

## カメラ付き携帯電話を用いた MPEG-7 による電子図鑑システムの検討 A Study of MPEG-7 Image Retrieval System with Camera Phone

工藤 大樹† 関口 俊一† 山田 悦久† 西川 博文† 加藤 嘉明†

Daiki Kudo Shun-ichi Sekiguchi Yoshihisa Yamada Hirofumi Nishikawa Yoshiaki Kato

### 1. はじめに

近年、デジタルカメラやカメラ付き携帯電話が急速に普及している。同時に、デジタルカメラ用ストレージデバイスも大容量化し、大量の画像、あるいは長時間の動画を個人ユーザーが撮影することが可能となった。さらには、その蓄積先である PC においても数百 GB の大容量 HDD が搭載され、個人がマルチメディアコンテンツライブラリを作成、所有できる時代となりつつある。

その一方で、実際に視聴したいコンテンツや、もう一度見たい写真の保存場所は、ユーザーの記憶や、ファイル名、撮影した日付などに頼らざるを得ない状況であり、マルチメディアコンテンツ検索技術の登場が囑望されている。

著者らは、マルチメディアコンテンツ記述の国際標準規格 MPEG-7 のビジュアル記述子[1]を用いた画像の検索に関する研究を行い[2], [2]に基づいて MPEG-7 による電子図鑑システムを試作した。本論文では、試作したシステムの概要、並びに同システムにおける画像検索の高速化・高精度化に関して検討を行ったので報告する。

### 2. 試作したシステム

ここで、今回試作したカメラ付き携帯電話を用いた MPEG-7 利用電子図鑑システムの概要を図 1 に示す。まず、カメラ付き携帯電話で花画像を撮影する。次に、撮影した花画像を赤外線通信を用いて検索サーバーに送信する。検索サーバーでは画像の MPEG-7 特徴量を抽出し、特徴量データベース中の参照画像の特徴量と類似度計算を行う。類似度が高い画像から、検索結果の Compact html を作成する。さらに、作成した Compact html の URL を記したメールを撮影した携帯電話のメールアドレスに送信する。携帯電話側では受信したメールからサーバーの URL にアクセスし、検索結果を得る。

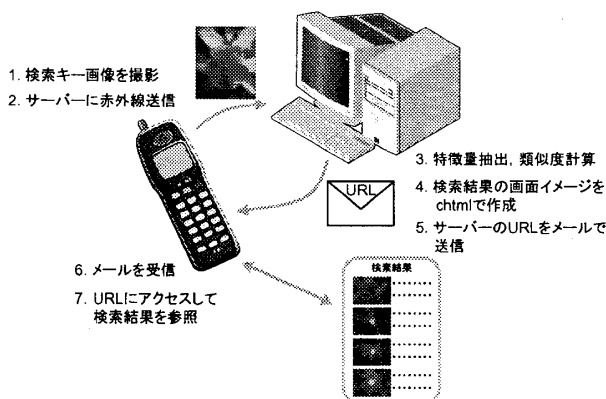


図1 システムの概要

携帯電話を用いて画像検索[3][4]を行うことで、様々な画像検索 Web サービスが考えられる。また、花画像の検索[4][5]に関しては、教育用途向け電子図鑑などのアプリケーションが考えられる。

### 3. 試作システムにおける類似度計算方法と課題

試作したシステムにおける類似度計算アルゴリズムは[2]に準ずる。使用する MPEG-7 のビジュアル記述子は Dominant Color, Color Layout, Edge Histogram の 3 種である。まず、Color Layout を抽出し、それから被写体領域の推測を行う。ここで、推測した被写体領域のみから Dominant Color を抽出することで特徴量抽出の高速化・検索精度の向上を図っている。さらに被写体領域の情報から Edge Histogram の類似度計算時に重み付けを行い、検索精度を向上させた。この方法により、画像の全領域から Dominant Color を抽出する場合に比べ、計算時間が平均約 20%短縮したほか、検索精度も向上した[2]。その一方で、この方法には以下の問題点が挙げられる。

- (1) Dominant Color 抽出高速化の効果が画像に依存する。
- (2) 大まかな被写体抽出しかできない。
- (3) Edge Histogram は被写体以外の部分も抽出・類似度計算の対象としている。

(1)に関しては、計算負荷が十分小さい場合、携帯電話やデジタルカメラ内で特徴量の抽出を行うことができる。前述の通り、現行のシステムでは画像データを検索サーバーに送信しており、通信データ量がやや大きい。画像データではなく、画像特徴量のみを検索サーバーに送信することで、通信データ量の大幅な削減が可能となる。さらには、携帯電話やデジタルカメラ内で、内臓メモリから画像検索も可能となる。

(2), (3)は、検索精度に大きく関わる問題である。携帯電話の画面サイズや通信帯域を考えた場合、検索結果の Compact HTML に掲載する情報は限られており、サムネイルを用いても多くの画像を載せることはできない。そのため、検索精度を高める必要があり、試作したシステムで用いた検索アルゴリズム[2]の検索精度をさらに向上させる必要がある。

### 4. Dominant Color 抽出の高速化

#### 4.1 高速化方法

まず、課題(1)Dominant Color 抽出の高速化に関して検討を行った。Dominant Color の抽出には GLA アルゴリズムによるクラスタリングが用いられることが多い[6]。クラスタリング処理は一般に計算負荷が重く、本システムにおいても特徴量抽出に要する計算時間の大部分を占めている。この Dominant Color 抽出の高速化が、携帯電話・デジタルカメラ内部での特徴量抽出・類似度計算を目指す場合は大きな課題となる。

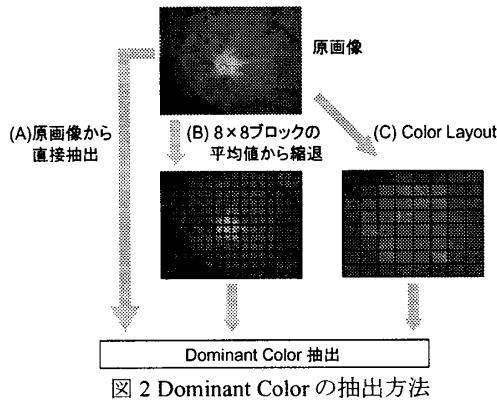


表1 Dominant Color抽出時間

	(A)画面全体から	(B)縮退画像を利用	(C)Color Layoutを利用
平均処理時間(VGA, msec)	1598	29.5	0.838
(A)を1とした場合の比	1	1/53	1/1900

高速化の方法を図2に示す。画像を8×8のブロックに分割し、ブロックごとに画素の平均値を用いて縮退させた画像からDominant Colorの抽出を行う。これにより、画素数が  $1/8^2 = 1/64$  になり、計算時間の短縮が期待できる。JPEG、MPEGなどの一般的な画像圧縮においては、画像を8×8のブロックに分割し、DCTを行う方法が用いられている。DCTの直流成分は各ブロックの平均値であり、これを直接利用することもできる。さらに、画像全体を8×8に分割した平均色であるColor Layoutの8×8 pixelsの画像からDominant Colorを抽出する方法も試み、比較を行った。これらの方法は携帯電話の限られた処理能力を使う上で有効といえる。以降、画像全体からDominant Colorを抽出する場合を(A)、縮退画像から抽出する場合を(B)、Color Layoutから抽出する場合を(C)とする。

#### 4.2 高速化の効果と特徴量に与える影響

実際の計算時間を表1に示す。VGAサイズの画像30枚を処理した際の1枚当たりの平均時間である。表1の結果はDominant Color抽出の時間のみで、Color Layoutの抽出時間や、縮退画像の生成時間は含まない。なお、計算時間の計測にはPentium 4 (2GHz)搭載のPCを用いた。表1から、(A)に比べ、(B)では計算時間が1/53、(C)では1/1900と大幅に短縮されているのがわかる。

画像サイズから、画素数比は、  
 (A)  $640 \times 480 = 307200(\text{Pixels})$   
 (B)  $640 / 8 \times 480 / 8 = 4800(\text{Pixels})$   
 (C)  $8 \times 8 = 64(\text{Pixels})$   
 (A) : (B) : (C) = 1 : 1/64 : 1/4800

となる。(A)、(B)に関しては、おおよそこの画素数比に比例した結果である。(C)に関しては、画素数比に比べ長い時間がかかっている。これは、(C)においては処理対象の画素数が非常に少なく、クラスタリング処理以外の時間が無視できないためであると考えられる。

現行の携帯電話に搭載のCPUの動作周波数を100MHz、Pentium4(2Ghz)との速度比を1/20と仮定した場合、  
 (B) Pentium4(2Ghz) : 29.5msec → 携帯電話CPU : 590msec  
 (C) Pentium4(2Ghz) : 0.838msec → 携帯電話CPU : 17msec

表2 Dominant Color抽出結果

	(A)		(B)		(C)	
	[R G B]	(%)	[R G B]	(%)	[R G B]	(%)
代表色1	*1[247 42 91]	32%	*1[247 42 91]	32%	*1[247 33 83]	23%
代表色2	*2[75 116 50]	23%	*2[75 116 50]	23%	*2[66 116 50]	10%
代表色3	*3[33 66 17]	19%	*3[33 66 17]	23%	*3[33 66 17]	32%
代表色4	*4[9 25 1]	13%	*4[17 25 1]	13%	[116 66 50]	6%
代表色5	*5[190 33 75]	3%	*5[198 33 75]	6%	*5[198 42 83]	6%
代表色6	*6[91 25 25]	3%未満	*6[99 42 33]	3%	[157 75 66]	6%
代表色7					[50 25 9]	3%

と、(B)においても1秒以下となる。

次に、実際に抽出したDominant Colorに関して述べる。表2に実際に抽出した代表色を示す。表2では、RGBの各値と、その割合を示している。また、同一、もしくは非常に近い色を\*と番号で示す。なお、抽出元の画像は赤い花と緑の葉の画像である。ここでは、画面全体から抽出した(A)を基準に考える。(A)と(B)では、抽出した代表色、割合の一部に若干の差が見られるものの、大きな違いはない。従って、類似度計算を行う場合においても大きな問題にはならないと考えられる。一方で、(C)の場合は、(A)に比べ、代表色・割合ともに若干の差異が見られる。特に、(A)にある代表色に相当する代表色が一部なく、逆に(A)・(B)に見られない代表色を抽出しており、類似度計算への影響も考えられる。これは、大きなブロックで色の平均化を行っている影響と考えられる。このことから、正確な代表色を求める必要がある場合、画面全体からクラスタリングを行って抽出することが望ましい。しかし、1/64程度の縮退であれば大きな問題はなく、十分正確な代表色を抽出可能と考えられる。

### 5. 被写体領域抽出改善による検索精度向上

#### 5.1 被写体抽出

次に、課題(2)の被写体領域の抽出に関して検討を行った。花画像の検索において、被写体である「花」の領域を抽出することは検索精度に大きな影響を与える[2][4][5]。Color Layoutからの抽出[2]では、大まかな被写体領域の特定しかできない。また、ユーザー入力によって花画像の特徴を学習させる方法[5]もあるが、全て自動で行うことが望ましい。そこで、Dominant Colorとして抽出された代表色で原画像の画素を置換することで減色した画像を用いて被写体領域の特定を行った。まず、原画像からDominant Colorを抽出する。次に、抽出したDominant Colorを用いて、原画像の画素を置換することで、減色処理を行う。そして、画素置換により減色した画像から、被写体の判定を行う。判定の方法としては、減色した画像において、中心付近に多く存在する色と外側に多く存在する色の割合から被写体色を判定する。被写体色と判定された画素の領域を仮の被写体領域とする。仮の被写体領域から、被写体色の最も多い部分を被写体の中心とし、そこから被写体色が連続している部分を被写体領域とする。実際の例を図3に示す。

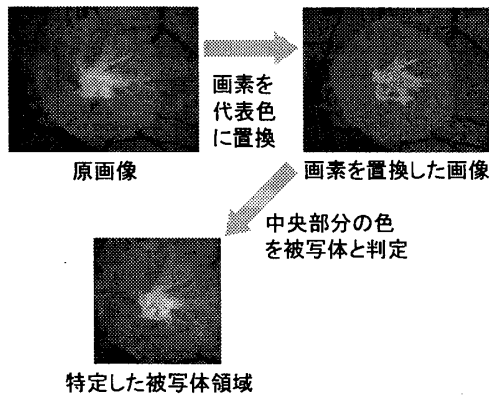


図3 被写体抽出例

表3 被写体抽出の評価

	Dominant Color抽出方法		
	(A)	(B)	(C)
効果的な被写体抽出ができた画像の割合	76.7%	75.3%	75.3%
被写体抽出ができたが、不十分な画像の割合	12.7%	14.1%	14.1%
被写体抽出ができたなかった画像の割合	11.3%	11.3%	11.3%

この被写体抽出法と、前述の Dominant Color の高速抽出と組み合わせることで、高速かつ正確な被写体領域の抽出が可能である。そこで、高速抽出による Dominant Color 特徴量の差と、特定する被写体領域の違いを調べるため、(A)、(B)、(C)それぞれの方法で抽出した Dominant Color を用いて被写体抽出を行っている。73 枚の画像から被写体抽出した結果の主観評価を行った。抽出した結果を、

- (1) 効果的な被写体抽出ができた
- (2) 被写体抽出ができたが、不十分である
- (3) 被写体抽出ができなかった

に分類し、それぞれの割合（該当する画像数／全体の画像数）を調べた。結果を表3に示す。表3からわかる通り、およそ75%の画像で効果的な被写体抽出を行うことができた。また、(B)もしくは(C)による高速抽出を行った Dominant Color を用いた場合においても、最終的に特定した被写体領域は、ほとんどの画像で差は見られなかった。このことから、被写体抽出のために Dominant Color を用いる場合においては、(B)もしくは(C)の高速抽出を用いても問題ないといえる。

## 5.2 被写体抽出結果を用いた画像検索

抽出した被写体領域を用いて課題(3)の Edge Histogram の抽出領域改善を行い、画像検索を行った結果を示す。まず、検索の条件を示す。カメラ付き携帯電話で撮影した画像を検索キー画像とする。データベース中には正解となる画像を1枚用意する。この正解画像の順位で評価を行った。なお、キー画像は41枚、データベース中には532枚の画像を用意した。なお、検索結果の比較のために、

- ① 被写体抽出を用いずに画面全体から Edge Histogram を抽出、類似度計算。
- ② 被写体抽出を用いずに画面全体から Edge Histogram を抽出。Color Layout による被写体抽出[2]で重み付けを行って Edge Histogram の類似度計算。

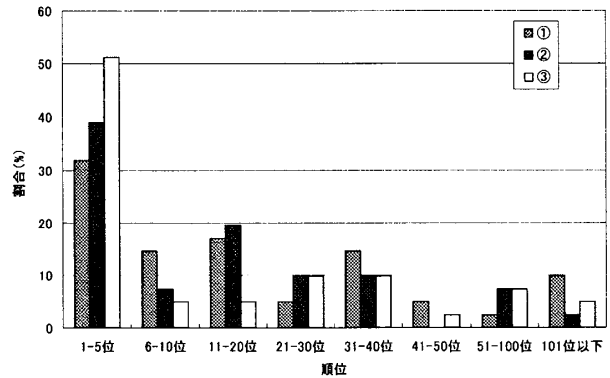


図4 Edge Histogram 検索結果

- ③ 本論文で提案した Dominant Color を用いた方法で被写体抽出を行い、被写体領域のみから Edge Histogram を抽出し、類似度計算。

の3つの方法で検索を行い、それぞれの検索結果を比較する。携帯電話での通信・画面表示を考えた場合、1-5位以内に正解画像の順位が入ることが一つの目標となる。

図4に正解画像の順位を示す。図4の横軸は順位、縦軸はその順位に入る正解画像の割合である。高い順位に多くの正解画像が入るほど良い。図4から、③の検索方法において、半数の画像で正解画像が1-5位以内に入ることがわかる。このように、本論文による③の方法が最も検索精度が高いことがわかる。同様に、被写体抽出を他の特徴量による類似度計算においても用いることで、より高精度の画像検索が可能であると思われる。

## 6. まとめ

カメラ付き携帯電話と検索サーバーによる MPEG-7 ビジュアル記述子を用いた画像検索システムを試作した。システムの課題として、高速化による携帯電話内での特徴量抽出をとりあげ、Dominant Color 抽出の高速化を行った。縮退画像・Color Layout を用いた Dominant Color の高速抽出とその速度、さらには原画像からの直接抽出との違いの評価を行った。その結果、抽出時間を大幅に短縮することができた。また、抽出結果も大きな差はないことを確認した。また、検索精度に関しても目標として、正解画像が類似度の上位1-5位以内に入ることを目指し、改善を検討した。Dominant Color による画素置換を用いた被写体抽出を行い、75%の画像で良好な被写体抽出結果を得た。得られた被写体抽出結果から、Edge Histogram を用いた類似度計算を行い、半数の画像で目標である正解画像が1-5位以内を達成した。今後、さらなる検索速度・検索精度の向上を図る。

## 参考文献

- [1] ISO/IEC 15938-3 MPEG-7 Part-3
- [2] 工藤 他, 電子情報通信学会2004年総合大会, D-11-134
- [3] 曾根 他, 情報処理学会第66回全国大会(2004), 3Z-9
- [4] 関口 他, 映像情報メディア学会2000年年次大会, 12-3
- [5] 岡村 他, 情報処理学会第66回全国大会(2004), 3Z-3
- [6] ISO/IEC 15938-8 MPEG-7 Part-8