

科学研究費助成事業 研究成果報告書

平成 26 年 5 月 28 日現在

機関番号：16101

研究種目：基盤研究(B)

研究期間：2009～2013

課題番号：21300036

研究課題名(和文) 脳波に基づく生体情報を利用した意味的・感性的マルチメディア検索に関する研究

研究課題名(英文) Semantic and Affective Multimedia Retrieval using EEG-based Biological Information

研究代表者

北 研二(KITA, Kenji)

徳島大学・ソシオテクノサイエンス研究部・教授

研究者番号：10243734

交付決定額(研究期間全体)：(直接経費) 12,300,000円、(間接経費) 3,690,000円

研究成果の概要(和文)：本研究では、画像、映像、音楽等のマルチメディア・コンテンツ検索において、脳波に基づく生体情報から得られる感性的な特徴量を用いてユーザの感性や嗜好に基づく検索を実現するための研究を行った。ユーザが未知のコンテンツを見るたびに脳波の計測を行うことは現実的に不可能であるため、正準相関分析に基づく線形回帰法を用いて、コンテンツから計算的に得られる特徴量から感性的な特徴量を推定する方法を提案した。また、応用として、類似画像検索により得られた検索結果に対して、ユーザの好みに応じ、画像の好悪分類を行う感性フィルタリング手法を提案し、実験により有効性を示した。

研究成果の概要(英文)：In this study, we exploited methods for affective information retrieval for multimedia contents such as images, movies, and music by estimating user preferences/tastes from biological information such as EEG. It is, however, difficult to measure the EEG when a subject visualizes an unknown contents. We proposed a solution where a linear regression method based on canonical correlation is used to estimate the cross-correlation affective features from contents features. As an application, we also proposed a sensitivity filtering method for eliminating unpleasant images from the results obtained with similar image retrieval methods, and confirmed that the proposed method was efficient as an affective retrieval system.

研究分野：総合領域

科研費の分科・細目：情報学 ・ メディア情報学・データベース

キーワード：マルチメディア 情報検索 脳波情報

1. 研究開始当初の背景

計算機の高性能化や記憶容量の大容量化にともない、情報のマルチメディア化が急速に進行しており、このような背景のもと、マルチメディア・コンテンツに対する知的な情報検索技術の必要性がますます大きくなってきている。マルチメディア・コンテンツに対する検索は、大別すると、コンテンツに付与されたキーワード等のメタ情報に基づき検索を行うものと、コンテンツそのものから計算的に得られる特徴量に基づき類似検索を行う内容型検索の2種類がある。

我々は、平成17年度～平成20年度に、基盤研究(B)「意味空間への情報マッピングに基づくクロスメディア検索に関する研究」において、質的に異なるさまざまなマルチメディア・コンテンツを意味空間と呼ばれる仮想的な空間にマッピングすることにより、コンテンツの意味に基づく検索を可能とする手法の研究開発を行った。この手法は、コンテンツに付与されたキーワード等のメタ情報と計算的に得られる特徴量との間の確率的な相関関係を機械学習しておくことにより、メタ情報の付与されていない未知コンテンツをキーワードと同一の意味空間上にマッピングすることで、マルチメディア・コンテンツの意味的な検索を行うというものである。この研究は、ある程度の成果を得た。たとえば、画像コンテンツ検索の場合、「花火」や「夕焼け」のような、キーワードと画像特徴量との間に明らかな相関関係が存在するものに対しては、キーワードに基づき対象の画像コンテンツを高い精度で検索することが可能である。しかし、我々が研究開発した手法は、人間の持つ認知・知覚情報処理機能には、到底、及ばない。人間は、抽象化されたり、デフォルメされた画像を見ても、その中に描かれている対象物が何であるかを瞬時に判断することができるが、このような機能の実現には、メタ情報と特徴量との相関関係を学習するというだけでは不可能である。さらに、マルチメディア・コンテンツ検索では、人間の持つ感性を何らかの形でモデル化する必要がある。たとえば、「楽しい絵」、「綺麗な絵」、「ある映像に適したバックグラウンド・ミュージックを探したい」などの知的な検索を行うためには、感性的な観点からの検索が必要不可欠である。

我々は、マルチメディア・コンテンツを人間に提示したときに誘発される生体的・生理的情報を利用することにより、知的で高精度な意味的・感性的検索を実現することができると考えた。特に脳波は、計測時に強い姿勢拘束を課さない等、ユーザ負担が少ないために、マルチメディア検索にとり有望な情報源となりうるのではないかと考え、本研究に取り組んだ。

2. 研究の目的

(1) 脳波信号情報をマルチメディア・コンテ

ンツのインデックスにどのように組み入れれば高精度の感性検索を実現できるかという点を明らかにする。脳波情報を単独に用いただけでは、高い検索精度を達成することは困難であると想定されるので、従来型の特徴量との組み合わせ方や、特徴量間の最適な距離尺度等について研究する。

(2) マルチメディア・コンテンツに対しては、さまざまな観点からの特徴量が考えられる。たとえば、脳波に基づく感性的特徴量、メタ情報等に基づく意味的特徴量、コンテンツから計算的に得られる特徴量などである。検索対象となるすべてのコンテンツに対し、感性的特徴量や意味的特徴量を付加することは現実的に不可能であるので、特徴量相互の相関等に基づき、お互いの特徴量を補完できるような仕組みを備えた検索モデルを設計する。

(3) テキスト(文書)を対象とした検索の場合とは異なり、画像、映像、音楽のようなマルチメディア・コンテンツの場合には、コンテンツの捉え方がユーザごとに異なる場合が多々ある。このため、マルチメディア・コンテンツ検索では、ユーザの嗜好・趣味や解釈の違いにより、期待される検索結果もユーザごとに変わると考えられる。ユーザの嗜好・趣味・解釈の特性をモデル化することにより、ユーザ個人に適用可能な柔軟な情報検索技術の研究開発を進める。

3. 研究の方法

(1) まず、脳波の測定と脳波からの特徴量の推定について説明する。

本研究では、マルチメディア・コンテンツを被験者に提示したときの脳波を測定するために、国際10-20法で定められている電極位置に対する脳波測定を行った。この中で右耳にあたる A_2 電極を基準電極、頭頂にあたる Cz 電極をグラウンド電極とし、残りの14個の電極から脳波を得る。14個の電極のうち、10個の電極から得られた値を用いて特徴量を求めた。脳波は、デルタ波(～4ヘルツ)、シータ波(4～8ヘルツ)、アルファ波(8～13ヘルツ)、ベータ波(13～30ヘルツ)に分かれ、デルタ波は眼球の動きによる影響を受けるので除外し、本研究では、5～20ヘルツの脳波を分析の対象にした。45個($_{10}C_2$)の相関係数の値をこれら3つの周波数帯域の脳波について計算し、合計135次元の特徴量を一定時間ごとに求めたものを感性的な特徴量とした。相関係数に基づく特徴量であるため、以下では、相関特徴量と呼ぶ。

(2) マルチメディア・コンテンツ検索では、一般的に、コンテンツから計算的に得られる特徴量を多次元ベクトルなどで表現する。検索対象となるすべてのコンテンツに対し、脳波測定を行うことは実際上不可能であるの

で、コンテンツから得られる特徴量から（脳波から得られるであろう）相関特徴量を推定する。このようにして推定された特徴量を以下では、感性特徴量と呼ぶ。

本研究では、正準相関分析と線形回帰法を用いて画像特徴量から感性特徴量を推定する。正準相関分析とは、2つの変数群に対する測定値が与えられている場合に、変数群ごとに合成指標を作成し、その合成指標の相関（正準相関係数）が最も大きくなるような重みを求めることで、変数群間の相関構造を探るための手法である。コンテンツから得られる特徴量から、正準相関係数を用いて線形回帰を行うことにより、感性特徴量を推定する。

(3) 本研究では、マルチメディア・コンテンツとして、画像、映像、音楽などのデータを対象に研究を行ったが、画像コンテンツの検索は、色やテクスチャなどの画像特徴量を用いる単純な類似画像検索との比較が容易である。

本研究における画像特徴量としては、Color Correlogram（以下、CC と略記）および、Local Binary Pattern（以下、LBP と略記）を用いた。CC は主に色情報に着目した特徴量であり、LBP はテクスチャに着目した特徴量である。CC は、Jing Huang らによって提案された近隣の色の共起頻度に基づく特徴量である。本研究では、共起頻度を抽出するピクセル間の距離 d を7として、256次元のCCの抽出を行った。また、LBP は、一般に、画像のテクスチャ特徴として用いられる。濃度変化に対して不変であり、計算コストが少ない特徴量である。画像中の注目ピクセルを中心として画像を n 方向に分割し、各方向に含まれるピクセルとの色相の差を求めることにより特徴を得る。本研究では、画像全体において求めた各ピクセルのLBP値の頻度であるLBPヒストグラムを使用した。分割数 n を8とした場合のLBPヒストグラムの横軸は、注目ピクセルの周辺に現れるエッジ数0~8と、エッジがまとまっていない場合を加えた10個の値を持つ。同様に分割数 n を16とした場合、横軸が18個のヒストグラムが生成できる。今回の実験では分割数 n が8と16の場合に加えて、注目ピクセルからの距離が異なる場合のLBP特徴量も利用した。最終的に生成された3つのLBPヒストグラムを組み合わせることで、46次元のLBP特徴量を抽出した。

4. 研究成果

(1) 感性フィルタリングの提案

マルチメディア・コンテンツ検索において、ユーザの感性的な側面からコンテンツの好き嫌い（好悪）を分類する感性フィルタリングの技術を提案した。

一般に、類似画像検索では、ユーザの入力した画像と、色・構図・模様などが類似していると思われる画像をデータベースの中か

ら検索する。検索結果の中には、ユーザの好きな画像もあれば、ユーザの嫌いな画像もあるであろう。感性フィルタリングとは、ユーザの好悪をモデル化することにより、ユーザが嫌いと思われる画像を検索結果の中から自動的に取り除く技術である。ここで、我々は、画像から想起されるユーザの感性は、ある程度、画像の色・構図・模様などの画像的な特徴に左右されており、画像特徴量から推定された感性特徴量を用いてフィルタリングを行う。具体的には機械学習手法として最大エントロピー法やサポートベクターマシンなどを用いて好悪分類モデルを作成する。

本研究で提案する感性フィルタリングは、画像から抽出した画像特徴量をもとに推定した感性特徴量を入力とし、好悪分類モデルにより、検索結果の各画像に好悪のラベル付けを行うものである。画像特徴量から相関特徴量を推定する理由は、画像自体に好悪のラベルが付いている場合、画像特徴量から好悪を分類することはできるが、もし色や構図が似ているが、好悪が異なるような画像の場合、画像特徴量を比較するとその差異は小さいと予測される。しかし、実際に画像を見ている人はその小さな差によって画像の好悪を分類する。そこで画像特徴量から感性特徴量を推定することで、画像特徴量では小さかった差が感性特徴量では大きな差になるのではないかと考えた。こうして画像特徴量を直接利用するのではなく、推定した感性特徴量を利用することで、より性能の良い好悪フィルタリングを行なうことができる。感性フィルタリングを実現するための感性特徴量の推定および、感性特徴量からの好悪分類モデル作成の流れを図1に示す。

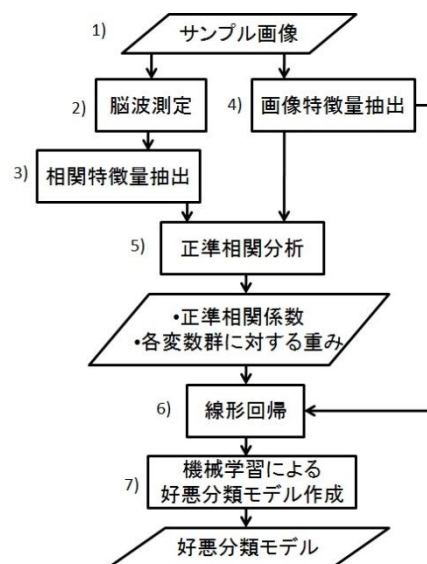


図1. 好悪分類モデルの作成

(2) 感性特徴量に基づく好悪分類

画像コンテンツを対象に、正準相関分析に基づき画像特徴量から推定した感性特徴量と、実際に脳波から得られた相関特徴量との

コサイン類似度を計算してみたところ、CC から推定した感性特徴量との平均類似度が -0.0054、LBP から推定した感性特徴量との平均類似度が、0.034 となった。この結果から、推定精度として高いとは言い難いため、脳波から直接抽出した相関特徴量を用いて好悪分類モデルの作成を行うと、画像から推定した感性特徴量をうまく分類できない可能性がある。そのため、推定した感性特徴量を学習用データとして作成した好悪分類モデルにより分類実験を行うことで、推定された感性特徴が画像の好悪分類に有効かどうかを確認する実験を行った。

本研究では、機械学習を行う上で Classias-1.1 というソフトウェアを使用した。このソフトウェアは、最大エントロピー法(Maximum Entropy Method ; MEM)やサポートベクターマシン (Support Vector Machine ; SVM)などの学習モデルをサポートしている。それぞれの学習モデルを用いて構築した分類器が、感性フィルタリングのための好悪分類に適したものであるかということの確認、および、各特徴量を用いた場合の比較の性能評価を行うための予備実験として、画像の好悪分類の精度評価を行う。実験は学習用画像として 200 枚 (好き:100、嫌い:100) を使い、分割数 200 で交差検定を行った。学習用画像についてはあらかじめ、被験者が画像を見ているときの脳波を測定しており、被験者に対し好悪分類アンケートを行った。なお、好悪のどちらでもないものは実験対象から除外する。

分類器の性能比較を行うための実験として、各画像特徴量(CC および LBP)から作成した好悪分類モデルを用いてそれぞれ MEM、SVM に基づく分類器により分類した実験結果と、あらかじめ収集したアンケートの結果を比較し、2 つの結果が同一になっている場合を正解、それ以外の場合はすべて不正解とし、正解率の計算を行った。図 2 にさまざまな特徴量に対する実験結果を示す。使用した特徴量は次に示す 5 種類である。 CC 特徴量、

LBP 特徴量、 感性特徴量、 CC から推定した感性特徴量、 LBP から推定した感性特徴量 (図 2 中の から はそれぞれこれらの特徴量と対応する)。

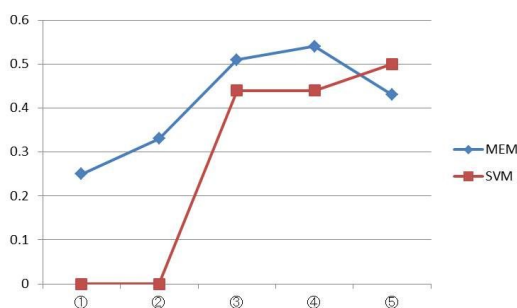


図 2 . 好悪分類正解率

図 2 を見ると、MEM と SVM を用いた場合、 の結果と比べ の方が正解率が上昇している。また と の結果を見ても、 より正解率は上昇している。このことから MEM と SVM を用いて好悪分類を行う場合、画像特徴量をそのまま使うより、推定した感性特徴量を使う方が分類の精度は向上することが分かる。MEM を用いた場合、 の結果が SVM と比べ の結果より悪くなっているが、これは画像特徴量から感性特徴量を推定する時に 100%の精度で感性特徴量を推定できていないことが原因であると考えられ、また より の正解率は向上しているため、フィルタリングに用いることに問題はないと判断する。

次に LBP から推定した感性特徴量を使用した場合の各分類器での詳細な分類結果の再現率と適合率を表 1 に示す。また好への分類を正例、悪への分類を負例とする。

表 1 . 分類結果の適合率と再現率

分類結果	正例		負例	
	再現率	適合率	再現率	適合率
MEM	25/100	25/63	62/100	62/137
SVM	99/100	99/198	1/100	1/2

表 1 を見ると、SVM を使用した場合、正例の再現率は非常に高いが、適合率を見ると正例に偏った分類を行っている。これより、SVM はフィルタリングに用いる分類器としては適さないと判断できる。

以上の結果から、画像の好悪分類に感性特徴量を用いることは有効であり、各画像特徴量をそのまま用いるより、推定した感性特徴量を用いて好悪分類が行う方が分類精度は向上することが分かる。また、本研究では実験の際、学習モデルとして MEM、学習アルゴリズムに L-BFGS 法を利用する。

(3) 感性フィルタリングの評価

画像特徴量として CC、LBP のそれぞれを用いた場合において感性フィルタリングの評価実験を行った。具体的には、入力画像である好きな画像に対し、各画像特徴量を用いた類似画像検索を行い、その検索結果に対して、感性フィルタリングを適用する。画像間の類似度には、各特徴量間のユークリッド距離の逆数を用いた。検索結果上位から嫌いな画像の削除に成功した数、類似度上位 10 位までにおける適合率などから感性フィルタリングの性能評価を行う。なお、検索の際、入力画像は評価用画像データ (検索対象) から除外することとする。また、使用した学習用画像データは 200 枚、評価用画像データ (検索対象) は 200 枚 (好:100、嫌:100)、評価用入力画像: 100 枚 (好きな画像のみ) とした。なお、脳波測定の被験者は 1 名である。

評価には、各検索結果の類似度上位 10 件までにおける平均適合率を用いた。また、好きな画像が削除されなかった場合および、嫌いな画像が削除できた場合を、フィルタリングが成功しているとみなす。好きな画像が削除されてしまった場合（過削除）および、嫌いな画像が削除されなかった場合（未削除）をフィルタリングが失敗しているとみなす。感性フィルタリング適用前と後の、CC と LBP それぞれを用いた検索結果の平均適合率を表 2 に示す。CC を用いた場合、適合率は減少してしまっただが LBP を用いた場合、感性フィルタリングを行うことによって平均適合率が 12.4%向上している。

表 2 . フィルタリング前後の平均適合率

	CC	LBP
フィルタリング前	62.7%	52.8%
フィルタリング後	61.0%	65.2%

また、二種類の画像特徴量を組み合わせることで以下の条件でも実験を行った。結果を表 3 に示す。

CC で類似画像検索を行った結果に対し、LBP により推定した感性特徴量で感性フィルタリング(以下 CC-LBP)

LBP で類似画像検索を行った結果に対し、CC により推定した感性特徴量で感性フィルタリング(以下 LBP-CC)

結果として、実験 の両方で平均適合率は向上した。推定した感性特徴量から感性フィルタリングを行う場合、感性特徴量の推定に利用する画像特徴量は CC より LBP の方が効果的であると分かる。

表 3 . 画像特徴量を組み合わせた際のフィルタリング前後の平均適合率

	CC-LBP	LBP-CC
フィルタリング前	62.7%	52.8%
フィルタリング後	64.5%	55.3%

以上のように、評価実験により、本研究において提案した感性フィルタリング手法は、感性的な画像検索に有効な個人に適應したフィルタリングを行うことができる。さらに検索結果に含まれる嫌いなコンテンツをユーザがチェックすることにより、好悪分類モデルを再学習するようなフィードバック機能を追加すれば、ユーザの嗜好の変化や使用状況・使用場面に応じて、より個人に特化した感性フィルタリングや感性的検索が実現できると思われる。

5 . 主な発表論文等

(研究代表者、研究分担者及び連携研究者に

は下線)

〔雑誌論文〕(計 2 件)

Keranmu Xielifuguli, Akira Fujisawa, Yusuke Kusumoto, Kazuyuki Matsumoto, Kenji Kita: Pleasant/Unpleasant Filtering for Affective Image Retrieval Based on Cross-Correlation of EEG Features, Applied Computational Intelligence and Soft Computing, 査読有, Vol. 2014, 2014, pp.1-10
<http://dx.doi.org/10.1155/2014/415187>

Wenbin Zhang, Haoze Lu, Yasuo Horiuchi, Satoru Tsuge, Kenji Kita, Shingo Kuroiwa: Text-Independent Speaker Identification Based on Reducing Intersession Variability of Speech Feature Using PCA Transformation, Journal of Signal Processing, 査読有, Vol. 15, No. 4, 2011, pp.275-278

〔学会発表〕(計 7 件)

藤澤日明, 松本和幸, 北研二: キャラクタの顔を用いたイラスト画像の作者識別, HCG シンポジウム, 2013 年 12 月 18 日, 松山市総合コミュニティセンター (愛媛県)

Motoyuki Suzuki, Shoei Nakagawa, Kenji Kita: Emotion Recognition Method Based on Normalization of Prosodic Features, APSIPA Annual Summit and Conference, 2013 年 10 月 29 日, 85 Sky Tower Hotel (Taiwan)

Yunong Wu, Kenji Kita, Fuji Ren, Kazuyuki Matsumoto, Xin Kang: Exploring Emotional Words for Chinese Document Chief Emotion Analysis, The 25th Pacific Asia Conference on Language, Information and Computation, 2011 年 12 月 17 日, 南洋理工大学 (シンガポール)

Qingmei Xiao, Satoru Tsuge, Kenji Kita: Music Retrieval Method Based on Filter-Bank Feature and Earth Mover's Distance, 2011 Seventh International Conference on Natural Computation, 2011 年 7 月 27 日, Galaxy Hotel (上海, 中国)

柘植覚, 三好真人, 肖清梅, 北研二: フィルタバンク特徴量と Earth Mover's Distance を用いた音楽検索, 情報処理学会音楽情報科学研究会, 2011 年 2 月 12 日, 九州大学 (福岡県)

岸本諭，楠本祐介，多田圭介，獅々堀正幹，北研二：感性特徴と画像特徴の相関に関する研究，電気関係学会四国支部連合大会，2010年9月25日，愛媛大学（愛媛県）

Le Zhang, Kenji Kita: A Skeleton Extraction and Pruning Algorithm based on Bisector Angle, The Sixteenth Korea-Japan Joint Workshop on Frontiers of Computer Vision, 2010年2月5日，安芸グランドホテル（広島県）

〔図書〕（計 0件）

〔産業財産権〕

出願状況（計 0件）

取得状況（計 0件）

〔その他〕

ホームページ等

6. 研究組織

(1) 研究代表者

北 研二 (KITA, Kenji)
徳島大学・大学院ソシオテクノ
サイエンス研究部・教授
研究者番号：10243734

(2) 研究分担者

獅々堀 正幹 (SHISHIBORI, Masami)
徳島大学・大学院ソシオテクノ
サイエンス研究部・教授
研究者番号：50274262

松本 和幸 (MATSUMOTO, Kazuyuki)
徳島大学・大学院ソシオテクノ
サイエンス研究部・助教
研究者番号：90509754

柘植 覚 (TSUGE, Satoru)
大同大学・情報学部・准教授
研究者番号：00325250

鈴木 基之 (SUZUKI, Motoyuki)
大阪工業大学・情報科学部・准教授
研究者番号：30282015

土屋 誠司 (TSUCHIYA, Seiji)
同志社大学・理工学部・准教授
研究者番号：70452654