

色相を考慮した画像検索手法 Image Retrieval Method considering the Hue

荒川 貴弘†
Takahiro Arakawa

清水 郁子†
Ikuko Shimizu

1. まえがき

現在一般的に使われている画像検索は、キーワードによる画像検索である。これに対し、与えられたクエリ画像によって、類似する画像を検索する類似画像検索がある。クエリ画像に似た画像を検索するのが類似画像検索であるが、一言で「似た画像」と言っても、これを定量化するのは非常に難しい。人が直感で感じることのできる「似ている」とは、とても抽象的で曖昧だということが言え、画像検索をする際にもこの曖昧さをどのように吸収するのが問題になる。

2. 従来研究

画像検索によく使われる特徴量に SIFT (Scale Invariant Feature Transform) [1]があげられる。SIFT は画像中の輝度の勾配から計算される特徴量であり、エッジなど輝度の変化する部分の特徴を捉えることができ、幾何学的な特徴量ともいえる。スケールや照明の変化に不変の特徴量で頑健性が高いので、物体認識において最もよく使われる特徴量である。

一方で、SIFT ではグレースケール画像に対して処理を行う。そのため色の変化には非常に弱い。例えば夕焼けの画像を検索することを考えてみよう。図 1 に示すような夕焼けの画像は全体的に赤やオレンジ色であることが多い。さまざまなカテゴリの画像の中からこの画像を検索したい場合は色を使用することが有効であることは容易に想像がつく。



図 1 夕焼けの画像[2]

この理由により、SIFT に色情報を加えた SIFT 特徴量についても様々な研究がなされている。最も有名なものに CSIFT[3]が挙げられる。これは反射スペクトラムをモデルにした Kubelka-Munk 理論[4]に基づいた色の情報を用いている。色の情報が特徴量に加わるため、特徴量の次元は SIFT の 128 次元に対して、その 3 倍の 384 次元となるが、これによって計算速度が落ちてしまう。また、CSIFT では特徴ベクトルとして色情報を保持しているため、色の扱いが難しいという問題もある。例えば色の傾向があまりなく、

形状特徴の方が重要であるカテゴリを考えてみる。この場合、色情報は特に重要ではなくなってくるが、特徴ベクトル内には色情報を用いているため、分類の際に色情報に引っ張られてしまい、単なる SIFT を用いた分類よりも分類性能が落ちてしまう問題がある。

そこで本論文では、画像から色情報を高速により精度を高く計算し、画像検索に応用する手法を提案する。画像からは色情報として画素から得られる色相を用いる。

3. 色相を考慮した画像検索手法

3.1 画像検索の概要

本手法では予めカテゴリ分けされた複数の教師画像を使って学習を行い分類器を作成しておき、この分類器を用いてテスト画像を分類する。SIFT と、後述する色情報を用いて総合的に分類結果を評価するといったパラメトリックな手法である。SIFT と色の情報は統合されずに別々に扱われる。

形状特徴とも言える SIFT のスコアを計算する分類器の学習には Bag-of-Keypoints[5]を用い、分類には Naive Bayes モデルを用いる。Bag-of-Keypoints とは、文書分類に使われている Bag-of-Words を画像分類にも応用したものであり、文書中 (画像中) に現れる単語 (特徴量) の頻度のみを用い、単語の現れる順番 (特徴点の位置) などは無視するという考え方である。Naive Bayes モデルとは、ベイズ理論に基づく確率的な分類手法であり、単語の頻度の類似度によって分類スコアを計算するといった手法である。Bag-of-Keypoints と Naive Bayes モデルを組み合わせることで、似た特徴が多く含まれる画像を検出することができる。

また、これとは別のプロセスにおいて色情報のみによる色スコアを計算するが、これは後述する。SIFT から得られた形状スコアと色情報から得られた色スコアを総合的に判断することで、総合スコアを計算し、分類精度を上げるが、これも後述する。

3.2 色傾向の定量化

本手法の目的は、類似画像検索において色情報を効率的に用いて検索精度を向上させることである。

色情報を用いる際には、「赤っぽい」や「青っぽい」といった特徴を捉えるために、RGB 色空間ではなく HSV 色空間を用いる。RGB 色空間が光の 3 原色である赤、緑、青の 3 種類のパラメータから構成されているのに対し、HSV 色空間では色相、彩度、明度の 3 種類のパラメータから構成される。色相とは、赤やオレンジ、黄色、緑、青などの色の種類を示しており、0~360 度で循環する値である。彩度とは、色の鮮やかさを示しており、彩度が高いと色がより鮮やかになり、彩度が低くなると色が薄くなり白っぽくなる。明度とは明るさを示しており、明度が低いと色が黒っぽくなり、明度が高いと色が鮮やかになる。これを図で

† 東京農工大学

示すと図 2 のようになる。「赤っぽい」や「青っぽい」といった量は HSV 色空間における色相に現れるため、本手法では色相の値を用いることにする。

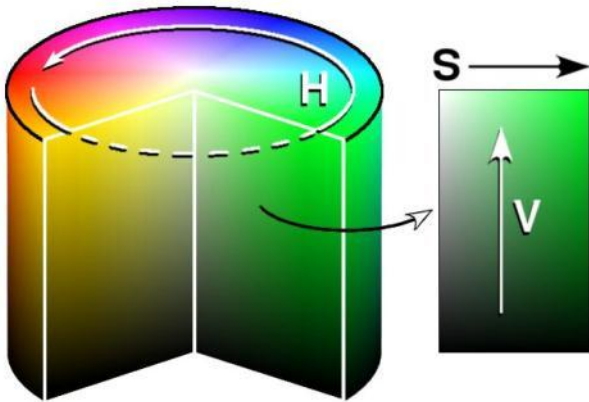


図 2 HSV 色空間[6]

1 つの画像に対する色の傾向を求める手法を述べる。画像の全画素から色相値を取得し、その平均と標準偏差を求める。色相平均によって画像が全体的に何色の傾向があるのかが分かり、色相標準偏差によって色のばらつきを定量化できる。なお、色相は $0\sim 360$ 度で循環する値なので、平均を求める際には単位ベクトル合算法を用いる。

各カテゴリに対する色の傾向は次のようにして求める。カテゴリに属する全ての画像の色相平均値の平均と標準偏差を求める。これをカテゴリの色情報とする。カテゴリに属する多くの画像が「赤っぽい画像」であったら、ここで計算されたカテゴリ色相平均は「赤」の位置になり、カテゴリ標準偏差は小さくなると思われる。これに対して、カテゴリ内に多くの色を含んだ画像がある場合、カテゴリ色相平均は何らかの値を取り、カテゴリ標準偏差は大きな値になると考えられる。カテゴリ色相平均とカテゴリ標準偏

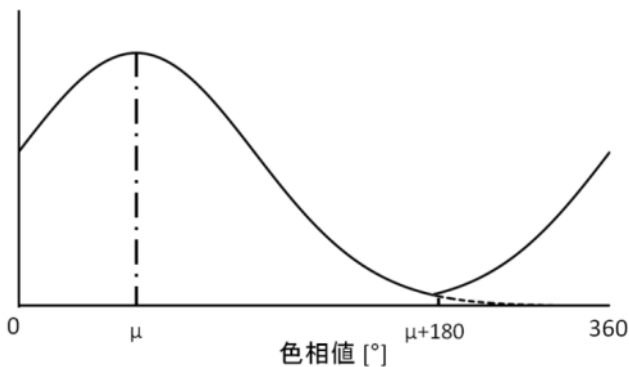


図 3 色相正規分布の例
差の 2 つを以って、カテゴリの色情報とする。

3.3 色スコアの算出

色スコアを計算するには、前節で述べた色情報を用いる。画像内における色相ヒストグラムは正規分布と仮定し、色情報である色相平均と色相標準偏差から正規分布を作成する。なお、この際にも色相は $0\sim 360$ 度で循環するので、色相平均-180 度から色相平均+180 度の範囲の正規分布を作

る。図 3 に、色相平均が μ 、色相標準偏差が σ のときの正規分布を示す。本来の正規分布であれば色相値が $\mu+180$ より大きくなると破線のように減衰していくが、先述の通り色相は循環するため実線のような分布になる。

分類対象の画像とカテゴリに属する画像の色相平均の差が小さければ、対象画像とカテゴリ内の画像の色が似ていると判断して色スコアを高くする。また、カテゴリに属する画像に色の傾向が強い場合は、カテゴリ色相平均と画像の色相平均の差がよりスコアに反映されやすくする。カテゴリ標準偏差が大きい場合と小さい場合では、カテゴリ色相平均と画像色相平均の差が同じときでも、前者の方が色スコアが高くなる。

これを実現するために、分類対象画像と分類先カテゴリの色相分布同士を積分し、これに重みを付けたものを色スコアとして計算する。この様子を図 4 に示す。実線が対象画像の色相分布、破線がカテゴリの色相分布だった場合、それぞれの分布の積分値である斜線部の面積が色スコアになる。これに加え、先述のカテゴリ色相標準偏差の値を考慮するため、積分値にこれを重みとする係数をかける。

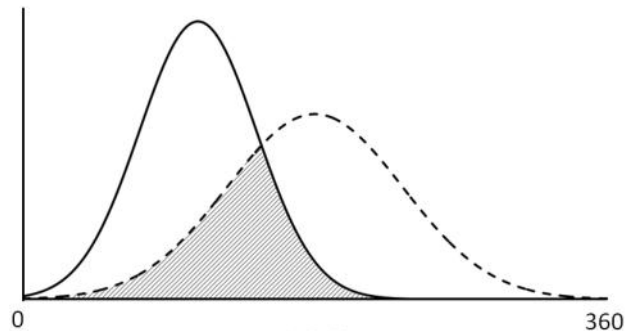


図 4 色スコアの計算手法

3.4 総合スコアの算出

SIFT によって求まる形状スコアと色情報によって求まる色スコアに重み α をかけたものを足し合わせることで総合スコアとする。 α の値は経験的に求めるとするが、色情報をより重要視したい場合は大きく、重要視しない場合は小さくすることで、色情報の影響を制御することができる。

4. 実験と考察

提案手法の有効性を検証するため、予めカテゴリ分けしている画像に対して、色情報の算出と、これを利用した画像分類を行った。カテゴリは「自転車」「バイク」「自動車」「猫」「犬」「飛行機」「夕焼け」の計 7 つとした。使用画像は Caltech Image Database と The PASCAL VOC 2006 から上記のカテゴリに属する画像を選択し、評価を行った。教師画像にはそれぞれ異なる 50 枚の画像を用い、分類テスト画像には教師データとは重複のない異なる 50 枚の画像を用いた。

4.1 色傾向の定量化

各画像に対して色情報を計算し、カテゴリの色相平均と標準偏差を求めた。結果のグラフを図 5 に、カテゴリ画像例とその色情報の結果を図 6 に示す。

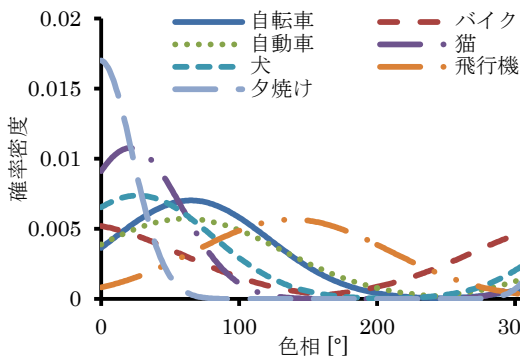


図 5 カテゴリごとの色相正規分布

図 5 を見ると、「夕焼け」カテゴリでは色相平均が赤色付近であり、色にあまりばらつきがないことが定量的に見ることができる。また、「飛行機」カテゴリでは空が写っている画像が多いため、カテゴリ色相平均は水色付近であるが、必ずしも空が写っている画像だけではないために、



図 6 カテゴリに属する画像例

標準偏差は大きめで緩やかな分布になっているのが分かる。図 6 を見ると、「自動車」カテゴリでは色の平均が緑色付近で標準偏差は大きいことが分かる。これは、自動車の背景にある空が影響しているのだと考えられる。しかし「自動車」カテゴリに属している画像は、自動車に重きが置かれていると考えられるので、背景を除去するといったような処理が必要かもしれない。

4.2 画像検索

SIFT 特徴量と色情報を用いて画像検索を行った。なお、総合スコア算出時の色スコアの重みは $\alpha=100$ とした。分類結果を表 1 に示す。この結果を見てみると、自転車、車

のカテゴリでは高い精度で分類を行うことができていて一方で、猫や犬といったカテゴリではあまり精度が高くない。猫や犬の画像の分類は SIFT のみであっても結果が悪かったので、SIFT 特徴量の抽出の時点で、特徴的な情報を捉えられていないのではないかと考えられる。

5. まとめ

本稿では、色相の値を用いて色情報を定量的に扱う手法を提案した。色情報として用いる際には、RGB 空間ではなく HSV 空間の色相の値を用いることで、「赤」や「青」といった色を 0~360 度の循環する値で表すことができる。これを利用し、画像中の全画素から色相の平均と標準偏差を求め、平均が画像全体の色、標準偏差が画像中の色のばらつきを示すパラメータとして、画像分類に応用した。画像分類を行う際には、形状特徴と呼ばれる SIFT による形状スコアと、色情報である色スコアを別々に計算し、これを足し合わせることで総合スコアとした。本手法により、分類精度が高まることを確認できた。

なお、現在は全画素から画素値を取得して平均と標準偏差を計算しているが、例えば「自動車」カテゴリのような対象物体がある画像では対象物体を検出し、その周辺の画素値のみを使った方がよいと思われる。一方で、「夕焼け」カテゴリのような画像では画像全体から画素値を取得した方がよいと思われる。今後は、この処理を自動で行えるようにしていきたい。

文 献

- [1] D. G. Lowe, Object recognition from local scale-invariant features, In Proceedings of the International Conference on Computer Vision, pp.1150-1157, 1999
- [2] http://fomalhaut.cocolog-nifty.com/fomalhaut/2006/08/_caa9.html
- [3] Alaa E. et.al, CSIFT: A SIFT Descriptor with Color Invariant Characteristics, Proceedings of the IEEE Computer Society Conference on CVPR, pp.1978-1983, 2006
- [4] P. Kubelka, New contribution to the optics of intensely light scattering materials, the Journal of Optical Society of America, pp.448-448, 1948
- [5] C. D. Manning, and H. SchFutze, Foundation of statistical natural language processing, The MIT Press, 1999.
- [6] <http://ja.wikipedia.org/wiki/HSV%E8%89%B2%E7%A9%BA%E9%96%93>

表 1 分類結果

	自転車	バイク	車	猫	犬	飛行機	夕焼け	合計
正解数	41	31	41	12	19	34	36	214
誤分類	16	26	17	19	45	5	8	136
未分類	9	19	9	38	31	16	14	136
F 値	0.707	0.492	0.701	0.202	0.262	0.648	0.667	0.799