

コミック画像のコマにおける類似性を用いた著作権画像の識別

因 雄亮*¹, 中村 健太郎*², 樋口 政和*³, 川崎 秀二*⁴, 小池 淳*⁵
村上 仁己*⁵

Comic Identification of the Copyrighted Images Using Frame Similarity

Yusuke In*¹, Kentarou Nakamura*², Masakazu Higuchi*³, Syuji Kawasaki*⁴, Atushi Koike*⁵,
Hitomi Murakami*⁵

ABSTRACT : The recent development of the information society leads to many multimedia data on the Web. Especially in Japan, there are a lot of comic images on the Web. However, such a phenomenon causes copyright problems. In order to develop useful information society, the need for comic managing system identifying the author of comic images and ignoring the copied images is increasing. Such a system is indispensable for our future information society. In this paper, we propose a system for identifying the image of comics, and evaluate its performance against conventional methods. Furthermore, we examine the performance with subjective evaluation result by specialists of cartoonist

Keywords : comic, image processing, similarity retrieval, copyright

(Received September 20, 2011)

1. はじめに

情報化社会の進展によって、Web上にはマルチメディアデータである画像が多く存在する。特に日本ではコミック画像が多く存在し、それらの多くは違法にUploadされた画像である。近年、この著作権問題が深刻となっている。そのため、膨大な数の画像を人の識別能力によって判断するのではなく、自動的に蓄積・管理する画像データベースシステムの必要性は、今後ますます高まると思われる。この手段の一つである、自動的なコミックの識別を行うにあたって、有効な検索方法は提示されていない。

コミック画像は同じ作品でも様々な絵柄が存在し、実在シーンと違う点も多い。例えば、白黒画像で表現されていること、濃淡がトーンで表現されていること、そして正面を向いている顔と横を向いている顔において、キ

ャクターは同じでも目の大きさや口の位置や、形が大きく変わることが多々存在する。代表的な、コミック画像の例を図1に示す。



図1 代表的なコミック画像の1ページ。キャラクターが同じでも、目の大きさや口の位置・形が大きく異なる例

*¹: 理工学研究科理工学専攻大学院生

*²: 情報科学科学部生

*³: 情報科学科博士研究員

*⁴: 情報科学科客員研究員

*⁵: 情報科学科教授 (hi-murakami@st.seikei.ac.jp)

これまで、コミックの画像の検出・識別は、コミックのページ単位で処理するのが基本であった。しかしながら、ページ単位での処理では、各コマが小さく表示され、正しい画像の検出・識別特性の低下が想定される。そこで我々は、コマ単位で人物画像の検出・識別を行う手法を考案し、その特性評価を行った。本論文では、まずコマ単位で人物画像の検出・識別を行うための、コマ分割手法を述べる。次に、ここで提案された方式により、ページ単位での手法に比ベドの程度の効率で検出率の向上が図れるか、実験を行った。さらに、これらの結果を、セミプロの漫画家による人物画像の検出・識別力にくらべドの程度達成されたかを、主観評価実験により確認した。

2. コミック画像のカテゴリ識別

一般的に大量の画像からユーザが必要とする画像を検索するには、以下に示す二つの手法が用いられる。

● 検索語に基づく手法

TBIR(Text-Based Image Retrieval)

● 画像特徴量に基づく手法

CBIR (Context-Based Image Retrieval)

TBIRを用いる場合、各画にキーワード情報を付与する必要がある。しかし、キーワードのみでは対象となる画像を表現する事が難しい等の理由で、良い検索結果が得られないことが多い。一方、CBIRでは、画像から得られる特徴量によって検索を行う仕組みにより、画像の類似度を定量化することができる。本稿はこのCBIRを用いてコミック画像を作品別にカテゴリ識別する方式を検討するものである。

G. Csurka らによって提案された画像を識別する手法であるbag-of-keypoints[1]を応用する。コミック画像をページ単位でbag-of-keypointsを用いる手法と、コミック画像にコマ分割[2-3]を行い、コマ単位で用いる手法を比較、検討を行った。また、コマ分割手法は一般に流通しているPCで1ページあたり、144dpiの画像に対し約60msで実行可能であり、多くの処理時間を必要としない。このカテゴリ識別のフローチャートを図2、3に示す。

2.1 コマ分割手法

コマ分割は、情報が存在する領域を検出した後、直線検出を行い、画像を再帰的に二分割する手法を用いている。以下に具体的な手順を述べる。

コミックのコミック画像の情報が存在する領域を輪郭走査によって検出し、得られた閉じた領域の輪郭線を明確にするために、式(1)を用いて重みづけを行う。

$$C(x, y) = \begin{cases} a_0 : \text{輪郭画素} \\ a_1 : \text{他の画素} \end{cases} \quad (1)$$

ただし、(x, y) は画像上のあるピクセルの座標を表す。また、経験的に $a_0=1$, $a_1=0.75$ を用いている。これにより、後のコマ分割線検出において、情報が存在する領域と余白領域を分ける直線が検出されやすくなる。また本稿での輪郭走査とは、ラスタースキャン後に、輪郭追跡処理を行うことをさす

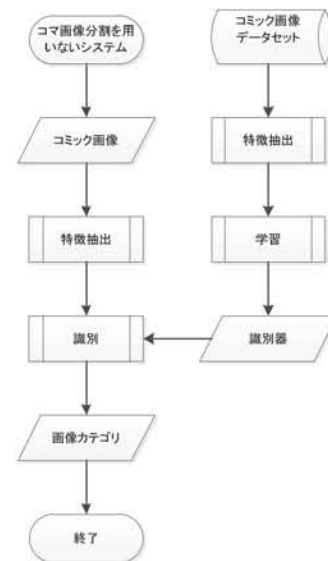


図2 コマ分割を用いない手法でのフローチャート

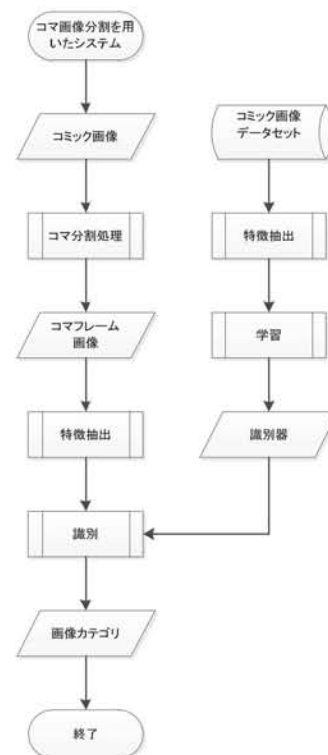


図3 コマ分割を用いた手法でのフローチャート

分割線における領域分割の過程を説明する。コミックの特性上、各コマは線によって区切られている。そのため、入力画像から直線らしい線を検出するため morphological operator を用いて、細線化を行い、ハフ変換 [4] を用いて検出した。ハフ変換ではハフ空間上のある一点 (ρ, θ) が求まり、それに対応する直線は式 (2) を用いて表現される。

$$\rho = x \cos \theta + y \sin \theta \quad (2)$$

ここで ρ は原点から直線に引いた垂線の長さを表し、 θ は垂線と x 軸とのなす角を表す。

検出された直線に対して、複数の直線パラメータの θ が k° 以内かつ、 ρ の差が m 以内、のものを同一の直線と見なし、統一する。ここでは、 $k=1, m=10$ を用いた。これはコマ間の余白領域で極端に狭いものは無視するためである。得られた直線の画素点の集合を $L(\rho, \theta)$ とする。また、 $L(\rho, \theta)$ と、その近傍画素を合わせて $M(\rho, \theta)$ とし、 $M(\rho, \theta)$ に含まれる画素点 (x, y) における濃度勾配の式 (3) を用いて計算した。

$$g_\theta(x, y) = g_x(x, y) \cos \theta + g_y(x, y) \sin \theta \quad (3)$$

$g_x(x, y)$: 水平方向の濃度勾配

$g_y(x, y)$: 垂直方向の濃度勾配

たとえば図のようにコマの外側の余白領域にコマ内の絵柄やセリフの一部がはみ出している場合やコマの形が多角形でない場合、コマ分割線がうまく検出されない。これに対処するために $L(\rho, \theta)$ の近傍 $M(\rho, \theta)$ を考える。

また、経験上、画像の中心付近の直線はコマ分割線になりやすい。そのため、ガウス関数式 (4) により、図 4 に示すとおり、検出された画像上の直線に対し、重みづけを行った。

$$G(\rho) = \exp\left(-\frac{\rho^2}{\sigma^2}\right) \quad (4)$$

σ : 画像の縦幅を四分した長さ

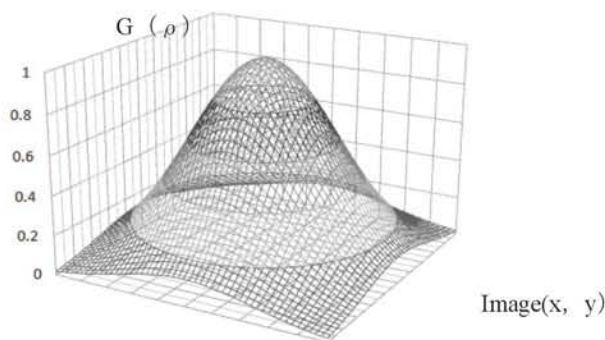


図 4 重みづけのためのガウス関数

この重み付けにより画像の中心付近の直線には大きな重みが付与され、中心から離れた直線には小さな重みが付与される。

以上のパラメータを用いて、直線 $L(\rho, \theta)$ に対する分割線判定値 $A(\rho, \theta)$ を式 (5) で求める。

$$A(\rho, \theta) = G(\rho) \sum_{(x, y) \in M(\rho, \theta)} \{g_\theta(x, y) C(x, y)\} \quad (5)$$

ここで、 θ が 45° 以上かつ 135° 以下ならば横分割線の候補とし、条件を満たさないならば、縦分割線の候補とする。またコマに沿った直線では $C(x, y)$ により、大きな重みが加算されるので、画像の中心付近でコマに沿った直線がコマ分割線候補として検出されることになる。これを繰り返すことにより、高精度なコマ分割を実現できる。

実際のコマ分割処理としては、まずコミック画像全体を入力とし、上記で検出された分割線候補の中で $A(\rho, \theta)$ の値が一番大きな直線で画像を分割する。この際、対象となる分割線が、画像全体を四角形で囲んだ領域を通っているかどうかで分割か非分割かを決定する。横分割線なら、長方形領域の上端と下端、縦分割なら、左端と右端の間を通っているならば、着目している分割線で画像を二分する。そうでなければ、分割せずに終了とする。分割された領域に対して、情報のある領域を再び矩形で囲み、上記のプロセスを再帰的に用いて、分割線を決定し、コミック画像を二分していく。

図 5 に図 1 のコミック画像をコマ分割した結果を示す。黒い枠で囲まれた矩形がそれぞれのコマ領域を示しており、余白を除去し、情報のある領域のみを抽出できた事が確認できる。



図 5 図 1 のコミック画像のページをコマ分割した結果

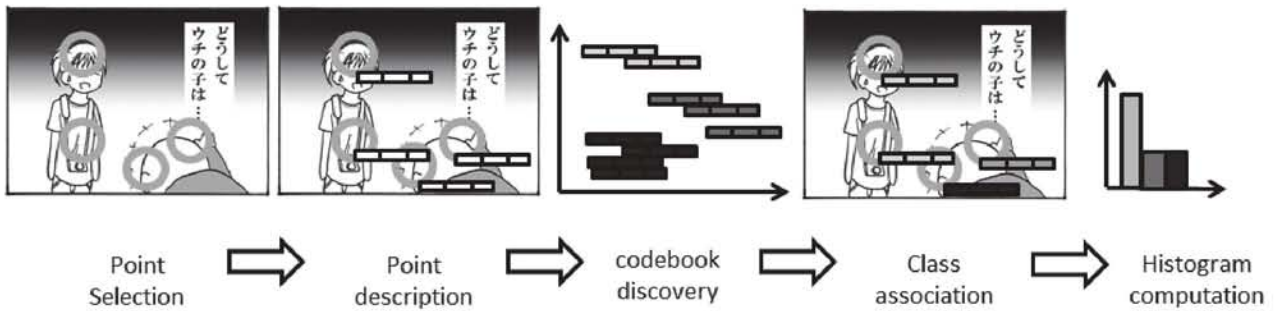


図7 Bag of keypointsモデルのフロー

コマ分割のフローチャートを図8に示す。

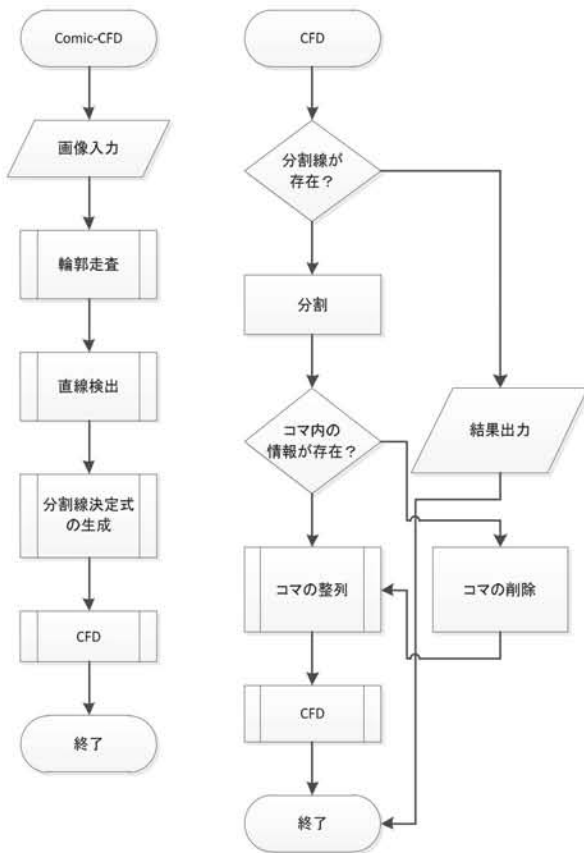


図8 コマ分割システムのフローチャート

2.2 特徴抽出手法

bag-of-keypointsモデルを用いている。これは画像内の位置情報を無視して画像を局所特徴の集合として捉えることで、画像のカテゴリ化を実現する手法である。この、特徴抽出にはSIFT[6]を用いた。SIFT特徴抽出の概要は、以下に述べる。

- ガウシアン差分を最大にすることによりスケールを決定
- 主要な勾配方向として局所方向を決定
- 勾配方向ヒストグラムを計算し、128次元ベクトル

($4 \times 4 \times 8$) で表現

●変化に対してロバストにするため記述子を正規化
この特徴量は比較的スケール変化に強固であり、安定して類似性が検出できる。

量子化するに当たっては、k-means[5]を用いて、クラスタリングを行っている。SIFT特徴点検出、SIFT特徴量記述、codebook生成、クラスの称号、そして今回使用する特徴量の記述といった一連のモデルの生成の流れを示したものを図7に示す。

SIFTは特徴点検出部と特徴点周辺の特徴量記述部に分かれている。特徴点検出部ではDOGを用いて原画像を数段階に縮小したマルチスケール画像群から濃淡勾配の変化が大きい点を複数検出し、特徴量記述部において、その特徴点周辺の局所領域をさらに、 4×4 の領域に分け、それぞれにおいて8方向のヒストグラムを作成するため、128次元特徴ベクトルで表現する。

SIFTによって得られた特徴ベクトルをk-meansにより、クラスタリングを行いcodebookと呼ばれる記述子を生成した。各クラスを元に出現頻度別に投票を行った。また、抽出される特徴点の数は画像により異なるため、出出現頻度を特徴点の数で正規化を行う。codebookはk次元のヒストグラムとして表され、今回はk=1000を用いた。

2.3 識別手法

codebookを元に、学習画像から特徴ベクトルを抽出し、識別器に訓練させる。同様に、カテゴリ化画像からも特徴ベクトルを抽出した。この値を識別器によりどのカテゴリに属するか決定した。識別器にはマルチクラス対応のSVM[7-8]を用いた。

3. 提案方式の結果と評価

3.1 実験データベース

実験用のデータベースには、現在日本の漫画マーケッ

トで広く読まれている 10 種類のコミックから選出した、942 ページを使用した。コミックには、前述した、コマ分割を用いてコミックのページ画像から、4145 枚のコマ画像を得ることができた。この際、100%コマ分割が正しく行われていることを確認している。

3. 2 評価方法

コミック画像のカテゴリ識別の精度を検証した。実験データセットを作品ごとに 10 個に分け、一つのグループを学習データとして利用し、識別器を訓練させた。

評価には、適合率を表す Precision(6)、再現率を表す Recall(7)、精度を表す F-measure(8)をそれぞれ用いた。

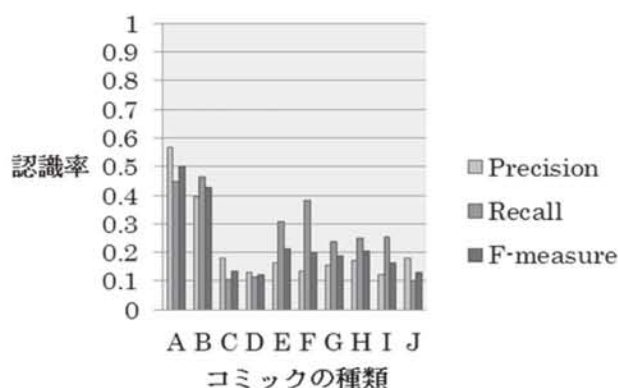
$$Precision = \frac{\text{正しく識別されたデータ集合}}{\text{識別されたデータ集合}} \quad (6)$$

$$Recall = \frac{\text{正しく識別されたデータ集合}}{\text{識別されるべきデータ集合}} \quad (7)$$

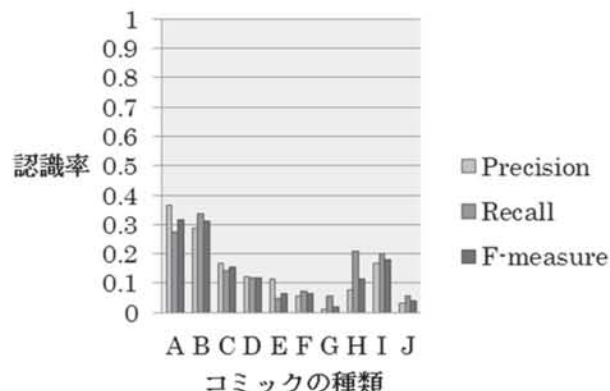
$$F\text{-measure} = \frac{2}{\frac{1}{Recall} + \frac{1}{Precision}} \quad (8)$$

3. 3 実験結果

図 9 にコミック画像のカテゴリ識別の成功率を示す。



(a) ページ単位でのカテゴリ識別成功率



(b) コマ分割単位でのカテゴリ識別成功率

図 9 コマ分割単位でのカテゴリ識別成功率

3. 4 考察

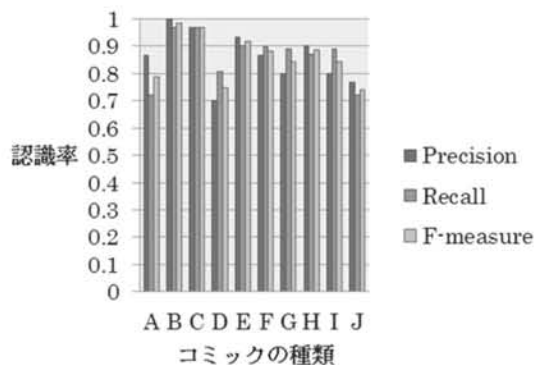
コマ分割を用いない場合においては、F-measureはもっとも大きいもので 31%であり、平均で 19%の識別率であった。コマ分割を用いた場合、F-measureは最も高いもので 50%であり、平均で 30%の識別率であった。コマ分割を用い場合の方が、10~20%ほど高い結果となった。この差の原因としてページ単位でvisual wordsを作成する際には、量子化誤差が生じたことや、特徴抽出の際の特徴点に違いが生じてきたことが原因であると考えられる。

4. セミプロによる主観評価結果との比較

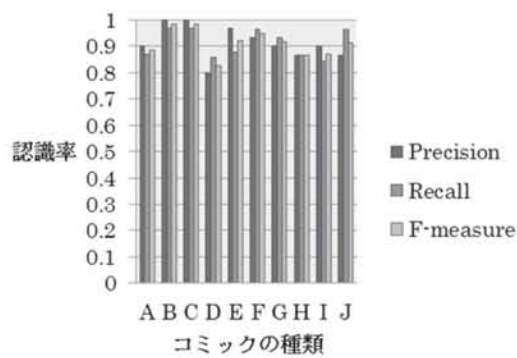
提案したコマ単位の識別手法が、人間が識別評価した結果とどの程度差異があるかを比較するため、成蹊大学の漫画研究会メンバーによって、識別を行った。評価者はコミックに十分知識と経験をもつ熟練者である。そのため、得られた結果は人間が行う識別としてはほぼ最高の結果であると想定される。ここでは実験に使ったコミックから、ページ、コマをそれぞれ 3 つ抽出し、これに対して、10 人が評価を行った。

4. 1 実験結果

図 10, 11 にコミックの作品別識別の主観評価結果を示す。



(a) ページ単位での主観評価識別成功率



(b) コマ分割単位での主観評価識別成功率

図 10 コマ分割単位での主観評価識別成功率

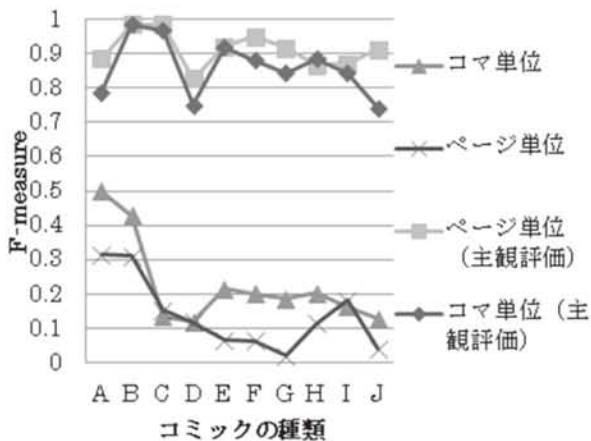


図 11 提案手法と主観表結果の比較

4. 2 考察

用いた画像を十分熟知しているセミプロが評価を行うため、ページ単位とコマ分割探知での識別成功率に関し、大きな差を生ぜず、80%以上の成功率であった。これは、人の場合、コンテンツの内容を理解しながら判別しているためだと、推測できる。具体的には、全体の絵柄だけでなく、吹き出し内にある文字や、人物像から内容を理解している。[9-10]

また、提案手法との差はF-measureの値にして、平均60%程度の差が生じている。

5. まとめ

本稿では、コミックの画像データにおいてコマ分割に基づく作品別のクラスに分類するシステムを提案し、コミック画像の作品別における分類実験を行った。

実験では10種類942ページの画像データセットを用い、カテゴリ分類を行った。コミックのページ画像をそのまま用いた場合、分類率は平均で19%だったが、提案手法を用いることで、30%の分類率まで向上させることができた。さらに、入力画像の無駄な領域を削減することにより、分類速度を25%程度高速化することができた。

セミプロが主観的に分類した場合と比べると、F-measureでは、平均60%低い結果であった。セミプロの行う主観的手法に対しては依然差があるが、コミック画像において、コマ分割を用いたカテゴリ分類手法の有効性を確認できた。

6. 今後の課題

コミック画像に対して、特徴抽出法や量子化手法、さらにはマルチクラス識別手法にさらなる改善を行うと

ともに、吹き出し内にある文字などのコンテンツの内容理解という要素を加味していく必要がある。

本研究は、文科省戦略的研究基盤形成支援事業によって行われた。記して、感謝する。

参考文献

- [1] Csurka, G., Bray, C., Dance, C. and Fan, L. "Visual categorization with bags of keypoints", Workshop on Statistical Learning in Computer Vision, European Conference on Computer Vision, pp. 122, 2004.
- [2] Yusuke In, et al., "Fast frame decomposition and sorting by contour tracing for comic images" [in Japanese], ITE Technical Report Vol. 34, No. 10, pp. 73~76, Feb. 2010.
- [3] Yusuke In, et al. "Using Fast Frame Decomposition and Sorting by Contour Tracing Mobile Phone Comic Imaging System". International Journal of Systems Applications, Engineering & Development Issue 2, Volume 5, pp.216-223. Faro, Portugal, Nov. 3-5. 2010
- [4] Leandro A. F. Fernandes, Manuel M. Oliveira, "Real-time line detection through an improved Hough transform voting scheme", Pattern Recognition (PR), Elsevier, 41:1, 2008, pp. 299-314.
- [5] Lowe, D. "Distinctive Image Features from Scale-Invariant Keypoints", International Journal of Computer Vision, Vol. 60, No. 2, pp. 91-110. 2004.
- [6] T. Joachims, "SVMmulticlass", [http://www.cs.cornell.edu/People/tj/svmlight/svmmulticlass.html](http://www.cs.cornell.edu/People/tj/svm/svmlight/svmmulticlass.html).
- [7] Burdick, HE "Digital Imaging theory and application", McGraw-Hill, New York, 1997.
- [8] T. Joachims, "SVMlight Support Vector Machine", <http://www.cs.cornell.edu/People/tj/svm/svmlight/>.
- [9] Yusuke In, et al. "Similarity Detection of Comic Images ~An Application of Image Local Features for Decomposed Comic Images ~" [in Japanese], ITE Annual Convention Report 10-2, Aug. 2011.
- [10] Yusuke In, et al. "Comic Image Category Classification using Local Features", Proceeding of the 2nd International Conference on Circuits, Systems, Control, Signals. pp. 55-60. Praha, Czech, Sep. 26-29, 2011