

色情報を用いた類似画像の検索技術に関する研究

範 雋偉[†] 西原 功[‡] 中野 慎夫[‡]

[†] NTT アイティ (株) 〒231-0032 横浜市中区不老町 2-9-1 関内ワイズビル

[‡] 富山県立大学工学部情報システム工学科 〒939-0398 富山県射水市黒河 5180

E-mail: [†] han.juni@ntt-it.co.jp, [‡] {nishihara, nakano}@pu-toyama.ac.jp

あらまし 現在インターネット上に膨大な画像情報があり、検索者が欲しい画像を素早く入手することは困難である。本研究では、全ての画像を 13 色の基本色に減色し、その色割合を計算する。さらに人間の視覚特性を考慮して、①代表色の色割合を使った類似度分析方法、②暖色の色割合を重視した 2 種類の類似度分析方法を提案する。実験と評価により、これらの類似度分析方法は画像検索の速度、精度において効果があることを確認した。

キーワード 類似画像検索, 色情報

Study on Similar Image Retrieval Technology Using Color Information

Juni HAN[†] Isao NISHIHARA[‡] and Shizuo NAKANO[‡]

[†] NTT-IT Corporation 2-9-1 Oizuchou, Naka-ku, Yokohama-shi, Kanagawa-ken, 231-0032 Japan

[‡] Toyama Prefectural University 5180 Kurokawa, Imizu-shi, Toyama-ken, 939-0398 Japan

E-mail: [†] han.juni@ntt-it.co.jp, [‡] {nishihara, nakano}@pu-toyama.ac.jp

Abstract It is difficult searching image which a user wants. In this study, all images are subtractive in 13 colors as basic colors and calculated ratio of colors. In consideration of human seeing sense, we proposed two methods, I Similar analysis method using typical color ratio, and II Similar analysis method using warm colors and cold colors. The experiment confirmed that these analysis methods are effective.

Keyword Similar Image Research, Color

1. まえがき

近年、インターネット技術、画像圧縮技術、記録装置の大容量化によって、インターネット上で画像が活発に利用されている。この状況において、膨大な画像情報から、希望のコンテンツを検索することは困難である。本研究では、参照画像を用いる類似画像検索において、特に全体的な画像の印象を表現する色に着目し、より高速、高精度な画像検索方法を提案する。

2. 類似度分析による類似画像検索方法の提案

2.1. 類似画像検索における前処理

類似画像検索とは、画像の内容（画像の色情報、輪郭情報）において似ている画像を探すことである。これまで様々な類似画像検索手法が提案されてきた。

本研究では、まず画像の全体の色情報を対象として $L^*a^*b^*$ 色空間における JIS で定義されている 13 種類の基本色（赤、黄赤、黄、黄緑、緑、青緑、青、青紫、紫、赤紫、白、灰色、黒）に減色処理を行う。減色方法は、ある画像のすべての画素の色と各基本色との距

離を求め、距離が最小のものを最も近い色と判定し、その色に変換する処理を繰り返すことで実現する。そこで、減色された画像が元画像と同じような色割合を持つ手法として、誤差拡散法を用いた。さらに、式(1)よりそれぞれの色の割合を求める。これらの色の種類と色割合の特徴を利用し、画像の類似度分析を行う。以下図 1 に画像の色割合の抽出結果を示す。



図 1 減色された画像の色割合の抽出結果

$$R_i = \frac{A_i}{\sum_{n=1}^{13} A_n} \quad (1 \leq i \leq 13) \quad \dots (1)$$

R_i : i 番目の色の割合、 A_i : i 番目の色の画素数。

2.2. 類似度分析方法の提案

本研究では、2種類の画像の類似度分析方法を提案する。まず方法1として、比較参照用として用いる全ての色情報を用いた類似度計算法を示す。次に提案する方法として、代表色のみを用いた方法2と寒色暖色を考慮した方法3について述べる。

方法1: 全ての基本色の色割合を使った類似度分析方法。式(2)より同じ色の割合の差(D)の和であるSを求める。そして、式(3)より、求めたSの正規化処理を行い、類似度Rを求める。以下図3に図2に示す画像を対象とした場合の方法1の計算例を示す。



図2 実験用の画像

$$S = D(\text{赤}) + D(\text{オレンジ}) + \dots + D(\text{黒}) \dots (2)$$

$$R1 = \left(1 - \frac{S}{200}\right) \times 100 \dots (3)$$

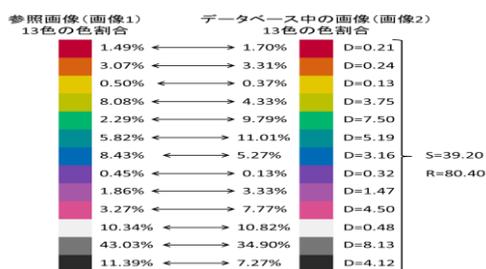


図3 方法1の計算例

方法2: 代表色の色割合を使った類似度分析方法を提案する。代表色の決定方法は13色の色割合を降順にランキングし、その上位から数えて色割合の80%に含まれる色を抽出することである。求めたSを式(4)より正規化処理を行う。以下図4に方法2の計算例を示す。

$$R2 = \left(1 - \frac{S}{\text{代表色の色割合} \times 2}\right) \times 100 \dots (4)$$

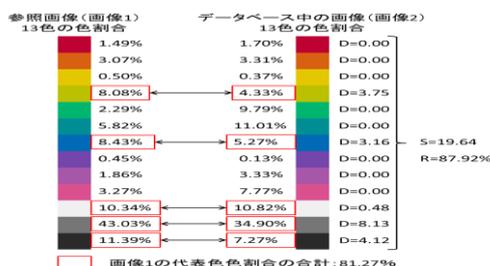


図4 方法2の計算例

方法3: 人間は代表色に加えて、画像の中で暖色がある部分に注目しているという考え方から、暖色の色割合を重視した類似度分析方法を提案する。暖色につ

ける重要度 α は暖色の面積比によって決定する。求めたSは式(5)より正規化処理を行う。重要度 α の面積比からの計算方法については、考察に述べる。以下図5に方法3の計算例を示す。

$$R3 = \left(1 - \frac{S}{(\text{他色の色割合} + \text{暖色の色割合} \times \alpha) \times 2}\right) \times 100 \dots (5)$$

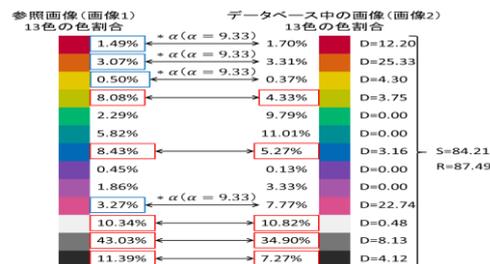


図5 方法3の計算例

3. 実験

3.1. 実験環境

今回、アンケート用として、20枚の参照画像はを画像カテゴリーより、「自然」11枚、「物」6枚、「人」3枚を選択した。そして、画像データベースには13,134枚の画像を用意した。以下図6にデータベース中の画像の一部を示す。



図6 データベース中の画像の一部

3.2. 実験方法

20枚の参照画像、3種類の画像類似度分析方法、10人を対象に、アンケートを行った。図7に示すように一枚の参照画像に対して、類似度分析の結果最も類似度が高いものから10枚の画像を、そのランキングを示さずに提示する。そして4段階評価を設定した。それは、①「とても似ている」②「どちらかというと似ている」③「どちらかというと似ていない」④「似ていない」である。各画像について①②を選択した人数を正しい類似結果であるとして集計し、画像の全枚数に対する①②の選択枚数の割合を計算し、チェックした枚数の割合とする。このチェックした枚数の割合が多いほど、正しい類似画像が検索されたとみなせる。図7に図1に示した実験用の参照画像についての類似画

像検索結果例を示す。

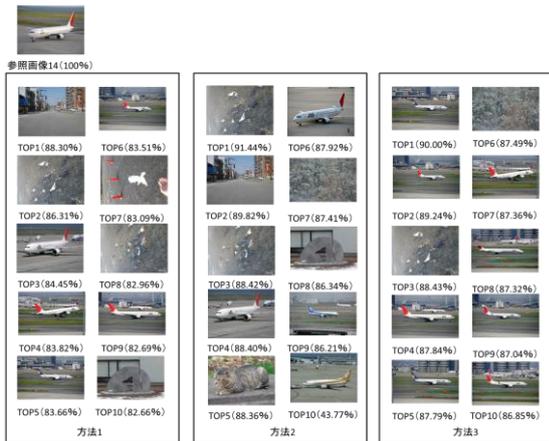


図 7 類似画像検索結果例

4. 評価

4.1. 方法 1 による類似度計算の評価

図 8 に方法 1 によるアンケート結果を示す。これより、方法 1 の場合では 20 枚の参照画像に対して「自然」、「物」、「人」3つのカテゴリの中では、「自然」のカテゴリが一番多く類似画像が検索されたことが分かった。

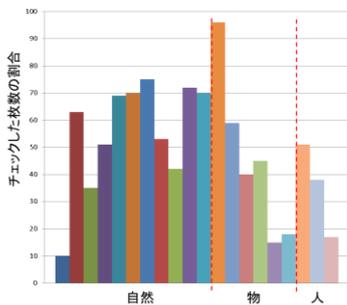


図 8 方法 1 の検索結果

4.2. 方法 2 による類似度計算の評価

図 9 に方法 2 の結果を示す。これより、方法 2 の場合では、20 枚の参照画像に対して、「物」のカテゴリが一番多く類似画像が検索されたことが分かった。

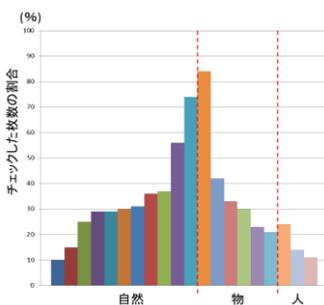


図 9 方法 2 の検索結果

4.3. 方法 3 による類似度計算の評価

図 10 に方法 3 の結果を示す、方法 3 の場合では、20 枚の参照画像に対して、「自然」と「物」の画像のカテゴリで多くの類似画像が検索されたことが分かった。

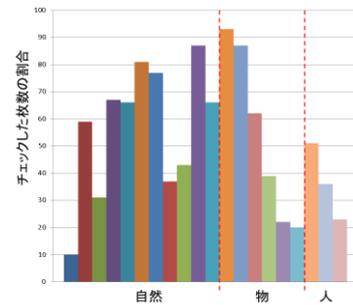


図 10 方法 3 の検索結果

5. 考察

本研究では、方法 1 の画像のすべての基本色の色割合を使った類似度分析方法を基準として、実験を行った。方法 2 では、人間がある画像を見る時に頭の中に残った印象はその画像の全体の色割合が大きい部分であると仮定している。今回の提案で色割合が大きい部分の情報抽出方法は、まず画像を 13 種類の基本色に減色し、各基本色の割合を求める。そして、13 種類の色割合より降順でランキングを作成する。最後に、上位から数えて色割合の 80% に含まれる色を抽出することである。図 11 に方法 2 の場合で類似度評価が一番高かった参照画像を示す。また、図 7 に示した結果は類似度評価が一番低いことが分かった。

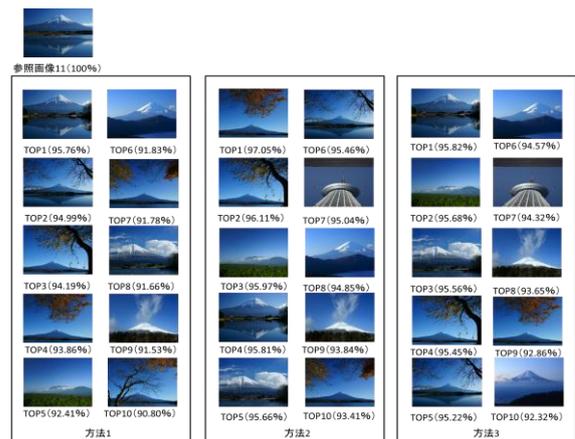


図 11 方法 2 で高評価であった参照画像に対する検索結果

図 11 に示す参照画像の 13 種類の基本色の色割合を図 12 の左に示す。この参照画像の代表色の色割合に含まれる色は、青 40.57%、黒 26.94%、青紫 26.82%である。そこで代表色はこの 3 色となり、その割合の合計

は 94.33%である。図 7 に示す参照画像の 13 種類の基本色の色割合を図 12 の右に示す。図により、代表色の色割合に含まれる色は、灰色 43.03%、黒 11.39%、白 10.34%、青 8.43%、黄緑 8.08%、青緑 5.82% の 6 色である。そして代表色の割合の合計は 81.27%である。以上のことより、代表色に含まれる基本色が少ない時に方法 2 の効果があり、画像の全体の印象が表現できていることが分かった。

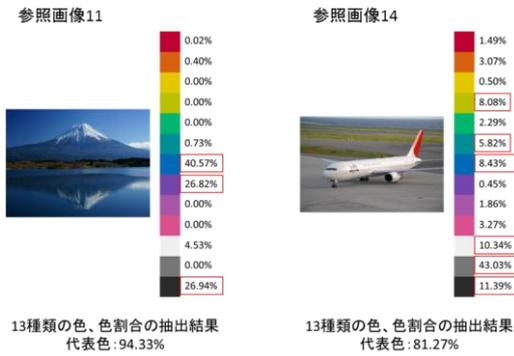


図 12 各参照画像の色割合

方法 3 では人間は代表色に加えて、画像の中で暖色がある部分に注目するという考え方から、暖色の色割合に重要度をつける類似度分析方法を提案した。重要度の推測方法を以下の図 13 に示す。

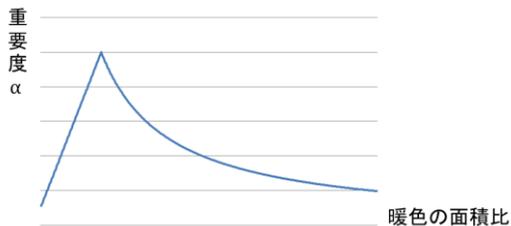


図 13 参照画像の暖色の面積比と重要度の関係

本研究において、重要度は、参照画像の中の暖色の面積比によって決定した。今回の実験では、暖色の色割合が 10%前後になった時に、暖色の重要度 α が 10 倍となるように設定した。これは事前に行った予備実験結果を基に決定した。以下に α の計算式を示す。

$$\alpha = 1 + \min\left(C_{\text{暖色}}, \frac{C_{\text{他}}}{C_{\text{暖色}}}\right) \dots (6)$$

$C_{\text{他}}$ = (代表色であって、暖色でない色) の割合

類似度分析の精度を高めるためには、ここに挙げた重要度の計算式について更に詳細に検討することが必要である。

6. おわりに

今回の実験の全ての方法の結果を図 14 に示す。今

回提案した 3 方法の中では、方法 3 の暖色を重視した画像類似度分析方法の総合評価が一番高い結果となった。図 15 に、方法 3 において計算対象となった色数を示す。方法 3 は全ての色を用いる方法 1 に比べて、約半分の色数で計算されており、同様に約半分の計算量で処理可能であるといえる。しかし、今回の検証実験では、限定された画像データベースと少数の参照画像による評価であるため、更にデータベースの数を増やした、普遍的な評価を行う必要がある。

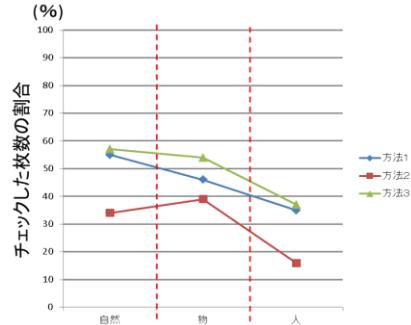


図 14 3 方法の比較結果

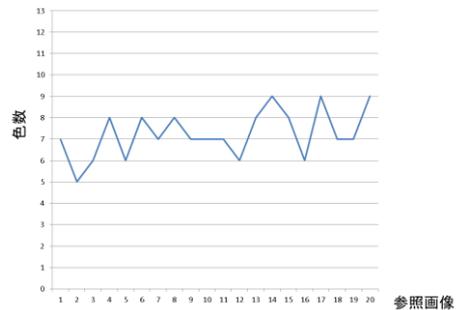


図 15 方法 3 における計算対象となった色数

文 献

- [1] 長谷山 美紀：“画像・映像意味理解の現状と検索インタフェース”，電子情報通信学会誌，Vol.93.No.9.2010
- [2] 大山 正・斉藤美穂：“色彩学入門 色と感性の心理”，東京大学出版会
- [3] “色彩用語事典”，日本色彩学会，東京大学出版会
- [4] JOHN CANNY：“A Computational Approach to Edge Detection”，IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL PAMI-8, NO.6 (NOVEMBER 1986)
- [5] 青木 功介ほか：“農作物育成診断のための均等色空間小領域における色補正手段”，画像電子学会誌 Vol.37 No.5 2008 年 9 月
- [6] A.Mochizuki, et al.：“An Image Retrieval Method for Video Database”，International Workshop on Image Technology 2003 (IWAIT 2003) Jan.21-22(2003)
- [7] 望月 綾音ほか：“映像データベースにおけるイメージ情報検索技術の研究”，富山県立大学，修士論文 2004