

Classificação de mangás por gênero

ARISTÓTELES TAVARES DE FREITAS JUNIOR, JESSICA PALOMA CARDOSO¹

¹UFMA – Universidade Federal do Maranhão
DEINF – Departamento de Informática
{aristoteles2992, palomasousa15}@gmail.com

Resumo: Mangás são parte importante da cultura pop mundial, tendo uma crescente de lucros com vendas digitais e físicas e um público extenso por possuir vários gêneros. Propõe-se então a classificação dos mangás em shounen e seinen. Segmentando as páginas em frames, é possível procurar patches contendo screentones e utilizar a Gabor wavelet transform e, a partir destas, criar um vetor de características que definirá a página. Pode-se então utilizar o SVM e o cross validation para classificar os mangás com base em seu gênero.

Palavras Chaves: classificação, bag of words, mangá, screentone, wavelet.

Introdução

Mangá é o nome dado a qualquer história em quadrinhos no estilo japonês criado no final do século XVIII. Atualmente, mangás tem seus capítulos lançados em mensais e, posteriormente, coleções de capítulos são agrupados em volumes dedicados, chamados *tankobon*. Suas características únicas os diferem de outras histórias em quadrinhos, como a expressividade de seus personagens, a ordem de leitura diferente da ocidental (a leitura acontece da direita para a esquerda). Existem também as metáforas visuais para os estados emocionais e físicos das personagens, linhas paralelas para indicar o movimento, frames para delimitar os desenhos e os *screentones*. (Figura 1)

Por ser uma indústria enorme e extensa, costuma-se dividir os mangás por gêneros, tendo *shounen*, *shoujo*, *seinen*, *josei* e *kodomo*, *yuri*, *yaoi*, *mecha*, *ecchi*, *hentai*, entre outros.

O mercado de mangás está em uma crescente, principalmente nos meios digitais. Em 2016, houve crescimento de 0,4% do mercado. As vendas de exemplares físicos caiu em 7,4%, comparado aos 27,5% de crescimento do mercado digital [1]. Assim, se faz necessário ferramentas que auxiliem a classificação dos mangás visando melhores algoritmos de busca ou de sugestões, por exemplo.

Este artigo propõe um método de classificação de mangás com base em seu gênero, sendo apresentado da seguinte maneira: Frames e *Screentones*, discutindo

sobre esta técnica e o seu uso no trabalho; Bag of *Screentone*, falando sobre o agrupamento de patches em clusters; Personagens, aplicando HOG e SVM para construir um classificador de faces; Resultados, mostrando e discutindo os resultados obtidos; Conclusão e Referências.



Figura 1: Página de um mangá

Frames e *Screentones*

Os frames são os encapsuladores da história e dos desenhos de um mangá. Para extrair quaisquer características, deve-se inicialmente segmentar os frames de uma página. Xufang Pang, et al. apresentou em [2] um método robusto para extração de frames de mangá. Detectando *connected components* de *pixels* pretos e os agrupando em regiões candidatas, Xufang Pang consegue segmentar até frames sem bordas e dividir blocos de frames ligados por balões. Seu código

para segmentação de frames foi disponibilizado para uso pessoal e em classe. (Figura 2)

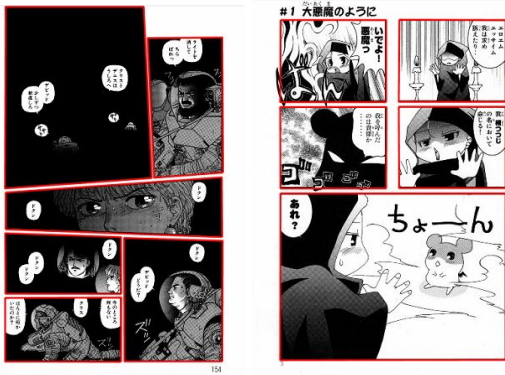


Figura 2: Frames segmentados

Os mangás possuem várias características que os difere de outras histórias em quadrinhos, entretanto buscou-se características pertinentes a um determinado gênero ou demografia de mangás. Wei-Ta Chu, et al. utilizou em [3] características de screentone, de frames e dos personagens para criar um estudo sobre os estilos de mangás, tendo como foco a discriminação do artista, do movimento artístico e do período de publicação dos mangás. Neste artigo optou-se por utilizar apenas o *screentone* como característica.

Screentone é uma técnica muito utilizada em mangás, consistindo de texturas e sombras aplicadas aos desenhos, sendo comumente usadas como alternativa ao hachurado. Convencionalmente, o screentone é adicionado ao mangá através de lâminas pré-impresas, entretanto também pode ser adicionado por meio de computação gráfica. Esta técnica é utilizada pois poupa o tempo do artista, permitindo a aplicação das texturas rapidamente onde uma sombra à mão levaria tempo, e talvez não resultasse em algo visualmente bom.

Existem vários estilos de *screentone* e cada autor possui um estilo característico, que é usado durante toda a história do mangá, e, ainda assim, mangás de mesmo gênero possuem *screentones* parecidos. Independente do autor ou gênero, o *screentone* é aplicado em locais semelhantes como dobras da roupa ou o cabelo das personagens.

Após a segmentação dos frames foi iniciada a busca pelo *screentone*. Achar as áreas correspondentes à busca, começou com a binarização de cada pixel do frame, checando se sua intensidade era maior que o threshold (embora Wei-Ta Chu tenha proposto 200, aqui escolheu-se 250). Os pixels de baixa intensidade tendem a apresentar *screentone*. Sobre esta nova imagem aplicou-se a dilatação, filtrando linhas de movimento. Criou-se uma máscara a partir da inversa da imagem

resultante e, sobre essa e na imagem original, aplicou-se uma operação aritmética. Para finalmente chegar à imagem final, somou-se à imagem gerada da operação aritmética a inversa da máscara. (Figura 3)

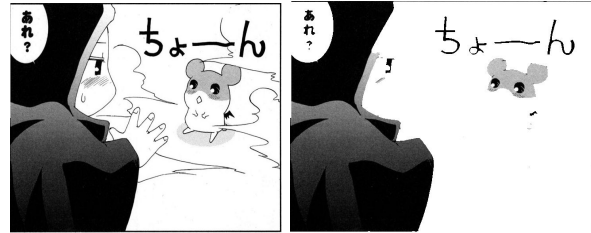


Figura 3: Frame original e *screentone*

Bag of *Screentone*

Duas características foram utilizadas: a razão de área de screentone pela área do frame e a Bag of *screentones*. Para extrair a segunda característica, são usadas as imagens contendo screentone geradas a partir dos frames. Cada imagem desta foi dividida em patches 40x40, sem sobreposição. Se mais de 80% dos pixels em um patch forem de baixa intensidade, então este patch é um patch de *screentone*. (Figura 4)

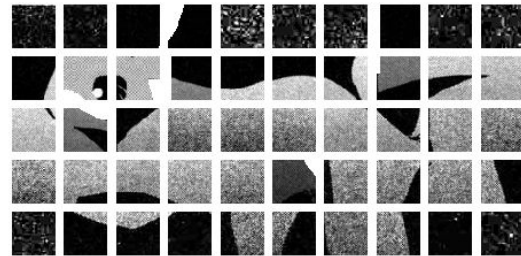


Figura 4: Conjunto de patches de *screentone* obtidos sobre os a área de screentone da Figura 3.

Sobre cada patch (Figura 5a) foi aplicada a *Gabor wavelet transform* com quatro orientações e três escalas, totalizando doze *wavelets* (Figura 5b). As *wavelets* de Gabor minimizam as informações carregadas, nos domínios de tempo e de frequência, servindo de base para a transformada de Gabor. Sobre o resultado, tirou-se a média e o desvio padrão em cada faixa de frequência, totalizando vinte e quatro características por patch de *screentone* de uma página.



Figura 5a: Patch de *screentone*

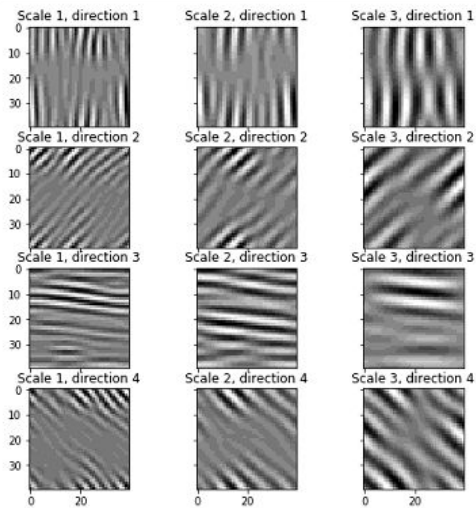


Figura 5b: Resultado da transformada sobre o patch de screentone (Figura 5a).

Após isso foi construído um dicionário visual composto das características de cada patch e, neste dicionário, foi aplicada uma clusterização utilizando o algoritmo *affinity propagation*, agrupando os patches em *clusters* com base em sua similaridade. Do algoritmo, extraímos os centróides (em forma de codebook) dos *clusters*, que seriam os patches que melhor representam seu conjunto. (Figura 6)



Figura 6: Clusters

Finalmente, criou-se um vocabulário a partir do dicionário visual e do codebook, que é o resultado da quantização de cada patch de screentone. Deste vocabulário fez-se um histograma, para armazenar a frequência com que determinada palavra do dicionário aparece.

Personagens

Outro tipo de informação no mangá se encontra no personagem, então há a necessidade de se fazer a extração desses elementos. Para essa tarefa foi desenvolvido um detector de faces, para isso foram coletados amostras negativas e positivas de faces do Manga109 [4], extraído o histograma de gradientes orientados (HOG) deles e usado um SVM para construir um classificador de faces.

Dado um frame, é feita uma binarização em cima da imagem e todas as regiões de pixels conectadas são separadas como um objeto de blob (Figura 7). Tipicamente faces possuem fundo branco, então um método básico de extração de blob é aplicado. É

considerado apenas blobs que correspondam a no mínimo 10% do tamanho do frame. Sobre cada blob é aplicado o classificador a fim de filtrar apenas as regiões de faces.



Figura 7: Detecção de regiões de blobs

Mangas de mesmo gênero tendem a possuir estilos de tracejados parecidos, a forte presença de bordas nas imagens de mangá faz do HOG um descritor desejável para extração de características.

Para cada mangá foram extraídas trinta faces, no conjunto de dez mangás, sendo cinco shounens e cinco seinen. Sobre essas faces é aplicado o histograma de gradientes orientados (HOG), conforme ilustrado na Figura 8.

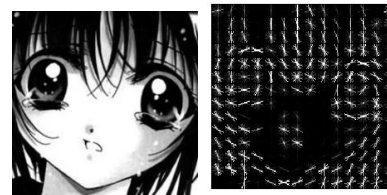


Figura 8: Histograma de gradientes orientados

Experimentos e Resultados

Como explicado anteriormente, a base é composta por amostras de trinta páginas e trinta faces para cada mangá, sendo 150 shounens e 150 seinen, totalizando 300 imagens.

Ambas as características, de screentone e de HOG foram treinadas no classificador SVM linear. Os dados foram divididos na proporção 70% para treino e 30% para teste. Na Tabela 1 é ilustrado os resultados individuais para classificação dos mangás pelo gênero.

Os resultados nos mostram a diferença entre abordagens, sendo que a utilização de faces puramente com o hog, e o uso do Bag de screentones tiveram resultados semelhantes. A combinação de ambos no

então, não obteve sucesso, com o SVM falhando em aprender em conjunto.

	Sensibilidade	Especificidade	Precisão
HOG	68,85%	80,85%	77,5%
Bag screentones	63%	81%	80%
HOG + Bag screentones	100%	0%	50%

Tabela 1: Comparações com resultados aproximados

Conclusão

Foi apresentado neste artigo a extração de características de screentone para a classificação de mangás de acordo com seu gênero. Embora possua bons resultados existe possibilidade de melhorias e novos trabalhos com outros gêneros. Existem poucos trabalhos sobre o estudo de estilo e gênero de mangás, mas com o mercado crescente e novas formas de os ler e obter novos trabalhos neste tema devem surgir no futuro, com melhores e diferentes métodos de estudo.

Referências

- [1] Japan's Manga Market Grows 0.4% in 2016, Digita Sees 27.5% Increase. <https://goo.gl/2t7xBc>, em AnimesNewsNetwork.
- [2] Pang, Xufang, et al. "A robust panel extraction method for manga." *Proceedings of the 22nd ACM international conference on Multimedia*. ACM, 2014.
- [3] Wei-Ta Chu, and Wei-Chung Cheng. "Manga-specific features and latent style model for manga style analysis." *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on*. IEEE, 2016.
- [4] Fujimoto, Azuma, et al. "Manga109 dataset and creation of metadata." *Proceedings of the 1st International Workshop on coMics ANalysis, Processing and Understanding*. ACM, 2016.