

平成20年度 卒業論文

論文題目

# エッジ抽出による画像検索精度絞込み

神奈川大学 工学部 電気電子情報工学科

学籍番号 200302878

梶間 浩幸

指導担当者 木下 宏揚 教授

# 目次

第1章	序論	5
第2章	基礎知識	7
2.1	代表的画像検索技術	7
2.1.1	TBIR	7
2.1.2	CBIR	9
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	11
2.2.3	ガンマ補正	13
2.3	エッジ(輪郭線)	13
2.4	1次微分フィルタ	14
2.4.1	Prewittフィルタ	14
2.4.2	Sobelフィルタ	14
2.5	2次微分フィルタ	15
2.5.1	ラプラシアンフィルタ	15
第3章	提案方法	16
3.1	検索概要	16
3.2	画像データ類似度検索案	18
3.2.1	比較画像との画素面積割合比較	18
3.2.2	比較画像との同画素割合比較	18
3.3	対象画像検索システム	18
第4章	評価	19
4.1	エッジ抽出画像の作成	19
4.2	絞込み用画像データ群収集	21
4.3	エッジ抽出画像とデータベース内画像比較	21

第 5 章 結論	24
謝辭	25
質疑応答	27

# 目 次

2.1	画像検索エンジン GazoPa	10
2.2	画像の2値化	11
2.3	中間値法スペクトル	12
2.4	単純平均法スペクトル	12
2.5	加重平均法スペクトル	12
3.1	エッジ生成検索絞込み	16
3.2	元画像処理ルート	17
3.3	パネル型クエリ生成インターフェース検索システム	18
4.1	類似画像サンプル4枚	20
4.2	エッジ類似画像比較器	21
4.3	データベース内類似画像 TOP50 と類似度順位:サンプル1	22
4.4	データベース内類似画像 TOP50 と類似度順位:サンプル2	22
4.5	データベース内類似画像 TOP50 と類似度順位:サンプル3	23
4.6	データベース内類似画像 TOP50 と類似度順位:サンプル4	23

# 表 目 次

2.1	対象画素への Prewitt フィルタ . . . . .	14
2.2	対象画素への Sobel フィルタ . . . . .	14
2.3	対象画素への Laplacian フィルタ . . . . .	15

# 第1章 序論

近年、通信技術の発達によるネットワークの普及に伴い、Web上には大量の画像データが存在している。画像データの活用方法も多岐に亘り、SaaS(Software as a Service)[1]など画像データを媒介とした事業・コンテンツや、蔵書や商品を画像データを用いてユーザーにそれらを円滑に提供するための環境の構築が成され、会社内に限らず、個人レベルで膨大な画像データを保有し、データベースより所望する画像データを入手する機会が増えてきている。特にネットオークションやネット通販などの売買関係においては対象商品の画像データによる色やサイズ、状態の確認は非常に重要な役割を果たしている。また、情報公開、情報発信の場として急速な成長を遂げている電子図書館においても蔵書画像データの閲覧、検索ができるサイトが増えている。一般に画像検索は、キーワードに基づく画像検索であるTBIR(Text Based Image Retrieval)と、画像の内容に基づくCBIR(Content Based Image Retrieval)に分類される[2]。TBIRでは検索対象情報にキーワードを、CBIRでは画像の色や構図を検索対象情報として検索を行う。

しかし、オークションや電子図書館など画像データを利用する要因が増える一方で、ユーザーが所望する画像に類似した画像データが作者、使用者の意図に依らず増大してきている。特に電子図書館における蔵書データにおいては、ユーザーは書物に限定して検索を行うことが多いため、類似した蔵書画像データが検索対象として挙げられることでユーザーへの検索時における負担が増加してしまう。また、ユーザーの所望する画像データが曖昧なものであると、画像検索における検索条件を特定できず、所望する画像データの取得が困難となる。形が似ていても実際は全く異なる物であったり、体裁は分かるがタイトルが思い出せない、といった漠然とした検索対象を膨大なデータからより正確に、より円滑に取り出す画像検索技術が必要とされている。

本稿では画像データの検索・絞り込みにあたり、曖昧で漠然としたユーザーの想像と類似した画像を所望する画像データの検索条件として提出してもらい、その類似画像に存在するエッジ(輪郭線)を取り出して画像データの検索対象情報として抽出し、それらを検索対象情報の類似比較に用いた画像検索精度の絞り込みを提案する。

## 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、ユーザー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、テキストベースの検索方法であるTBIR、画像全体の特徴量を用いた検索方法であるCBIRとがある。

#### 2.1.1 TBIR

TBIR(Text Base Image Retrieval)はデータベース内の含まれる画像データに、画像のタイトルや製作者、内容を形容するキーワードを索引として画像に付加し、それを基にしたキーワード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となっている。

例として、前述した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、神奈川大学図書館の蔵書検索[3]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挙げられる。

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(Content Base Image Retrieval) は色や構図など、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」[5]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、 unnecessary 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 [6] 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。

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

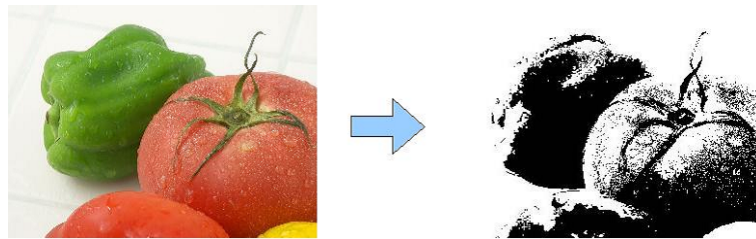


図 2.2: 画像の2値化

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[7]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[8]。

- 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。

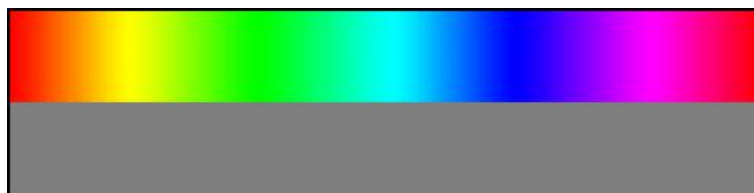


図 2.3: 中間値法スペクトル

- 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。

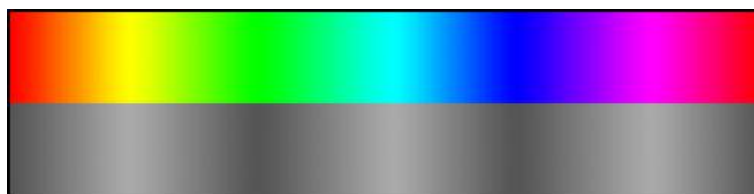


図 2.4: 単純平均法スペクトル

- 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。

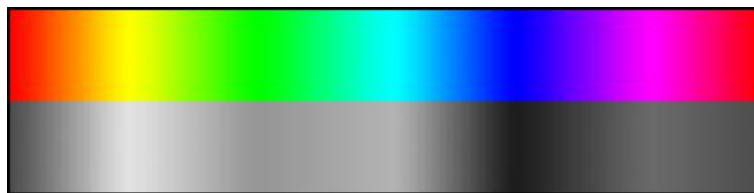


図 2.5: 加重平均法スペクトル

### 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 [6][9] 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 エッジ(輪郭線)

エッジ [10] とは物体の外縁を現す線、または画像を特徴付ける線要素である。画像データに色が存在する限り、エッジは必ず存在する。画像内の物体と物体、あるいは物体と背景の境目がエッジであり、画像の濃度や色が急激に変化している部分にエッジが現れる。エッジを機械的に見つけるには、ある画素に対してその周囲の画素との比較を行えばよい。これらの比較には、関数の変化分を取り出す微分演算が主に利用される。デジタル画像のデータは一定間隔をおいて飛び飛びに並んでいるため、演算は隣接画素同士の差分で微分を近似する。それら微分演算を行うための、隣接画素同士の演算を表現する係数の組を「微分フィルタ」と呼び、形式によって複数の種類がある。微分フィルタには1次微分フィルタ(グラディエント)と2次微分フィルタ(ラプラシアン)が存在する。

## 2.4 1次微分フィルタ

ある対象画素を中心とした上下左右9つの画素値に対して、表1に示すような係数をそれぞれ乗算し、結果を集計する。それを垂直方向、水平方向の2つの係数行列を用いて処理を行う。対象画素の画素値は $x$ 、 $y$ 方向成分がそれぞれ(2.2)、(2.3)式で表され、変化の強さは(2.4)式で求められる。

$$f_x(i,j) = (f_{(i+1,j)} - f_{(i-1,j)})/2 \quad (2.2)$$

$$f_y(i,j) = (f_{(i,j+1)} - f_{(i,j-1)})/2 \quad (2.3)$$

$$z = \sqrt{f_x^2 + f_y^2} \quad (2.4)$$

### 2.4.1 Prewittフィルタ

$f_x$ 方向では縦に、 $f_y$ 方向では横に勾配を作り、局所的ノイズ抑えたエッジ抽出を行う。

表 2.1: 対象画素へのPrewittフィルタ

-1	0	1	-1	-1	-1
-1	0	1	0	0	0
-1	0	1	1	1	1
fx 方向			fy 方向		

### 2.4.2 Sobelフィルタ

$f_x$ 、 $f_y$ 方向それぞれに対象画素と隣接する画素係数の重みを大きくし、勾配を大きく計算することでエッジが強調される。Prewittに比べて検出感度は高くなるが、それと同時にノイズに対する感度も高くなってしまふ。

表 2.2: 対象画素へのSobelフィルタ

-1	0	1	-1	-2	-1
-2	0	2	0	0	0
-1	0	1	1	2	1
fx 方向			fy 方向		

## 2.5 2次微分フィルタ

上下左右のそれぞれの変化の変化率、2つの1次微分を組み合わせて差分を取ることで2次微分フィルタとなる。1次微分フィルタでは隣り合う画素の変化分の大ききなところ強調して検出するが、エッジを構成する線の両隣の画素の差が常に大きいとは限らないので、2次微分では差分変化の大ききところをピックアップする。

### 2.5.1 ラプラシアンフィルタ

縦横方向の2次微分の結果を足し合わせて計算される [11]。これにより、方向に依存しないエッジが取得できる。ラプラシアンの画素成分は (2.5) で表される。フィルタには主に対象画素に隣接した画素の差分を取る4近傍と、更に範囲を広げた8近傍がある。

$$f(x, y) = \partial^2 f(x, y) / \partial x^2 + \partial^2 f(x, y) / \partial y^2 \quad (2.5)$$

表 2.3: 対象画素への Laplacian フィルタ

0	1	0		1	1	1
1	-4	1		1	-8	1
0	1	0		1	1	1
4近傍				8近傍		

セキュリティで用いられる指紋認証の基礎となる指紋の検出には、少ない画像情報の中から個々の指紋の特異点を得るために主に Sobel フィルタが扱われている。また、医療分野における X 線 CT 検査を用いた画像診断で病巣の陰影解析にラプラシアンフィルタが用いられている。



## 第3章 提案方法

### 3.1 検索概要

類似画像の検索手順を以下に示す。

1. 提示された類似画像に画像処理を施し、エッジを抽出する。
2. 抽出されたエッジ画像を基にデータベース内の画像と比較、検索する。
3. 対象とする画像検索システムにおいて同類似画像で検索を行い、画像データ群を入手する。
4. 抽出されたエッジ画像を基に画像データ群と比較、検索を行う。
5. 類似度の高い画像を所望する画像と類似しているか視覚的に比較する。

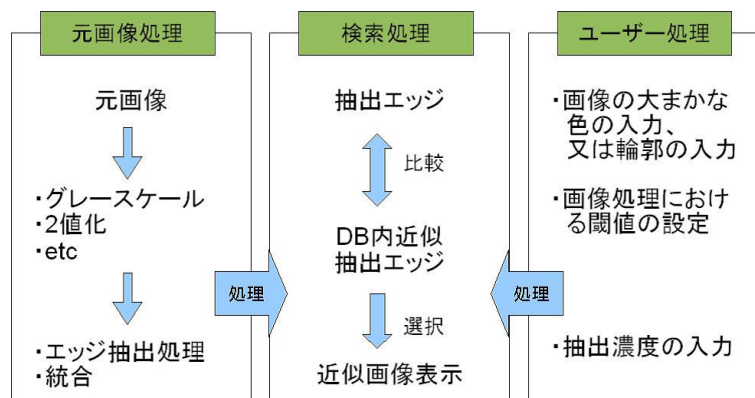


図 3.1: エッジ生成検索絞込み

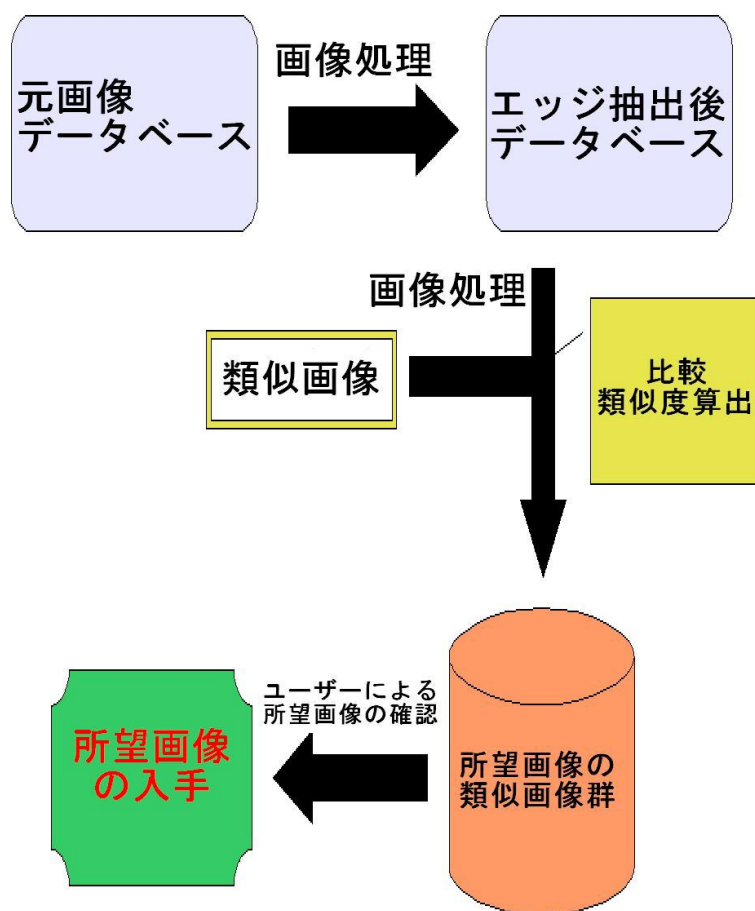


図 3.2: 元画像処理ルート

図(3.1)内の左半分の処理項を図(3.2)に示す。手順1で画像データをグレースケール化した後2値化し、エッジ抽出を行う上でノイズとなる無関係なエッジが極力入らないよう画像に処理を施す。また提示された所望画像についても同様の処理を施す。

## 3.2 画像データ類似度検索案

### 3.2.1 比較画像との画素面積割合比較

提示されたエッジ画像とデータベース内の比較画像の画素面積における白黒比において比較を行い、それぞれの画像の視覚的な明るさ、暗さの近い画像を求める。

### 3.2.2 比較画像との同画素割合比較

上記比較に合わせ、それぞれの画像の同画素における配色の合致比較を行い、それぞれの画像のエッジの近似している画像を求める。

## 3.3 対象画像検索システム

対象画像と同じ縦横比のキャンパスを格子状に $9 \times 6$ で分割したパネルを用いる。ユーザーはカラーパレットを用いて色を作成し、 $9 \times 6$ のパネルに所望の画像を描き、クエリを生成する。クエリと各画像の類似度をマッチング処理により算出する。

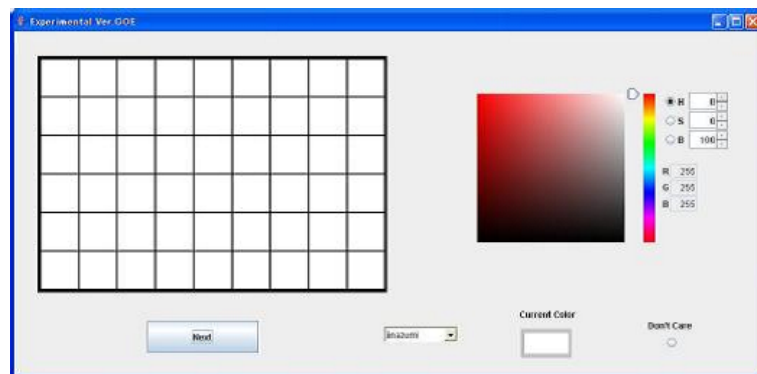


図 3.3: パネル型クエリ生成インターフェース検索システム

## 第4章 評価

### 4.1 エッジ抽出画像の作成

検索画像データベース(1250枚)の中から類似画像をランダムに4枚を選出し、それら画像データに対して各種画像処理、エッジ抽出を行う。今回は作成した画像処理装置(図4.1)より、グレースケール処理、2値化の閾値入力を手動で行った。

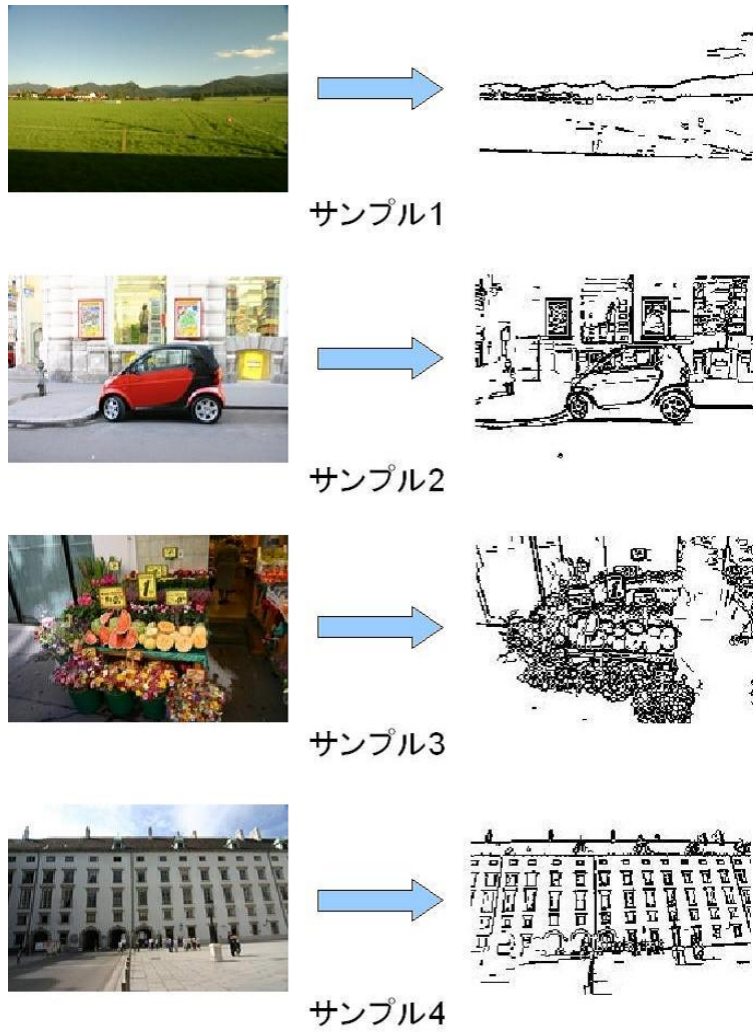


図 4.1: 類似画像サンプル4枚

## 4.2 絞込み用画像データ群収集

前述した画像検索システムを用いて、検索画像データベースから類似画像サンプルの類似画像データ群を50枚ずつ収集、それぞれに画像処理を施し、エッジ画像50枚×4枚分を作成する。

## 4.3 エッジ抽出画像とデータベース内画像比較

4枚のエッジ抽出画像を元に画像データ群を比較器に通し、検索システム単体のみ用いた場合と作成した比較器を通した場合の類似順位と類似度を比較する。また、双方の順位付けられた画像データ群から類似画像サンプルとの類似画像を視覚的に選出し、所望する画像がどれだけ上位に存在するか比較する。

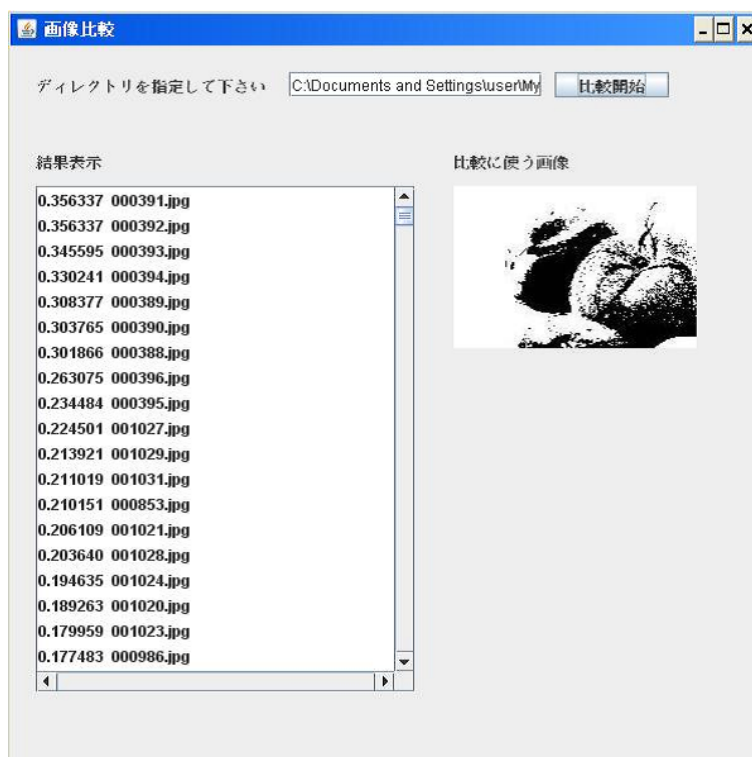


図 4.2: エッジ類似画像比較器

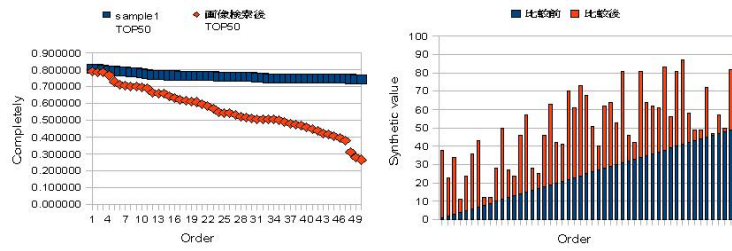


図 4.3: データベース内類似画像 TOP50 と類似度順位: サンプル1

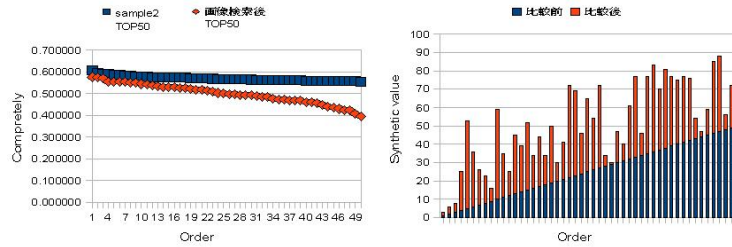


図 4.4: データベース内類似画像 TOP50 と類似度順位: サンプル2

サンプル1、2共に視覚的に似た画像の類似度が高く評価され、それらが上位を埋める形となった。サンプルはどちらも屋外の画像であるが、屋内の画像が若干比較に引っかかった。エッジのレイアウトが類似しているため、視覚的に類似していない画像でも高順位になってしまう点に、この検索システムの不確実性が表れてしまった。

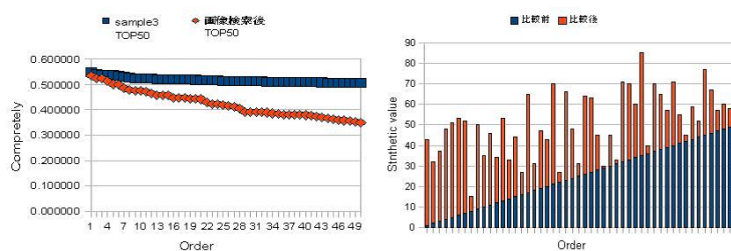


図 4.5: データベース内類似画像 TOP50 と類似度順位: サンプル 3

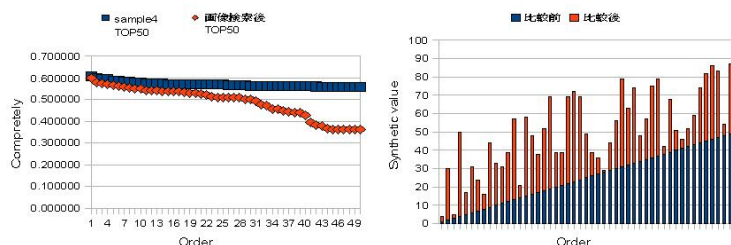


図 4.6: データベース内類似画像 TOP50 と類似度順位: サンプル 4

サンプル4は先のサンプル1、2と同様、数値的にも視覚的にも類似された画像が表れたが、サンプル3においては抽出されたエッジが入り組み、類似に値する画像データの選出が少なかった。視覚的にも隣接する色は際立って一見エッジが取り安そうだが、細かなエッジがノイズになってしまい、検索精度の低下に繋がってしまった。



## 第5章 結論

本稿では、画像に存在するエッジを類似画像に適用し、ユーザーの所望する曖昧な情報の画像を入手するための画像検索システム中における検索精度の向上、絞込みを行った。類似画像比較サンプルとして行った4枚の画像データは高位類似としてほぼ数値的、視覚的にも類似に値する結果であった。別途に第三者によるサンプル画像の提示と同検索の画像比較を行ったが、検索精度を大きく減する点は見受けられなかった。しかし、異なるサンプル画像の類似画像群の中で類似度の範囲に同じ画像が検索に掛かっているものが見られた。これは類似条件の一致が偶発的に表れたものであると考えられる。最終的な視覚的比較においてサンプル画像とは異なる画像であったため軽視したが、今後の課題として異なるサンプル画像における対象画像の合致は度重なることで検索精度の低下に繋がる危険があるので、プログラム起動後で検索を行った類似画像群同士における間引きを行えるよう検討する。また、画像処理とエッジ抽出画像の抽出、画像群の収集が手動であったこと、検索使用時におけるシステム速度の遅延が著しかったことを踏まえ、上記問題点を改善したプログラムの修正とユーザー負担の軽減を行う。

## 謝辞

本研究を進めるにあたり、終始熱心に御指導していただいた木下宏揚教授に心から感謝致します。良き研究生活を送らせていただいた木下研究室の方々、忙しい中作業に付き合っ頂いた皆様に深く感謝致します。

電気電子情報工学科4年 梶間浩幸

## 参考文献

- [1] ”IT用語辞典” < <http://e-words.jp/w/SaaS.html> >
- [2] ”タグの共起と類似画像を利用したタグ付け支援システム”  
< <http://www.ieice.org/iss/de/DEWS/DEWS2008/proceedings/files/e10/e10-5.pdf> >
- [3] ”神奈川大学図書館蔵書検索” < <http://opaclib.kanagawa-u.ac.jp/> >
- [4] ”Dublin Core について”  
< <http://avalon.slis.tsukuba.ac.jp/sugimoto/Lect2005/DL05/DL111105.pdf> >
- [5] ”画像検索エンジン GazoPa” < <http://www.gazopa.com/> >
- [6] 田村秀行 ”コンピュータ画像処理” , オーム社 (2002)
- [7] ”osakana.factory グレースケールのひみつ”  
< <http://ofo.jp/osakana/cgtips/grayscale.phtml> >
- [8] ”Le Premier Soleil グレースケール”  
< <http://www.geocities.co.jp/Milkyway/4171/graphics/002-6.html> >
- [9] ”osakana.factory ガンマって何?” < <http://ofo.jp/osakana/cgtips/gamma.phtml> >
- [10] ”輪郭を抜き出す”  
< <http://mikilab.doshisha.ac.jp/dia/research/person/shuto/research/0605/bibun.html> >
- [11] ”Firter 処理” < <http://www.gifu-nct.ac.jp/elec/yamada/iwata/filter/index.html> >

## 質疑応答