

Line-Based Drawing Style Description for Manga Classification

Wei-Ta Chu
National Chung Cheng University
Chiayi, Taiwan
wtchu@cs.ccu.edu.tw

Ying-Chieh Chao
National Chung Cheng University
Chiayi, Taiwan
raff1219@hotmail.com

ABSTRACT

Diversity of drawing styles of mangas can be easily perceived by humans, but are hard to be described in text. We design computational features derived from line segments to describe drawing styles, enabling style classification such as discriminating mangas targeting youth boys and youth girls, and discriminating artworks produced by different artists. With statistical analysis, we found that drawing styles can be effectively characterized by the proposed features such that explicit (e.g., density of line segments) or implicit (e.g., included angles between lines) observations can be made to facilitate various manga style classification.

Categories and Subject Descriptors

I.4.7 [Image Processing and Computer Vision]: Feature Measurement—*Feature representation, Size and shape, Texture*; I.5.2 [Pattern Recognition]: Design Methodology—*Feature evaluation and selection*

General Terms

Design, Measurement, Performance

Keywords

Line features; statistical analysis; manga style classification

1. INTRODUCTION

As large amounts of artistic images have been digitized and shared online, artwork analysis has attracted increasing attention recently. For example, Li et al. [5] designed brushstroke features and conducted statistical analysis to distinguish von Gogh's paintings from his contemporaries'. Garneiro et al. [3] collected a monochromatic artistic image database associated with global semantic annotation, local compositional annotation, and pose annotation. Artistic image annotation and retrieval were formulated as optimization problems, and several features and algorithms were compared to show the computer vision community several interesting challenges. Arora et al. [2] studied classification methods and fea-

tures to achieve artistic genre classification, such as Renaissance, Baroque, and Impressionism.

The aforementioned works were mainly proposed for oil or watercolor paintings. However, *comics*, as another important artistic type, has not been well studied from neither the aspect of feature design nor classification. In addition to text-based metadata, drawing style is an important characteristic for readers to retrieve or manage comic images. Comic style classification is thus an important foundation for users to access comics in a content-based manner. Different from paintings constituted by color regions, characters/objects in comics are mainly concretized by lines. In this paper we specifically focus on monochromatic Japanese comics, or *manga*, which is a major part of Japanese publishing industry and builds a market of 5.5 billion US dollars at least in year 2009. There are no color regions in manga¹, and thus global distributions of conventional visual features are not expected to work well. Human beings can easily recognize different artists' drawing styles, but the mechanism and features for us to do this are not clear. In view of this, we attempt to propose features specifically designed for manga, and investigate how these features benefit manga style analysis. Contributions of this paper are summarized as follows.

- Manga-specific feature design: We design manga-specific features extracted from line segments. Density and magnitude of lines, relationship between nearby lines, and many other line properties constitute an artist's drawing style.
- Feature analysis: We investigate how the proposed features can be used to discriminate different styles by statistical analysis. From this process we reveal implicit characteristics to describe an artist's style or specific topics.
- Applications: With the proposed features and statistical analysis, we enable several applications that have not been developed before. For example, drawing styles and content in youth boy magazines (shonen manga) and in youth girl magazines (shojo manga) are quite different and can be discriminated based on the proposed features. We also investigate how to discriminate different artists' works. These novel applications facilitate large-scale manga browsing and retrieval.

2. PREPROCESSING

Panel segmentation. Given a comic page, we first segment panels in this page so that features can be extracted from individual panels later. The method proposed in [1] is adopted and modified to build this segmentation module. First, connected components in this page are found and labeled, and a minimum bounding box

¹In the following, we will use manga and comic interchangeably.

covering each connected component is then determined. The boxes of sizes smaller than a predefined threshold or are inside other bigger boxes are removed. Finally, overlapped boxes are merged, i.e., a larger minimum bounding box covering all overlapped boxes is determined, and each of the remaining box corresponds to a panel. The top row of Figure 2 shows an example of the process.

Speech balloon detection. Because speech balloons are only slightly relevant to drawing styles, we would like to segment balloon regions and exclude them from the feature extraction process. A speech balloon has the following two characteristics. First, the region around a balloon’s boundary is usually blank. Second, words in the balloon are aligned regularly, either in the left-to-right and then top-to-down order (western reading style) or in the top-to-down and then right-to-left order (Chinese and Japanese reading styles). Based on the initial minimum bounding boxes extracted in the beginning of the panel segmentation module, by the first characteristic we detect candidates of balloons by finding the boxes with blank regions nearby box boundaries. For each candidate box, intensity values of pixels inside are respectively projected onto the horizontal and vertical axes, and a horizontal profile and a vertical profile are obtained. We then calculate autocorrelation between profile values in each profile. If the highest five autocorrelation values are all larger than an adaptive threshold, the corresponding bounding box is viewed to contain a speech balloon. In this work the threshold is set as 1.1 times the average profile value. The bottom row of Figure 2 shows examples of speech balloon detection, in the representation of red rectangles.

Character detection. We cannot describe comic styles without character information. However, detecting characters in comics is not trivial, because current face detection methods are mainly developed for real human’s face and cannot be directly employed. Therefore, we instead detect eyes first, which is relatively easier in comics because eye regions are nearly round with unique texture, and then extend the eye regions outward to determine a region that covers the character’s face and hair. To develop the eye detector, we collect positive and negative eye samples, extract histogram of oriented gradient (HOG) from them, and use the support vector machine (SVM) to build a eye classifier. Given a panel where balloons have been removed, the multi-scale sliding window approach is adopted to detect eyes by the SVM-based eye classifier. Sliding windows of three different sizes, i.e., $0.1W \times 0.1W$, $0.2W \times 0.2W$, and $0.3W \times 0.3W$ (W is width of the panel), are respectively constructed in the detection process. Finally, the region covering a character’s head (mainly including face and hair) is determined by extending eye regions outward.

3. LINE FEATURE EXTRACTION

We can easily observe difference of strokes in mangas of different styles or in different topics. Characteristics of strokes may include density, magnitude, length, orientation, and so on. They are all roughly related to *lines* drew to represent characters. Therefore, based on the character region extracted by the character detection module, we design six line features to describe configurations of strokes and edges.

From the character region, the Canny edge detection algorithm is first adopted to calculate edge magnitude of each pixel, and then the edge linking algorithm implemented by Kovessi [4] is used to detect edges. Because the detected edges would be curved or very long, we further adopt the line segment fitting algorithm, also implemented by Kovessi, to segment edges into straight line segments. The set of straight line segments is the foundation for line feature extraction. Figure 2 shows two sample results of line segment detection, from which we can visually inspect that different drawing



Figure 1: Examples of preprocessing results. Top row, left to right: original comic page, initial connected components, results of panel segmentation. Bottom row: results of speech balloon detection (red rectangles) and character detection (blue rectangles) in two selected panels.

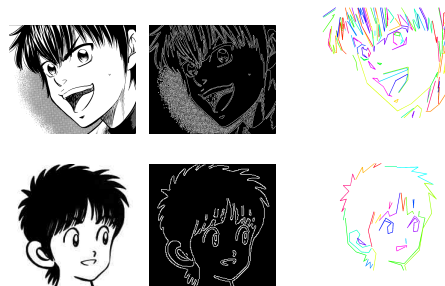


Figure 2: Sample results of line segment detection. Left: source image; middle: edge detection results; right: line segment results, where different line segments are shown in different colors (better seen in color version).

styles may yield different line segment characteristics in terms of number of lines, density, and so on. In the following, we describe details of line features.

Included angle between lines (ℓ_1). For any two spatially adjacent lines, we calculate the included angle between them by $\theta = \arccos \frac{\langle u, v \rangle}{\|u\| \|v\|}$. Include angles between line segments in a panel are then equally quantized into twelve ranges, and thus the feature ℓ_1 of a panel can be represented as a 12-dimensional histogram.

Line orientation (ℓ_2). Orientation of a line segment is defined as the included angle between it and the horizontal axis. Orientations of line segments in a panel are also equally quantized into twelve ranges, and thus the feature ℓ_2 can also be represented as a 12-dimensional orientation histogram.

Density of line segments (ℓ_3). Density of line segments is an important factor of drawing style. We first calculate the centroid $c = (x_c, y_c)$ of each line segment by $x_c = \frac{1}{\|\mathcal{L}\|} \sum_{x_i \in \mathcal{L}} x_i$ and $y_c = \frac{1}{\|\mathcal{L}\|} \sum_{y_i \in \mathcal{L}} y_i$, where \mathcal{L} is the set of pixels in the line segment, and (x_i, y_i) is the coordinate of the i th pixel in the segment. A line segment \mathcal{L}_b is said to be in the neighborhood of the line segment \mathcal{L}_a if the distance between their centroids is less than the adaptive threshold $t = \min(W, H) \times 0.04$, where W and H are

width and height of the panel, respectively. For each line segment in a panel, we calculate the number of lines in its neighborhood, and the information over all line segments are gathered to form the feature ℓ_3 , in the representation a 20-dimensional histogram, i.e., densities of nearby lines are quantized into twenty ranges.

Orientation of nearby lines (ℓ_4). Similar to ℓ_2 , orientations of a line segment’s nearby lines are calculated, and such information over all line segments in a panel are gathered to form ℓ_4 , in the representation of a 12-dimensional orientation histogram (with appropriate normalization).

Number of nearby lines with similar orientation (ℓ_5). To a line segment \mathcal{L} , we calculate the number of its nearby lines that have similar orientation to \mathcal{L} , i.e., their orientations are quantized into the same range. Such information over all line segments in a panel are again gathered to form a 20-dimensional histogram.

Line strength (ℓ_6). Comics designed for boys and girls usually have different styles in strokes, in the representation of density as well as strength of line segments. To capture variations of line strengths, we embed a design in the framework of Canny edge detection, which includes noise reduction, finding edge gradient, non-maximum suppression, and thresholding. At the thresholding stage, all possible thresholds can be thought to constitute a threshold space. This space is equally divided into twenty subspaces, indicating twenty different settings of the high threshold and the low threshold. Twenty threshold settings are respectively applied to the edge gradients corresponding to a panel, yielding twenty different edge maps $\mathcal{E}_1, \mathcal{E}_2, \dots, \mathcal{E}_{21}$, showing all detailed edges (\mathcal{E}_1) to strong edges only (\mathcal{E}_{21}). Suppose that the sums of gradient for $\mathcal{E}_1, \mathcal{E}_2, \dots, \mathcal{E}_{21}$ are E_1, E_2, \dots, E_{21} , respectively, the feature ℓ_6 is a 20-dimensional vector $(\frac{E_2}{E_1}, \frac{E_3}{E_1}, \dots, \frac{E_{21}}{E_1})$.

4. STATISTICAL ANALYSIS

Based on the aforementioned features, we would like to investigate *how the proposed features benefit manga style classification?* In the current work, we focus on two style classification tasks: discriminating mangas in youth boy magazines (shonen manga) and in youth girl magazines (shojo manga), and discriminating artists who produced mangas of the same topic, i.e., baseball games. To examine whether a feature is statistically significant in discriminating mangas of different styles, we set the null hypothesis that two groups have the same average in this feature, and use a two-sided permutation test to compute the p-value [5]. Smaller p-values indicate stronger evidence against the null hypothesis, i.e., statistics of this feature in two groups are statistically different.

Discriminating between shonen mangas and shojo mangas. We collected shonen mangas from *Shonen Jump*, *Shonen Magazine*, and *Shonen Sunday* (three most popular shonen manga magazines in Japan), and collected shojo mangas from *Bessatsu Margaret*, *Bessatsu Friend*, and *Sho-Comi*. Totally 240 panels were extracted from the collected shonen mangas and shojo mangas, respectively, to conduct the following statistical analysis. To analyze line features, features $\ell_1, \ell_2, \dots, \ell_6$ are calculated from each panel, and panels in shonen mangas and shojo mangas are compared by the two-sided permutation test. A feature is said to differ between two groups at significance level α if its corresponding p-value is below a threshold α . In many statistical studies as well as this work, the threshold α is set as 0.05.

Table 1 shows p-values of different features for discriminating shonen mangas and shojo mangas. The features with p-values below 0.05 are especially shown in bold face. We see that include angles (ℓ_1), nearby lines with similar orientations (ℓ_5), and line strength (ℓ_6) are more statistically significant. To get insights from these statistical results, we show sample characters from shonen

Table 1: P-values of different features for discriminating between shonen mangas and shojo mangas.

	ℓ_1	ℓ_2	ℓ_3	ℓ_4	ℓ_5	ℓ_6
p-value	0.040	0.415	0.152	0.430	0.017	0.011



Figure 3: Sample characters from shonen mangas (left three) and shojo mangas (right three).

mangas and shojo mangas in Figure 3, where we can visually see that 1) more lines with similar orientations, especially in hair regions, are in shojo mangas (larger ℓ_5); and 2) more strong strokes, and thus large line strengths, are in shonen mangas (larger ℓ_6). The statistical results match with the visual inspection, and the corresponding features are expected to contribute more in discriminating shonen mangas from shojo mangas.

We respectively randomly selected fifteen panels from shonen mangas and shojo mangas, extracted features, and then calculated the normalized Euclidean distance between features. Figure 4(a) and Figure 4(b) show the distance matrices calculated based on ℓ_5 and ℓ_6 , respectively, where the (i, j) th entry denotes the distance between the i th panel and the j th panel; the first fifteen panels come from shonen mangas, and the last fifteen panels come from shojo mangas. We can see that distances between the panels from the same type of mangas are generally small. This trend can be easily seen especially in Figure 4(b).

Discriminating between artists drawing baseball mangas. We further study whether the proposed features are able to discriminate different artists’ drawing styles. To eliminate the influence of semantics variations, we limit this study to baseball mangas, which are quite popular in Japan and several famous artists have related artworks. In this work, we respectively collected eighty panels from *Ace of Diamond* by Terajima Yuji (TU), *MIX* by Mitsuru Adachi (MA), and *Big Windup* by Mizushima Tsutomu (MT). Two-sided permutation tests are then conducted to examine statistical significance of features in discriminating three pairs of artists, i.e., TU vs. MA, TU vs. MT, and MA vs. MT.

Table 2 shows p-values of different features in discriminating artists. As drawing styles would differ from different perspectives, we see statistically significant features change when we focus on different artist pairs. For example, TU’s and MA’s artworks mainly differ in the features of included angles (ℓ_1), density (ℓ_3), similar orientation (ℓ_5), and line strength (ℓ_6), while TU’s and MT’s artworks differ in the features of line orientation (ℓ_2) and similar orientation (ℓ_5). Sample characters from three artists’ works are shown in Figure 5 for visual inspection of effective features. For example, we can clearly see that MA’s works have low line density, yielding smaller values in ℓ_3 and ℓ_5 . This can also be seen from the line segment results shown in Figure 2, where the top row shows a TU’s character and the bottom row shows a MA’s character.

We respectively randomly selected ten panels from TU’s mangas and MA’s mangas and then calculated the normalized Euclidean distance between features. Figure 4(c) and Figure 4(d) show the distance matrices calculated based on ℓ_3 and ℓ_5 , respectively, where we can see panels can be clearly categorized into two groups. Distances between panels of the same artist’s work are generally low, indicating the consistency of drawing style of an artist.

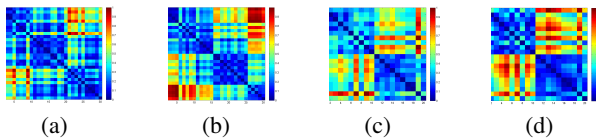


Figure 4: Distance matrices between panels in different styles. Left two: Distance matrices calculated by comparing shonen mangas and shojo mangas based on ℓ_5 and ℓ_6 , respectively. Right two: Distance matrices calculated by comparing TU's mangas and MA's mangas based on ℓ_3 and ℓ_5 , respectively.

Table 2: P-values of different features for discriminating between artists.

	ℓ_1	ℓ_2	ℓ_3	ℓ_4	ℓ_5	ℓ_6
TU vs. MA	0.037	0.183	0.000	0.277	0.000	0.000
TU vs. MT	0.105	0.006	0.075	0.007	0.199	0.074
MA vs. MT	0.325	0.091	0.161	0.061	0.011	0.000

Figure 6 is an interesting illustration showing different artists' drawing styles in terms of line features. We calculate the third order moment (skewness) of features $\ell_1, \ell_2, \dots, \ell_6$. Although the skewness value is just a coarse representation of a high-dimensional histogram, after appropriate shifting and normalization, from Figure 6 we can clearly see difference between artists and demonstrate effectiveness of the proposed features.

5. APPLICATIONS

Discriminating between shonen mangas and shojo mangas.

Based on the proposed features, we construct a classifier to discriminate shonen and shojo mangas based on the support vector machine. According to the statistical test results shown in Table 1, we respectively construct SVM classifiers based on ℓ_1 only, ℓ_5 only, ℓ_6 only, and combination of all proposed features. The five-fold cross validation scheme is adopted in training and testing, and classification accuracies based on different features are shown in Table 3. Although shonen (shojo) manga magazines contain various mangas produced by different artists and thus the intra-class variations are large, Table 3 shows that the proposed features yield encouraging classification performance. The best classification accuracy can be achieved when all features are combined.

Discriminating between artists drawing baseball mangas. According to Table 2, effective features discriminating different artist pairs are different. Therefore, we respectively construct SVM classifiers based on individual features, and construct an SVM classifier based on all features as well. The five-fold cross validation scheme is also used in training and testing. Table 4 shows classification accuracies of discriminating different artist pairs based on different feature settings. Because Ma's drawing style is apparently different from others, classification accuracies of the first and third pair are quite high. Because Ma's artworks often have little variation on line strength, the feature ℓ_6 solely can be a strong clue for classification. On the other hand, to discriminate TU from MT, combining all features yield the best classification performance.

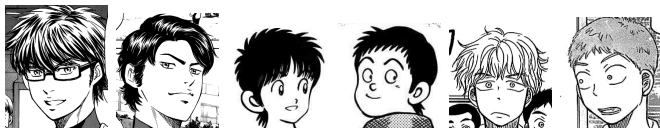


Figure 5: Sample characters from *Ace of Diamond* (left two), *MIX* (middle two), and *Big Windup* (right two).

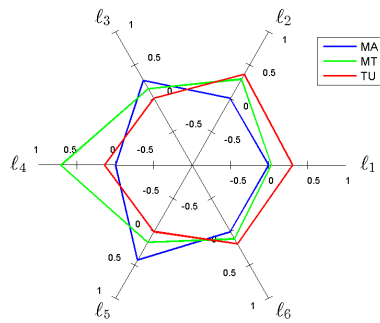


Figure 6: The spider chart showing different artists' drawing styles in terms of line features.

Table 3: Classification accuracy of discriminating between shonen mangas and shojo mangas.

	ℓ_1	ℓ_5	ℓ_6	combine
Accuracy	73.7	70.3	74.5	82.0

6. CONCLUSION

In this paper we propose manga-specific features to describe drawing styles, enabling style classification such as discriminating shonen and shojo mangas, and discriminating artworks produced by different artists. With statistical analysis, features effective to describe various styles are investigated and are used to construct SVM classifiers for style classification. We observe several interesting facts about manga styles, and obtain encouraging classification performance even with simple features. With such study, style-based comics browsing or retrieval would be developed in the future. Furthermore, more elegant features are to be proposed, and manga styles can be discussed in more detail.

Acknowledgement The work was partially supported by the Ministry of Science and Technology in Taiwan under the grants NSC101-2221-E-194-055-MY2 and MOST103-2221-E-194-027-MY3.

7. REFERENCES

- [1] K. Arai and T. Herman. Method for automatic e-comic scene frame extraction for reading comic on mobile devices. In *International Conference on Information Technology*, pages 370–375, 2010.
- [2] R. Arora and A. Elgammal. Towards automated classification of fine-art painting style: a comparative study. In *ICPR*, pages 3541–3544, 2012.
- [3] G. Carneiro, N. de Silva, A. Del Bue, and J. Costeira. Artistic image classification: an analysis on the printart database. In *ECCV*, pages 143–157, 2012.
- [4] P. Kovsi. School of computer science and software engineering. <http://web.csse.uwa.edu.au/>.
- [5] J. Li, L. Yeo, E. Hendriks, and J. Wang. Rhythmic brushstrokes distinguish van gogh from his contemporaries: findings via automated brushstroke extraction. *IEEE TPAMI*, 34(6):1159–1176, 2012.

Table 4: Classification accuracies of discriminating between artists.

	ℓ_1	ℓ_2	ℓ_3	ℓ_4	ℓ_5	ℓ_6	combine
TU vs. MA	74.28	64.20	70.00	62.14	77.85	90.71	90.00
TU vs. MT	65	72.85	62.85	72.85	50.00	69.28	76.11
MA vs. MT	71.42	67.14	64.29	68.57	74.28	86.42	86.11