

大規模データベースからの線画の部分的複製検出

Partial Copy Detection of Line Drawings from a Large-Scale Database

孫 維瀚 黄瀬 浩一

Weihan Sun, Koichi Kise

大阪府立大学大学院工学研究科

Graduate School of Engineering, Osaka Prefecture University

E-mail: sunweihan@m.cs.osakafu-u.ac.jp, kise@cs.osakafu-u.ac.jp

Abstract

部分的複製とはオリジナル素材を部分的に切り出すことによって作成される複製である。不正な利用者は、発見を困難にするため、部分的複製を改変すると共に、他の素材の中に埋め込む形で使うことがある。本稿では、このような場合にも対処可能な複製検出法を提案する。対象としては、漫画のような線画を取り上げる。線画の場合、改変の手法として手書き複製が考えられるが、これが問題をより困難なものとしている。この問題に対処するため、本手法では、著作権保護対象の線画に対する局所特徴量の照合を考える。具体的には、局所領域の抽出手法としてMSERを、特徴量の記述子としてHOGを用いる。また、11,603ページの漫画のデータベースを用いた実験の結果から、提案手法が、印刷線画の部分的複製ばかりでなく、複雑な背景に埋め込まれた手書き線画に対しても有効性があることを示す。

1 はじめに

近年、コンテンツの不正コピーによる著作権の侵害が社会的な問題となっている。この問題を解決するため、不正コピーを自動的に検出する技術が求められている。保護すべきコンテンツには様々なものがあるが、その中でも図面や漫画などの線画は重要なものの一つである。線画は、濃淡のない直線又は曲線だけで描く図形であり、単色が基本となる。線画の不正利用者は、オリジナルをそのまま使うことは希であり、オリジナルの一部(部分的複製)を、自分の線画の一部として使うことが多い。また、漫画などで部分的複製を作成する場合には、オリジナルの切り抜きをそのまま使うだけでなく、手書きで複製するなどの改変を伴う場合が考えられる。それ故に、印刷の複製だけでなく、手書きの複製も考えなければならない。

以上のような問題に対処するための著作権保護技術

には、電子透かしと画像検索という2つがある。

電子透かしは、対象とするコンテンツに著作権の情報を埋め込む技術である。カラー画像のような冗長性の高いコンテンツについては、人間に知覚できないように著作権情報を埋め込むことができる。この中には、幾何学的な変換に対する耐性を持つ手法もある。例えば、Basらは、画像の特徴点をマーカとして利用し、幾何学的な変換に対する不変な電子透かしを提案している[1]。しかし、カラー画像に比べて線画は冗長性の極めて少ないコンテンツであり、人間に知覚できないような埋め込みが困難といえる。

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存しておき、疑わしい画像(被疑画像)を検索質問としてデータベースを検索することにより、複製を検出する。特に画像については、各種変換や部分的複製に対して安定である局所特徴量と呼ばれるものがよく用いられる。この中で、SIFT(Scale-Invariant Feature Transform)[2]は最も良く知られた局所特徴量であり、画像検索への有効性が示されている。また、KeらはSIFTを改良したPCA-SIFT(Principal Component Analysis SIFT)[4]を用い、ある程度改変された画像を手がかりに、そのオリジナルが検出できることを示している[3]。さらに、Kiseらは、PCA-SIFTを用いて大規模画像検索が可能であることを示している[5]。

これまで、画像検索に基づく手法が手書き変換を伴う線画の検索に有効かどうかは示されていなかったが、最近、孫らが一定の有効性がある手法を提案した[6]。この手法では、局所特徴量を用いるという方針のもと、MSER(Maximally Stable Extremal Region)[7]とHOG(Histogram of Oriented Gradients)[8]という2つの技術を組み合わせて画像を索引付けする。ただし、この手法はデータベースの規模を大きくしようとすると、メモリ使用量と処理時間の双方に問題が生じてしまうという欠点を持つものである。

そこで本稿では、文献[6]の索引付け方法を改良する

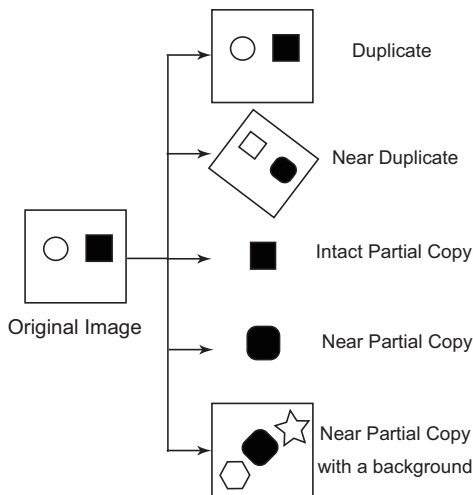


図 1 不正コピーの種類 .

ことにより、大規模化にも耐え得る線画の部分的複製検出法を提案する。大規模化のため、提案手法では、特徴ベクトルの次元削減を行う。具体的には、孫らの手法 [6] で提案された 2916 次元の特徴ベクトルを主成分分析により 100 次元まで削減する。さらに本稿では、1 万ページを超える漫画のデータベースを用いた実験を通して、以下のことを示す。

- 印刷とスキャンを経た部分的複製については、SIFT を用いた複製検出法が極めて有効であり、提案手法も同程度に有効である。
- 手書きによる部分的複製については、SIFT では殆ど検出できないが、提案手法ではある程度検出できる。
- 提案手法では、回転や拡大縮小を受けた部分的複製であっても検出できる。
- SIFT を用いる手法に比べて、提案手法では使用メモリ量が少なくなると共に、検索時間も短くなる。

2 タスク定義

最初に複製検出のタスクを定義しておく。

オリジナルの複製とその利用は、図 1 に示す 5 種類に分類できる。本稿において複製 (duplicate) とは、オリジナルを改変せず直接利用する場合を指す。オリジナル全体に対して改変処理を施した複製は、オリジナルに近い複製 (near duplicate) と呼ばれる。オリジナル全体ではなく、その一部分を用いる場合は、partial copy (部分的複製) と呼ばれる。特に、部分的複製が殆ど改変を受けていない場合 (intact partial copy) と改変を受けている場合 (near partial copy) を区別する。部分的複製の場合には、他の素材を背景とし、その中に埋め込まれて使われることが多い。

複製検出の難易度という観点から見ると、最も容易

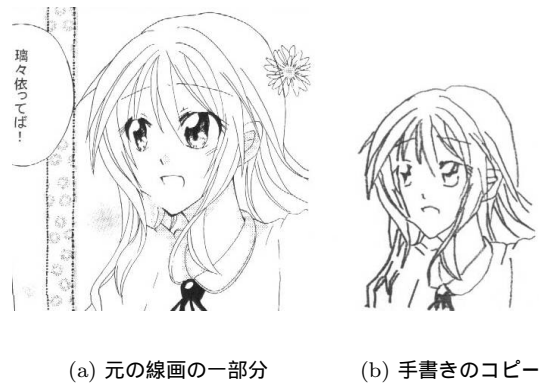


図 2 線画と手書きのコピーの例 .

なのは duplicate であり、図 1 の下に向かうほどより難しくなる。特に、漫画などの線画の場合には、手書きによる near partial copy の可能性がある。例を図 2 に示す。この図の手書き複製では、オリジナルにある台詞の吹き出しと花が略されているとともに、個々の線についても筆記具の違いなどからオリジナルとの間に微妙な差があり、スケールも違う。本研究では、こういった場合であっても複製を検出すること、特に漫画の中に漫画の部分的複製を埋め込むといったような、複雑な背景を伴うような場合であっても検出することを目標とする。

3 提案手法

上に述べたそれぞれのタイプの複製を検出する一つの方策は、各々に適した手法を用いることであろう。しかし、実際に利用する場面を考えると、全種類の不正利用を統一的に検出できる手法の方が利便性が高い。提案手法は、このような観点から開発されたものであり、図 1 のすべてのタイプの不正コピーに適用可能なものである。

3.1 処理の流れ

提案手法の処理過程を図 3 に示す。処理はデータベース処理、クエリ処理という 2 つの部分からなる。

データベース処理では、著作権を保護すべき画像 (著作権保護画像と呼ぶ) を集めて局所特徴量を抽出し、データベースを作る。局所特徴量の抽出は、領域検出器 (region detector) による局所領域の抽出と、特徴量記述子 (feature descriptor) による特徴ベクトルとしての記述からなる。画像のサイズにもよるが、一般に抽出される特徴ベクトルの数は画像あたり数百から数万と膨大になる。データベースには、特徴ベクトルと著作権保護画像の ID の組を記録しておく。

一方、クエリ処理では、著作権侵害の可能性のある画像 (被疑画像と呼ぶ) をクエリ画像としてデータベースに問い合わせ処理を行う。まず、データベース処理

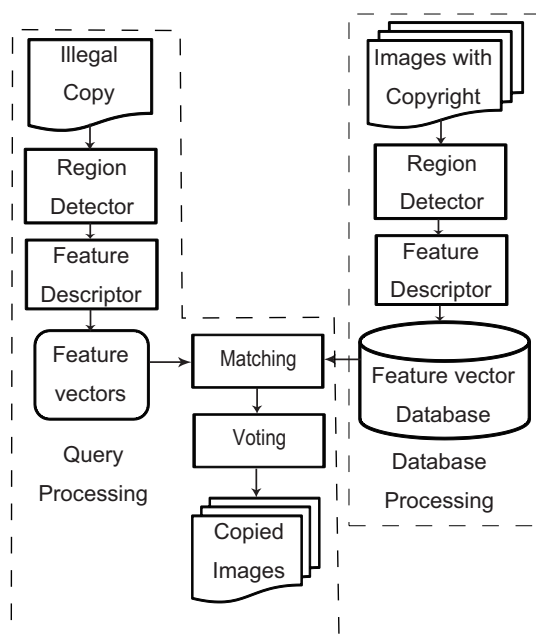


図 3 提案手法の処理過程 .

で用いたものと同じ領域検出器と特徴量記述子を用いて、特徴ベクトルを抽出する。それから、最近傍探索によって、クエリ画像の特徴ベクトルとデータベースの特徴ベクトルを照合し、対応する著作権保護画像に投票を行う。その結果、得票数が最大の画像から一定数を著作権侵害の可能性として報告する。

線画の部分的複製を検出するため、提案手法では、MSER を領域検出器として、HOG を特徴量記述子として用いる。マッチングの時間を短縮するため、ANN(Approximate Nearest Neighbor search)[9]を用いて、最近傍の特徴ベクトルを計算する。さらに、処理時間とメモリ量の双方を削減するため、主成分分析を用いて特徴ベクトルの次元を削減する。

以下、領域検出器、特徴量記述子、次元削減、照合 (Matching) と投票 (Voting) の順に述べていく。

3.2 領域検出器: MSER

Matasらは、wide-baseline マッチングのため、MSER と呼ばれる手法を提案している [7]。多重解像度表現を用いることによって、違う視点から得られた画像を照合している。

概略を以下に述べる。極値領域 (Extremal region) という領域を考える。この領域は、領域内の全部の画素の明暗度は周りの画素の明暗度より高い又は小さいとして定義された領域である。この極値領域の中で、安定な面積が一番大きい領域を MSER という。この領域の共分散行列を対角化して、MSER を楕円領域に変換する。図 4 に抽出された楕円領域の例を示す。この楕円 MSER はアフィン変換と輝度変化に耐性を持つ。手書きの複製では、原画像の線の明暗と濃淡が変化するので、MSER の特性は、手書き複製から安定な領域を



(a) 印刷線画から抽出した MSER



(b) 手書き線画から抽出した MSER

図 4 MSER の例 .

抽出するために重要となる。

線画への適用に際しては次の点に対処が必要となる。

まず、スクリーントーンへの対処である。漫画などの線画では明暗を表現するためスクリーントーンを使うことが多い。これが MSER を求める障害となるので、まずスクリーントーンのパターンを簡単な画像処理によって削除する。具体的には、ガウシアンフィルタで平滑化して消去する。

次に、有効な MSER の取舍選択である。小さい MSER (楕円の面積が小さいもの) から得られる特徴ベクトルは、手書き複製による変動の影響を受けやすいため、複製の識別にはあまり役に立たない。したがって、利用する楕円のサイズを一定以上の大きさを持つものに限る。

最後に、識別性の向上である。複製の識別のためには、楕円領域内に多くの線分が含まれていることが望ましい。そこで本手法では、楕円のサイズを一定割合 (実験では 6 倍) だけ拡大する。それから、楕円の長軸を画像の y 軸に平行となるように回転し、局所領域を得る。図 5 に、拡大、回転された MSER の領域の例を示す。

3.3 領域記述子: HOG

HOG (Histogram of Oriented Gradients) は、人検出のため、Dalalらによって提案された特徴量である [8]。基本的な考え方は、局所の勾配あるいはエッジの方向の分布で、物体の見えや形状を表現するというもので

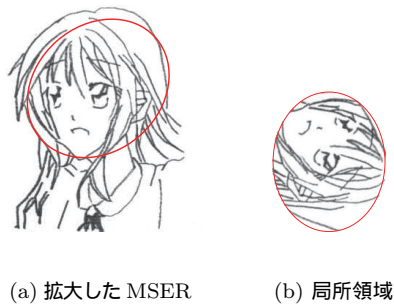


図 5 局所領域の抽出 .

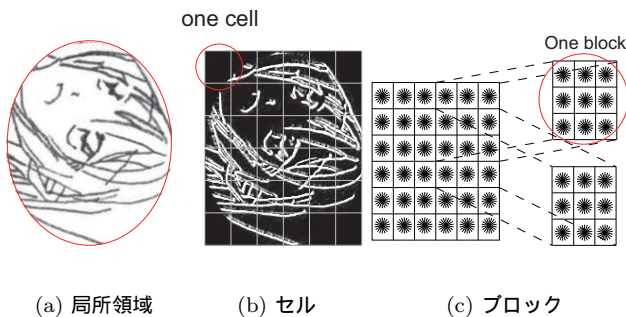


図 6 HOG 特徴量計算領域の構造 .

ある．提案方法では，MSER で定義された各々の局所領域に対して HOG 特徴を抽出する．

図 6(b) に示すように，まず，各ピクセルの輝度から勾配強度と勾配方向を計算する．このとき，勾配方向を 9 方向に離散化する．そして，局所領域を囲む最小矩形を，セルと呼ばれる小領域に分割する．次に，各セル領域において，輝度の勾配方向ヒストグラムを作成して，9 次元のベクトルを得る．それから，図 6(c) に示すように， 3×3 の合計 9 個のセルを組み合わせるブロックを構成する．このとき，ブロックは互いに重なることを許す．次に，セルのベクトルを結合して正規化することによりブロックを表すベクトルを得る．最終的に得られる特徴ベクトルは，全てのブロックのベクトルから構成される高次元ベクトルである．1 つのブロックは 9 個のセルを含み，局所領域からは 6×6 個のブロックが得られるため，特徴ベクトルの次元数は $9 \times 3 \times 3 \times 6 \times 6 = 2916$ 次元となる．

3.4 主成分分析による次元削減

以上のような高次元の特徴ベクトルをそのまま処理に用いると，照合に時間がかかる，記録のために大きなメモリが必要である，などの問題が生じる [6]．この問題に対して，本手法では，主成分分析を適用し，特徴ベクトルの次元を削減することによって対処する．

具体的な処理は以下の通りである．まず，実験に用いる画像とは別の画像を用意し，特徴ベクトルを抽出する．本手法では現在のところ 6 万の特徴ベクトルを

用いている．そして，その特徴ベクトルを主成分分析にかけて次元を削減する．用いる次元数としては，現在，100 次元としている．

3.5 照合と投票による複製検出

上記のプロセスを経て得られる特徴ベクトルは多数となるため，その照合には効率化が必要である．本手法では，近似最近傍探索の一手法である ANN (Approximate Nearest Neighbor)[9] を使って，照合の高速化を図る．ANN では，k-d 木を用いてデータを記録する．探索時には，クエリ画像となる特徴ベクトル q を用いて k-d 木を根から葉に向かって辿ることにより，暫定的な最近傍となる特徴ベクトル p を得る． q と p のユークリッド距離を r とするとき，真の最近傍は q から半径 r の超球中に存在するので，その空間を探索すればよい．ANN ではこのとき，近似のパラメータ ε を用いて縮小した半径 $r/(1+\varepsilon)$ の超球内を探索することにより，高速化を図るものである．縮小のために真の最近傍が得られないというリスクを負うものの，処理は劇的に高速化される．

一般に，照合の結果として得られる対応関係の中には，誤ったものも多数含まれる．誤対応を除去する簡便な方法は，距離の比を用いるものである．いま，クエリ画像の特徴ベクトルを q ，その最近傍となるデータベース中の特徴ベクトルを p_1 ，第 2 位となる特徴ベクトルを p_2 とするとき，

$$\frac{d(q, p_1)}{d(q, p_2)} < T \quad (1)$$

となる，すなわち 2 位との差が顕著である場合にのみ対応するとみなす．ここで， T は閾値である．

以上の処理により，被疑画像から得た各特徴ベクトルについて，対応するデータベース中の特徴ベクトルを得ることができる．データベース中の特徴ベクトルに対しては，どの著作権保護画像から得たものであるかがわかっている．これをその画像に対する投票と考えることにより，最大得票数を得る画像として，対応する著作権保護画像を検出することができる．

ユーザにとって少数の画像を検査することはそれほど苦にならないと考えられるため，本手法では，得票数第 R 位までの画像を出力する．

4 実験

4.1 実験条件

本手法の有効性を検証するため，漫画を対象として実験を行った．

まず，データベースに格納する著作権保護画像としては，10 種類の漫画から得た 11603 ページの単色画像を用いた．

クエリ画像については、次のような手順により作成した。まず、データベースの画像から 101 枚の画像を選択した。次に、各々について、約 1/5 程度の面積を持つ部分領域を切り出した。部分画像には、人物や顔の一部、建物や花の一部が含まれている。

near partial copy としては、この部分領域を用いて次のように 2 種類を作成した。一つは、部分領域の画像を印刷してからスキャンして、再度、画像化することにより得られる部分的複製（以後、印刷と呼ぶ）である。もう一つは、部分領域の線画をまねて手書きで複製し、スキャンして画像化することにより得られる部分的複製（以後、手書きと呼ぶ）である。両者とも、元の部分領域画像の 3/4 の大きさとなるようにスケール変動を加えた。そしてこれらの画像を、データベースには存在しない漫画（背景）に埋め込むことによって、クエリ画像とした。背景はデータベースに登録していない漫画からランダムに選択した。

クエリ画像の例を図 7 に示す。クエリ画像としては、背景となる漫画を持たないものに加えて、部分的複製の面積を基準として、背景の面積を 2 倍、4 倍、10 倍の 3 通りに変化させ、合計 4 通りのクエリ画像を作成した。

画像の照合では、ANN の ϵ を 5 に設定した。実験に用いた計算機は、CPU が Opteron 2.8GHz、メモリが 64GB のものである。

4.2 実験 1

まず、上記のクエリ画像を用いて複製検出実験を行った。本実験では、提案手法のほか、比較手法として SIFT [2] を用いた複製検出法も試した。比較手法では、局所特徴量として SIFT を用いる以外は、基本的には提案手法と同じである。距離の比の閾値 T としては、二つの手法で、0.85 という値を設定した。

実験の結果を図 8 と図 9 に示す。横軸は得票数の順位 R 、縦軸は R 位までに正解が含まれている割合（累積検出率と呼ぶ）を表す。図 8 を見ると、印刷の複製については、SIFT は提案手法を多少上回る結果を得ていることがわかる。両者に共通して言えることは以下の 2 点である。下位の画像まで見ることにより、累積検出率はわずかではあるが改善する。また、背景がない場合の累積検出率が最も高く、背景の面積が大きくなるにつれて精度が少しずつではあるが低下する。

一方、図 9 を見ると、手書きの複製については、SIFT は全く有効に動作しないことがわかる。これは SIFT 特徴が手書きのような微妙ではあるが複雑な変動に対して非常に脆いことを意味している。これに対して、提案手法は、印刷の場合に比べて低下はするものの一定の累積検出率を得た。具体的には、例えば、背景のない手書きの場合で 5 位の累積検出率が 89%、10 倍の背景を伴う場合で 75% という結果であった。これは、MSER+HOG という組み合わせが手書きの変動に対して有効である

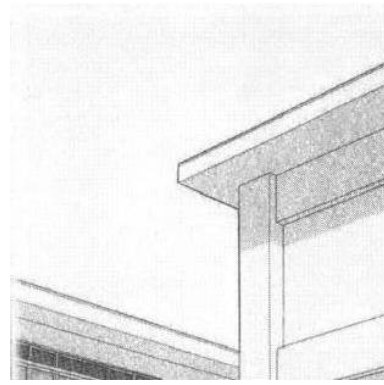


図 10 提案手法で失敗した印刷の部分的複製。



(a) 類似パターンから抽出した局所領域



(b) 文字領域から抽出した局所領域

図 11 失敗原因の例：類似部分。

ことを示している。

印刷の部分的複製については、提案手法で検出に失敗した画像は図 10 に示す一例のみであった。これは、この図のような線が少ない線画からは、検出に有効な特徴ベクトルが抽出できないためと思われる。

手書きの部分的複製について、提案手法で検出に失敗した原因は、次の 2 つであった。

一つは、違う漫画であっても似ている部分が多いことである。図 11(a) に示すように、テクスチャとして似た領域からは類似の特徴ベクトルが得られる。また、図 11(b) に示すように、文字を多く含む部分領域からも類似の特徴ベクトルを得ることが多く、これが原因となり誤投票が増加した。

これらは手書きの部分的複製に対して背景となる部分である。印刷の部分的複製の際には、特徴ベクトルが互いに似ているため、数多くの対応が得られる。その結果、上記のような似ている部分の悪影響がそれほど顕著に表れない。一方、手書きの部分的複製の際には、対応すべき特徴ベクトルの距離が大きくなるため、相

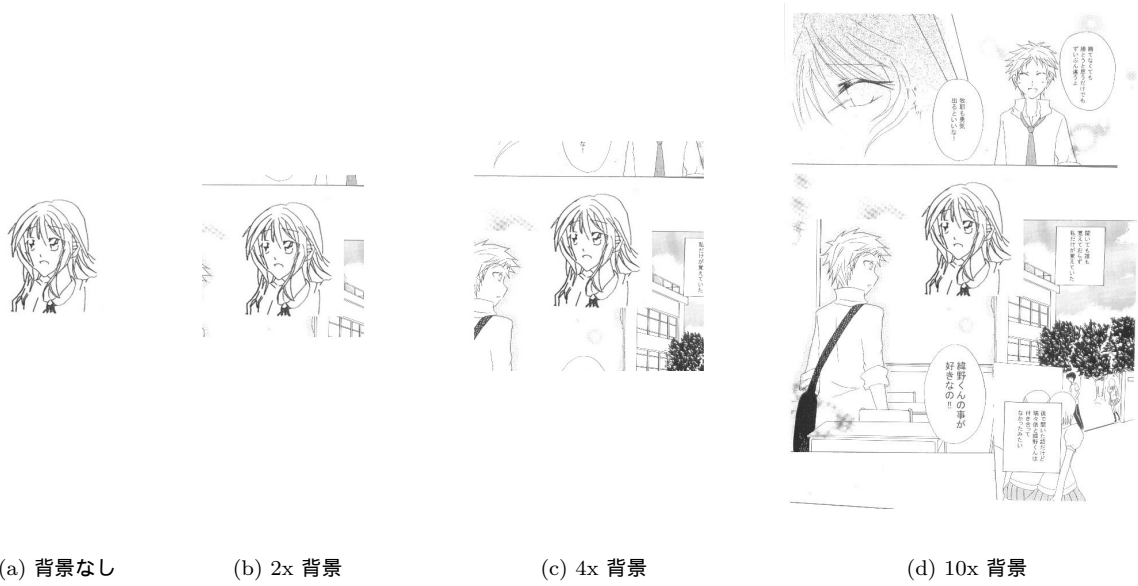


図 7 クエリ画像の例 .

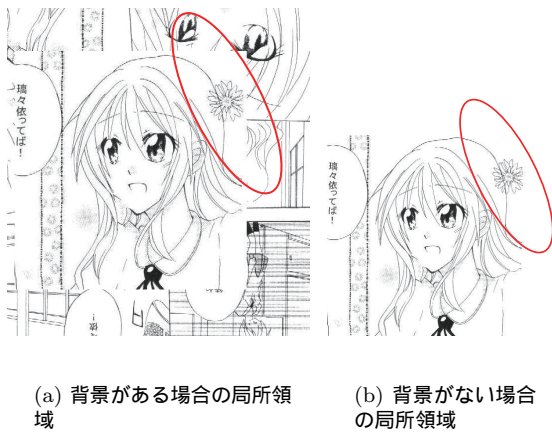


図 12 失敗原因の例：背景の影響 .

対的に上記の似ている部分の悪影響が顕著になる . これが , 失敗の原因となっている .

もう一つの原因は , 局所性の不足によるものである . 図 12(a) に示すように , クエリ画像から得られた局所領域には , 背景の一部にまたがるものがある . このような場合 , データベース中の特徴ベクトルと類似しない特徴ベクトルが得られ , 誤投票につながった .

メモリ量と処理時間については以下の通りである . SIFT を用いた手法では 17GB のメモリを要したのに対して , 提案手法では 8.2GB となり , 半分以下であった . また , 表 1 に二つの手法で要した処理時間を示す . 提案手法では , SIFT を用いた手法より処理時間が 1/6 ~ 1/3 となっており , 高速であることがわかる .

4.3 実験 2

次に , 提案手法に対して , 回転とスケール変換に対する耐性を検査した . クエリ画像としては , 印刷 , 手

表 1 処理時間 (一クエリ画像の平均検索時間) .

size of background	1x	2x	4x	10x
method using SIFT	2,376ms	4,386ms	9,476ms	18,361ms
proposed method	382ms	1,021ms	2,730ms	5,615ms

表 2 回転実験の結果 .

rotation degree	cumulative detection rate ($R = 5$)	
	printed	handwritten
0°	99%	89%
30°	99%	90%
45°	99%	80%

書きの双方について , 背景のない部分的複製を用いた . 具体的には , これらを 30° , 45° に回転したものや , 大きさを 2 倍 (原画像の 3/2 倍) にしたものをを用いた .

5 位累積検出率を表 2 と表 3 に示す . ここで , 表 2 の 0° , 表 3 の 3/4 とは , 実験 1 と同様のクエリ画像を表す . この結果より , 提案手法は , スケール変化や回転についても一定の耐性を持つことがわかる . また , これらの表からは , 回転やスケール変化に対する耐性は , 手書きの場合により得にくいこともわかる . なお , スケール変換の実験において , 拡大した場合の検出精度が高いのは , より多くの局所領域が抽出できるためである . 多くの局所領域が抽出できれば , それだけ投票の回数が増え , 信頼性が増すものと考えられる .

4.4 実験 3

最後に , 検索時間と検索率の関係を示す . 本実験では , 背景なしの手書き部分的複製をクエリ画像とした . 結果を表 4 に示す . この結果より , 近似のパラメータ

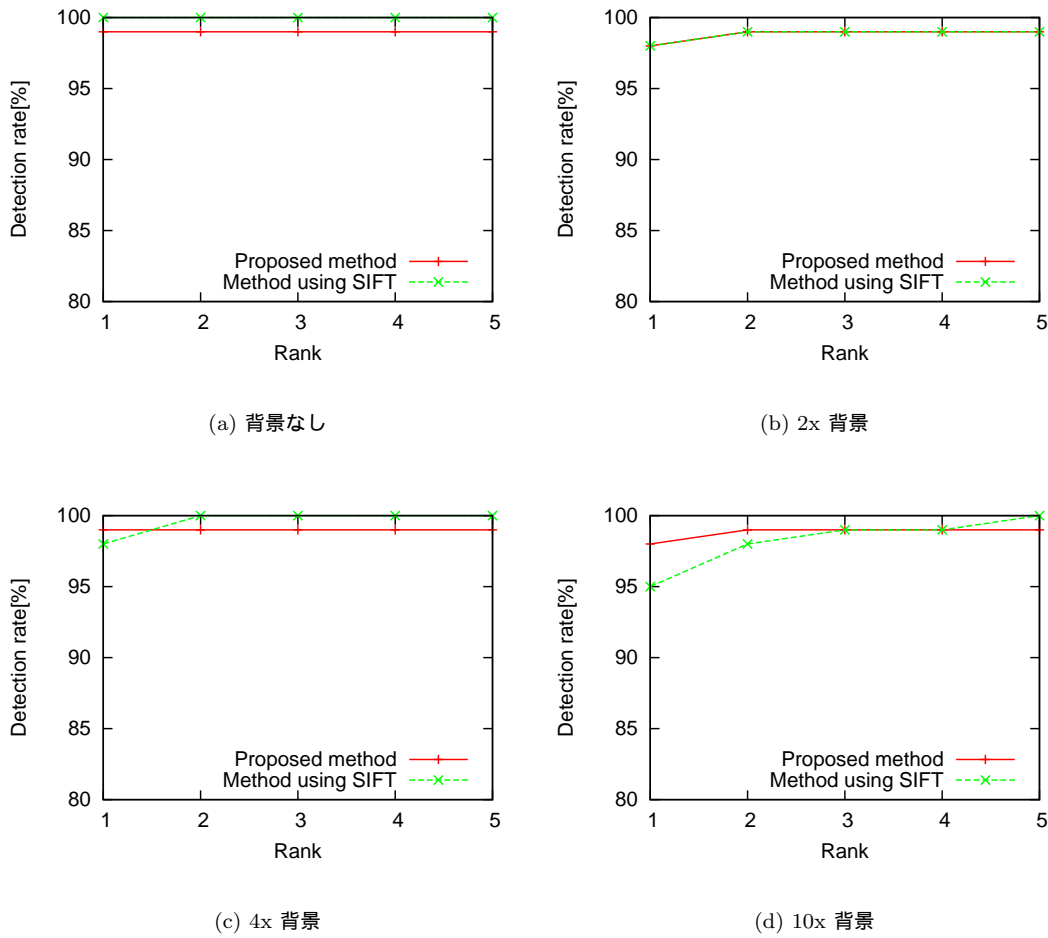


図 8 印刷に対する検出結果 .

表 3 スケール変換実験の結果.

scale	cumulative detection rate ($R = 5$)	
	printed	handwritten
3/4	99%	89%
3/2	99%	98%

ε を小さくして正確さを増せば、検出率は向上するものの、膨大な処理時間が必要であることがわかる。逆に言えば、処理時間を抑えるためには、累積検出率をある程度、犠牲にしなければならない。

どの程度のパラメータの値が良いのかは、処理の目的によるが、この結果は十分満足できるレベルとはいえず、さらに処理時間を短縮するとともに検出率を向上させる仕組みが必要であると考えられる。

5 まとめ

本稿では、局所特徴量の照合による線画の部分的複製の検出手法を提案した。提案手法では、MSER で抽出した局所領域を HOG で記述し、主成分分析により

表 4 ε に対する検索時間（一クエリ画像の平均処理時間）と累積検索率 ($R = 5$) .

ε	cumulative detection rate ($R = 5$)	detection time
1	93%	64,995ms
5	89%	336ms
10	84%	84ms

次元削減することにより、複雑な背景を伴う場合でも、印刷した線画と手書きの不正コピーを一定の精度で検出可能である。

SIFT を用いた手法との比較実験により、以下のことがわかった。

1. SIFT を用いる手法、提案手法は、共に印刷した線画の検出に対して高い有効性がある。
2. SIFT を用いる手法は、手書きの線画に対してほぼ無力である。一方、提案手法では、検出率は低下するものの 5 位の累積検出率で 75–89% を得、有効性が示された。

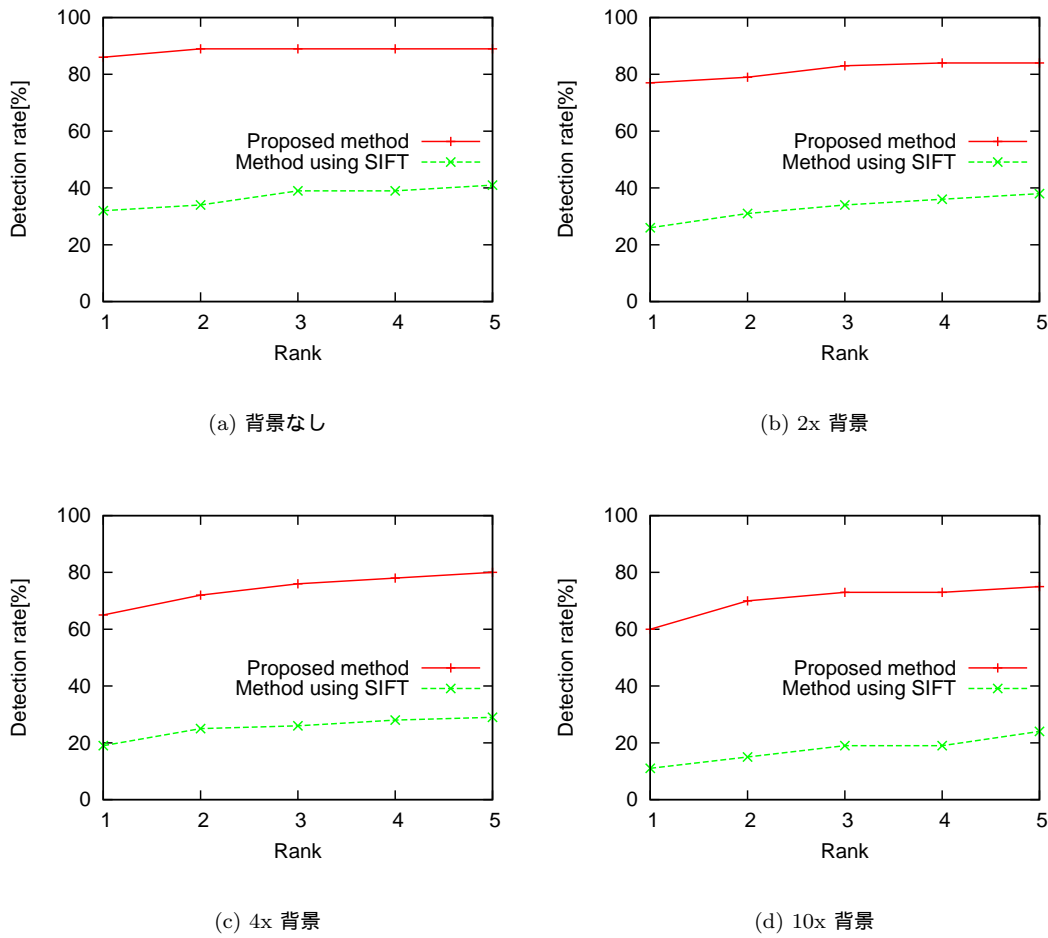


図9 手書きに対する検出結果 .

- 提案手法には，ある程度の回転とスケール変換に対する耐性がある .
- SIFT を用いる手法より，提案手法では，メモリ量が少なく，検索時間も短くなる .

今後の課題は，さらなる検出精度の向上と処理時間の短縮などである .

参考文献

- [1] 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.
- [2] D. G. Lowe, “Distinctive image features from scale-invariant key-points”, Int. J. Comput. Vis. 60(2), pp.91-110, 2004.
- [3] Y. Ke, R. Sukthankar, and L. Hustion, “Efficient near-duplicate detection and sub-image retrieval”, MM, pp. 869-876, 2004.
- [4] Y. Ke and R. Sukthankar, “PCA-SIFT: A more distinctive representation for local image descriptors”, Proc. CVPR, Vol. 2, pp.506-513, 2004.
- [5] K. Kise, K. Noguchi and M. Iwamura, “Simple Representation and Approximate Search of Feature Vectors for Large-Scale Object Recognition”, Proc. BMVC, pp.182-191, 2007.
- [6] 孫 維瀚, 黄瀬 浩一, “局所特徴量の照合による線画の部分的複製検出”, 電子情報通信学会技術研究報告, PRMU2008-228, pp.121-126, 2009.
- [7] J. Matas, O. Chum, M. Urban and T. Pajdla, “Robust Wide Baseline Stereo from Maximally Stable Extremal Regions”, Proc. BMVC, pp.384-393, 2002.
- [8] N. Dalal and B. Triggs, “Histograms of Oriented Gradients for Human Detection”, Proc. IEEE CVPR, vol.1, pp.88-893, 2005.
- [9] S. Arya, D. Mount, R. Silverman and A. Y. Wu, “An optimal algorithm for approximate nearest neighbor searching”, Journal of the ACM, 45, 6, pp.891-923, 1998.