

平成21年度 卒業論文

論文題目

自由度のあるパネル型クエリ生成  
インタフェースを用いた画像検索システム

神奈川大学 工学部 電子情報フロンティア学科

学籍番号 200602827

小林 恭之

指導担当者 木下 宏揚 教授

# 目次

第1章	序論	4
第2章	基礎知識	6
2.1	代表的画像検索技術	6
2.1.1	TBIR	6
2.1.2	CBIR	8
2.2	画像処理技術	10
2.2.1	2値化	10
2.2.2	グレースケール	10
2.2.3	ガンマ補正	12
2.3	画像検索の流れ	13
第3章	パネル型クエリ生成インタフェースとその改良	14
3.1	パネル型クエリ生成インタフェースとその問題点	14
3.1.1	ユーザビリティを考慮したパネル分割数	15
3.1.2	平均色を用いた類似度検索	15
3.1.3	問題点	16
3.2	透明パネルの提案	17
第4章	改良インタフェースと評価	18
4.1	改良インタフェース	18
4.2	評価	19
4.2.1	従来システムでの検索結果	20
4.2.2	改良システムでの検索結果	21
4.3	今後の課題	22
第5章	結論	23
	謝辞	24

参考文献	24
質疑応答	26

# 目 次

2.1	画像検索エンジン GazoPa	9
2.2	画像の2値化	10
2.3	中間値法スペクトル	11
2.4	単純平均法スペクトル	11
2.5	加重平均法スペクトル	11
2.6	画像検索の流れ	13
3.1	パネル型クエリ生成インタフェース	14
4.1	改良インタフェース	18
4.2	従来システムでの検索結果	20
4.3	改良したシステムでの検索結果	21

# 第1章 序論

近年，デジタルカメラの普及によって個人でも多くの画像を所有することが多くなってきている。また，インターネット上にも多くの画像データが存在している。そのため，蓄積されていく多くの画像データの中から，ユーザの所望する画像データをより効率よく検索する技術が重要となっている。

一般に画像検索は，テキストに基づく画像検索（以下，TBIR:Text-Based Image Retrieval）[1]と，画像の内容に基づく画像検索（以下，CBIR:Content-Based Image Retrieval）[2][3][10]に分けられる。TBIRはキーワードをメタデータとして検索を行うのに対して，CBIRは画像の色や形，テクスチャ，構図等をメタデータとして検索を行う。

TBIRでは，キーワードをクエリとするため人の感覚に近い検索ができる。しかしながら，所望の画像に付与されたメタデータと検索するユーザのクエリが必ずしも一致するとは限らず，ユーザは所望する画像を得ることが困難となる。また，人が手動であらかじめ画像にメタデータを付与しなければならないため，膨大な数の画像データが蓄積されたデータベースにおいては非現実的である。

一方，CBIRは画像データの色や形を計算機によって自動的に算出できるため，画像に対するメタデータの付与が比較的容易に行える。クエリ生成手法としては，思い描いている所望の画像をユーザ自身で生成する方法が提案されている。しかしながら，クエリ生成方法によっては，クエリの生成手順が必要以上に多いとユーザの負担になる，検索するユーザが思い描いている所望の画像を適切にクエリとして生成できない，ユーザが生成したクエリからユーザの検索意図をうまく汲み取ることができないといった問題点がある。

スケッチによるクエリ生成手法では，ユーザがキャンパスに自由にクエリを生成することができる。文献において，ユーザはスケッチしたクエリに空，草原などのオブジェクト情報を組み合わせることにより，人の思考に合ったクエリを生成することが可能である。概略図によるクエリ

生成手法では，ユーザが簡単な図形を組み合わせて所望の画像をある程度大まかに描くことにより，ユーザの負担を軽減している．

しかしながら，これらのクエリ生成手法を用いる検索システムでは，ユーザによって生成されたクエリから色情報，形状，構図等の特徴量を抽出し，多数の特徴量からなるメタデータとマッチングを行うため，ユーザの検索意図がうまく検索システムに伝わらず検索精度が低下してしまう可能性がある．文献では，ユーザがクエリ生成の際に色，面積，位置，形状，配置といった特徴に重み付けをすることでユーザの検索意図を明確にし検索精度の向上を図っているが，ユーザ自身がどこにどの程度の重み付けをしていいかわからない場合がある．したがって，ユーザが思い描いている所望の画像を負担なく適切にクエリとして生成できることが望ましい．

過去に研究室でパネル型クエリ生成インタフェースによる画像検索システムが開発された．しかし，従来のパネル型クエリ生成インタフェースでは，何色でもよい場合というパネルが用意されていない．よって，ユーザの検索意図をうまく汲み取れなかったり，制限されたりしてしまう．本稿では，パネル型クエリ生成インタフェースを改良し，何色でも良い場合のパネルを用意することで，よりユーザの検索意図を汲み取りやすくする検索システムを提案する．

## 第2章 基礎知識

### 2.1 代表的画像検索技術

大量の画像データに対応するため、蓄積された多数の画像データに対して、ユーザの所望するデータをより早く、効率的に取り出せる画像検索技術の実用化が必要となる。現在の検索方式には代表的なものとして、テキストをメタデータとした検索方法であるTBIR、画像の内容をメタデータとした検索方法であるCBIRがある。

#### 2.1.1 TBIR

TBIRはデータベース内の画像データに、画像のタイトルや製作者、内容を形容するキーワードをメタデータとして画像に付与し、それに基づいてユーザの求めるキーワードとのマッチングにより検索を行う方法である。検索範囲がタイトルや製作者といった一様に統一された分類内容の検索に適しており、静止画像検索に限らず、動画像検索においても主流である。例として、電子図書館における蔵書検索が挙げられる。蔵書検索を行うに当たって、所望する蔵書の名称や特徴を定義したものを検索情報として扱う。これらは多様な情報資源を表す共通な要素として基本15エレメント[4]から構成される。

1. タイトル:情報資源に与えられた名前
2. 作成者:情報資源の内容の作成に主たる責任を持つ実体
3. キーワード:情報資源の内容のトピック
4. 内容記述:情報資源の内容の記述
5. 公開者:情報資源を利用可能にすることに対して責任を持つ実体

6. 寄与者:情報資源の内容への寄与に対して責任を持つ実体
7. 日付:情報資源のライフサイクルにおける何らかの自称に対して関連付けられた日付
8. 資源タイプ:情報資源の内容の性質もしくはジャンル
9. 記録形式:物理的表現形式ないしデジタル形式での表現形式
10. 資源識別子:与えられた環境において一意に定まる情報資源に対する参照
11. 出处:現在の情報資源が作り出される源になった情報資源への参照
12. 言語:当該情報資源の内容の言語
13. 関係:関連情報資源への参照
14. 時空間範囲:情報資源の内容が表す範囲あるいは領域
15. 権利管理:情報資源に含まれる,ないしは関わる権利に関する情報

この基本15エレメントはデータベースの内容に準じて項目に加減が施されて,神奈川大学図書館の蔵書検索[5]にもこれらの項目が見受けられる.通常,Web上の画像データはHTML文書からリンクがなされており,文書には画像の内容を説明するキーワードが含まれている場合が多いため,画像データをリンクしているHTML文書を解析することでTBIRによるWeb上の画像検索を行うことが可能である.

しかし,TBIRには次のような問題点がある.

画像内容考慮 付加されるキーワードは画像の内容に依らず,HTML文書に含まれるキーワード検索しか行わないため,所望する画像データの他にそのキーワードの含まれる画像データも検索されてしまう.

手作業による索引付け 各画像データに索引を手作業で付加することになるため,膨大な画像データを取り扱う場合においてデータベース所有者への負担が増えてしまう.

追加画像の索引付加 未完成のデータベースに画像データが追加される場合,追加された画像データにおいても索引付けを行わなければならない,更なる負担が伴う.



索引の主観性 データベース所有者の主観によってキーワードの付加が行われるため，ユーザーがデータベース所有者の主観と一致しない場合，効果的な検索を行うことができない．

コスト 大規模なデータベースを管理する場合，上記負担を軽減するための管理費，人件費を考えるとコストがかさんでしまう．

既に膨大な画像データが存在し，増え続けている今日の情報化社会において，これらの問題は致命的であり，所望する画像データを効率的に入手するための画像検索技術としては満足のできるものではない．

### 2.1.2 CBIR

CBIRは色や形，構図など，画像の特徴量を利用した検索方式である．前述したTBIRでは手作業によってあらかじめ全ての画像データにキーワードを付加する必要があったが，CBIRでは画像データに既に存在する画像特徴を計算機によって自動判別して画像間の類似度を判定し，ユーザーが指定した画像に類似した画像を検索することが可能である．また，画像データを視覚的に捉えることができ，よりユーザーの感覚に近い画像の検索が可能である．

用例として，日本発画像検索エンジン「GazoPa」[6]を挙げる．GazoPaの検索システムはユーザーから提供される画像のアップロードやネットサーフィン中に見かけた画像を検索に利用できる．所望する色や形に近い画像を見つけ，似たようなデザインをもっと探したい時に，その見つけた画像をそのまま検索画像として利用できるのである．更にFlashを利用したdrawerを使用することで，ユーザーが作成した画像を検索画像に転用できるため，ユーザーの想像を直に検索システムへ伝えることができる．



図 2.1: 画像検索エンジン GazoPa

しかし、例にも適用されるがCBIRには次のような問題点が挙げられる。

画像間のギャップ ユーザーの所望する想像上の画像と実際の画像の特徴の間にギャップが存在し、その二つが完全に一致することは非常に困難である。

検索速度の遅延 Webを用いるシステムの場合、不必要な画像のダウンロードを行って検索速度が著しく遅くなる場合がある。

## 2.2 画像処理技術

### 2.2.1 2値化

白と黒だけによる画像の表現処理 [7] である。各画素の明るさを一定の基準値により、白色と黒色の2つの値に変換する処理を行う。この一定の基準値を閾値という。閾値より小さい値を持つ画素を白、閾値より大きい値の画素を黒として表現する。通常、画像の各画素は0~255のRGB値を持っていて、RGB値の平均値が各画素における明るさとなる。2値化処理は、画素 $(x,y)$ の濃淡画像を $f(x,y)$ 、閾値を $T$ とすると、一般的には次項に従って処理される。

$$\begin{cases} \text{白色} & f(x,y) \leq T \\ \text{黒色} & f(x,y) > T \end{cases}$$

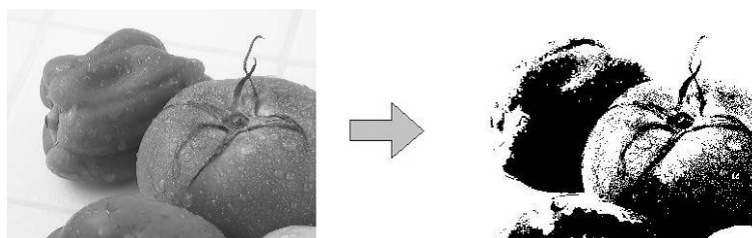


図 2.2: 画像の2値化

基準となる閾値をの値を変更することで2値化処理後の画像が異なり、2値化処理を行うことで画像からの検索情報の抽出が容易になり、また判定処理なども高速に実行できる。

### 2.2.2 グレースケール

2値化では白と黒の2色のみで表現したが、グレースケール [8] は画像を白から黒までの明暗だけで色の情報は含まずに表現する。灰色を何階調で表現するかをビット数によって表す。1ビットの場合は白と黒のみで中間色がない状態、つまり前述した2値化と同義となる。8ビットなら(白と黒を含めて)256階調、16ビットなら65536階調の灰色で表現する。

グレースケール処理はその計算方式によりいくつかの計算方法が存在する。

- 中間値法:出力画素=(最大成分+最小成分)/2

対象画素における最大最小成分の平均を画素値とする方法である。カラーパターンが灰色一色に変換され、それぞれの色の違いが全て潰れてしまっている。非常に直感的で簡単ではあるが、意図して使う以外は適切ではない。



図 2.3: 中間値法スペクトル

- 単純平均法:出力画素=(R成分+G成分+B成分)/3

対象画素のRGB各成分の平均値を画素値とする方法である。RGB成分の値によってグレーに濃淡ができてはいるが、異色である青と緑が処理後では同色であったりと、人間が色の違いによって感じる明るさの違いを無視してしまっている。



図 2.4: 単純平均法スペクトル

- 加重平均法:出力画素=(0.299\*R成分+0.587\*G成分+0.114\*B成分)

RGB各成分に重み付けを加え、視覚感度を考慮した計算方法。単純平均法で問題のあった青と緑にも処理後の差異が生じ、処理後の多色化が成される。



図 2.5: 加重平均法スペクトル

### 2.2.3 ガンマ補正

(ガンマ)とは画像の明るさの変化に対する電圧換算値の変化比である。画像の色データと画像が実際に出力される際の信号の相対関係を調節して、より自然に近い表示を取得し、元画像データに忠実な画像の表示を再現するための補正をガンマ補正[7]という。補正の数式を(2.1)に示す。

$$= 255 * ( /255)^{1/\gamma} \quad (2.1)$$

画像データと出力データとの値が正比例している時、 $\gamma$ 値は1になるが、実際には正比例しない。スキャナなどの入力装置やプリンタなどの出力装置はそれぞれ特有の $\gamma$ 値を持っているため、スキャナで入力した画像をそのまま出力してしまうと $\gamma$ 値が1に近づかないため色合いが違ってしまう。広義に解釈すると、現在のほぼ全ての画像データは特定の出力環境に合わせられて作られていると言ってよい。出力環境が変わることでその時出力した画像データは完全な状態で表示することができないということである。そのため、出力機器の違いによる差異を緩衝し、より正確な画像データ取得においてユーザーはガンマ補正を行う必要がある。

## 2.3 画像検索の流れ

画像検索の流れは、まずキーとなる画像データを用意する。キーとなる画像データとは、ユーザが所望する画像を検索するときの元となる画像データである。キーとなる画像データに画像処理を施し、検索の元となるメタデータを抽出する。また、検索される画像データベースの画像にも処理を施し、メタデータを抽出する。キーとなる画像のメタデータと、画像データベースのメタデータとの類似度を比較し、類似度の高い順に順位付けしたものを所望画像の類似画像群とする。画像検索システムによる画像の検索はここまでで完了である。最終的にはユーザの目による画像の確認によって、所望の画像を入手することができる。

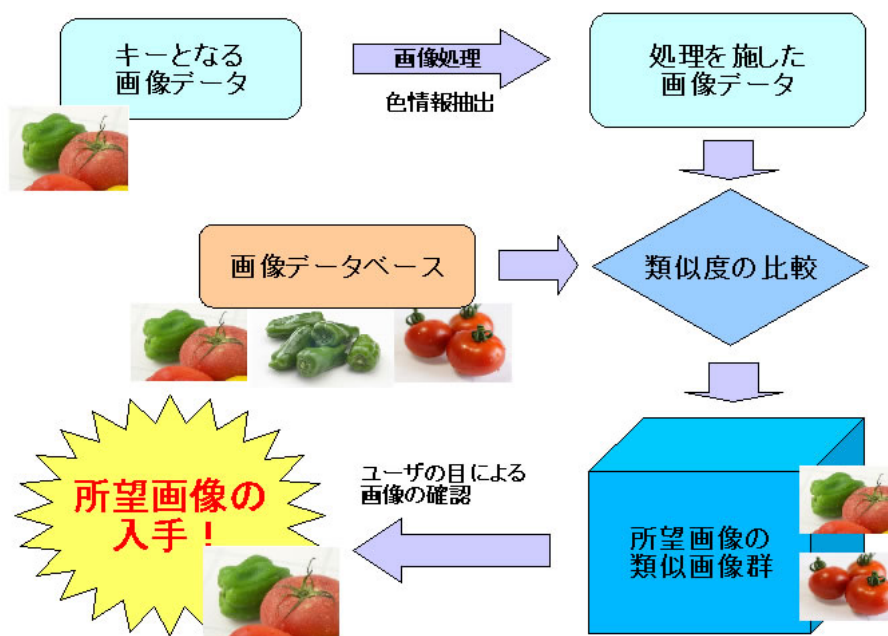


図 2.6: 画像検索の流れ

## 第3章 パネル型クエリ生成インタフェースとその改良

2006年に野沢ら [9] は、パネル型クエリ生成インタフェースを用いた画像検索システムを提案した。本章では、このパネル型クエリ生成インタフェースの問題点を述べ、その改良を提案する。

### 3.1 パネル型クエリ生成インタフェースとその問題点

パネル型クエリ生成インタフェースを用いた画像検索システムとは、図3.1のように、ユーザが検索したい画像の大まかな色をパネルに塗り、それをクエリとする検索システムである。本節では、野沢らが開発した画像検索システムを解説したのち、その問題点を述べる。

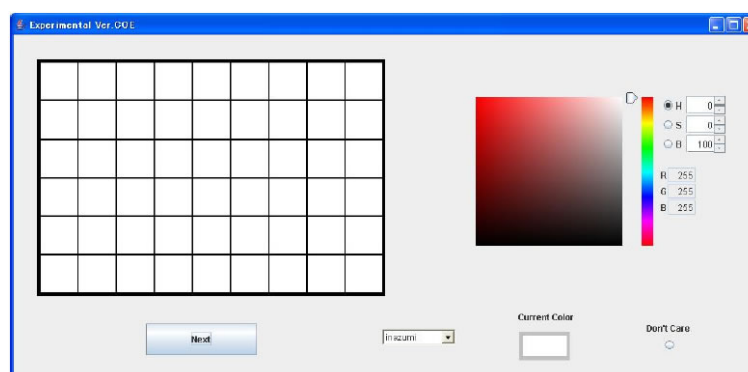


図 3.1: パネル型クエリ生成インタフェース

### 3.1.1 ユーザビリティを考慮したパネル分割数

$M \times N$ のパネルによって生成されたクエリは所望の画像と同じ縦横比であるため、パネル単位で独立に色を扱うこと、パネルの位置に基づいた色情報を扱うことを可能にしている。それゆえ、ユーザは思い描いている所望の画像のどこがどのような色であるか明確に検索システムに伝えることができる。

しかしながら、キャンパスの分割数が極端に多い場合や極端に少ない場合は、ユーザが思い描いた所望の画像を再現できない、ユーザのクエリ生成の際の負担になるといった問題点がある。そのため、ユーザの検索意図が明確に検索システムに伝わらず検索精度が低下してしまう可能性がある。したがって、ユーザが負担なく思い描いている所望の画像を適切にクエリとして生成できるパネル数が望ましい。

ここでは、過去の実験データの結果を基に最適だと判断されている $9 \times 6$ のパネルを使用している。

### 3.1.2 平均色を用いた類似度検索

パネル型クエリ生成インタフェースは $9 \times 6$ に分割したパネルを用いているため、マッチング処理をする際には検索対象となる画像も $9 \times 6$ に分割しパネル単位で独立に色情報を扱う。

データベースにおける全ての画像に対してメタデータを算出する。原画像を $9 \times 6$ に分割したそれぞれをブロックと定義する。メタデータは分割した1ブロックを原画像の対応する画素の平均色にしたものとする。

マッチング処理は、対応するクエリの各パネルと各ブロックでユークリッド距離を求めるものと定義し、類似度は算出されたユークリッド距離の平均とする。

また、マッチング処理では、色情報を基にしたものと構図情報を基にしたものがある。色情報を基にしたマッチング処理では、メタデータは平均色を用いているため、ユーザが印象に残った色を選択しても異なってしまう可能性がある。そこで、まず1ブロック当たり2つの代表色を候補メタデータとし算出する。構図情報に基づくマッチング処理では、パネルに対応する画像の境界部分において色差分を算出する。



### 3.1.3 問題点

このパネル型クエリ生成インタフェースでは、パネルの色の初期値は白となっており、たとえ色を何も塗らなかったとしても色を指定しない検索といったものを行えない。よって、ある部分は何色でも良い場合という検索をすることができないので、生成されるクエリがユーザの検索意図をうまく汲み取れない場合がある。

## 3.2 透明パネルの提案

3.1.3節で述べた問題点を解決するために、透明パネルを提案する。透明パネルとは、色を指定しないパネルのことであり、透明で塗ったパネルの部分はあらゆる色の場合でも検索されるようになっている。

例として、中心が赤く、周りは何色であってもいい国旗を検索したいとする。従来のパネル型クエリ生成インタフェースでは、中心を赤く塗っただけでは中心が赤く周りが白い画像しか検索できない。他の国旗を検索したい場合は、自分で周りの色を緑や黄色を指定しなくてはならない。透明パネルを用いた場合、中心を赤く、周りを透明で塗ることによって中心が赤、周りは白や緑、黄色などの場合の画像をすべて検索することができ、よりユーザの検索意図を汲み取りやすくなると考えられる。

## 第4章 改良インタフェースと評価

### 4.1 改良インタフェース

パネル型クエリ生成インタフェースを改良し、透明パネルを実装させた。透明パネルは任意の色すべての場合を検索できるパネルである。図4.1が改良したインタフェースである。図4.1では、すべてのパネルを透明パネルで塗っている。透明パネルは薄い灰青色で表示される。

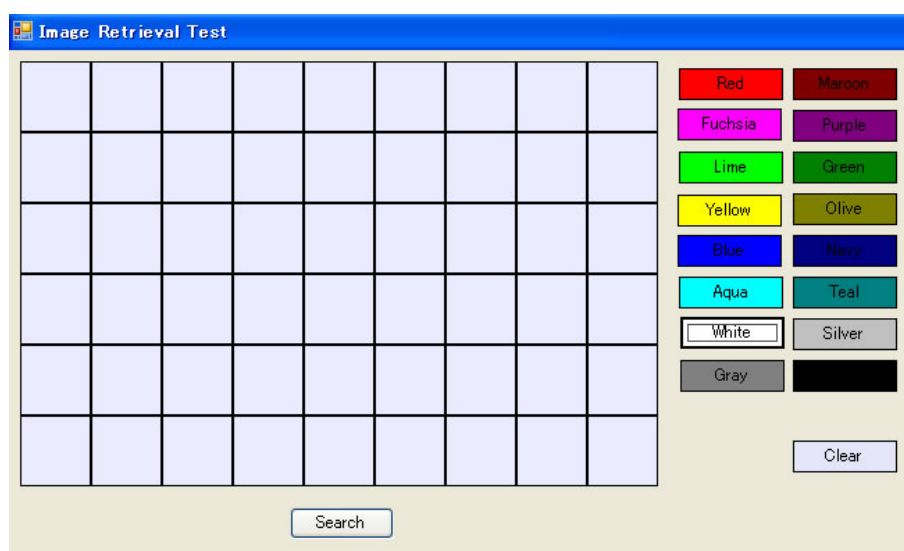


図 4.1: 改良インタフェース

## 4.2 評価

従来のパネル型クエリ生成インタフェースの検索結果と、改良したパネル型クエリ生成インタフェースの検索結果を比較する。比較するに当たって、クエリはどちらも中心のみを赤くぬることにする。中心以外のパネルは、従来のパネル型クエリ生成インタフェースでは初期値（白）、改良したパネル型クエリ生成インタフェースでは透明パネルで比較を行う。データベース内の画像は1240枚とし、上位4枚で比較を行う。

### 4.2.1 従来のシステムでの検索結果

従来のパネル型クエリ生成インタフェースでは、パネルを塗らなかった場合、初期値の白として検索される。検索結果を見てみると、中心の赤より周辺の白のほうがパネル数が多いため、それと類似した画像、つまり周辺が白い画像のほうが上位にきている。

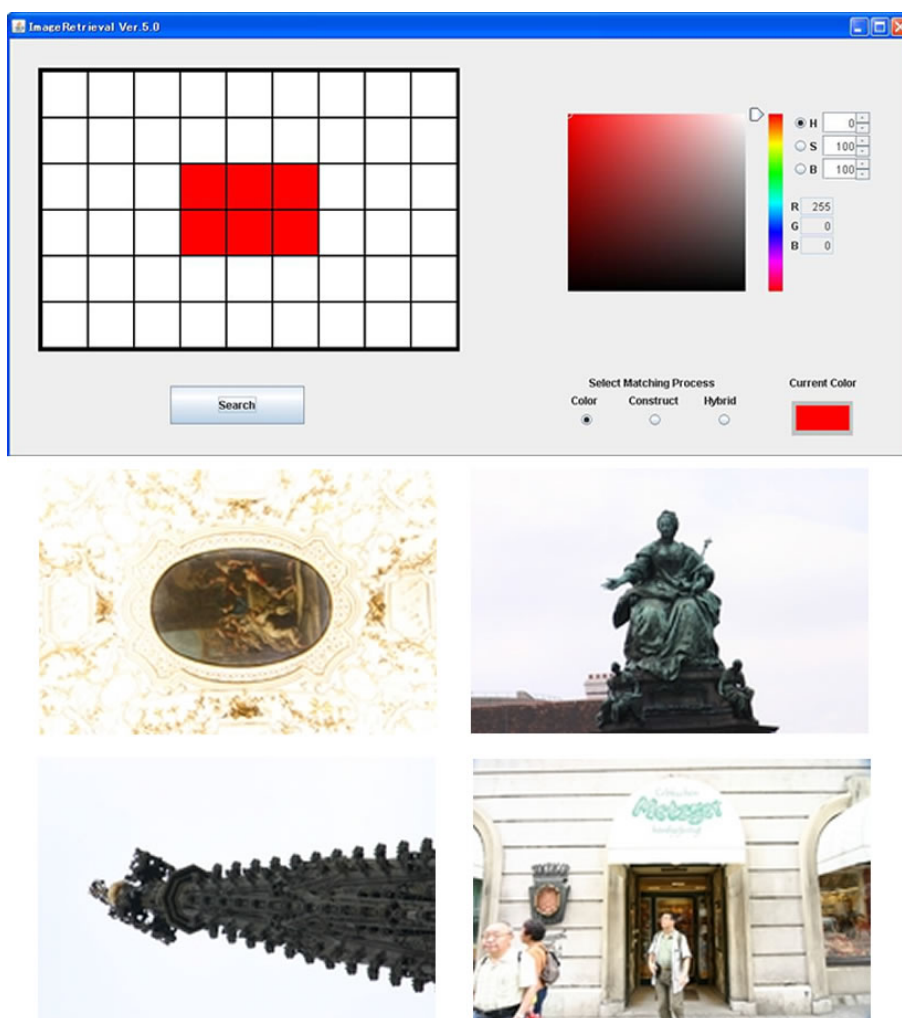


図 4.2: 従来のシステムでの検索結果

### 4.2.2 改良したシステムでの検索結果

改良したパネル型クエリ生成インタフェースでは、周辺は透明パネルで塗られているため、画像データベースのメタデータが何色でも検索される。検索結果を見てみると、中心が赤い画像が検索されている。周辺の色は何色でも良いため、様々な場合が検索されているのがわかる。

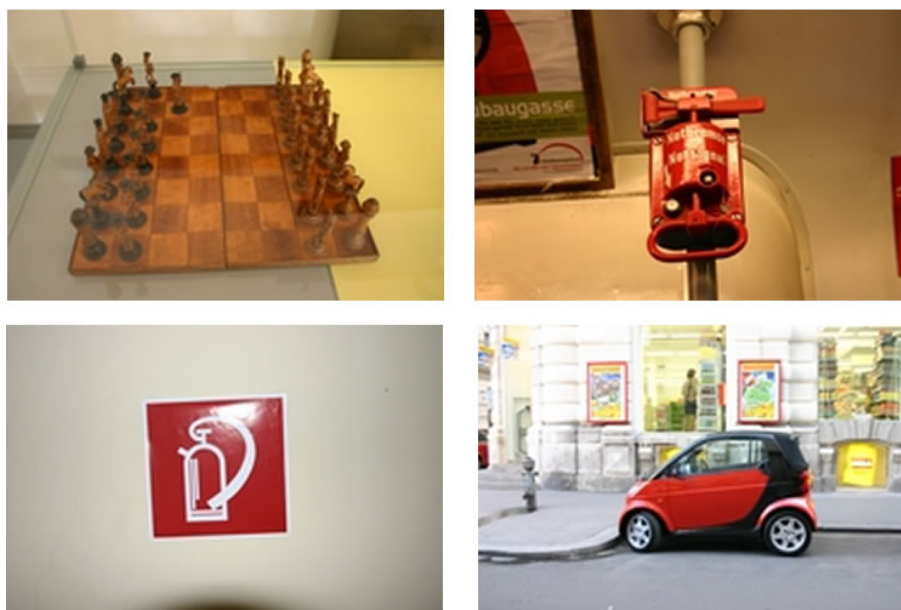
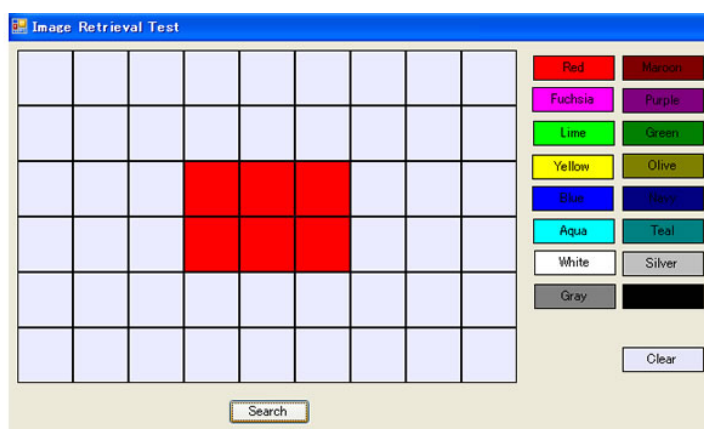


図 4.3: 改良したシステムでの検索結果

### 4.3 今後の課題

今後の課題としては、画像データベースのメタデータの生成を検索毎に行っているので、画像の検索に時間がかかることが挙げられる。これは、一度検索されたことのある画像のメタデータを保存しておくことで解消できると考えられる。

また、画像データベースのメタデータの生成をRGB値によって行っているが、メタデータとクエリを比較するときの閾値の設定によって検索結果に大きな差が出てしまうことがわかった。これはクエリを生成する際の色の設定が細かく指定できない点にも問題があると考えられる。閾値を絞りすぎると、クエリのRGB値と合わなくなり全く検索されなくなる。逆に閾値を広くすると、クエリのRGB値と近くない色まで検索されるようになり、精度が下がってしまう。これは、クエリを生成する際に指定できる色の種類を増やすことで解決できると考えられる。また、閾値をユーザが指定できるようにすることで、検索精度の絞り込みをユーザが自由に行えるようにできるのではないかと考える。

## 第5章 結論

本稿では、パネル型クエリ生成インタフェースを改良し、よりユーザの所望する画像を検索しやすくするため、検索の自由度を考慮した検索システムの開発を行った。従来のパネル型クエリ生成インタフェースでは、ユーザの検索意図が「中心が赤い画像」だったとしても、周辺のパネルが白いことによって生成されるクエリの検索意図が「周辺が白い画像」にすり替わってしまうことが起きることがわかった。そこで、白ではなく任意の色を意味する透明パネルを導入することにより、この問題を解決することができた。改良したパネル型クエリ生成インタフェースでは、透明パネルを実装することによって、よりユーザの検索意図を汲み取りやすくなったと考えられる。

今後の課題としては、画像データベースのメタデータの生成を検索毎に行っているため、検索に時間がかかってしまうことが挙げられる。これは一度検索したメタデータを保管しておくことで解消できると考えられる。また、画像データベースのメタデータの生成をRGB値によって行っているが、メタデータとクエリを比較するときの閾値の設定によって検索結果に大きな差が出てしまうことがわかった。閾値をユーザが指定できるようにすることで、検索精度の絞り込みをユーザが自由に行えるようにできるのではないかと考えられる。



## 謝辞

本研究を行うにあたり、終止熱心にご指導して頂いた木下宏揚教授と鈴木一弘氏、ご多忙の折研究室に足を運び様々な面で有益なご助言をして頂いた森住哲也氏に深く感謝いたします。さらに、公私にわたり良き研究生活送らせて頂いた木下研究室の方々に感謝いたします。

## 参考文献

- [1] 小野淳史, 天野督士, 斗谷充宏, 佐藤隆, 坂内正夫: "状態遷移モデルとシーン記述言語による自動キーワード付与機能をもつ画像データベースとその評価," 信学論(D-II), vol.J79-D-II, no.4, pp.476-483, 1996年.
- [2] 斎藤剛史, 金子豊久: "花と葉による野草の自動認識," 電子情報通信学会論文誌, Vol.J84-D-II No.7, pp.1419-1429, 2001年7月.
- [3] 松崎公紀, 櫻村雅章, 小沢慎治: "概略図を入力とした特徴量グラフに基づく絵画画像検索," 電子情報通信学会論文誌, Vol.J87-D-II No.2, pp.521-533, 2004年2月.
- [4] "Dublin Core," < <http://www.kanzaki.com/docs/sw/dublin-core.html> >
- [5] "神奈川大学図書館蔵書検索," < <http://opaclib.kanagawa-u.ac.jp/> >
- [6] "画像検索エンジン GazoPa," < <http://www.gazopa.com/> >
- [7] 田村秀行: "コンピュータ画像処理," オーム社, 2002年.
- [8] "osakana.factory グレースケールのひみつ," < <http://of0.jp/osakana/cgtips/grayscale.phtml> >
- [9] 野沢理倫, 稲積泰宏, 木下宏揚: "パネル型クエリ生成インタフェースにおけるユーザビリティの検討," 第1回イメージメディアクウォリティとその応用ワークショップ, 2006年9月.
- [10] 南部藤太郎, 八村広三郎: "類似判定基準の自動設定を行う画像検索システム," 「人文科学とコンピュータシンポジウム」, 平成14年9月.

## 質疑応答

- パネルの色について，赤色以外でも検証を行ったか（松澤先生）

赤色以外についても検証は行った．所望の画像のイメージに近い色で細かく塗り分けることによって，目的の画像をうまく検出することができた．また，色がわからない部分に透明パネルを使うことで，より所望の画像を検索しやすくすることができる．