて、システムがその内部での計算方法に変更を加えて、ユーザの意図に適合した結果を出力する。この方法によって、検索過程を繰り返しながら、システムが学習を行って、検索精度を向上させることが可能になる。

4-4-6 個人適合

感性検索結果を各個人に合ったものにするための研究も行われている。検索過程において学習を行えば、その結果、検索者個人に適合した検索結果が得られる。このような方法とは別の方法として、因子分析を用いる感性検索において、すべての検索対象に対してすべての感性語対の評価値を表す個人ごとの評点行列から、各個人の個性を表現する個性行列を求めて、検索

に利用する方法が提案されている¹⁰⁾。

4-4-7 将来展望

これまでに扱われてきた感性は、感性語、画像、音楽などで表現するものであった。それらは、いわゆる感性であるが、必ずしも幅広い感性を十分に表現できるものではない。今後は、より日常的で生活に密着したモノやコトの感性を表現することが重要であり、これによって感性検索が日常的に普及していく可能性が開ける。例えば、「ハワイの空の色のように明るいTシャツ」、「早春に着るセーター」、「フレンチレストランに行くときに履く高価そうに見える靴」などのような感性表現をして、欲しい商品を探し出すビジネスモデルが考えられる。感性検索は、商品知識が少ない人や自分の要求を言葉では的確に表現できない場合でも使いやすいという特徴がある。このような感性検索の特徴を利用する検索システムは、株式会社アルベルトなどが開発を進め、導入されている。今後は、感性を手軽に、かつ、より適切に表現する方法の研究が求められる。また、人を取り巻くモノ、コトについて、その感性表現方法を探索する必要がある。このために、感性コーパスのようなものを構築して研究すると有効であろう。感性コーパスとは、言葉、画像、音楽だけでなく、「もの」や「価格」などの日常のモノやコトなどを幅広く集めて、感性を表現した大規模な事例集、サンプル集である。非専門家が、より幅広く、日常的なものを適切に感性表現できることは今後の課題である。

■参考文献

- 1) 柏崎尚也, 岡本歌織: “感性パラメーター法による楽曲のジャンル分け—感性情報の数値的一般化に関する研究(1),” 日本感性工学会研究論文集, vol.7, no.2, pp.243-249, 2007 Dec.
- 2) 宝珍輝尚, 都司達夫: “印象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.SIG2 (TOD13), pp.69-79, 2002 Mar.
- 3) 原田将治, 伊東幸宏, 中谷広正: “感性語句を含む自然言語文による画像検索のための形状特徴空間の構築,” 情報処理学会論文誌, vol.40, no.5, pp.2356-2366, 1999 May.
- 4) 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,” 情報処理学会論文誌, vol.33, no.11, pp.1373-1838, 1992 Nov.
- 5) 椋木雅之, 田中大典, 池田克夫: “対義語対からなる特徴空間を用いた感性語による画像検索システム,” 情報処理学会論文誌, vol.42, no.7, pp.1914-1921, 2001 Jury.
- 6) 山際央明, 柴田滝也: “データベースを用いたインテリア・コーディネート・エージェントに関する研究—リビングに調和する3次元家具データ検索システム—,” 日本感性工学会研究論文集, vol.7, no.4, pp.685-692, 2008 June.
- 7) 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.3201-3212, 2001 Dec.
- 8) 坂本 崇, 梶川嘉延, 野村康雄: “音楽感性空間における非線形判別分析を用いた曲印象別グループの分割,” 情報処理学会論文誌, vol.40, no.4, pp.1901-1909, 1999 Apr.
- 9) 杉原太郎, 森本一成, 黒川隆夫: “m-RIK: 個人の感性特性に対応可能な音楽検索システム,” 情報処理学会論文誌, vol.46, no.7, pp.1560-1569, 2005 July.
- 10) 宝珍輝尚, 熊切健夫, 井田俊博, 都司達夫, 樋口 健: “感性マルチメディア検索における擬逆行列を用いた個人適応法,” 感性工学研究論文集, vol.4, no.1, pp.27-30, 2004 Mar.