Proceedings of the IIEEJ Image Electronics
and Visual Computing Workshop 2012
Kuching, Malaysia, November 21-24, 2012

FACE DETECTION AND FACE RECOGNITION OF CARTOON CHARACTERS USING FEATURE EXTRACTION

Kohei Takayama† Henry Johan‡ Tomoyuki Nishita†

†The University of Tokyo ‡Nanyang Technological University

ABSTRACT

In this paper, we propose methods for face detection and face recognition of cartoon characters which appear in cartoons, comics, games, etc. These methods can be applied to various applications, such as character search, automatic character classification, and character image editing. Previous researches on face detection and face recognition of real people have been abundant, on the other hand, there are little researches on face detection and face recognition of cartoon characters. Using previous methods for real people, cartoon character faces can hardly be detected and recognized because the face features of cartoon characters differ greatly from those of real people in terms of size and shape. Our methods solve these problems by considering the face features of cartoon characters. Using our methods, face detection and face recognition for cartoon characters search can be performed with good accuracy.

1. INTRODUCTION

Face detection is the technique to discover people's face, and face recognition is the technique to detect and identify people's face in images and movies. Many previous methods for face detection and face recognition of real people have been proposed (e.g., [1, 9]). Those methods have been already implemented in our daily devices and have a wide range of applications, such as face detection capability of digital camera or personal authentication by face recognition. On the other hand, there are only a few researches of face detection and face recognition for drawn characters, such as those appearing in cartoons, comics, games, etc. We collectively call these drawn characters as "cartoon characters". Cartoon characters are world-widely spreading, and we believe that face detection and face recognition for cartoon characters are also important research topics because various applications using them can be considered.

There are many applications of face detection and face recognition system of cartoon characters. Character search is one of the most important applications. An image containing a cartoon character is given as an input, and it is possible to automatically search images containing the same cartoon character from Web or database. This function is currently not available in existing search systems (for example, Google Image Search [10]). This function, however, is useful for copyright holder, animator, and illustrator because they can find web pages which contain unapproved use of copyright images, they can investigate whether a cartoon character is popular, and they can confirm whether a similar character exists.

In this paper, we present methods for face detection and face recognition for cartoon characters. Previous techniques for face detection and face recognition for real people's face can hardly be applied for cartoon characters because cartoon characters faces differ from real people faces greatly in respect of organs position and organs size and color shade. Thus, it is necessary to design face detection and face recognition methods specialized in cartoon characters considering the face features of cartoon characters.

In this paper, we first present a technique for face detection for cartoon characters based on feature-based approach. The goal is to identify the positions of faces in the input images. We next present a technique for face recognition for cartoon characters based on model based approach. The task is to extract the features, skin and hair colors as well as the amount of hair, and retrieve similar faces from a database.

2. RELATED WORK

Face detection methods for real people can roughly be classified into two approaches [2]: feature-based approaches and image-based approaches. Feature-based approaches are used for searching faces based on the intrinsic invariant features of faces [1]. These techniques search the location of faces by finding the universal features of face, and defining the conditions which determine a face using the features. On the contrary, image-based approaches search the locations of faces by generating the face definitions automatically from a lot of image data. Imager::Anime Face [3], which is one of image-based approaches, is a face detection method of cartoon characters. This method judges whether face region candidate of input images are faces or non-faces using a function made from 300 million or more image data. In our research, we adapt the feature-based approach when designing our method because feature-based approaches do not require sample images, and the conditions of detection can be changed according to user's preference.

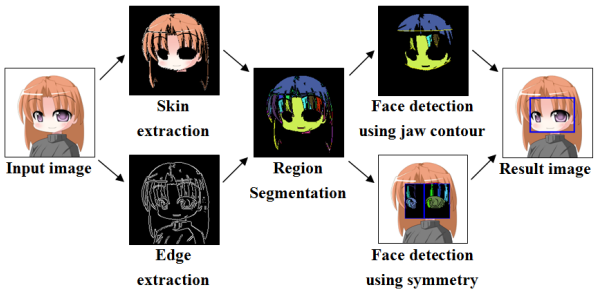


Fig. 1 *Our framework for cartoon character's face detection.*

Face recognition methods of real people can be also classified into two approaches [8]: model-based face recognition and image-based face recognition. Model-based face recognition works by extracting the features of face or interest points. Image-based face recognition recognizes using template matching of the whole face images. In our proposed method, since we have extracted the features of face during the face detection process, we make use of the extracted features and employ the model-based approach for cartoon character face recognition.

3. FACE DETECTION

In designing our method, we consider the characteristic of how the face of a cartoon character is drawn. Since cartoon characters differ in face organs positions and sizes greatly for every character, it is difficult to directly use them as find common features. On the contrary, for cartoon characters, it is easy to extract skin color and edges in the image, so our face detection uses these as features.

Our method works with images with the following conditions. First, input images are not monochrome images but color images, because our technique uses skin color and hair color of characters. Second, the character's skin color has to be near real people. Other colors, such as green faces are not the target of our technique. Third, we target front or somewhat leaning faces. We do not target side or back view of faces.

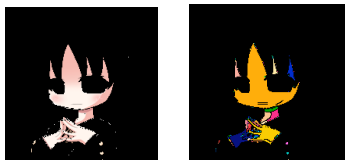
The steps of our proposed method are described in Fig. 1. First, skin color regions and edges are extracted from an input image. Next, region segmentation is performed using them. The resulting regions are considered as face region candidates. Then, based on two criteria, jaw contour and symmetry, we judge whether every face region candidate is a face region. Finally number of faces, their positions, and sizes are returned as a result.

3.1 Skin Color and Edge Extraction

Jamie and Shaogang [5] described that the hue values of the skin color (in HSV color space) of real people concentrate on the region from 6 to 38 degrees. Therefore, taking into consideration some variations, skin color region of cartoon characters is defined as hue of 0 to 40 degrees and value of more than 75% in HSV pixels whose colors are considered as skin colors are extracted from the input image. Moreover, the pixels on the edges (edge pixels) are extracted by using the Canny method [4] (Fig. 2).

(a) Input image (b) Skin color (c) Edge

Fig. 2 *Example of skin color and edge extraction, (a) an input image, (b) skin color extraction from (a), (c) edge extraction from (a).*



(a) Skin color (b) Segmentation

Fig. 3 *Example of face region candidate extraction, (a) skin color pixels, (b) segmentation results with different color for every region.*

3.2 Face Region Candidate Extraction

Region segmentation of skin color pixels is carried out by using skin color pixels and edge pixels. The segmentation is performed by using the flood-fill approach. We start from a skin color pixel and we gradually grow the region until we reach an edge or non-skin color pixels. Region segmentation is performed by repeating this until we process all the skin color pixels. The regions resulted from segmentation are considered as face region candidate. However, regions with very few pixels or too long and narrow are not considered as face candidate regions (Fig. 3).

3.3 Face Detection using Jaw Contour

For every face candidate region, we judge whether it is a face. Jaw contour and symmetry are used as conditions for face detection in our approach. Face region candidates which fulfill the two conditions are considered as a face.

Based on our observation, we realize that in most cartoon the jaw of the characters is clearly visible (not hidden by hair or other elements). We use this finding in our face detection algorithm, extracting the lower contour of the segmented region and trying to fit it as a quadratic curve. If it fits, we assume it corresponds to the jaw contour of a character face, if not, we assume it is not. The detailed algorithm is as follows.

First, the lowest pixel $(x_i, y_i)(i = 1, \ldots, n)$ for every x-coordinate of a region is extracted, and collectively they are considered as n (n is the number of pixels in the width of the region) jaw pixels (Fig. 4 (a)). In most cases, the jaw contour of a cartoon character has a shape which is close to a quadratic curve (Fig. 4 (b)). Since it is thought that these pixels form a jaw contour when this region is a face, these pixels are approximated to a

(a) Jaw pixels *(b) Quadratic curve*

Fig. 4 *Example of face detection using jaw contour, a face region candidate in blue color, (a) extracted jaw pixels in white color, (b) fitted quadratic curve in yellow color.*

quadratic curve using the least-square method.

Coefficient of determination R^2 determines whether the calculated quadratic curve represents the jaw contour of a face. y_a is the average value of y_i. S is the sum of squared residuals. We consider the approximation is good when R^2 is near to 1. When R^2 is larger than 0.85, the contour is judged as maybe a jaw contour.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - y_a)^2}{S}. \quad (1)$$

3.4 Face Detection using Symmetry

Face is symmetrical and especially eyes and hair exist in the left side and right side of the face. First, non-skin color pixels contained in a face region candidate bounding-box are extracted, and the region is segmented using the method described in Section 3.2. The main colors in each region are clustered by using the k-means method [7]. In our experiments, k is 3. When eyes and hair exist symmetrically, the main colors of the regions containing them are almost the same. Therefore we check whether there are two segmented regions in which their two or more color clusters are similar. If such two regions exist, the face region candidate is considered as maybe a face.

4. FACE RECOGNITION

Two purposes for face recognition for cartoon characters in this paper are as follow. The first purpose is the features extraction from the detected face and determination of the individuality of the face. The second purpose is character search. Character search is a task which looks for images in a database which contain the same cartoon character as an input image, and returns the cartoon character images as results. In order to achieve this goal, face recognition is performed using the following technique. The steps of our proposed method are described in Fig. 5. First, the skin color is extracted (same as the skin color extraction in face detection). Second, the features of hair color and hair quantity which are important components of cartoon characters are extracted. Feature vector of a cartoon character's face is determined by these three features. Third, the feature vector of the cartoon character in the input image is compared with other cartoon character's feature vectors in the database, and a similarity value is calculated. We then return a list of characters from the database with high similarity values.

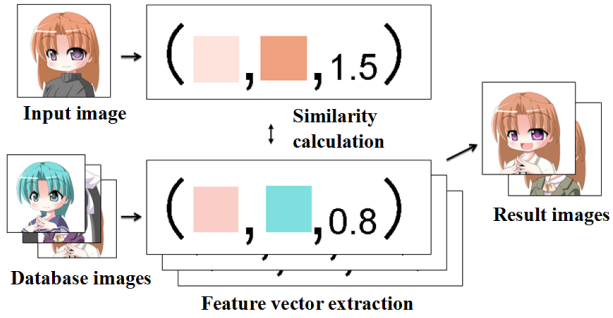


Fig.5 *Our framework for cartoon character's face recognition.*

(a) Input image *(b) Hair extraction*

Fig. 6 *Example of hair extraction, (a) input image, (b) hair extraction from (a).*

4.1 Hair Extraction

Generally we can consider that hair exists in the surroundings of face region in cartoon characters. Therefore, the color which exists in top region and side regions of the face is extracted, and the extracted main color is considered as the hair color. First, we extract the pixels which are located at the top and side regions of the face, and these pixels colors plot to a Lab color space. Second, the plotted colors are clustered by k-means method, and the color with the most number of pixels is considered as hair color. Based on the extracted hair color, the hair region is determined using the following method.

The pixels which exist around the face region and have color similar to the extracted hair color are added to the hair region. We compute the Euclidean distance in the Lab color space to determine whether the pixel color is similar to the extracted hair color. We start from one pixel and repeatedly grow the hair region by checking the neighboring pixels.

The hair color is then re-calculated based on the hair region, and pixels which do not have similar colors to the re-calculated hair color are removed from the hair region. As a result, the hair color and the hair region are determined (Fig. 6).

4.2 Face Similarity Calculation

We describe the method for calculating the similarity of feature vectors of two faces and judging whether these faces are the face of the same character. The distance of two faces is calculated using the skin color, hair color, and hair quantity. When the distance value is small, we determine that the similarity of the two faces is high and the two faces are of the same character face.

For skin color and hair color, the Euclidean distance between two colors in Lab color space are calculated as distances d_1, d_2. For hair quantity, we consider the ratio between the hair region and the face region and, the difference of these values is distance d_3. As shown in Equation (2), our method calculates the total, applying weight (w_1, w_2, w_3) to each distance. The total is the distance of two faces D_{sum}. In this paper, w_1 (skin color) is 0.125, w_2 (hair color) is 0.625, w_3 (hair quantity) is 0.25. The smaller value of D_{sum}, the more two faces are similar.

$$D_{sum} = \sum_{i=1}^{3} w_i d_i . \quad (2)$$

4.3 Character Search

Character search is a task of searching images which contain the same cartoon character as an input image from a database. This is realized with the following technique. First, feature extraction is performed on all the images in the database in advance, and the calculated feature vectors are saved. Second, when a cartoon character image is given as an input, we calculate its feature vector. Then, similarity calculation of the feature vector of the input image and the feature vector of all the images in database is performed, and D_{sum} is calculated. Search results are the images which fulfill one of the following conditions which can be selected by users.

- The specified number of images in an ascending order of D_{sum}
- All the images whose D_{sum} is below a constant value

5. EXPERIMENTS AND RESULTS

5.1 Face Detection

493 various cartoon character images are given as an input and comparative experiment of face detection rate is conducted. Imager::AnimeFace [3] and OpenCV Face Detection [6] are used as candidates for comparison. Both of these are image based approach which requires a large amount of images data beforehand. The method of OpenCV can change the kind of target face by changing the image database. So two kinds of image data, cartoon characters data and real people data, are prepared for OpenCV, and cartoon characters face detection rate is calculated based on each data. The results of face detection rate are shown in Table 1.

True Positive means having detected a right face region. False Positive means having detected wrong regions in addition to a right face region. False Negative means not having detected a right face region. As shown in Table 1, the proposed method achieves a higher rate of True Positive than AnimeFace and OpenCV, and a smaller rate of False Negative. Some of result examples are shown in Fig. 7. The detected face region is marked by a blue square.

Table. 1 *Comparison of face detection rate for every method.*

Method	True Positive	False Positive	False Negative
Proposed Method	74.2%	14.0%	11.8%
Anime Face	71.0%	0.0%	29.0%
OpenCV(Cartoon)	28.4%	0.6%	71.0%
OpenCV(Real)	13.3%	0.2%	86.4%

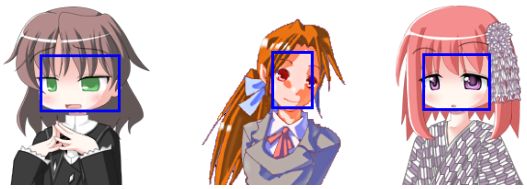


Fig. 7 *Examples of face detection, face region is marked by a blue square.*

5.2 Face Recognition

We performed the following two experiments. First experiment is the case where the database has only one image of the same character as an input image, and second experiment is the case where the database has two or more images of the same character as an input image.

First, the database used in the experiment is created. 300 images of cartoon characters where we performed face detection and feature extraction are prepared. It consists of 150 kinds of various cartoon characters, 100 characters have one image and 80 characters have two or more images. These images are considered as a database. Next, an image of cartoon character is given as an input, and face detection and feature extraction are performed. The distances between the face in the input image and all the faces in the database are calculated. If there are images with high similarity, these images are returned as results.

The following is the result of the first experiment. We prepare 100 input cartoon character images with only one image of the same character in the database, and investigate whether the same character image is returned. As result, 71% of the search input images are success (output images contain the same characters as input images), and 29% of the search are failure (output images contain different characters in 21% of the search, and feature extraction failed in 8% of the search). The reason for failure is that the three features of different characters happen to be similar.

The following is the result of the second experiment. *F-measure* as shown in Equation (3) was used as the measure of quality assessment. This is used for judging the usefulness of the search when the correct search results contain two or more images. R means the number of correct images in search results, N means the number of output images of search results, and C means the number of correct images (the same character images) which exists in the database. The higher *F-measure*, the better performance is the search method.

Table. 2 *The quality assessment of character search using our method.*

Index	Result Value
precision	0.476
recall	0.563
F-measure	0.516

(a) Input image *(b) Output images*

(a) Input image *(b) Output images*

(a) Input image *(b) Output images*

Fig. 8 *Examples of face recognition, (a) are input images, (b) are output images.*

$$precision = \frac{R}{N}.$$
$$recall = \frac{R}{C}. \quad (3)$$
$$F\text{-}measure = \frac{2R}{N+C}.$$

We prepared 80 input cartoon character images with two or more images of the same character in the database, and investigated whether the same character images are returned. The *precision*, *recall*, and *F-measure* were calculated on each input image. Their average values are shown in Table 2. It shows that our method can recognize many different characters. The result examples are shown in Fig. 8. The left images are input images, and the right images are output images from the database with high similarity to the input images. The first and second figures are successful examples where all search results contain the same character as the input images. The third figure is an example of a failure case where search results contain images of a character different from the input image.

6. CONCLUSION AND FUTURE WORK

In this paper, we have proposed a method of cartoon characters face detection by using features, such as skin color, edge, jaw contour, and symmetry. Using only an input image, our method is more accurate than previous methods. We have also proposed a face recognition method where, the face individuality of the cartoon character is determined by using features, such as skin color, hair color, and hair quantity. Using our face recognition method, character search can be performed successfully for many character images.

Future work for face detection is to increase the success rate by adding new conditions for deciding a face, and improving region segmentation. Future work for face recognition is to extract new features, such as eye color or eye shape, and increase the success rate by adding these features for recognizing a face.

REFERENCES

[1] R.Hsu, A.Mohamed, & K.Jain, ”Face detection in color images,” IEEE Trans. PAMI, Vol.24, No.5, 696-706, 2002.
[2] E.Hjelmas & B.K.Low, ”Face Detection: A Survey,” Computer Vision and Image Understanding, Vol.3, No.3, 236-274, 2001.
[3] Imager::AnimeFace anime face detection demo, http://anime.udp.jp/face-detect/
[4] J.Canny, ”A Computational Approach to Edge Detection,” IEEE Trans. PAMI, Vol.8, No.6, 679-698, 1986.
[5] J.Sherrah & S.Gong, ”Skin Color Analysis,” CvOnline, 2006 from http://homepages.inf.ed.ac.uk/cgi/rbf/CVONLINE/entries.pl?TAG288
[6] Face Detection OpenCV Wiki, http://opencv.willowgarage.com/wiki/FaceDetection
[7] S.P.Lloyd, ”Least squares quantization in PCM,” IEEE Trans. ISIT, Vol.28, No.2, 129-137, 1982.
[8] R.Brunelli & T.Poggio, ”Face recognition: Features versus templates,” IEEE Trans. PAMI, Vol.15, No.10, 1042-1052, 1993.
[9] Q.Chen, H.Wu, & M.Yachida ”Face detection by fuzzy pattern matching,” In Proc. ICCV, Vol.5, 591-597, 1995.
[10] Google Image Search, http://www.google.com/imghp