Similar Partial Copy Recognition for Line Drawings Using Concentric Multi-Region Histograms of Oriented Gradients

Weihan Sun
Osaka Prefecture University, Japan
sunweihan@m.cs.osakafu-u.ac.jp

Koichi Kise
Osaka Prefecture University, Japan
kise@cs.osakafu-u.ac.jp

Abstract

Since line drawings just employ simple lines to represent objects, similar drawings which represent the same objects can be created easily. Therefore, for protecting the copyright of line drawings, similar partial copy is a problem we have to consider. In this paper, we focus on similar partial copy recognition for line drawings and propose Concentric Multi-Region Histograms of Oriented Gradients (CMR-HOG) to increase the recognition rate. By the experiments of similar comic face recognition, the effectiveness of the proposed method has been proved.

1 Introduction

Line drawings are a type of images that consist of distinct straight and curved lines in monochrome or few colors placed against plain backgrounds. Unlike color pictures, line drawings emphasize form and outline, over color, shading, and texture, to express objects abstractly. Before the development of photography and halftones, they were the standard format for illustrations of print publications. Even nowadays, line drawings, such as logos, caricatures, and comics, are still popular in our daily life. Therefore, there is a requirement for protecting their copyright.

In practice, illegal users do not only duplicate the whole images directly, but also crop their interesting parts, which are called partial copies. Because of simplicities of line drawings, it is easy to duplicate them by handwriting. In addition, by gripping features of abstract objects, similar copies can also be created, which bring more challenge to their copyright protection.

For protecting the copyright of line drawings, we should consider the detection of similar partial copy from complex background. However, since the similar copies contain many differences in detail with the original images, some features which outperform for image retrieval are ineffective for similar copy detection. Such as SIFT (Scale-Invariant Feature Transform) [1], whose high performance for image retrieval has been proved [2], loses its effectiveness for detection of handwritten line drawings [3].

For the detection of illegal copies of line drawings, Sun and Kise have proposed a method using local feature matching and achieved the detection of not only printed but also handwritten partial copies of line drawings [3]. In addition, considering only the regions of interest (ROIs) need to be protected, a method using technique of object detection to detect ROIs and feature matching to recognize similar ROIs has been proposed for detecting the similar partial copies of line drawings [4]. However, it has not achieved a high recognition rate for similar ROIs.

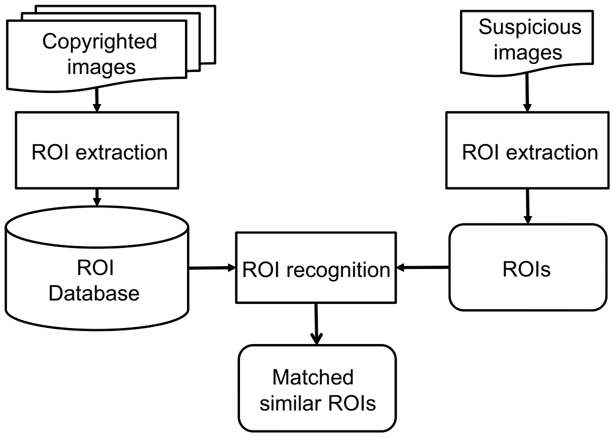


Figure 1. Processing of partial copy detection.

In this paper, we just focus on the similar partial copy recognition and propose Concentric Multi-Region Histograms of Oriented Gradients (CMR-HOG) to increase the ROI recognition rate. In the part of experiment, we prove the effectiveness of the proposed method by the recognition of similar comic faces.

The rest of this paper is arranged as: Section 2 takes a view of the processing of the illegal copy detection, which has been proposed in [4]. Section 3 introduces the proposed method of CMR-HOG. Experiments and results are shown in Section 4 Finally, Section 5 is conclusions and future work.

2 Processing of partial copy detection

The processing of our method is the same as [4], which is shown in Fig. 1. First, copyrighted images are collected. Through the processing of ROI extraction we can extract their ROIs and build a ROI database. On the other hand, suspicious ROIs are extracted from suspicious images by the same processing. Through the ROI recognition, suspicious ROIs are matched with ROIs in the database.

In the part of ROI extraction, we apply Viola-Jones detection framework [5], which has been proved to have the effectiveness for detecting the certain kind of line drawing patterns [4].

The main work of this paper is around the recognition of similar ROIs, which is described in the next section.

3 Similar ROI recognition

3.1 Range of ROI

By applying the Viola-Jones detection framework we can detect some regions which contain one certain category of objects as Fig. 2 shows. Normally the ob-

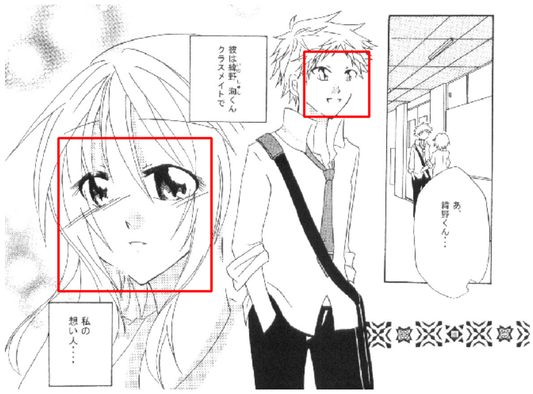


Figure 2. Examples of detected regions by Viola-Jones detection framework.

(a) A. (b) B. (c) C.

Figure 3. Example of similar ROIs. (A) Query image, (B) Similar image focus on small region around the center, (C) Similar image focus on large region around center.(Pattern A and C belong to the same character.)

(a) A. (b) B. (c) C.

Figure 4. Example of similar ROIs. (A) Query image, (B) Similar image focus on small region around the center, (C) Similar image focus on large region around center. (Pattern A and B belong to the same character.)

jects are in the center of these regions. However, since detected regions converge to some general features of one category of objects, it is difficult to classify them only considering the information extracted from detected regions. Therefore, we should enlarge the range of detected regions as our ROIs. With the increase of ranges, we can get more discriminative information. As shown in Fig. 3, if only considering a small region around the center of the images, we may get the answer of Fig. 3(b) as the similar matching with Fig. 3(a) since they have the similar face expressions, and if enlarging the ROI, Fig. 3(c) can be reported by considering the feature from hair style. However, larger regions can also increase noises. For example, as Fig. 4 shows, although Fig. 4(a) and Fig. 4(b) express the faces of the same character, the character around the main character of Fig. 4(b) decreases the similarities with Fig. 4(a). In contrast to it, Fig. 4(c) is more similar with Fig. 4(a) by considering the contours of the characters.

In addition, for similar object recognition, some sta-

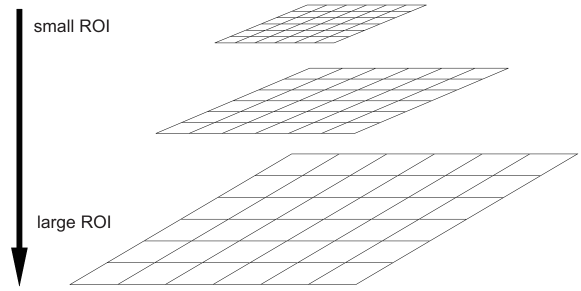
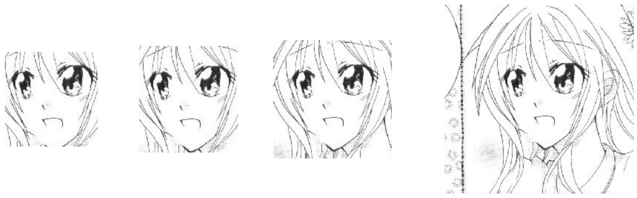


Figure 5. Concentric Multi-Region model to describe object.



(a) n=1 (b) n=2 (c) n=3 (d) n=8

Figure 6. ROI examples of different layers.

ble features should be described in detail, and unstable ones should be described coarsely. Usually the features near the centers of the objects are more stable than the ones far from the centers. For example, as shown in Fig. 4, the glasses of the character in Fig. 4(a) is the key to classify them. Although the glasses is much smaller than the character around in Fig. 4(b), it is more important. Therefore, we propose a Concentric Multi-Region model to describe the object for similar ROI recognition. As shown in Fig. 5, it has a pyramid structure, in which one layer represents one ROI, and the ranges of ROIs are increasing from the top to the bottom. Instead of searching the most discriminative ROI among them, we propose to search the most discriminative features from each layer and combine them into a new kind of feature.

The region detected by Viola-Jones method is treated as a unit ROI which is the top layer of the pyramid. The sides of under layers are set as $1 + (n-1)/8$ times as unit ones, which n represents the layer's level. Fig. 6 shows the examples of different layers. In this research we applied a 8 layers' pyramid, and get 8 ROIs for each detected region.

3.2 Feature descriptor

As the feature descriptor, we apply Histograms of Oriented Gradients (HOG) [6] to describe each ROI. As shown in Fig. 7, we first calculate the gradient magnitude and orientation at each pixel, and divide each ROI into 8×8 cells evenly. Then, the gradient orientation are quantized into 6 bins. For each cell, we calculate the gradient orientation histogram based on the gradient magnitude. After that, cells are combined into overlapped blocks as 3×3 cells per block. By normalizing the features in blocks we obtain $6 \times 3 \times 3 \times 6 \times 6 = 1,944$ HOG features for one ROI. These features are combined into a HOG feature vector which contains $1,944$ dimension.

Although the parameters of feature descriptor are the same for different ROIs, the fineness of descriptions are different because the sizes of ROIs are different.

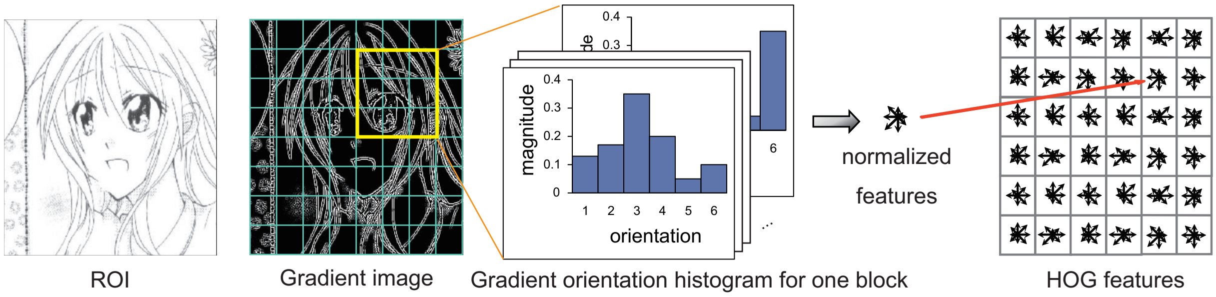


Figure 7. Calculation of HOG features.

Comparing to cells in bottom layers of the pyramid, cells in the top layers contain less pixels which cause a finer description.

To increase the calculation speed of HOG vectors, we apply the integral image to calculate the gradient orientation histograms as in [7]. For each bin we calculate one integral image of gradient magnitude. By 3 times' calculations, we can get the sum value of one bin for any cell. Therefore, the orientation histogram of any cell requires $3 \times 6 = 18$ times' calculations.

3.3 Feature selection

To select the most discriminative features, we apply the Adaboost algorithm [8]. AdaBoost is a machine learning method to combine weak classifiers into a strong classifier by an iterative algorithm. For each HOG feature, we build a decision tree as our weak classifier. By Adaboost method, we evaluate effectiveness of the features and choose the most discriminative ones for each ROI. The final feature vector is a concatenation of these selected HOG features, which is named as Concentric Multi-Region HOG vector (CMR-HOG vector). In this research, we applied the 300 features for each ROI, therefore, the final CMR-HOG vector contains $2,400$ dimensions.

3.4 Matching

In the part of matching, we just calculate the distance between the CMR-HOG vectors, and match the nearest pairs.

4 Experiments

To test the performance of our proposed method, we used faces of comic characters as our data. With the comic face detector applied in [4], we got $29,300$ patterns (faces or non-faces) from 23 kinds[1] of comics (Vol.1–3), and built the ROI database. As our queries we collected 221 face patterns from 17 kinds of comics (Vol.4) among the comics of our database. It means our queries are similar with patterns in database but not the same. As our training set for feature selection, we applied 809 faces of 4 characters of one comic and 171 non-faces. As comparing methods, HOG feature vectors ($1,944$ dimensions) of single ROIs were also applied for recognition. The experimental results are shown in Table 1. Only the matchings between the same characters of the same comic are treated as right answers.

Table 1. Recognition experimental results.

Region size	Recognition rate
$n = 1$	67.4%
$n = 2$	72.4%
$n = 3$	73.8%
$n = 4$	76.5%
$n = 5$	75.6%
$n = 6$	79.6%
$n = 7$	76.5%
$n = 8$	74.7%
multi-region	81.0%

Here, n represent the level of layers in pyramid which ROI belongs to. From the result, we can see: (1) for single ROI, $n = 6$ performed best, (2) our proposed method performed better than the best result of single ROIs.

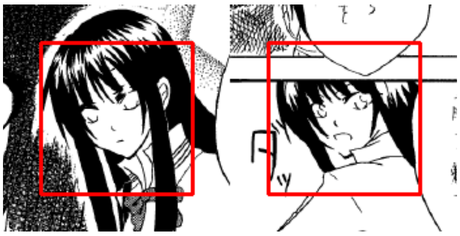
Some examples of successful matching images by the proposed method are shown in Fig. 8. The left side is query pattern and right side is ROI in the database. Only the regions inside the bonding box are considered in this research. We can see our method achieved the detection of similar parts, even if the query pattern contains noises and changes under a certain range.

The main reasons for failures are serious noises and changes which cause similar features with wrong answers. Such as shown in Fig. 9, the right part is an erroreously detected part; the character of left part is closing her eyes while covering her mouth with hand, which make it similar with the right pattern.

Comparing to the best result using single ROI, the proposed method performed better. As Fig. 10 shows, the single ROI just focuses on the pose of characters, which may cause some error matchings. In contrast to it, by considering the detail of small regions around the center, the proposed method compensates such weak points. However, because of the limited training set, it is difficult to cover all discriminative features of comic faces.

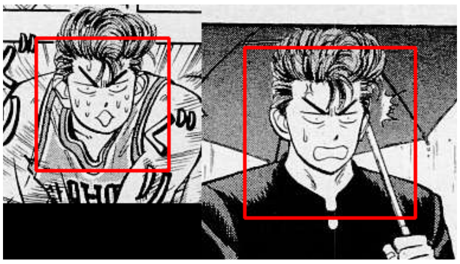
[1]The comics applied in experiments include "20th Century Boys", "Kare Kano", "Neon Genesis Evangelion", "Fighting Spirit", "Hoshin Engi", "H2", "Hunter × Hunter", "Phantom Blood", "Stone Ocean", "Lucky Star", "Master Keaton", "Maison Ikkoku", "Miyuki", "Monster", "Rozen Maiden", "Planetes", "Rosario + Vampire", "Rough", "SLAM DUNK" "Rurouni Kenshin", and 3 comics drawn for this research.

[2]Fig. 8(a): left part is from vol. 4, pp. 192 and right part is from vol. 2, pp. 141 of Naoki Urasawa, "Monster", Shogakukan Inc. Fig. 8(c): left part is from vol. 4, pp. 142 and right part is from vol. 2, pp. 122 of Takehiko Inoue, "SLAM DUNK", Shueisha Publishing Co. Fig. 8(d): left part is from vol. 4, pp.

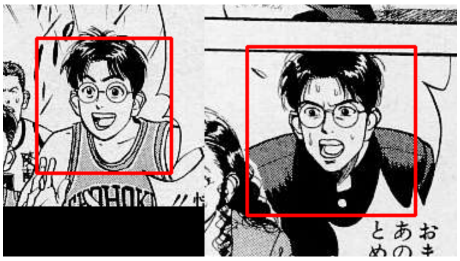


(a) A.

(b) B.



(c) C.



(d) D.

Figure 8. Examples of right matching.[2]

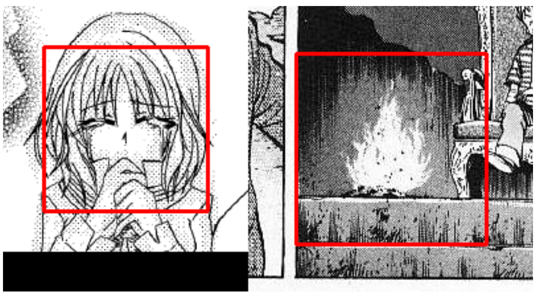


Figure 9. Example of error matching.[3]

5 Conclusions and future work

In this paper, we have focused on the similar partial copy recognition of line drawings and proposed CMR-HOG to increase the recognition rate. From the experimental results of similar comic face recognition, we have proved the effectiveness of the proposed method.

183 and right part is from vol. 2, pp. 186 of Takehiko Inoue, "SLAM DUNK", Shueisha Publishing Co.

[3]Right part is from vol. 3, pp. 114 of Naoki Urasawa, "Monster", Shogakukan Inc.

(a) ROI ($n = 6$).

(b) Multi-Region.

Figure 10. Example of correct matching by the proposed method but failed by comparing methods.

Our future work contains the following points: (1) enlarge the training set, (2) improve the matching method by considering the discrimination of different features, (3) compare the proposed method with other object recognition methods.

References

[1] D. G. Lowe, "Distinctive image features from scale-invariant key-points", International Journal of Computer Vision, vol. 60(2), pp.91–110, 2004.

[2] K.Mikolazcyk, and C. Schmid, "A Performance Evaluation of Local Descriptors", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, No.10, pp.1615–1630, 2005.

[3] W. Sun and K. Kise, "Detecting Printed and Handwritten Partial Copies of Line Drawings Embedded in Complex Backgrounds.", Proceedings of the 10th International Conference on Document Analysis and Recognition, J93-D, No. 6, pp. 909–919 2009.

[4] W. Sun and K. Kise, "Similar Partial Copy Detection of Line Drawings Using a Cascade Classifier and Feature Matching", International Workshop on Computational Forensics, pp. 121–132, 2010.

[5] P. Viola, M. Jones. "Robust Real-Time Face Detection", International Journal of Computer Vision vol. 57(2), pp. 137–154, 2004.

[6] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection", IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 88–893, 2005.

[7] Q. Zhu, S. Avidan, M. Yeh and K. Cheng, "Fast Human Detection Using a Cascade of Histograms of Oriented Gradients", IEEE Conference on Computer Vision and Pattern Recognition, vol. 2, pp, 1491–1498, 2006.

[8] Y. Freund and R. Schapire, "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting", Journal of Computer and System Sciences, vol. 55(1), pp. 119–139, 1997.