

平成24年度卒業論文

論文題目

イラスト著作権保護のためのHOGによる複製
検出

神奈川大学 工学部 電子情報フロンティア学科

学籍番号 200902762

高橋 涼

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	7
2.1	電子透かし	7
2.2	代表的技術検索	7
2.2.1	TBIR	8
2.2.2	CBIR	9
2.3	画像処理技術	11
2.3.1	2値化	11
2.3.2	グレースケール	11
2.3.3	ガンマ補正	13
2.4	画像検索の流れ	14
2.5	ウェーブレット変換	15
2.6	エッジ検出	17
2.6.1	グラディエント (1次微分)	17
2.6.2	ラプラシアン (2次微分)	17
2.7	対象とする画像の定義	19
2.8	特徴量	20
2.9	局所特徴量	22
2.10	使用する局所特徴量	24
2.10.1	HOG (Histograms of Oriented Gradients)	24
第3章	提案方式	26
3.1	従来の処理の流れ	26
3.2	改良点	28

3.3	結果	30
第 4 章	結論	31
4.1	まとめ	31
4.2	今後の課題	31
第 5 章	謝辞	32
第 6 章	質疑応答	1

目 次

2.1	画像の2値化	11
2.2	中間値法スペクトル	12
2.3	画像検索の流れ	14
2.4	グラディエントに用いるオペレータ	18
2.5	ラプラシアンに用いるオペレータ	18
2.6	輝度勾配と強度の算出	25
3.1	従来の複製検出の流れ	27
3.2	改良点	29
3.3	結果画像	30

第1章 序論

近年、インターネット上には音声や画像などのデジタルデータが大量に流通しており、個人レベルでもそれらびデジタルコンテンツを自由に扱えるようになってきている。

しかし、それらの発達に伴い不正なコピーによる著作権の侵害などの問題が発生している。著作権を保護すべきコンテンツとしては様々なものが考えられるが、その中でも画像コンテンツ、特にイラストコンテンツが、漫画などの電子書籍化、pixiv等の大型のイラスト投稿サイトの影響で急速に増加している。

イラストの不正利用者はオリジナルの画像をそのまま使うだけでなく、様々な処理によって改変を加えたり、線画部分をトレースして使用したり、線画を元に全く別の着色をすることで自分の作品として用いることが多い。

このような問題に対処するためにコンテンツの著作権の保護の手法として、電子透かしや画像検索に基づく複製検出などの技術が提案されている。

電子透かし [1] は、動画や音声、画像などのマルチメディアデータに、画質や音質にはほとんど影響を与えずに特定の情報を埋め込む技術のことである。著作権情報を埋め込むために利用されることが多い。何も描かれていないように見える部分に反対側から強い光を当てると像が浮かび上がる「透かし」技術は紙幣などに使われているが、これになぞらえて電子データ中に見ると見ることのできないデータを埋め込むことを電子透かしと呼ぶ。しかし、線画のような助長性の少ないコンテンツの場合は、人間に知覚できない形で埋め込むことが困難となっている。

一方、画像検索に基づく複製検出では、著作権保護の対象となる

コンテンツをデータベースに保持し、複製の疑いがある画像を検索質問として検索をし、複製検出を行う。このときデータベースに保持する情報として、各種変換や部分的複製に対して安定である局所特微量と呼ばれるものがよく用いられる。

局所特微量を用いたデータベースで複製を検出する研究として、孫氏らによって提案された線画著作権保護のための部分的複製検出法がある [2][4][5]。この方法では、局所特徴領域の検出器として MSER (Maximally Stable Extremal Gradients) [3] を、局所特微量の記述子として HOG (Histograms of Oriented Gradients) を用いるもので、更に、データベースの容量を減少するために、PCA (Principal Components Analysis) を用いて特徴ベクトルの次元数を削減すると共に、近似最近傍探索により検出時間を短縮したもので手書きによる線画の複製に対する有効性を示したものである。しかし、対象とする画像が線画像から線画像を複製したものであるため、トレース画像のようにカラーイラストからその輪郭線部分を複製したものや構図を複製したものについてはその有効性が検証されていない。

従来の研究 [21] では計算時間の早い SURF (Speeded Up Robust Features) 及び、線画の検出に有効とされる HOG を用いて部分的複製検出を行った。オリジナル画像や回転・大きさを変えた画像には有効性が示されたが、手書きなどによるトレース画像に対して精度の高い結果が得られなかった。

HOG 特微量は、画像の局所領域から輝度勾配・輝度強度を取り出す特微量で、物体のおおまかな形状を表現することが可能となる。そのため、歩行者や人工物などの物体を検出するのによく用いられる。

原因として局所特微量抽出の際、局所領域に十分に線分が含まれていなかった為、精度が落ちてしまったと考えられる。よってよりよい精度の検出法が必要である。

そこで本論文では、精度の低かったトレース画像に対してのみ研究を行う。特微量を抽出する前に画像にウェーブレット変換を行う。そしてウェーブレット変換を行った画像から高周波成分の多い領域を選択し、その領域に対して HOG 特微量を抽出部分的複製検出を

行う.

第2章 基礎知識

2.1 電子透かし

電子透かし [6] はデジタルコンテンツを不正利用から保護するために、人に気づかれないように著作権者を示すマーク（情報）を、画像や音などのデジタルコンテンツにそっと忍ばせる技術です。

電子透かしは、次のような意外と簡単な原理に基づいています。

1, キャリアデータの変更：キャリアの一部を電子透かし情報に基づき変更する。

2, 画像、音声などデジタルコンテンツの冗長性の利用：画像や音データの一部を変更しても、冗長性の高いコンテンツの場合は全体的な構成に影響を及ぼさないことを利用し、データの一部を書き換えて電子透かしとする。

3, 人間の感覚のルーズさ、生理学的特徴の利用：上記2で変更する場合、さらに人間の感覚の特性をうまく利用して、人が気づき難い部分のデータを変更していくことにより電子透かしとする。

2.2 代表的技術検索

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。

2.2.1 TBIR

TBIR[7]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、それを基にし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あげられる。蔵書検索を行うにあたって、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 [8] 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事象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位に定まる情報資源に対する
- 11, 出所: 現在の情報資源が作り出される源になった情報源への参照
- 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[9]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嬢の画像検索を行うことが可能である。

しかし、TBIRにも次のような問題がある。

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なわないため、所望する画像データのほかにそのキーワードの含まれる画像データも検索してしまう。

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おこなわなければならない、さらなる負担が伴う。

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はできない。

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。

既に大き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。

2.2.2 CBIR

CBIR[10][11][12]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挙げる。GazoPa[13]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。

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。

2.3 画像処理技術

2.3.1 2値化

白と黒だけによる画像の表現処理 [14] 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。図 2.1 は2値化処理を施した例である。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。

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

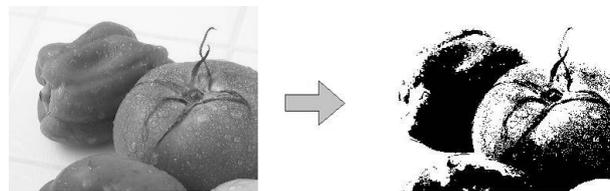


図 2.1: 画像の2値化

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。

2.3.2 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 [15] 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 2.2 に中間値法のスペクトルを示す。

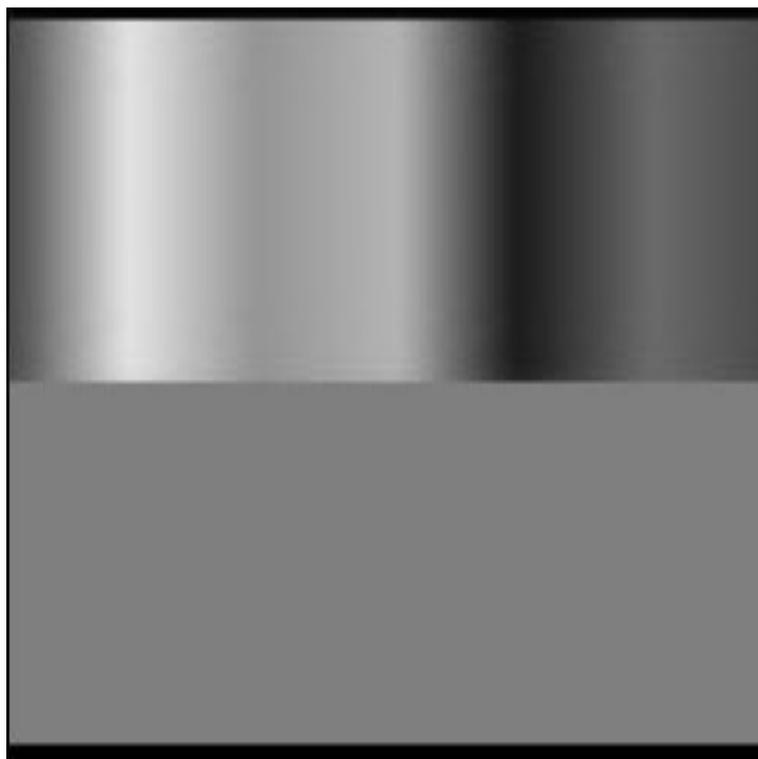


図 2.2: 中間値法スペクトル

2.3.3 ガンマ補正

γ (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 [14] という。 γ 補正の数式を (2.1) に示す。

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

画像データと出力データとの値が正比例している時、 γ 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 γ 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 γ 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。

2.4 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2.3は画像を検索する際の流れを示したものである。

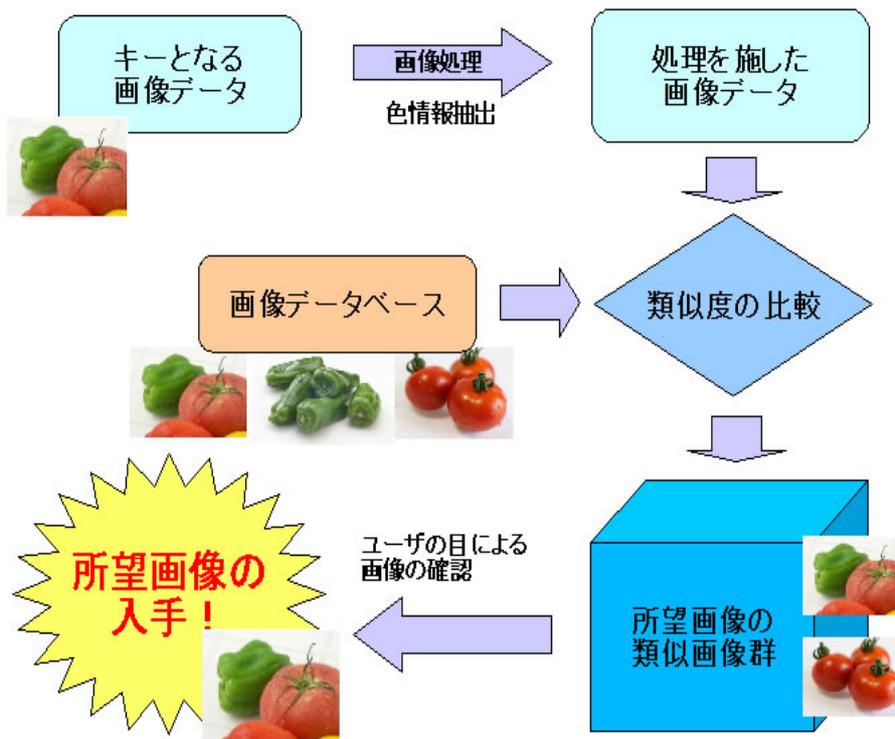


図 2.3: 画像検索の流れ

2.5 ウェーブレット変換

ウェーブレット変換 [17][18][19] とは、周波数解析の手法のひとつである。ウェーブレット変換は、時間的变化の特徴と周波数成分の混じり方を知るために用いられ、時間と周波数にかかわる信号情報を同時に抽出することができる。

ウェーブレット理論の誕生は1980年代初めで、フランスの石油探査技師 Morlet が、石油の埋蔵場所を特定するため、人工地震波信号の精密な時間（埋蔵場所の特定）と周波数（埋蔵資源の分別、石油かどうか）の情報を同時解析するために考案した。

つまり、ウェーブレットは、“時間的变化の特徴”と“周波数成分の混じり方”を知るために用いられ、時間と周波数にかかわる信号情報を同時に抽出することができる。

ところで、周波数解析の手法としては古典的なフーリエ変換が真っ先に挙げられる。フーリエ変換は周波数成分を解析するための手法であり、コンピュータ処理するための DFT (Discrete Fourier Transform) と、その高速計算バージョンである FFT (Fast Fourier Transform) が有名どころである。

例えば、人の声の性質（声色、声紋）をはじめとして、楽器の音色や画像、自然界のあらゆる信号はすべて、固有の周波数成分の混じり方を有している。こういった周波数成分の混じり方を解析できるのが、フーリエ変換である。

しかし、フーリエ変換では信号のもつ性質をすべて周波数情報に置き換えてしまうため、時間に関連する情報が失われてしまう。言い換えれば、時間とともに刻々と変動する信号であっても、フーリエ変換では周波数成分だけしか得られないので、大ざっぱな信号情報しか見いだせない。

そこで、時間を一定間隔ずつ切り出しては次々にフーリエ変換する手法として、短時間フーリエ変換 (short-time Fourier Transform) が考案された。ただ、解析する時間幅が一定に固定されていることが最大のネックとなり、時間と周波数を同時に解析するには物足り

ないものであった。

この物足りなさを埋め合わせるものとして、周波数に応じて解析する時間幅を変化させるというウェーブレットが発案されたのである。つまり、周波数の低い成分に対しては“解析する時間幅を長くする”，高い成分に対しては“逆に、短くする”のである。このように、解析する時間幅を周波数に反比例させることによって、多くの信号解析が合理的に行えるようになる。

ウェーブレットの基本式は関数の内積によって計算されるのが一般的である。すなわち、信号 $x(t)$ とウェーブレット ψ_ℓ の内積は、

$$d_\ell = \int_{-\infty}^{\infty} x(t)\psi_\ell^{\bar{}}(t)dt$$

で定義される。

2.6 エッジ検出

画像内のすべての画素は、色情報を持っている。その画像内の明るさが急激に変化する場所がオブジェクトの領域の境界、つまり輪郭となる。エッジ検出とは、画像の明るさの変化によりオブジェクトの輪郭を算出することを指す。明るさの変化値は、微分演算を利用することで算出ができる。

しかし、デジタル画像は連続ではないため、厳密には微分演算は行えない。そこで、隣接する画素の微分値の近似値を差分により算出する必要がある。隣接する画素の差分は、微分パラメータを用いて画素に重み付けすることにより算出ができる。この差分値が隣接する画素の微分値の近似値となる。

2.6.1 グラディエント（1次微分）

画面上の座標 (x,y) における明るさの勾配はグラディエントにより算出ができる。この微分値は、ベクトル $(\Delta x, \Delta y)$ として表現ができる。微分と強さを算出する式を以下に示す。

【微分】

$$\Delta x = f(x + 1, y) - f(x, y) \quad (2.2)$$

$$\Delta y = f(x, y + 1) - f(x, y) \quad (2.3)$$

【強さ】

$$\sqrt{\Delta x^2 + \Delta y^2} \quad (2.4)$$

グラディエントに用いるオペレータを図 2.4 に示す。

2.6.2 ラプラシアン（2次微分）

ラプラシアンはグラディエントをさらに微分したもので、単純に画像上の画素 (x,y) における明るさの差分 $\Delta(x,y)$ のみを表す。以下は

ラプラシアン of 式である.

$$\Delta(x, y) = 4 \cdot f(x, y) - \{f(x, y - 1) + f(x, y + 1) + f(x - 1, y) + f(x + 1, y)\} \quad (2.5)$$

ラプラシアンに用いるオペレータを図 2.5 に示す.

	Gradient	Prewittフィルタ	Sobelフィルタ
Δx算出 オペレータ	0 0 0	1 0 -1	-1 0 1
	0 1 -1	1 0 -1	-2 0 2
	0 0 0	1 0 -1	-1 0 1
Δy算出 オペレータ	0 0 0	1 1 1	-1 -2 1
	0 0 1	0 0 0	0 0 0
	0 -1 0	-1 -1 -1	1 2 1

図 2.4: グラディエントに用いるオペレータ

	ラプラシアン1	ラプラシアン2	ラプラシアン3
Δ(x,y)算出 オペレータ	0 -1 0	-1 -1 -1	1 -2 1
	-1 4 -1	-1 8 -1	-2 4 -2
	0 -1 0	-1 -1 -1	1 -2 1

図 2.5: ラプラシアンに用いるオペレータ

2.7 対象とする画像の定義

簡易的な改変(回転が大きさ)を加えた画像ではHOGによる複製検出で十分に有効な精度がでているので輪郭線や構図の手書きによるトレース画像のみを対象とする.

2.8 特徴量

特徴量とはその画像が持つ特徴を数式や数値によって表したものである。画像の特徴としては以下のようなものがある。

- ・ **面積** 2値画像中の図形 S の面積は S 中の画素の数として定義可能である。
- ・ **周囲長** 物体の輪郭線の画素数. 斜め方向はデジタル図形特有の誤差を生じるので $\sqrt{2}$ 倍して数える。
- ・ **円形度** 面積, 周囲長を基に, 形状の複雑さを測る特徴量. 円形になるほど値が1に近づく。
- ・ **伸長度** ある領域が凸であるというのは, その領域の任意の2点を結ぶ線分が必ずその領域内を通る. また, 任意の領域に対し, その領域を含む最小の凸領域を凸閉包と呼ぶ. 任意の領域に対し, その凸閉包から元の領域を差し引くとその領域の凹部が検出されることになる. この凹部がその領域の重要な形状測度となる。
- ・ **モーメント** 座標 (i, j) における画像 $f(i, j)$ の $(p + q)$ 次のモーメントは

$$m_{pq} = \sum_i \sum_j i^p j^q f(i, j) \quad (2.6)$$

で定義される。

2値画像の場合, 0次のモーメント m_{00} は面積を表すことになる。また, m_{10}/m_{00} および m_{01}/m_{00} は重心座標 $G(i_G, j_G)$ を表す。重心の周りのモーメント M_{pq} は, 特に重心モーメントと呼ばれ,

$$M_{pq} = \sum_i \sum_j (i - i_G)^p (j - j_G)^q f(i, j) \quad (2.7)$$

で与えられる。

また,2次モーメントは慣性モーメントとも呼ばれ,画像 $f(i, j)$ の2次重心モーメントは式(2,7)で $p = 2, q = 0$ ならびに $p = 0, q = 2$ で与えられ,

$$M_f = \sum_i \sum_j \{(i - i_G)^2 + (j - j_G)^2\} f(i, j) = M_{02} + M_{20} \quad (2.8)$$

となる.

式(2,8)で重心を原点とした時のx軸,y軸についての2次モーメントをそれぞれ μ_{02}, μ_{20} とすると,原点の周りの2次モーメントは

$$\mu_2 = \mu_{02} + \mu_{20} = \sum_i \sum_j j^2 f(i, j) + \sum_i \sum_j j i^2 f(i, j) \quad (2.9)$$

さらに,原点を通る傾き θ の直線

$$j = i \tan \theta \quad (2.10)$$

の周りの2次モーメントを求め,これが最小となる角が慣性主軸 θ で,次式で表される.

$$\theta = \frac{1}{2} \tan^{-1} \left\{ \frac{2\mu_{11}}{\mu_{20}\mu_{02}} \right\} \quad (2.11)$$

この θ は図形が伸びている方向を表す特徴量となる.

- ・ **フーリエ記述子** ある領域の境界線を追跡していくと一つの閉曲線が得られる.この閉曲線上の任意の点における接線の偏角を,ある基準点(出発点)からの距離 d の関数として $\alpha(d)$ と表すことにする. α_0 を基準点での偏角とすれば,この $\alpha(d)$ は $\alpha_0 + 2\pi$ まで変化する d の関数となる.すなわち,閉曲線の一周の長さを D とすれば, $\alpha(D) = \alpha_0 + 2\pi$ となる.ここで,閉曲線は周期性を持つことから

$$\alpha_N(d) = \alpha(d) - 2\pi \frac{d}{D} \quad (2.12)$$

とすると, $\alpha_N(d)$ は周期 D の周期関数となり, フーリエ級数に展開できることになる. $\alpha_N(d)$ のフーリエ級数展開を

$$\alpha_N(d) = \sum_{k=1}^{\infty} \left\{ A_k \cos\left(2\pi k \frac{d}{D}\right) + B_k \sin\left(2\pi k \frac{d}{D}\right) \right\} \quad (2.13)$$

とあらわすと, 係数 A_k, B_k は閉曲線の形状を表す特徴量となる. この係数をフーリエ記述子と呼ぶ.

しかし, 適用できるのは閉曲線だけであり, 開曲線には適応できないという欠点を持つ.

2.9 局所特徴量

画像の照合に用いられる特徴量は大きく大域特徴量(global features)と局所特徴量(local features)に分類できる. 大域特徴量は画像全体から抽出される特徴量で, 本研究で用いる局所特徴量は, 画像の局所領域から抽出される特徴量である. 最も著名な局所特徴量としてSIFTがあげられる.

局所特徴量を抽出するには, どの領域から特徴量を取り出すかを決定し, その領域から特徴量を抽出するという2段階の処理が必要となる. 前者を行うプログラムをdetector, 後者を行うプログラムをdescriptorと呼ぶ.

大域, 局所のどちらも特徴量はベクトル(特徴ベクトル)として表現され, これを照合することで画像の検索をすることが可能となる.

局所特徴量のdetectorに対応する処理が大域特徴量の抽出では不要となるので, 大域特徴量の抽出は計算的に有利であり, 特徴量が画像1枚に対し1つとなるため, 画像の索引付けに用いるデータ量としても大幅に少なく済むという利点もある.

しかし, 画像の一部が隠れるなどした場合, 大域特徴量では同じ値を得ることが不可能になってしまう. 局所特徴量を用いると, 隠れたことによって得られない特徴量があっても, 隠れていない部分から同じ特徴量を得ることができるという利点がある. この特性が, トレー

ス画像のように画像の輪郭線だけを用いるなど部分的な複製を検出する際に必要であると考えられる.

2.10 使用する局所特徴量

2.10.1 HOG (Histograms of Oriented Gradients)

HOG[16][20]はDalas氏らによって提案された、画像の局所領域から輝度勾配・輝度強度を取り出す特徴量で、物体のおおまかな形状を表現することが可能である。そのため、一般に歩行者の検出や人工物などの物体の検出に用いられる。

まず、局所領域を3*3のブロックに分割しそのブロックを5*5のセルに分割する。分割したセルごとに次式を用いて輝度Lから勾配強度mと勾配方向 θ を図2.6のようにセルの座標(x,y)から計算する。

$$m(x, y) = \sqrt{f_x(x, y)^2 + f_y(x, y)^2} \quad (2.14)$$

$$\theta(x, y) = \tan^{-1} \frac{f_y(x, y)}{f_x(x, y)} \quad (2.15)$$

$$\begin{cases} f_x(x, y) = L(x + 1, y) - L(x - 1, y) \\ f_y(x, y) = L(x, y + 1) - L(x, y - 1) \end{cases} \quad (2.16)$$

そして、0度から180度までを20度ずつ9方向に分割し輝度勾配のヒストグラムを作り、ブロックごとに正規化を行う。最終的にすべてのヒストグラムを統合して特徴量とする。

こうして得られた特徴量は、全ブロックのベクトルから構成される多次元ベクトルで、互いに重なりあったブロックから抽出されているので位置ずれに対する安定性が高いと考えられる。この特性は、トレース画像の手書きによる線のおれに対応し複製の検出を行う際に重要な性質であると考えられる。また、HOG特徴量は、手書き文字認識で用いられている局所方向ヒストグラム特徴とほぼ同じものであるため、トレース画像の検出にも有効性が期待できる。

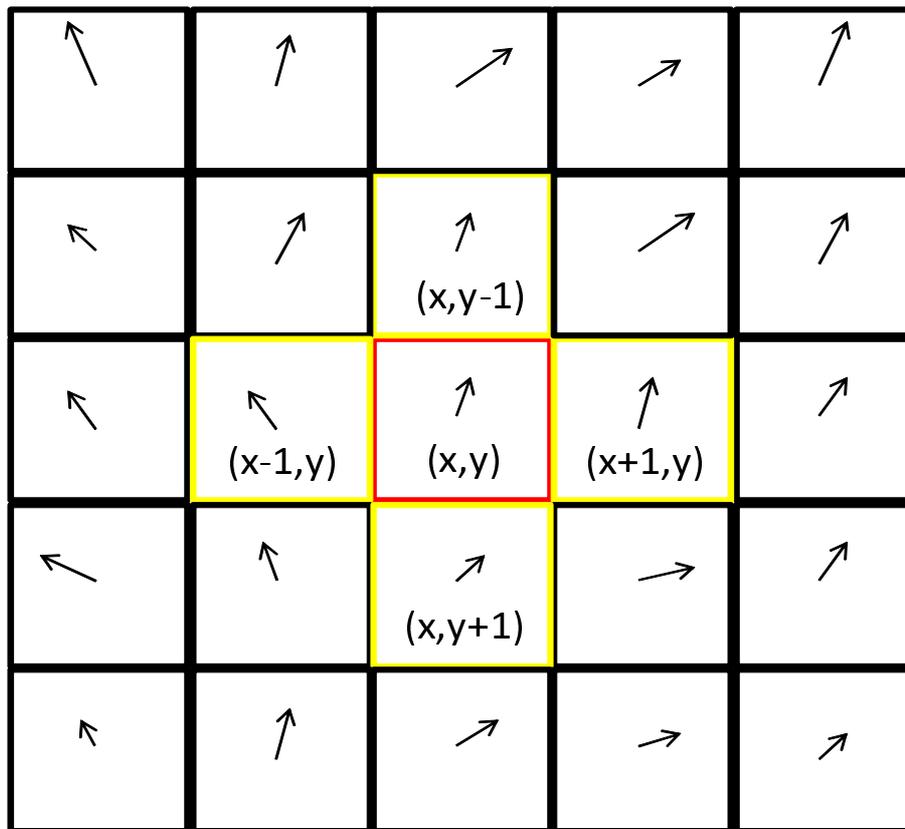


図 2.6: 輝度勾配と強度の算出

第3章 提案方式

3.1 従来の処理の流れ

従来の研究では，図3.1のように2つの処理からなる流れで複製の検出を行っている．

まず，データベース側の処理では著作権保護の対象となるオリジナル画像から線画部分を抜き出し，線画像から局所特徴量を抽出し，データベースに登録する．局所特徴量の抽出アルゴリズムとして，SURFとHOGを用いる．また，クエリ側の処理では著作権を侵害する可能性がある画像をクエリとしてデータベース側の処理と同様に局所特徴量を抽出して，クエリの特徴量とデータベース側の特徴量を照合し，著作権保護対象の画像に一致する特徴量の個数分投票を行う．

クエリとデータベースに登録された画像からは，局所特徴量の種類によるが数百から数万の特徴量が抽出され，その数だけ投票の処理が行われる．

投票処理の終了後，得票数が多い画像から一定数までを著作権侵害の可能性がある画像として報告する．

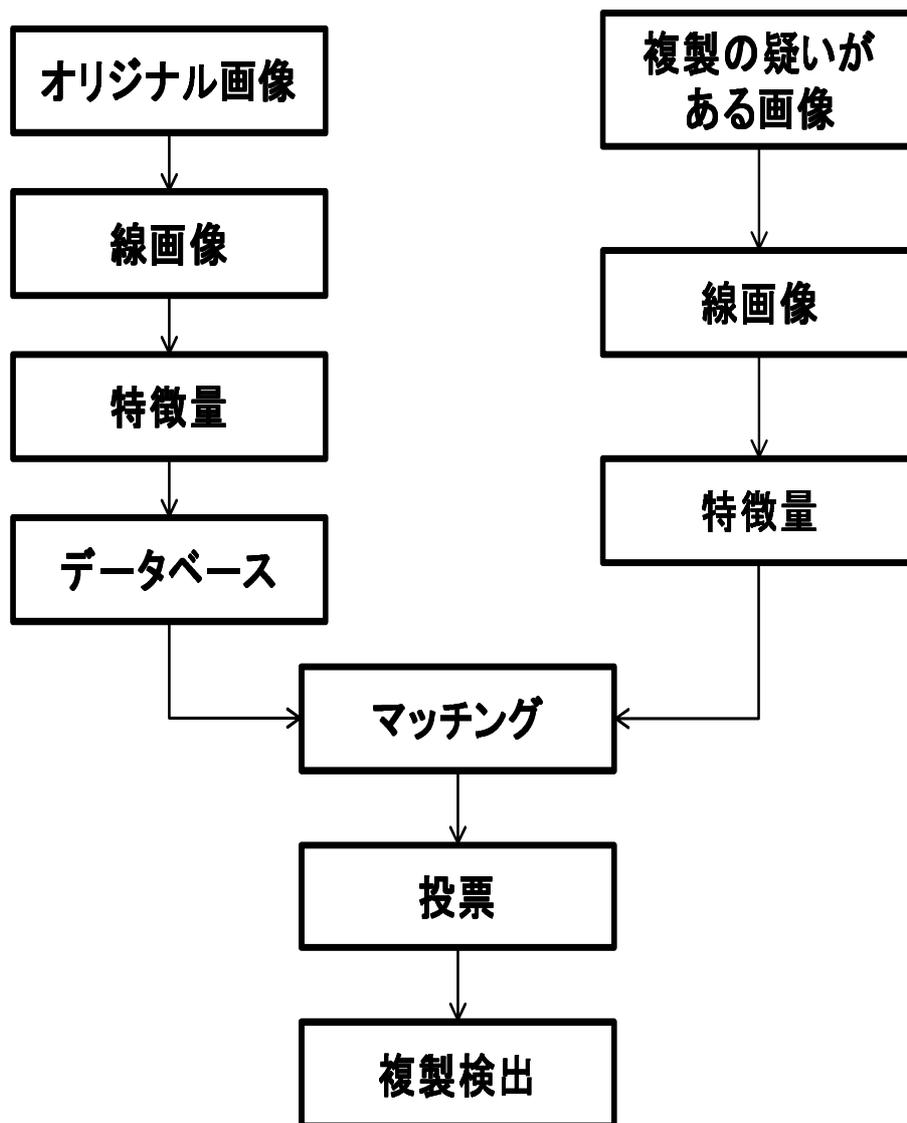


図 3.1: 従来の複製検出の流れ

3.2 改良点

この方法では手書きによるトレース画像に対して有効性は示されなかった。原因としては特徴量抽出の際、線分が十分に含まれている領域を正確に抽出できていなかったと考えられる。

- 1 : 画像にウェーブレット変換を用い周波数解析を行う。
 - 2 : ウェーブレット変換により、求められた x-y 方向高周波画像を 2 値化する。
 - 3 : その画像を大きさに応じて 16×16 , 8×8 画素の領域に分割する。
 - 4 : 分割した領域のピクセル値の和を求める。
 - 5 : 全領域のピクセル値の平均を求める。
 - 6 : 平均値を基準の値とし、平均値以上の値の領域を使用する。
- この方法で局所領域選択の際の問題を解決する。

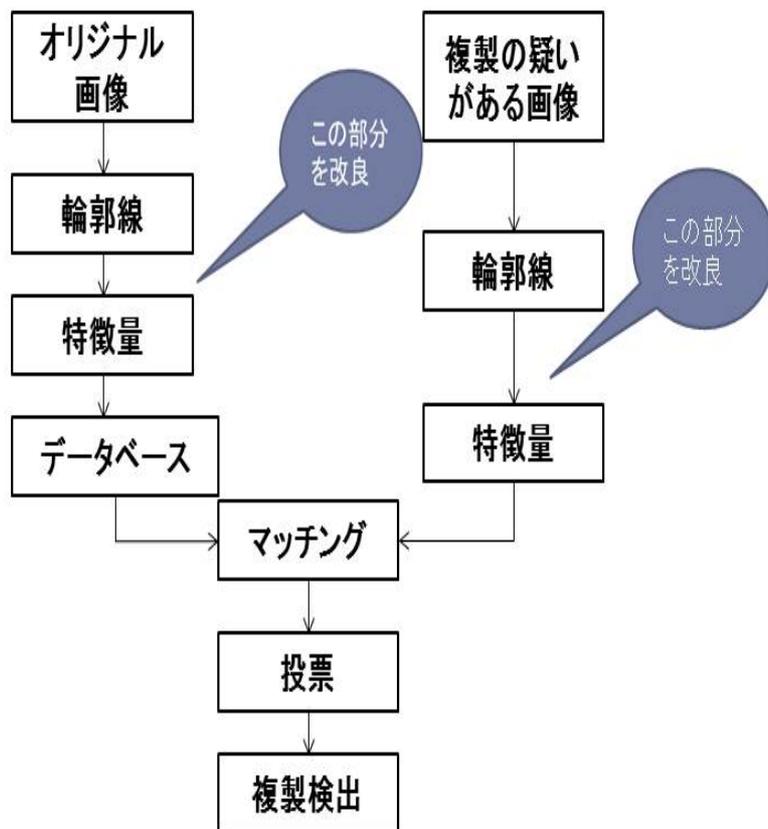


図 3.2: 改良点

3.3 結果

図 3.3 で示すようにウェーブレット変換を行い2値化した画像から、線分の多く含まれる領域を抽出することに成功した。

図 3.3 が抽出された画像である。この画像の白い部分が線分が多いと判定された領域である。この領域を複製検出に用いることで複製検出の精度が上がると思われる。

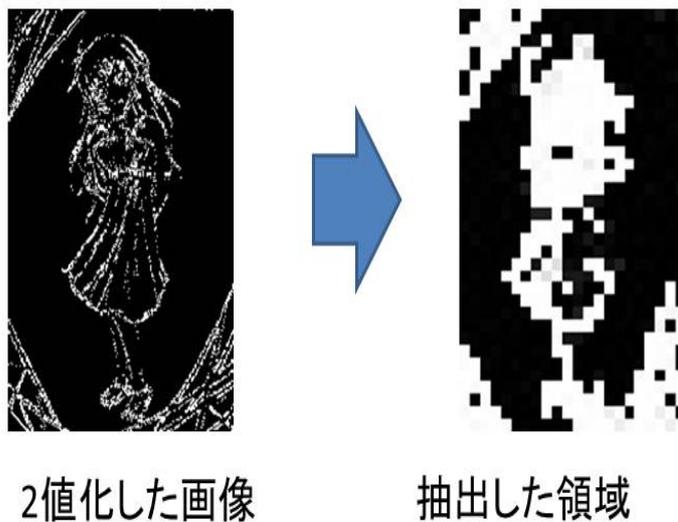


図 3.3: 結果画像

第4章 結論

4.1 まとめ

高周波成分の多く含む領域を抽出することで、イラストコンテンツにおいて重要な情報の線分が多く含まれるを抽出することができる。

従来の研究での精度の低さは局所領域の抽出が正確でなかったため精度が落ちたと考えられる為、先に線分の多く含まれている領域として抽出し複製検出に使用することで精度が上がると考えられる。

4.2 今後の課題

- ・実際に抽出できた領域を用いての複製検出を行う。
- ・データベースのメモリ量や計算時間が膨大なものになってしまう為、特徴量の次元数の効率のよい削減など、メモリ量・計算時間の削減手法の検討が必要である。

第5章 謝辞

本研究を行うにあたり，終始熱心に御指導していただいた木下宏揚教授と宮田純子助手に心から感謝致します．また，公私にわたり良き研究生生活を送らせていただいた木下研究室の方々に感謝致します．

2013年 2月
高橋 涼

参考文献

- [1] IT 用語辞典 e-words ”<http://e-words.jp/w/E99BBBE5AD90E9808FE3818BE38197.html>”
- [2] Weihan Sun, and Koichi Kise, “線画の著作権保護のための部分的複製検出法”, “電子情報通信学会論文誌”, D Vol. J93-D No.6 pp.909-919, 2010.
- [3] J. Matas, O. Chum, M. Urban, and T. Pajdla, “Robust wide baseline stereo from maximally stable extremal regions”, BMVC, pp.384-393, 2002.
- [4] D.G.Lowe, “Distinctive image fetures from scale-invariant key-points”, Int.J.Comput.Vis., vol.60, no2, pp.91-110, 2004.
- [5] P.Bas, J-M Chassery, and B.Macq, “Geometrically invariant watermarking using feature points”, IEEE Trans. Antennas Propag., vol.11, no.9, pp.1014-1028, 2002.
- [6] 小野東, “電子透かしとコンテンツ保護, オーム者開発局”, オーム社, 東京, 2001年.
- [7] 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.
- [8] “Dublin Core ”, < <http://www.kanzaki.com/docs/sw/dublin-core.html> >

- [9] “神奈川県図書館蔵書検索”, ” < <http://opaclib.kanagawa-u.ac.jp/> >
- [10] 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.
- [11] 松崎公紀, 檜村雅章, 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.
- [12] 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.
- [13] “画像検索エンジン GazoPa”, < <http://www.gazopa.com/> >
- [14] 田村秀行: “コンピュータ画像処理”, オーム社, 2002年.
- [15] osakana.factory グレースケールのひみつ,
< <http://ofo.jp/osakana/cgtips/grayscale.phtml> >
- [16] Navneet Dalal and Bill Triggs, “Histograms of Oriented Gradients for Human Detection”, Proc.of IEEE Conference on Computer Vision and Pattern Recognition(CVPR), pp.886-893, 2005
- [17] 三谷政昭 “やりなおしのための通信数学”, CQ出版社, 2008年.
- [18] 伊藤光, 仁木雅, 工藤健慈, 木場俊暁, 金丸隆志, 関根優年, “ウェーブレット変換の多重解像度表現を利用した画像認識システム”.
- [19] Charles E. Jacobs, Adam Finkelstein, David H. Salesin, “Fast-Multiresolution Image Querying”, in Proceedings of SIGGRAPH 95, August, 1995年.
- [20] 藤吉弘亘, “HOG 特徴量と Boosting を用いた人検出”, 2009年

-
- [21] 工藤護, “イラスト著作権保護のための HOG 特徴量による複製検出”, 2012

第6章 質疑応答

- Q. 取り出したいのは輪郭線ですか？

A.HOGの複製検出を行う際、余分な領域も含み特徴量を取り出してしまっていると考えられるので輪郭線の多い領域を先に抽出しその領域に対して複製検出を行うことで精度をあげる。