MANGA-SPECIFIC FEATURES AND LATENT STYLE MODEL FOR MANGA STYLE ANALYSIS

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

• Introduction
• Manga Features
• Style Model
• Applications
  • Style-based Artist Retrieval
  • Style-based Art Movement Retrieval
  • Style-based Artwork Period Retrieval
• Conclusion
Introduction

• Motivation: Many mangas (Japanese comics) are published every year, building a big market and conveying knowledge and culture
• Goal: Novel access scenarios based on *manga styles*
  • Manga styles can be described in a style space constituted by style elements, such as line drawing, screentone, and panel arrangement
Introduction

• Contributions
  • *Manga-specific features*: line features, screentone primitive, panel arrangement features
  • *Latent style model*: describe manga pages as documents, adopt the latent Dirichlet allocation to discover style elements
  • *Applications*: style-based artist retrieval, style-based art movement retrieval, artwork period retrieval
Preprocessing

• Panel segmentation
  • We adopt the panel extraction method implemented by Pang et al. to segment panels from manga pages

Screentone Features

• Screentone is a technique to apply texture or shade to objects or scene.
• Different artists have different habits in using screentone.
• Screentone detection
  • 1. Binarize each pixel by checking intensity
  • 2. Pixels with lower intensity values are applied the erosion and dilation operations
  • 3. Extract patches from the screentone areas
Screentone Features

• Screentone features
  • 1. $s_1$: The ratio of screentone area to the whole panel area
  • 2. $s_2$: bag of screentone primitives
    • Apply the Gabor wavelet transform (4 orientations, 3 scales) to each
      screentone patch
    • Average and standard deviation of transform coefficients in each
      frequency band are concatenated as patch’s texture feature vector
    • Apply the affinity propagation algorithm to cluster feature vectors
      (codebook construction)
    • Quantization each screentone patch into one of the primitives
    • A manga page can then be represented as the bag of screentone
      primitives

Cluster1  Cluster2  Cluster3  Cluster4
Screentone Features

Top row: sample manga pages from three different artists. Bottom row: the BoP distributions corresponding to these artists.
Panel Features

• How several images are placed inside a page also presents artistic styles.

From bounding box of each panel, we extract features to describe characteristics of layout.

1) $p_1$: average panel height (derived from bounding boxes)
2) $p_2$: average panel width
3) $p_3$: standard deviation of $p_1$
4) $p_4$: standard deviation of $p_2$
Panel Features

5) $p_5$: the ratio of total panel area to the whole page
6) $p_6$: average panel area
7) $p_7$: standard deviation of $p_6$
8) $p_8$: average slope of vertical panel boundaries
9) $p_9$: average slope of horizontal panel boundaries
10) $p_{10}$: standard deviation of $p_8$
11) $p_{11}$: standard deviation of $p_9$
Panel Features

Top row: sample manga pages from three different artists.
Bottom row: panel feature distributions corresponding to these pages.
All Features

• The proposed screentone features and panel features are concatenated with line features proposed in [5] to form a feature vector describing a manga page.
• The line features are used to describe a character’s face. We detect the largest and frontal face in a manga page to extract line features.

Latent Style Model

• We develop a style model based on LDA
• Each manga page is described by a feature vector.
• Use PCA to reduce dimensionality from 314 into 20, and employ the K-means clustering algorithm to construct the visual vocabulary.
• Each manga page, therefore, can be represented as a visual word through quantizing the corresponding feature vector.
Latent Style Model

- View several manga pages of the same artist as a document, view each manga page as a word, and view the discovered latent topics as style elements.
- A document $d_i$ is represented as a bag of $N_i$ visual words, denoted by $d_i = \{v_1, v_2, ..., v_{N_i}\}$
- Documents are assumed to be characterized by $K$ style elements. The latent style model assumes the generative process for a corpus consisting of $M$ documents each of lengths (number of words) $N_i$. 

\[ \begin{array}{c}
\alpha \\
\theta_i \\
\{z_{i,j}, v_{i,j}\} \\
N \times M \\
\phi_k \\
\beta
\end{array} \]
Latent Style Model

- Given a set of documents \( D = \{d_1, d_2, \ldots, d_M\} \) with the observed visual words, we can efficiently learn the model by the Gibbs sampling algorithm.
- Style probabilities of a document can be estimated, which enable us to represent a document as a distribution of style elements.
Style Element Distribution

Top: sample manga pages from three different documents.
Bottom: style element distributions corresponding to these documents.
Style-Based Artist Retrieval

- The manga collection of the same artist is randomly divided into subsets, each of which consists of 20 manga pages. Each subset is viewed as a manga document.
- The proposed style model is used to discover style elements of an artist.
- Given a query document, find documents that were produced by the same artist who produced the query.
- MAP@10 = 0.806
Style-Based Artist Retrieval

• Sample manga pages produced by three different artists.
Style-Based Art Movement Retrieval

- An art movement is a tendency or style in art with a specific common philosophy or goal, followed by a group of artists.
- Styles of mangas produced by artists coming from the same studio are correlated.
- Our dataset: 8 artists belonging to 3 art movements
Style-Based Art Movement Retrieval

- View 20 manga pages as a manga document. Learn style elements distribution of each art movement.
- Given a query, find manga documents produced by the artists of the same movement.
- MAP@10=0.854
Style-Based Artwork Period Retrieval

- Some popular mangas had been published for more than twenty years.
- Since the first volume published in 1987, *JoJo’s Bizarre Adventure* has been published for three decades. There are eight parts consisting of more than 110 volumes.

*Fig. 7:* Left to right: sample manga pages from *JoJo’s Bizarre Adventure* Part 1 (1987–1988), Part 3 (1989–1992), and Part 8 (2011–).
Style-Based Artwork Period Retrieval

- Given a query manga document from JoJo8, for example, we would like to retrieve documents that are also from JoJo8.
- MAP@10 = 0.73
- Sample results
Conclusion

• Feature design: screentone features and panel features
• Style model construction: Based on LDA, implicit style elements are discovered.
• Novel applications at the style level
  • Artist retrieval
  • Art movement retrieval
  • Artwork period retrieval
• Future works
  • Large-scale experiments
  • More innovative ways to access mangas
QUESTIONS?

Wei-Ta Chu

National Chung Cheng University, Taiwan
wtchu@ccu.edu.tw