Animation character identification from color images

Alexis Vallet

University of Technology of Belfort-Montbéliard

October 8, 2013

- (Semi) supervised classification of animation character images.
- Dealing with variations in character posture, occlusion, drawing style, exaggerations.
- Application domain: web artist communities such as Pixiv, deviantArt.

Figure : Images illustrating variations for a single character.

- Preprocessing: removing outlines, switching color space.
- Segmentation to isolate parts of interest - hair, clothes, face...
- Classification by comparing segmentation against training set.

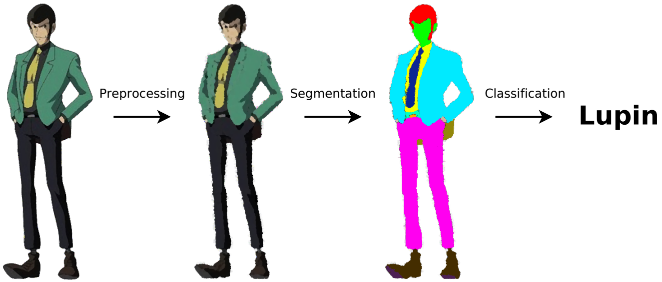
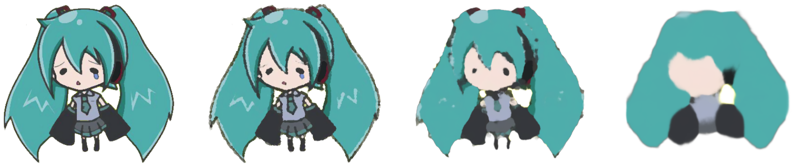


Figure : Diagram depicting how preprocessing, segmentation and classification interact.

- Consider 4 square windows around the pixel to filter.
- Compute mean color and variance in lightness (L in HSL) for each window.
- Assign mean corresponding to smallest variance.



(a) Before filtering. (b) Small window. (c) "good" window. (d) Large window.

Figure : Results of Kuwahara filtering with varying window size.

Felzenszwalb' segmentation [1]

- Graph method based on Kruskal's algorithm.
- Efficient: $O(n \log(n))$ time with 4-connected neighborhood.
- Accurate: neither too "coarse" nor too "fine".
- But depends on a scale parameter k which controls the size of segments.

(a) Original image (b) $k = 100$. (c) $k = 1000$.

- Compute 4/8-connected graph on the pixels of the image.
- Edges weighted by euclid distance in color space:

$$w(u_1, u_2) = ||(l_1, a_1, b_1) - (l_2, a_2, b_2)||$$

- Consider edges in ascending weight order.
- Fuse segments C_1, C_2 related by edge (u_1, u_2) if $w(u_1, u_2) \leq MInt(C_1, C_2)$.

$$MInt(C_1, C_2) = \min(Int(C_1) + \tau(C_1), Int(C_2) + \tau(C_2))$$

Where $Int(C) = \max_{\{u,v\} \in MST(C,E)} w(u, v)$, and $\tau(C) = \frac{k}{\sum_{v \in C} d_v}$.

- Post processing by merging segments with close hue.
- Allows varying segment sizes and non locally connected segments.

(a) Original image.

(b) Before merging.

(c) After merging.

Spectral classification method

- For segmentation S consider features $(f_i : S \to \mathbb{R}_i^q)_{1 \leq i \leq m}$. (average color, gravity center, size...)
- For each feature f_i, compute K-nearest neighbor graph G_i on S with weights $w(u,v) = e^{-\frac{||f_i(S_u) - f_i(S_v)||^2}{\sigma_i^2}}$ and Laplacian L_i.
- Classifying images according to the eigenvectors of the Laplacians.

(a) Example of graph on S

$$L_i(u,v) = \begin{cases} \sum_{u' \text{ adjacent to } u} w(u,u') & \text{if } u = v \\ -w(u,v) & \text{if } u \text{ and } v \text{ are adjacent} \\ 0 & \text{otherwise} \end{cases}$$

(b) Laplacian matrix definition.

Method from Wilson, Hancock, Luo [2], with some changes:

- Compute k eigenvectors corresponding to smallest non zero eigenvalues of Laplacian $L_i \in \mathbb{R}^{n \times n}$.
- Invariance by vertex permutation using symmetric polynomials $P = (P_1, ..., P_n)$.
- Difference in number of vertices by padding with zeros.
- Classifying concatenated vectors E_j with SVM.

$$\Phi = (\sqrt{\lambda_1} e_1 ... \sqrt{\lambda_k} e_k) \in \mathbb{R}^{n \times k}$$

(a) e_j is eigenvector of L_i corresponding to λ_j

$$E_j = \begin{pmatrix} signum(P_1(\Phi_j)) \ln(1 + |P_1(\Phi_j)|) \\ ... \\ signum(P_n(\Phi_j)) \ln(1 + |P_n(\Phi_j)|) \end{pmatrix}$$

(b) Where Φ_j denotes the j^{th} column of Φ

Results and analysis

- Low recognition rate (close to random).
- Graphs do not encode enough information about individual segments.
- Deals poorly with different number of segments.
- Could be salvaged with dimensionality reduction?

Segment matching classification

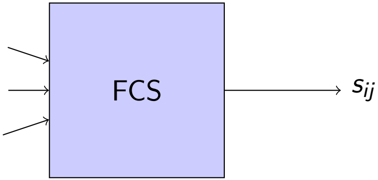
- Consider 3 features for each segment: average $L^*a^*b^*$ color, gravity center, and area.
- Measure similarity between segments using a fuzzy system.
- Find a one to one relation between similar segments of 2 images.

Figure : Original images (left) and corresponding relation (right). Segments with the same color are matched together.

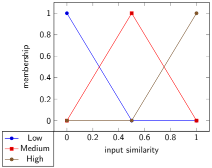
$$c_{ij} = e^{-\frac{||c_i - c_j||}{\sigma_c^2}}$$

$$x_{ij} = e^{-\frac{||x_i - x_j||}{\sigma_x^2}}$$

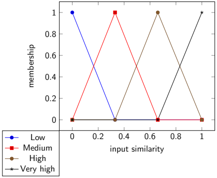
$$a_{ij} = e^{-\frac{|a_i - a_j|}{\sigma_a^2}}$$



(a) Inputs and output.



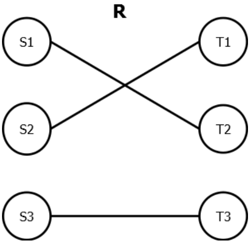
(b) Inputs membership functions.



(c) Output membership functions.

Figure : Segment similarity fuzzy control system.

- Measure overall similarity $sim(S, T)$ between segmentation S and T by sum of matching segments similarity weighted by segment areas.
- Classify by nearest neighbor.

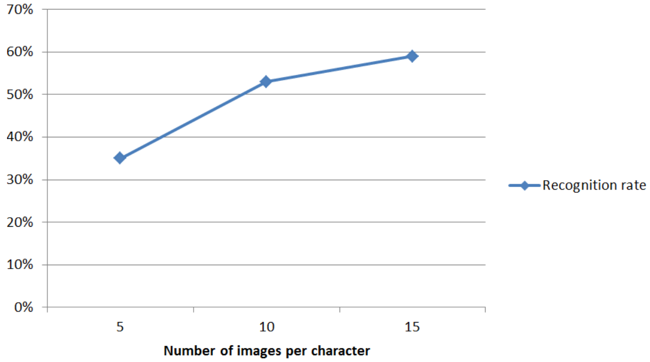


$$sim(S, T) = \sum_{(S_i, T_j) \in R} (|S_i| + |T_j|) s_{ij}$$

Where s_{ij} denotes the similarity between segments S_i and T_j computed by the fuzzy control system.

Results and analysis

- 59% recognition rate for dataset with 12 characters and 15 images per character for a total of 180 images.
- Recognition rate scales well with size of dataset.
- Has trouble with characters sharing similar color palette.



Possible extensions:

- Color palette issues: determining a (possibly non-linear, or high-dimensional) color space ideally separating training data, with some (semi) supervised embedding method [3] ?
- Background extraction: detecting important character features (face, hair, clothes) using method inspired by the face detection algorithm from Viola and Jones [4] ?
- Also using segmentation graph, as in works from Bach and Harchaoui [5] ?

References

Pedro F Felzenszwalb and Daniel P Huttenlocher. "Efficient graph-based image segmentation". In: *International Journal of Computer Vision* 59.2 (2004), pp. 167–181.

Richard C Wilson, Edwin R Hancock, and Bin Luo. "Pattern vectors from algebraic graph theory". In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 27.7 (2005), pp. 1112–1124.

Kiichi URAHAMA. "Semi-supervised classification with spectral projection of multiplicatively modulated similarity data". In: *IEICE transactions on information and systems* 90.9 (2007), pp. 1456–1459.

Paul Viola and Michael J Jones. "Robust real-time face detection". In: *International journal of computer vision* 57.2 (2004), pp. 137–154.

Zaïd Harchaoui and Francis Bach. "Image classification with segmentation graph kernels". In: *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*. IEEE. 2007, pp. 1–8.