

代表色領域の位置関係に着目した大容量画像からの 類似部分画像の高速抽出

五味 愛 伊藤 貴之

お茶の水女子大学大学院

E-mail: {gomiai, itot}@itolab.is.ocha.ac.jp

概要

筆者らは、実写撮影画像の視線情報を高速検索するシステムVIEWGLEの研究に従事し、その中で大容量画像から類似部分画像の高速抽出アルゴリズムを用いている。本論文では、この類似部分画像抽出アルゴリズムを一般化した手法を提案し、その処理時間や正確さについて考察する。本手法は3ステップから構成される。最初のステップでは大容量画像に対して前処理を施し、その結果を蓄積しておく。続いて入力画像を提示されたときに、2つめのステップとして大雑把な類似度判定を行い、大容量画像中から少数の候補部分画像を抽出する。続いて3つめのステップとして、それら候補部分画像と入力画像の類似度を算出し、応答曲面法を用いて類似度が最大となる最適部分画像を出力する。本手法では、厳密に確実に画像を検索できる保証はない。しかし本論文の実験結果から、本手法が高速に、ある程度の満足のできる類似画像を抽出できることが示されている。

Fast Extraction of Similar Partial Images Based on Positional Relationship of Representative Color Regions

Ai Gomi* Takayuki Itoh

Graduate School of Humanities and Sciences, Ochanomizu University

E-mail: {gomiai, itot}@itolab.is.ocha.ac.jp

Abstract

We engage in Study of VIEWGLE which quickly retrieves viewing data of digital images, large-capacity images, and use an algorithm to quickly extract the similar partial image in there. In this paper, we suggest the general method using the algorithm and examine the processing time and accuracy. Our method composes 3 steps. The first step is a preprocessing for large-capacity images and stores the result. After bringing a input image, the second step extracts candidates of target parts by rough similarity estimation. Finally the third step calculates similarity values between the candidates and input images and discovers optimal target parts by searching for the maximum similarity in the response surface method. We have not guaranteed the accuracy and exactitude of the partial image retrieval yet, however the result of this paper, the technique realizes nearly quick retrieval from large-capacity images.

1. はじめに

デジタルカメラの目覚ましい発展に伴い、現在では、大容量画像を簡単に作成、使用できるようになり、大量の画像に対する画像検索は欠かせないものとなってきた。そして近年では主にインターネット上にて、膨大な画像情報の中から希望する画像を抽出するためのツールとして、画像検索エンジンが提供されている。

インターネット上で使われている大半の検索エンジンは、キーワードに基づくものや、単純な色ヒストグラム等を使ったものである。キーワードに基づく画像データベースの多くは、各々の画像に手作業でキーワードを与えているため、大規模画像データベースの構築には非常に大きな工数を要する。またキーワードだけに頼った画像データベースでは、言語での表現が困難な情報をもとに画像を検索することはできない。また色ヒストグラムだけに頼った画像データベース

においては、画像に何が写っているのかを把握した上で画像を検索することができない。そこで近年では、CBIR (Context-Based Image Retrieval)という、画像の内容に基づく画像検索に関する研究が進んでいる。これらの研究の多くはウェブ上で稼動するシステムとして開発されている[1]。CBIRの技術は画像検索のために参照する特徴量で大別され、その多くは色、形状、テキスト等の特徴量として利用している。

CBIRに関する研究の多くは、ある入力画像に対して、画像全体どうしを比較することで類似画像を検索する、いわゆる全体画像検索という技術に相当する。それに対して、任意の大容量画像と入力画像を与えられたときに、入力画像への類似部分を大容量画像から抽出する、いわゆる類似部分画像検索は、現在も活発に研究発表が進んでおり、まだ研究の余地があるといえる。類似部分画像を抽出するための最も一般的で基本的な解決手法に、テンプレートマッチング法が知られ

ており、その改良手法[2,3,4]も報告されている。それとは別に、「物体検索」「内容検索」という考え方にに基づく類似部分画像検索手法も多く研究発表されている。

筆者らは、実写撮影画像の視線情報を高速検索するシステムVIEWGLE[5]の研究に従事している。このシステムでも、大容量画像から類似部分画像の高速抽出は重要な技術である。VIEWGLEの研究の過程にて筆者らは、類似部分画像抽出技術に以下の要件があることを議論した。

[要件1] 必ずしも最適な抽出結果でなくてもいいから、ある程度満足できる抽出結果を、高速に実現したい。

[要件2] 入力画像中に写る物体の、幾何変換、撮影条件の変化、形状的特徴の変化などに対して、ある程度頑強な抽出を実現したい。

[要件3] 複数の入力画像に対して、同一の大容量画像を反復的に用いる場合には、この大容量画像に対して予め前処理を施すことで、高速化を図りたい。

本論文では、文献[5]で示した類似部分画像抽出手法を一般化することで、上述の[要件1~3]を満たす類似部分画像抽出手法を提案する[6]。本手法は以下の3ステップで構成されることを前提とする。

[ステップ0] 大容量画像に対する前処理。この前処理は1枚の大容量画像に対して一度だけ行えばよい。

[ステップ1] 入力画像を入力したときの前半処理。ここでは大雑把な類似判定により、大容量画像中から少数の候補部分画像を抽出する。

[ステップ2] 入力画像を入力したときの後半処理。候補部分画像と入力画像の類似度を算出し、応答曲面法を用いて類似度が最大となる最適部分画像を出力する。

本論文では、以上の3ステップを含むアルゴリズムを実装し、その計算時間や正確さについて考察する。現時点での筆者らの実装では、ステップ0にて大容量画像に対し、減色処理による領域分割を行う。続いてステップ1にて、入力画像にも同様に領域分割を行い、その代表色領域の位置関係が大容量画像の中の領域と類似している部分を、候補類似部分画像として抽出する。最後にステップ2にて、複数の候補類似部分画像と入力画像の類似度から、応答曲面法を用いて最適部分画像を抽出する。

CBIRに関するシステムのサーベイ論文[1]によると、現存するCBIRシステムが用いている特徴量を比較すると、画素の色統計量を用いている事例が最も多く、続いて周波数特性(テクスチャ)と物体境界線が使われており、少数ながら物体の位置関係に着目したシステムがある、とのことである。本手法は物体を意味する代表色領域の位置関係に着目することで、CBIRを高速化する一手法、と考えることができる。

2. 関連研究

本論文における「部分画像」とは、大容量画像の中から任意の位置、任意の大きさの長方形領域を切り取って得られる画像のことである。任意の入力画像を与えられたときに、この大容量画像に類似度の高い部分画像を抽出することを、本論文では「類似部分画像抽出」と称する。

よく知られている類似部分画像抽出手法に、テンプレートマッチング法が挙げられる。この方法では、部分画像と入力画像の対応画素間の画素値の差分を累積することで、部分画

像と入力画像の類似度を算出する。しかし、大容量画像から抽出しうる全ての部分画像において類似度を算出していたのでは、莫大な計算時間を必要としてしまう。そこでテンプレートマッチングを高速化した一般的な方法として、SSDA(Sequential Similarity Detection Algorithm)法[2]や粗密探索法[3]、アクティブ探索法[7,8,9]が提案されている。これらの高速化手法よりもさらに高速な手法を確立するためには、類似度判定の回数を劇的に減らす手法や、同一の大容量画像を複数回用いるという限定的条件下での高速化手法の確立、などが必要であると考えられる。またテンプレートマッチング法をベースにした手法は、入力画像中に写る物体の大きさ・形状的特徴・光学的変化に対する頑強さに欠ける傾向があり、この問題も同時に解決する必要がある。また、同一の大容量画像を複数回用いることを前提として、前処理としてその特徴をあらかじめ抽出することで、入力画像を入力してからの処理時間を削減する手法[4]が提案されている。しかしこれらの手法において実験に用いられている画像は、デジタルカメラの画像サイズに比べて小さいものが多く、日常生活に用いる規模の大容量画像において高速検索が得られるかは不明である。

入力画像中に写る物体の動作や変形に対する頑強さを得るためには、テンプレートマッチング法が用いている対応画素単位の局所的な類似度判定よりも、画像の大局的な特徴をとらえた類似度判定手法の導入が必要であると考えられる。このような考え方は、容量画像から特定物体を抽出する、いわゆる「物体検索」や「内容検索」と呼ばれている画像検索技術に導入されている例が多い。一例としてVisualSEEK[10]、QBIC[11]では、画像の色情報に基づいて領域分割を行い、物体領域を切り出している。また、エッジ情報を利用して物体領域を分割する手法も提案されている[12,13]。

また、近年の類似部分画像検索手法の特徴として、入力画像に写っている物体から形状や輝度情報の特徴を取り出し特徴点のマッチングを行う手法が増えている。代表的なものとして、Difference-of-Gaussian関数を使用した手法[14]、Gaborフィルタを使用した手法[15]、コーナー検出法を使用しアフィン変換を適用した画像マッチングを行った手法 [16]、Harris-Laplacian手法を使用して特徴点を取り出した手法[17,18]が挙げられる。これらのマッチングの精度は物体の回転、スケール、多少のノイズ等に頑丈で非常に安定している。しかし、これらの方法においては、グレースケールの画像を使用しているので、似たような形状をした異なる色の物体でも検出されてしまう恐れがある。

なお本研究の対象からは外れるが、顔などの特定物体の発見に焦点をあてた部分画像抽出手法[19-21]なども、興味深い研究と言える。しかし、これらの文献における実験結果もまた、一般的なデジタルカメラの画像サイズのような大容量画像より小さい画像を扱っており、日常生活に用いる規模の大容量画像においても高速検索が行えるかは不明である。

3. 提案手法の概要

3.1 アルゴリズムの概要

まず、本論文が提案する類似部分画像抽出手法を、以下のよう形式づける。

- 入力画像を I 、大容量画像を I' とする。本論文では大容量画像を、100万画素以上の画素数を持つ画像とする。
- I' から任意の長方形領域を抽出したものを部分画像と称し、 I'' とする。このとき、 I'' の中心を I' 上の座標値

(x, y) とし、 I' の幅を w 、高さを h とする。

- I と I' の類似度を計算し、類似度が最大となる I' を見つけ出す。

本論文では、以下の 3 ステップにより、類似部分画像を高速に抽出する手法を提案する。

[ステップ 0] 大容量画像 I' の特徴を抽出するための前処理を行う。この時点で入力画像 I は与えられていないとする。

[ステップ 1] I が与えられると、大雑把な類似度判定によって、 I' の中から少数の候補部分画像 I'' を抽出する。以下、ステップ 1 における類似度判定値を s_1 と称する。

[ステップ 2] ステップ 1 で抽出した各々の候補部分画像 I'' に対して、従来手法と同様な手法を用いて I との類似度を判定する。以下、ステップ 1 における類似度判定値を s_2 と称する。続いて応答曲面法を適用して、 s_2 値を補間する曲面を生成する。この曲面が最大値を示す位置における部分画像 I'' を、最適部分画像を近似する解であるとみなして出力する。ステップ 1 で抽出される候補部分画像が少数であれば、テンプレートマッチング法に比べて類似度判定の回数を劇的に減らすことができる。

画像の類似度判定手法の多くは、特徴色、境界線、周波数特性の 3 種類の特徴を参照している。この 3 種類の特徴に対応するステップ 0 からステップ 2 までの処理は、表 1 のようになる。本論文では、この 3 種類の特徴のうち、特徴色に着目した類似部分画像抽出手法を提案するものである。

表 1.3 種類の特徴に対するステップ 0~2 の各処理

	特徴色	境界線	周波数特性
ステップ 0	I' から特徴色を抽出	I' から境界線を抽出	I' の周波数特性を算出
ステップ 1	特徴色分布が類似する複数の領域を大雑把に抽出	境界線分布が類似する複数の領域を大雑把に抽出	周波数特性が類似する複数の領域を大雑把に抽出
ステップ 2	特徴色の類似度が最適である部分画像を特定	境界線の類似度が最適である部分画像を特定	周波数特性の類似度が最適である部分画像を特定

筆者らは文献[5]において、提案手法の原型となる類似部分画像抽出手法を述べている。文献[5]に対する本論文の進展は以下の通りである。

- 大容量画像のための前処理を[ステップ 0]と位置づけることで、同一の大容量画像を反復的に用いる際の高速化を実現する。
- [ステップ 1]を改良することで、入力画像に写る物体の形状的特徴や光学的特徴の変化に対して、ある程度の頑強な抽出を実現する。
- 文献[5]ではほとんど述べていなかった、本手法の計算時間や正確さについて、分析結果を提示する。

3.2 本研究の応用例

部分画像検索の利用法は多々考えられるが、1章で前述した[要件 1~3]に関連する利用法の一例を以下に示す。

多くの風景を広角撮影して収めた大容量画像データベー

スがあるとする。任意の写真を入力し、この写真に類似した風景をデータベースから抽出したいとする。このような用途での要件は、以下の通りである。

- 1枚の大容量画像からの部分画像検索でさえ、従来手法ではある程度の大きな計算時間がかかるくらいであることから、多数の大容量画像からの部分画像検索には非常に大きな計算時間がかかると考えられる。そこで高速な部分画像検索技術は非常に重要であると考えられる。
- データベースに格納された画像と、入力する任意の画像では、撮影条件が異なることが予想される。このため、両者の画像の間に多少の形状的特徴や光学的特徴の差異があっても、ある程度は頑強に検索を成功させる必要がある。
- データベースに格納されている大容量画像は不変であるとする、大容量ゆえに時間のかかる処理は、前処理としてあらかじめ済ませておくことが望ましい。

以上の理由により、上述の用途で部分画像検索を利用する際には、1章で前述した[要件 1~3]を満たす提案手法が有用であることがわかる。一例として、5章にて後述する[実験 5]の実写画像を見ていただきたい。この画像のような記念写真を大量に保存したデータベースの中から、特定の服装を着用した人物を抽出する、というような目的において、本手法は特に有用であると考えられる。

4. 提案手法の実装

本章では、筆者らの現段階の実装について論じる。現段階の実装は、画像に減色処理を施して生成された特徴色領域に基づいている。物体境界線や周波数特性、あるいは特徴点に基づいた類似度判定処理は、まだ実装していない。そのため現段階での筆者らの実装では、画像の色彩にある程度の特徴を有するカラー画像に限り、ある程度確実な検索を実現できる。逆に言えば、現段階での筆者らの実装では、色相値を計算できない画像（例えばモノクロ画像）や、色相値の分布が狭い画像（例えばセピア画像）を対象とすることはできない。ただし例えば、文献[12,13]の考え方を併用した実装、あるいは文献[14-18]の考え方を併用した実装を実現すれば、さらに頑強な類似部分画像検索を得られると考えられるだろう。

以下、特徴色領域に基づく各ステップの処理について論じる。

4.1 [ステップ 0] 減色処理

筆者らの実装の[ステップ 0]では、K-means 法を用いて、大容量画像の画素を K 個のクラスタに分類する。これにより、類似する色をもつ画素を同一クラスタに収め、結果として画像を減色する。また、本研究では、明度を一変数として分離した表色系である YCbCr 空間を使用している。色空間数であるクラスタの数 K は、現時点での実装ではユーザが指定している。アルゴリズムの手順を以下に示す。

1. 色空間（例えば YCbCr 空間）に画像中の全ての画素を配置する。
2. クラスタの中心 K 個を適当に配置する。
3. 各画素を K 個のクラスタのいずれかに分類する。
4. クラスタの中心を再計算する。
5. 新しいクラスタの中心が、前のクラスタの中心と十分に近かったら終了。さもなければ 3. に戻る。

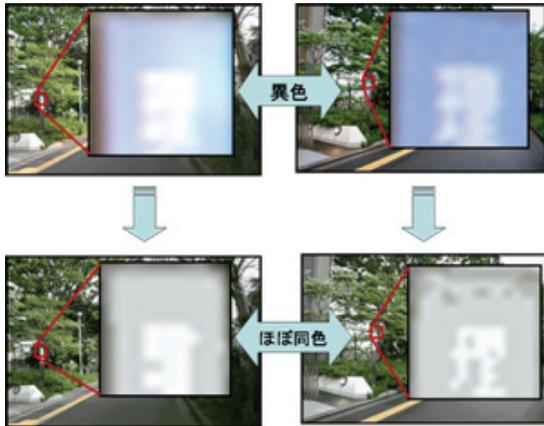


図1. K-means法の結果画像 (左上:晴天時・減色前, 右上:雨天時・減色前, 左下:晴天時・減色後, 右下:雨天時・減色後)

以上の過程により、わずかに異なった色調や、天候の違いによる色の変化を含む画像同士でも、対応する物体が写る領域を同じクラスタに分類することができる。図1では、色調の異なった画像間での減色操作による結果を示している。上部の画像は天候が雨と晴れの日撮影した画像であり、下部の画像は、それらを15色に減色した結果である。これらの画像を比較すると、下部の2枚の画像では、対応する物体が同じクラスタに分類されていることがわかる。クラスタの数Kは、現時点での実装ではユーザが、適切なクラスタになるように指定している。以後に述べる実験においては、いくつかのK値を実験した結果、最適なK値を適用している。その結果、それぞれの画像において、30~50色の減色処理を施している。近年では、適切なK値を自動的に決定するクラスタリング手法も研究されており、これらの適用は今後の課題であると考えられる。

[ステップ0]では続いて、各クラスタにラベリングという処理を施す。この処理では、同一のクラスタに属する隣接ピクセル群に対して、固有のIDを与えることにより、画像の領域分割を実現する。

5章に示す実行結果からもわかるように、以上の処理は提案手法全体の中でも、特に大容量画像において計算時間の大きな割合を占めることが多い。よって大容量画像における前処理を[ステップ0]と位置づけて使用前に済ませておくことは、実用性の観点で重要である。特徴色領域に基づく実装だけでなく、境界線や周波数特性に基づく実装においても、この利点は同様である。

4.2 [ステップ1] 候補部分画像の高速抽出

入力画像Iを得ると提案手法は、まず[ステップ0]と同様にIに対して大容量画像と同じK色で、減色処理を施す。この処理は[ステップ0]とは異なり、計算時間は非常に小さい。続いて以下の手順により、大雑把な類似度判定を行う。本研究では、多角形法と名づけている。

まず入力画像Iのラベリング結果から、特に大きな領域をn個抽出する。続いて、このn個の中心点を連結するn角形(n=2の場合は線分)を生成する。このn個のラベルの特徴色と同じ組み合わせの特徴色のラベルをI'からも抽出し、任意の組み合わせでn角形を生成する。I'からは非常に多くのn角形が抽出されるが、この中でIから生成したn角形と形状的に類似するn角形が生成された領域周辺は、Iと類似

する色分布を有すると判断し、この周辺を長方形に切り取ったものを候補類似画像I'の一つとする。

図2はn=3のときの例である。Iから大きな特徴色領域を3個抽出し、この中心点を連結する三角形を抽出する。また、これと同じ3色の組み合わせを有する三角形をI'から多数抽出する。I'から抽出された三角形の各々について、Iから抽出された三角との相似度を判定する。ここでの相似度の判定は、三角形の頂点座標の位置関係を考慮し、四則演算のみで計算し計算時間を極力減らす方法を取っている。以下、この相似度を s_1 とする。続いて、相似度の高い三角形の周辺をI'から長方形領域として抽出し、候補類似画像I'とする。

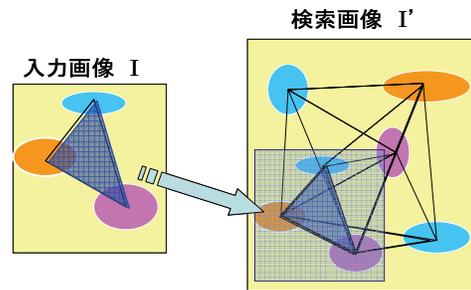


図2. 入力画像中のTと大容量画像中のT'の比較

5章で後述する筆者らの実行結果では、大容量画像I'から数千~数百万もの多角形を抽出し、相似度判定の結果として数枚から数十枚の候補部分画像を得ている。ここで抽出される多角形の数は膨大であるが、その各々の多角形の相似度判定の計算時間は非常に小さい。

なお一般的には、IとI'との間で、対応する物体や風景同士が同じサイズで写っているとは限らない。よってどれくらいの縮尺なのかわからない画像において、現時点での実装では以下の方法により、適切な大きさでI'を抽出している。

まずIから抽出したn角形をPとし、I'から抽出したPに相似度の高いn角形をP'とする。続いてPに外接する長方形をRとし、その幅と高さを w_p, h_p とする。同様に、P'に外接する長方形をR'とし、その幅と高さを $w_{p'}, h_{p'}$ とする。このとき提案手法はR'とRの大きさの比を $A = 0.5(w_{p'} / w_p + h_{p'} / h_p)$ とする。このAを用いて、Iの幅と高さ(w,h)をそれぞれA倍することにより、I'の幅と高さ決定する。

4.3 [ステップ2] 最適部分画像の近似生成

ステップ1で抽出した候補類似画像は、あくまでも大雑把な類似度判定によって抽出されたものであり、必ずしも候補類似画像の中に類似度が最大である画像があるとは限らない。著者らの観察では、大容量画像から複数の候補類似画像を抽出した位置の中間に、類似度が最大である部分画像が存在することが多い。

ここで、大容量画像からの部分画像について、従来の類似度判定手法[1-4]を用いた類似度算出結果は、大容量画像中の位置に対して滑らかな連続関数を形成する可能性が高いと仮定する(図3参照)。このような場合には、少数の類似度算出値を補間する曲面を生成することで、類似度が最大である位置を近似することができる。この考え方を利用して[ステップ2]では、[ステップ1]によって抽出された少数の候補部分画像I'_{i}に対して、既存の画像類似度算出手法を用いて入力画像Iとの類似度 s_2 を算出し、これを応答曲面法により補間する。以上の処理によって本手法では、類似度が最大である位

置を近似する。テンプレートマッチング法に代表される従来手法では、類似度 s_2 の算出回数が膨大になるのに対して、本手法では[ステップ 1]において候補部分画像を数枚〜数十枚に絞り込むことにより、 s_2 算出の計算時間の総計は非常に小さくなる。よって従来手法と比較して、大幅に計算時間を短縮できる。

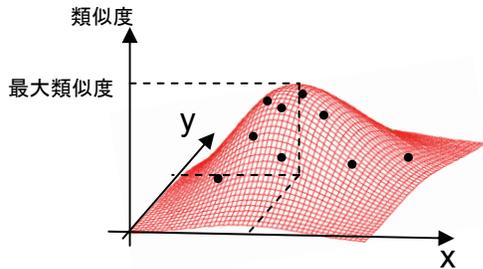


図 3. 応答曲面法

応答曲面法は、入力変数値と応答値の関係を「応答曲面」と呼ばれる近似関数で表すことにより、最適化問題を解く手法の一つであり、応答曲面として最小二乗曲面がよく使われる。ここで2次多項式を用いた場合、応答曲面は式(1)で表される。

$$s = \beta_0 + \sum_{i=0}^n \beta_i x_i + \sum_{i=0}^n \beta_i x_i^2 + \sum_{i < j} \beta_{ij} x_i x_j \quad (1)$$

これを本研究の問題におきかえると、以下ようになる。まず、各々の P^* に対して算出した類似度を式(1)に代入する。ただし式(1)において、 S は類似度であり、 $n=2$ であり、 x_0, x_1 は P^* 上の P^* の中心座標値 (x, y) の各数値である。この式に数組の入力変数値と応答値を代入して得られる連立方程式を解くことで、 β の各値を算出できる。この最小二乗曲面は一般的には画像類似度が最大となる1点をもつので、その点における (x, y) を算出することで、類似度が最大となる部分画像の中心座標値を近似する。

最後に本手法では、中心座標値を (x, y) とする部分画像 P^* を、幅および高さ (w, h) の値を数段階に設定して生成する。そして、その各々に対して I との類似度 s_2 を算出し、この値が最大であるものを出力画像とする。テンプレートマッチング法をはじめとする従来手法の多くは、入力画像と大容量画像中の対応する風景同士の縮尺が大きく異なるとき、うまく作動しないことがある。提案手法では、以上の処理により、この問題点を改善している。

5. 実験結果

本章では、提案手法を実装し、いくつかの画像について実験した結果を示す。筆者らは GNU gcc を用いて提案手法を実装し、COMPAQ Evo (Windows XP, CPU 2.79GHz, RAM 1.0GB) を用いて実行した。

[実験 1] 図 4(左)は本実験の入力画像 (858×1270 画素)、図 4(右)は大容量画像 (3456×2304 画素) である。ステップ 0 のクラスタ数は、50 個に指定した。図 5 は、入力画像と同サイズの出力画像(858×1270 画素)の生成結果である。本実験では、テンプレートマッチングの高速化手法である SSDA 法 [2] と、本手法との検索結果を比較した。SSDA 法における結果画像抽出位置に対する、本手法の出力画像抽出位置の差異は、x 軸方向に 3 画素、y 軸方向に 2 画素であった。これは入力画像および大容量画像のサイズと比較してほんの微小

な誤差であることから、本手法では SSDA 法とほぼ同じ類似部分画像を抽出できたといえる。



図 4. 実験 1 : (左)入力画像(右)大容量画像



図 5. 実験 1 の検索結果
(左)テンプレートマッチング(右)本手法

表 2. SSDA法と本手法の計算時間.

SSDA 法	本手法	
66 分 23 秒 26	ステップ 0	38 秒 67
	ステップ 1	1 秒 15
	ステップ 2	0 秒 45
	合計	40 秒 27

本手法と SSDA 法の計算時間を表 2 に示す。筆者らの実装した単純な SSDA 法では、非常に大きな計算時間がかかっている。もともとテンプレートマッチングは、画素数の小さいテンプレート画像との比較に向けた手法であり、本実験のように入力画像の画素数が大きい場合には必ずしも向いているとは言いが切れない。本手法では SSDA 法に比べて短時間で類似部分画像検索を実現できていることがわかる。また、それぞれのステップごとの計算時間を見てみると、前処理部分のステップ 0 にほとんどの計算時間が費やされており、ステップ 1 以降の計算時間は約 2 秒である。同一の大容量画像を反復的に用いる場合には、前もってこのステップ 0 を行っておくことで、高速検索が可能となる。

[実験 2] 図 6(左)は本実験の入力画像 (202×583 画素)、図 6(中)は大容量画像 (3456×2304 画素)、図 6(右)は出力画像 (400×1154 画素) である。この結果より本手法が、入力画像と出力画像が同じサイズでなくても、的確な検索結果を導くことを示している。

ステップ 0 の K-means 法では、35 個のクラスタに分類を行った。この検索のステップ 1 で生成された多角形の数は 7,560 個であった。このうち 8 個を用いて抽出した候補画像 P^* を図 7(上段)で示す。更に、一番左の画像が入力画像であり、下段は、抽出した 3 つのラベルの位置を示している。この 8 個の大容量画像中の位置を円で示したものを図 8 に示す。ここで円の半径は s_2 の大きさに比例して大きく表示されている。また赤い点は、応答曲面法によって得られた、類似度が最大とされる位置である。この結果より、図 6(右)に示す出力画像が検索されている。



図 6. 実験 2 の検索結果：
(左)入力画像(中)大容量画像(右)出力画像



図 7. 実験 2 のステップ 1 における候補画像 I'

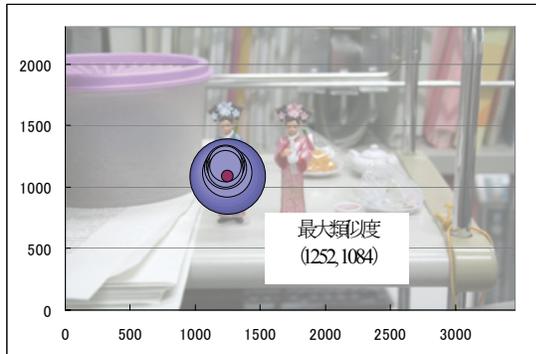


図 8. 実験 2 の候補画像 I' の類似度分布

[実験 3] 図 9(左)は本実験の入力画像 (700×1121 画素), 図 9(中)は大容量画像 (3456×2304 画素), 図 9(右)は出力画像 (500×800 画素) である. ステップ 0 の K-means 法では, 35 個のクラスタに分類を行った. この実験結果は, 大容量画像において類似する色を持つ物体を複数含んでいるにもかかわらず, 的確に検索ができたことを示している.

この検索で生成された多角形の数 は 2,417,580 個である.

[実験 2]と同様に, ステップ 1 で抽出された候補画像 I' を図 10 で示す. また, これらの I' 上での位置づけ, および応答曲面法による最適部分画像の中心座標の抽出結果を, 図 11 に示す.



図 9. 実験 3 の検索結果：
(左)入力画像(中)大容量画像(右)出力画像



図 10. 実験 3 のステップ 1 における候補画像 I'

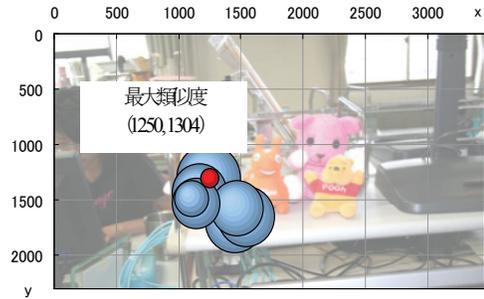


図 11. 実験 3 の候補画像 I' の類似度分布

[実験 4] 図 12(左)は本実験の入力画像 (430×383 画素), 図 12(中)は大容量画像 (2048×1360 画素), 図 12(右)は出力画像 (700×623 画素) である. ステップ 0 の K-means 法では, 35 個のクラスタに分類を行った. この実験結果は, 入力画像が携帯電話の内蔵カメラなどで撮影するような小容量画像であっても, 本手法による検索が可能であることを示している. 更に本手法では, 入力画像と大容量画像中の被写体の向きや, 表情等が多少異なっているにもかかわらず, 的確な検索が可能であることを示している. この検索で生成された多角形の数 は 94,500 個である. [実験 2,3]と同様に, ステップ 1 で抽出された候補画像 I' を図 13 で示す. また, これらの I' 上での位置づけ, および応答曲面法による最適部分画像の中心座標の抽出結果を, 図 14 に示す.



図 12. 実験 4 の検索結果：
(左)入力画像(中)大容量画像(右)出力画像



図 13. 実験 4 のステップ 1 における候補画像 I'

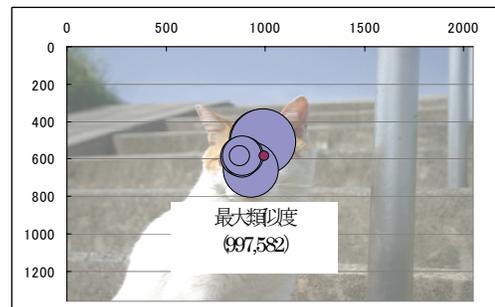


図 14. 実験 4 の候補画像 I' の類似度分布

[実験 5] ステップ 0 の K-means 法では, 30 個のクラスタに分類を行った. 図 15(左)は本実験の入力画像 (282×750 画素), 図 15(中)は大容量画像 (1664×2496 画素), 図 15(右)は出力画像 (700×1799 画素) である. この実験結果は[実験 4]と同様に, 人物においても, 表情や姿勢が多少異なっているにもかかわらず, 本手法による検索が可能であることを示している. この検索で

生成された多角形の数は、53412 個である。[実験 2,3,4]と同様に、ステップ 1 で抽出された候補画像 I' を図 16 で示す。また、これらの I' 上での位置づけ、および応答曲面法による最適部分画像の中心座標の抽出結果を、図 17 に示す。



図 15. 実験 5 の検索結果：
(左)入力画像(中央)大容量画像(右)出力画像



図 16. 実験 5 のステップ 1 における候補画像 I'

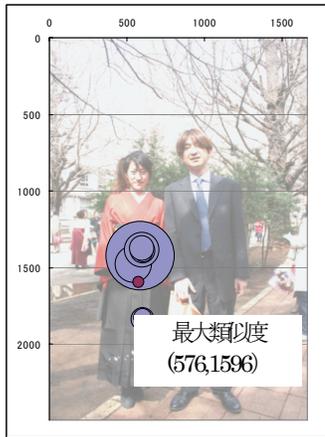


図 17. 実験 5 の候補画像 I' の類似度分布

6. 考察

筆者らの実行結果を総括して、以下に考察を行う。
まず計算時間の内訳について考察する。表 3 は、[実験 1]～[実験 5]のステップごとの計算時間をまとめたものである。この表から、どの実験においても[ステップ 0]に大半の計算時間を費やしており、[ステップ 1,2]の計算時間は非常に小さいことがわかる。このことから、同一の大容量画像を反復的に用いる際、前処理として[ステップ 0]を実行し、その結果を記録しておくことで、高速検索が可能となることがわかる。
続いて、本手法で設定すべきパラメータについて考察する。本手法の結果を最も左右するパラメータは、画像を減色する際のクラスタ数である。しかし、類似部分画像検索の目的において、常に最適なクラスタ数を自動決定するのは難しい。そこでやむをえず、筆者らの現時点での実装では、ユーザが手動でクラスタ数を決定している。このクラスタ数が適切でないと、適切な類似部分画像検索ができない場合がある。

表 3. 実験 1～5 の計算時間(秒)

	実験 1	実験 2	実験 3	実験 4	実験 5
ステップ 0	38秒67	46秒82	38秒14	12秒18	20秒18
ステップ 1	1秒15	0秒41	1秒23	0秒19	0秒47
ステップ 2	0秒45	0秒06	2秒21	0秒22	0秒41
合計	40秒27	47秒29	41秒59	12秒59	21秒06

一例として、図 18 の画像を用いて、クラスタ数の変化に伴う減色結果および検索結果の違いを、表 4 に示す。この表では、4.2 節における n の値を 3 とし、大容量画像の減色結果から最も大きな色領域 3 個を切り取った結果を示している。それと同時に表 4 では、入力画像においても大容量画像から切り取った色領域と同一の色領域 3 色を選んで切り取った結果、さらには検索結果となる出力画像を示している。

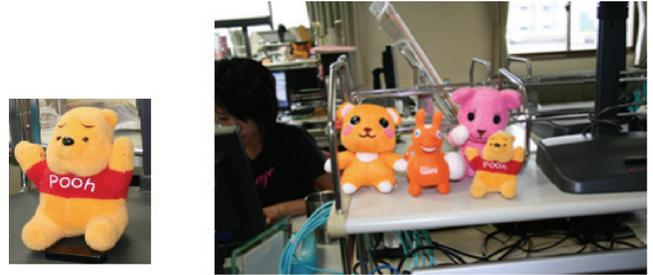


図 18. 入力画像(左)大容量画像(右)

表 4. 実験 3 におけるクラスタ数の違い

クラスタ数	37	40	45	50
入力画像 (435 × 553)				
大容量 画像 (3456 × 2304)				
出力画像				

この結果からわかるように、クラスタ数が 37 色および 40 色とした場合、クラスタ数が少なすぎるために、大容量画像において、検索したい物体の一部と後方にある物体（具体的にいうと、実験 3 の大容量画像における右端のぬいぐるみの赤い帯領域と、その左後ろに接しているピンク色のぬいぐるみの領域）とが同クラスタに属してしまっている。その結果として、意味のない出力画像が生成されているのがわかる。一方で、クラスタ数が 45 色および 50 色の場合、的確にクラスタリングができていたため、出力画像は的確な部分を抽出している。
このように、単純な K-means 法を適用したクラスタリングでは、必ずしも物体として意味のある画像領域分割を実現できるとは限らない。そのため今後の課題として、色分布のみに頼らない意味あるクラスタ方法の導入が必要であるといえる。一例として、減色処理と物体境界線抽出処理の両方の結果を参照する画像領域分割手法の導入が考えられる。また別の例として、いったん細かく画像を領域分割した後で、画像空間上で近接している類似色領域を併合するクラスタリン

グ手法も知られており、このような手法の導入も考えられる。

7. まとめ・今後の課題

本論文では、大容量画像の中から入力画像と類似している部分を高速に検索する一手法として、3ステップ式のアルゴリズムを提案し、その実行結果を示した。

現段階の実装は特徴色だけに基づいて画像の類似度を判定するので、色相的な特徴をもたないモノクロ画像やセピア画像などは対象外としているが、それ以外のカラー画像においても検索がうまくいかない事例が多々あると予想される。そこで今後の課題として、境界線や周波数特性などに基づいた類似度判定手法も導入していきたい。また、この問題を改善するために、6章でも論じたとおり、減色処理に基づく画像領域分割手法の改善も必要であると考えている。

また、現時点での筆者らの実装では、検索したい物体が大容量画像中に多数含まれていた場合に、その全ての物体を含む多数の類似部分画像を適切に検索することができない。解決策として、多角形法において検索された候補類似部分画像を、その位置でクラスタリングし、それぞれのクラスタに属する候補部分画像ごと応答曲面法を適応する、といった方法を今後取り入れていきたい。

本論文の冒頭にも論じたように、本手法は厳密に類似度が最大である部分画像を抽出しているとは限らない。実験1でも論じたように、テンプレートマッチングによって抽出された類似部分画像と、本手法によって抽出された類似部分画像には、多くの目的において差し支えないと思われる若干の誤差が観察される。しかし実際の用途において、この誤差の許容範囲は、数値的にというよりも主観的に決定されるものであり、またアプリケーションの種類にも依存すると考えられる。本論文では本手法のアプリケーションを具体的に特定していないが、本手法を主観評価するためには、具体的なアプリケーションを実際に開発した上で、被験者に対して検索結果を主観評価してもらう必要がある、と考えている。

謝辞

応答曲面法のプログラムを提供して頂いた京都大学小山田耕二教授、比戸将平氏に感謝の意を表します。

参考文献

- [1] R. C. Veltkamp, M. Tanase, Content-Based Image Retrieval Systems: A Survey, Rapport Technique, 2000.
- [2] E. I. Barnea and H. F. Silverman, A class of Algorithms for Fast Digital Image Registration, IEEE Transactions on Computers, Vol. C-21, pp. 179-186, Feb 1972.
- [3] M. Atiquzzaman, Coarse-to-Fine Search Technique to Detect Circles in Images, International Journal Advanced Manufacturing Technology, 15:96-102, 1999
- [4] A. Kimura, T. Kawanishi, K. Kashino, Similarity-based Partial Image Retrieval Guaranteeing Same Accuracy as Exhaustive Matching, IEEE 2004 International Conference on Multimedia & Expo. (ICME2004), 2004.
- [5] A. Gomi, T. Itoh, K. Koyamada, S. Hido, VIEWGLE: Fast Extraction of Similar Partial Images for Querying Viewing Parameters, NICOGRAPH International 2006.
- [6] 五味, 伊藤, 代表色領域の位置関係に着目した大容量画像からの類似部分画像の高速抽出, 第22回NICOGRAPH論文コンテスト, 2006.
- [7] 村瀬, V. V. Vinod, 局所色情報を用いた高速物体探索—アクティブ探索法—, 電子情報通信学会論文誌D, Vol. 81-D, No.

9, pp. 2035-2042, 1998.

- [8] 彌富, 萩原, 適応ファジー推論ニューラルネットワークとアクティブ探索法を用いた画像認識, 電子情報通信学会論文誌 D, Vol. J87-D2, No. 4, pp. 958-966, 2004.
- [9] 竹田, 加藤, 西田, 照合幅と照合順序を考慮した高速類似領域探索法, MIRU2006, 2006.
- [10] J. R. Smith, S. F. Chang, VisualSEEK: a fully automated content-based image query system, Proc. of ACM Multimedia, pp. 87-93, 1996.
- [11] M. Flickner, Query by Image and Video Content: The QBIC System, IEEE Computer, Vol. 28, No. 9, 1995.
- [12] 串間, 赤間, 紺谷, オブジェクトに基づく高速画像検索システム: ExSight, 情報処理学会論文誌, Vol. 40, No. 2, 1999.
- [13] 金原, 佐藤, 浜田, プリミティブ分解による多様な検索条件を扱うカラー画像検索, 情報処理学会論文誌, Vol. 37, No. 11, 1996.
- [14] D. G. Lowe, Distinctive image features from scale-invariant keypoints, International Journal of Computer Vision, Vol. 60, No. 2, pp. 91-110, 2004.
- [15] J. Mutch, D. G. Lowe, Multiclass Object Recognition with Sparse, Localized Features, IEEE Conference on Computer Vision and Pattern Recognition 2006 (CVPR '06), pp. 11-18, 2006.
- [16] T. Kadir, A. Zisserman, M. Brady, An affine invariant salient region detector, European Conference on Computer Vision (ECCV), pp. 404-416, 2004.
- [17] K. Mikolajczyk, C. Schmid, Indexing based on scale invariant interest points, Proc. of International Conference on Computer Vision (ICCV 2001), pp. 525-531, 2001.
- [18] 木村, 川西, 大塚, 柏野, 重み付き特徴点照合に基づく高速画像検索, 電子情報通信学会技術報告, PRMU2005-24 (DE2005-2), 2005
- [19] T. Hattori, H. Kitajima, T. Yamasaki, Face Pattern Recognition And Extraction From Multi Persons Scene, Proceedings of ICEIS 2003 (Fifth International Conference on Enterprise Information Systems), ACM, AAAI and IEEE, pp.92-99, 2003.
- [20] R. Fergus, P. Perona, A. Zisserman, Object Class Recognition by Unsupervised Scale-Invariant Learning, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 264-271, 2003.
- [21] S. Helmer, D. G. Lowe, Object Class Recognition with Many Local Features, Generative Model Based Vision 2004 (GMBV 2004).

著者紹介



五味 愛

2006年お茶の水女子大学理学部情報科学科卒業。現在お茶の水女子大学大学院人間文化研究科数理・情報科学専攻在学中。情報処理学会会員。



伊藤 貴之

1990年早稲田大学理工学部電子通信学科卒業。1992年早稲田大学大学院理工学研究科電気工学専攻修士課程修了。同年

日本アイ・ビー・エム(株)入社. 1997年博士(工学). 2000年
米国カーネギーメロン大学客員研究員. 2003年から2005年
まで京都大学大学院情報学研究科 COE 研究員(客員助教授相
当). 2005年日本アイ・ビー・エム(株)退職, 2005年よりお
茶の水女子大学理学部情報科学科助教授. ACM, IEEE
Computer Society, 情報処理学会, 芸術科学会, 画像電子学会,
他会員.