

RESEARCH

Open Access



# Cartoon copyright recognition method based on character personality action

De Li<sup>1</sup>, Lingyu Wang<sup>1</sup> and Xun Jin<sup>1\*</sup>

\*Correspondence:  
xunjn@ybu.edu.cn

<sup>1</sup> Department of Computer  
Science and Technology, Yanbian  
University, Yanji, China

## Abstract

Aiming at the problem of cartoon piracy and plagiarism, this paper proposes a method of cartoon copyright recognition based on character personality actions. This method can be used to compare the original cartoon actions with the action characteristics of pirated or copied cartoons to identify whether there is piracy or plagiarism. Firstly, an image preprocessing scheme for character extraction is designed. GrabCut interactive image segmentation algorithm was used to obtain cartoon characters, and then binarization and morphological processing were performed on the results. Secondly, a feature extraction scheme based on character profile, moving character and character pose is designed. By extracting the perimeter and area of the character contour, the length-to-width ratio of the smallest rectangle and the inclination angle of the contour, the character contour features are obtained. The three-dimensional coordinates are established by the central point position of the cartoon character in the two-dimensional image and the change of the character's zoom in and out, and the character's motion angle characteristics are calculated. By skeletonizing a character to obtain the pose characteristics, and using deburring operation to remove redundant branches, then extract the skeleton joint angle information. Finally, feature fusion is performed on the extracted features. The experimental results show that the proposed method breaks the limitation of the conventional single feature based recognition, and can better extract the character features including contour feature, motion feature and pose feature through multi-feature based extraction, so as to protect the cartoon copyright.

**Keywords:** Personality action, Copyright protection, Cartoon character, Motion feature, Skeleton joints

## 1 Introduction

With the development of the internet, the way of distributing comics has changed dramatically. In the past, paper comics occupied the mainstream of comic market. Nowadays, digital comics become the main way of dissemination of comics. Because of its easy dissemination characteristics, comic industry faces serious infringement problems such as piracy or plagiarism. Piracy is the act of redistributing a work without the permission or authorization of the copyright owner. Plagiarism is the act of copying or altering content to some extent. The diversified forms of expression of comics make it difficult to investigate the liability for infringement, and the frequent

phenomenon of piracy and plagiarism seriously hinders the development of the comic industry. However, due to the insufficient theory of legal protection and the hidden form of infringement by infringers, it is difficult to identify and prove cartoon infringement in China. Therefore, the research of cartoon copyright protection technology is urgent.

Only by improving its innovation and originality can comics attract new users, and only by innovating original works can they constantly refresh the classics and become classics. Originality is always an important prerequisite for the long-term and sustainable development of the comic industry. At the same time, as half of the animation industry, the cartoon industry has provided original materials and market guidance for animation, film and television for a long time. Original cartoon works play an important role in expanding the market of animation, games, films, advertising and other fields.

Cartoon is not only an art form that uses exaggeration, metaphor, symbol and other techniques to describe various phenomena in life or current events, but also plays many functions in communication, such as satire, praise and entertainment. According to the message of comics, comics are divided into philosophy comics, science comics, advertising comics and knowledge comics. The main profit point of comics is not from publishing, but from the spin-offs of the stories and characters they create. As a prominent feature of comic characters, the "individual action" of comic characters is particularly important for the copyright protection of comic characters. "Individual action" reflects the identity, inner activity, thought quality and other characteristics of the characters in the cartoon. It is the soul of the characters in the cartoon and also the painstaking efforts of many cartoon creators. Cartoon copyright originally refers to the right to copy cartoons, but with the change of its dissemination mode, the infringement behavior of infringers is increasingly diversified, piracy and plagiarism undoubtedly bring great difficulties to the determination of infringement. Therefore, this paper proposes a feature extraction method for personality actions of characters in cartoon images, to detect plagiarism by feature similarity. By comparing the features of original actions with those of pirated or copied actions, we can identify the infringement of personality actions and achieve copyright recognition.

In this paper, by studying the image pretreatment and image feature extraction, the characteristics of the character personality action are extracted in the aspects of contour feature, motion feature and pose feature. An interactive foreground extraction method of GrabCut algorithm is used to achieve the integrity extraction of cartoon characters. The extracted characters are binarized, then morphologic corrosion and expansion are carried out. To obtain comprehensive characteristics of personality action, we extract several features including the perimeter and area of the character contour, the length-width ratio and inclination angle of the smallest rectangle, the motion angle of the character in the adjacent cartoon pictures, and the joint angle of the character pose. Finally, feature fusion is performed on the extracted features. This method can be used to compare the characteristics of original cartoon actions with those of pirated actions to identify the existence of plagiarism.

The contributions of this paper can be summarized as follows: (1) we propose a copyright recognition framework for cartoon characters based on character personality action features. (2) With the contour, motion and pose features, we obtained

comprehensive characteristics of character actions. (3) A robust cartoon character recognition is achieved by the similarity detection on the proposed features.

## 2 Related works

### 2.1 Image feature engineering

Image feature extraction is a process of extracting valuable information from images. According to different image types and the purpose of feature extraction, the selection of extracted features will be different, and the specific content extracted from each feature will also be different. Because feature extraction is a very important step in many image algorithms, there are many algorithms about image feature extraction are studied and developed by many scholars [1–8]. To achieve plagiarism detection by image identification, several kinds of image features can be extracted for originality analysis, such as color feature, texture feature, shape feature, edge feature, point feature and spatial relationship feature.

Color is the most intuitive feature in the image. Color has little dependence on the size of the image and has high robustness. The common color features include color histogram, color moment, color aggregation vector and color correlation graph. Malviya [1] proposed a forensics technology for image copying and moving forgery operation based on automatic color correlation graph of pixel.

Texture describes the surface properties of the image, with rotation invariance and strong anti-noise ability. However, texture cannot reflect the essence of the image, when the image resolution changes, texture will also appear deviation. The methods of texture feature representation include local binary mode, gray co-occurrence matrix, Voronoi checkerboard, Markov random field model, Gibbs random field model and wavelet transform. Aouat [2] proposed a new texture segmentation method using a gray level co-occurrence matrix, which analyzed every pixel of the image and extracted the edges of the image using gradient detection. In [3], a texture analysis descriptor was proposed based on quaternion Fourier transform for color images.

Shape, as the essential feature of an image, can effectively extract valuable information. The shape is usually described by contour feature and region feature, in which the contour feature refers to the outer boundary and the region feature refers to the whole shape region. The methods to represent shape features include boundary direction histogram, shape parameters, contour similarity, Fourier shape description method, shape invariant moment method, finite element method and wavelet descriptor. The shape parameters include contour perimeter, area, spindle direction and similarity. Shaban [4] proposed a method for target recognition using invariant moments, which is achieved through pre-processing, feature extraction, invariant feature engineering and class calculation.

Edge is the symbol of region position in an image and is often used in image segmentation and image matching. Edge detection operators are divided into first-order differential edge detection operator and second order differential edge detection operator. The first-order differential edge detection operators include Roberts edge detection operator, Prewitt edge detection operator, Sobel edge detection operator and Canny edge detection operator. Second order differential edge detection operators include Laplace operator and Gaussian Laplace operator. In [5], a gesture recognition algorithm was proposed based on the improved Canny operator. The improved Canny operator is used to detect

the edge of the hand image. The convex hull and convex defect detection algorithms of geometric features are used to achieve effective fingertip tracking. Finally, a convolutional neural network is used to achieve fingertip ordering.

The feature points of images play an important role in comparing the consistency of features in two images. Common points in images include corner points and spots. Corner points are inflection points of objects or intersecting parts between lines in images. Spots refer to areas with color and gray level differences from surrounding areas. The general feature points for image identification include oriented brief (ORB) feature point, difference-of-Gaussian (DOG) feature point, Harris feature point, scale invariant feature transform (SIFT) feature point and speeded up robust feature (SURF) point. Rathi [6] proposed a blind image forgery detection algorithm based on dual clutch transmission (DCT) and SURE.

Spatial relation refers to the spatial position or relative direction relationship between multiple objects in an image. There are two common methods for extracting spatial relations: one is to segment the image and extract features from the segmented region; the other is to divide the image into uniform sub-blocks and extract features from each sub-block. By calculating the center point of the target object, Shotton [7] proposed a method to learn the features of the contour segment and used the relative positions between objects to carry out target detection based on contour. Bobick [8] used motion contour to describe the global characteristics of a human movement, and the motion energy diagram and motion history diagram were used to represent the changes of the specific implementation of the action.

It can be seen from the above research status that the edge, shape and spatial characteristics of objects have achieved good results and have been used in many fields. However, due to the hyperbolic description of cartoon characters, the detection results are poor when using a single kind of feature. Therefore, we propose to combine several kinds of image feature engineering technologies to extract the character features, including contour feature, motion feature and pose feature.

## 2.2 Motion recognition

The piracy or copying of cartoon character personality action is an important part of cartoon infringement. There are several achievements in motion recognition research at home and abroad. According to the scheme of feature extraction, it can be divided into traditional recognition methods and deep learning-based recognition methods [9].

Methods based on deep learning adopt neural network models such as recursive neural network, to model the nodes of each part of the human body in an end-to-end manner [10]. Converse3D convolutional network proposed by Tran [11] is a milestone of recent three-dimensional (3D) convolutional network-based motion recognition. To better reflect 3D scene by image denoising, a group-based nuclear norm and learning graph model was proposed for group-based image restoration [12]. In [13], a multi-view hash model was proposed to enhance the multi-view information by a hash learning method and a variety of multi-data fusion methods. It used a view stability evaluation method to explore the relationship between views. For deeper semantic analysis, a task-adaptive attention mechanism was proposed in [14]. In [15], a scheme of no-reference image quality assessment was proposed to analyze the features of image quality by distortion

identification and targeted quality evaluation. In [16], a multi-feature fusion and decomposition framework was proposed to extract discriminative and robust features. Fong [17] proposed a dynamic skeleton model of spatio-temporal graph convolution network, which improved expression ability and generalization ability through spatio-temporal pattern.

At present, deep learning-based human motion recognition method has been widely used in many tasks of computer vision, but it requires a large amount of training data, high cost and complex algorithm. Moreover, the research of cartoon character's personality action is aimed at the cartoon character rather than the human body. Cartoon character actions have the problems of wide range of action changes and insufficient data sets, which also bring new challenges to the recognition task. In view of the above reasons, this paper adopts the traditional action recognition method to extract the action features of cartoon characters from multiple angles, and finally carries out feature fusion.

The traditional action recognition method uses artificial features to recognize actions. Lena [18] extended the human appearance characteristic information from two-dimensional (2D) to 3D space and proposed a method of space-time body characteristics. On this basis, Ni [19] used the combination of original, forward depth and backward depth motion history maps to represent the motion history map of 3D changes. Chakraborty [20] proposed a spatio-temporal interest points (STIP) detection method which uses points of interest to achieve local descriptors with more repeatable, stable and unique human behaviors. Dalal [21] proposed a method to calculate the gradient direction eigenvalues of local images by histogram of gradient direction. Wang [22] proposed a method to extract dense trajectories by using optical flow field to track densely sampled points and obtain trajectories. Furthermore, a global smoothness constraint is applied between points in the dense optical flow field.

In the previous work [23], we studied cartoon character motion feature extraction based on motion direction. In this research, which is the extension of the previous work, we further studied pose feature extraction and feature fusion methods, and together with the multi-type of features, we achieved cartoon character copyright recognition by similarity detection with the features.

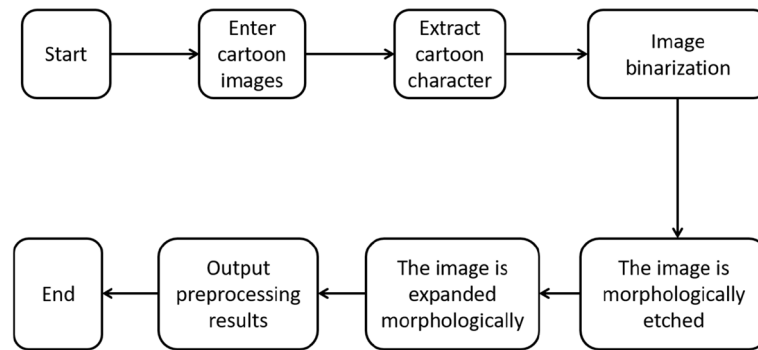
### 3 Copyright recognition based on character action

The proposed copyright recognition method for cartoon character is mainly composed of four steps, including cartoon image preprocessing, character feature extraction, feature fusion and similarity detection. The character feature extraction procedure can be further divided into three steps: contour feature extraction, motion feature extraction and pose feature extraction. Next, we provide a detailed explanation of each step.

#### 3.1 Cartoon image preprocessing

The cartoon image preprocessing uses GrabCut algorithm of interactive foreground extraction to obtain cartoon characters of binary images, as well as performing morphological corrosion and swelling operations on the binary images. The specific process is shown in Fig. 1.

The acquisition of cartoon characters takes into account the purpose of discarding background and extracting only foreground objects, as well as the characteristics of



**Fig. 1** Image preprocessing procedure



**Fig. 2** Image marking for separating foreground and background

cartoon in more free and rich form of expression, more intense visual effects, more personalized, single color and high saturation. GrabCut image segmentation method based on graph theory is selected among image segmentation algorithms such as threshold segmentation, region segmentation, edge segmentation and graph theory segmentation. However, part of the image segmentation results are not ideal, will appear to take the foreground as background, background as foreground. Therefore, aiming at the demand of cartoon character extraction, we remove the rectangular automatic segmentation box of GrabCut algorithm and use GrabCut interactive segmentation. Instead of using rectangle initialization, interactive segmentation goes directly into the mask image mode and adds two brushes to mark the foreground and background. Among them, the red brush marks the foreground and the blue brush marks the background as shown in Fig. 2.

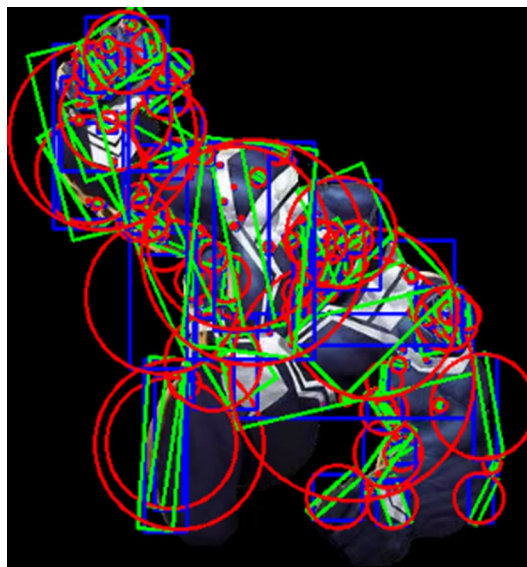
There are interference points and interference lines in the image after obtaining the cartoon character. Morphological corrosion and expansion operations are used to remove the white edge of the character and other interference pixels. However, image corrosion and expansion are mainly aimed at the thresholding value (0 or 255), thus image binarization is carried out to prepare for subsequent operations. For binary images, there are only two colors, black and white. Therefore, the amount of binary image data is small, and the image after binarization can highlight the contour of the character.

There are multiple connected regions in the image after segmentation, and these connected regions have a serious interference effect on the feature extraction of subsequent character actions. For instance, when selecting a character in polygon box, the small connected areas of interference points and interference lines will be also selected by polygon with the character as shown in Fig. 3. Therefore, after binarization, the segmented image is further processed by morphological operation of corrosion first and then expansion, in order to obtain the simply connected region of cartoon characters.

### 3.2 Contour feature extraction

For contour feature extraction of cartoon character, first the perimeter and area of contour are calculated from the preprocessed image. Next, using polygons to box the character. There are many peripheral polygons of contour, such as peripheral circle, peripheral ellipse, peripheral horizontal rectangle, peripheral minimum rectangle and fitting polygon. While the polygon selected in this procedure is the peripheral minimum rectangle. Then, the length–width ratio and tilt angle of the smallest rectangle are calculated. The extracted character contour features are established as feature vectors. The specific process of the character contour feature extraction is shown in Fig. 4.

When calculating the perimeter and area of the contour, the character contour pre-processed by the cartoon image should be obtained first, and then the perimeter and area of the contour can be obtained by calculating the number of all the pixels contained in the contour. The length-to-width ratio and tilt angle of the outline were calculated by using the outer minimum rectangle of the cartoon character. The character contour is obtained from the simply connected region of the preprocessed cartoon character. The contour is drawn with the smallest rectangle for obtaining the length, width and tilt angle of the rectangle. Therefore, the aspect ratio and tilt angle characteristics of the minimum rectangle surrounding the character can be extracted. The tilt angle refers to the angle formed by the clockwise rotation of  $x$  axis and the first edge of the encountered



**Fig. 3** Character extraction without morphological processing

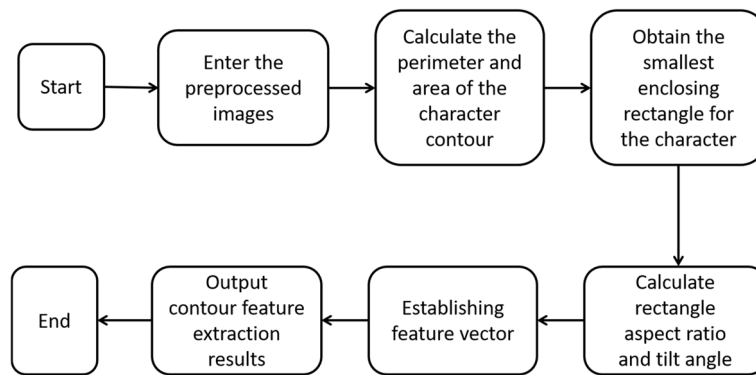


Fig. 4 Flowchart of character contour feature extraction

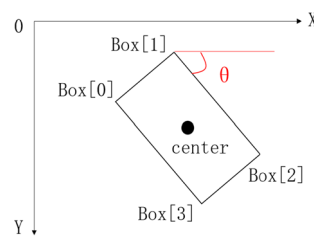


Fig. 5 Illustration for tilt angle of minimum enclosing rectangle

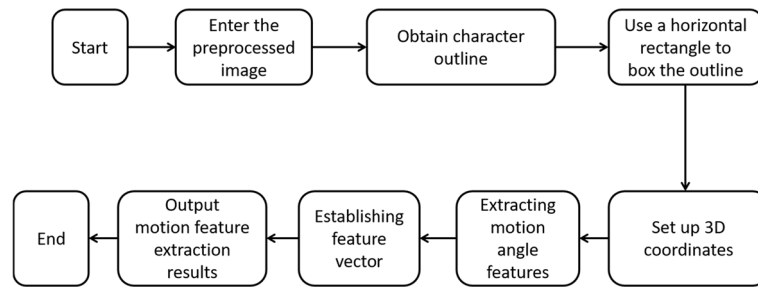
rectangle, as shown in Fig. 5. In the figure,  $\theta$  is the tilt angle of the smallest enclosing rectangle.

### 3.3 Motion feature extraction

For feature extraction of moving cartoon character, the outer horizontal rectangle is used to box the character contour of the preprocessed image. Calculate the coordinate of the horizontal rectangle center point as the coordinate of the central point of the character contour. Then, a 3D coordinate is established. The establishment method of the 3D coordinate is to take the  $x$  and  $y$  axes of the contour center as the  $x$  and  $y$  axes of the 3D coordinate, and the establishment of  $z$ -axis is calculated according to the proportion of the horizontal rectangle to the image and given a certain weight. After obtaining the 3D coordinates of the center point and four vertices of the horizontal rectangle, the coordinates are visualized in the 3D coordinate system. In a continuous moving character image, the 3D coordinates of its center point are calculated, respectively, to obtain the motion vectors of position changes of the character. The included angle of the vectors, namely the motion angle of the cartoon character, can be obtained by calculating the motion vectors. Finally, the angle features of the moving character are extracted to build feature vectors. The flowchart of the motion feature extraction is shown in Fig. 6.

After the 3D coordinate is established, the 3D coordinate of the contour center point can be obtained. The motion vector is defined by using the 3D center point of each action of the moving character. By calculating the angle of the adjacent vectors, the motion angle of the moving character from an action A to another action B is obtained. The motion angle features are extracted to build motion feature vectors. When calculating the motion angle, let  $\mathbf{m}$  and  $\mathbf{n}$  be two non-zero vectors, and the cosine value of the





**Fig. 6** Flowchart of motion character feature extraction

included angle between the two vectors can be calculated by formula (1). The vector is represented by coordinates,  $\mathbf{m} = (x_1, y_1, z_1)$ ,  $\mathbf{n} = (x_2, y_2, z_2)$ , the calculation method is shown in formula (2), the values of  $|\mathbf{m}|$  and  $|\mathbf{n}|$  can be calculated by formula (3) and formula (4), respectively, which can be substituted into formula (1) to obtain the calculation formula (5) of the included angle of 3D coordinate vectors:

$$\cos\langle m, n \rangle = \frac{\mathbf{m} \cdot \mathbf{n}}{|\mathbf{m}||\mathbf{n}|}, \quad (1)$$

$$\mathbf{m} \cdot \mathbf{n} = x_1x_2 + y_1y_2 + z_1z_2, \quad (2)$$

$$|\mathbf{m}| = \sqrt{x_1^2 + y_1^2 + z_1^2}, \quad (3)$$

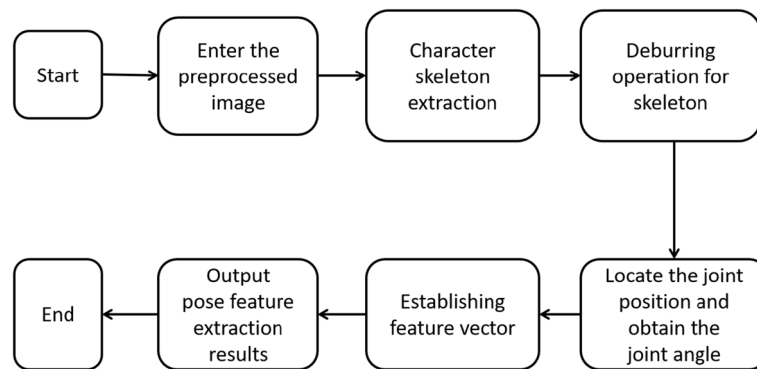
$$|\mathbf{n}| = \sqrt{x_2^2 + y_2^2 + z_2^2}, \quad (4)$$

$$\cos\langle m, n \rangle = \frac{x_1x_2 + y_1y_2 + z_1z_2}{\sqrt{x_1^2 + y_1^2 + z_1^2} * \sqrt{x_2^2 + y_2^2 + z_2^2}}. \quad (5)$$

### 3.4 Pose feature extraction

The pose feature extraction for cartoon character is achieved by extracting the character skeleton. First, the character skeleton is obtained through a refinement algorithm. Secondly, considering that the skeleton is prone to burr after refinement, it is necessary to carry out deburring operation. Then, the angle information of skeleton joint is extracted. Finally, the feature vectors of the character pose are established. Figure 7 shows the flowchart of the pose feature extraction.

Skeleton dynamic provides important information for motion feature extraction. In order to recognize the motion of cartoon characters, skeleton extraction for cartoon characters becomes a necessary step. Hilditch [24] algorithm is a classic refinement method, but when determining whether pixel points are edge points and whether they should be deleted, the conditions are not simple enough, which is easy to cause problems of slow processing speed and poor universality. Compared with Hilditch algorithm, the middle-axis transform refinement algorithm can accurately approximate the geometric



**Fig. 7** Flowchart of pose feature extraction

structure of the object and retain details. However, one-off processing of the skeleton will leave more skeleton branches, which is not conducive to the subsequent extraction of intersection points. Zhang proposed a skeletonization method (Zhang-Suen) [25] by iterative algorithm is more accurate. By comparing the Hilditch serial refinement, axial transformation refinement and Zhang-Suen parallel refinement, we found the Zhang-Suen algorithm outperforms the other methods. Therefore, we select Zhang-Suen algorithm as the method of skeleton extraction for cartoon characters. The skeleton extraction effects of Hilditch, middle-axis transform and Zhang-Suen algorithm are verified in the experiments.

There are burrs in the extracted character skeletons, which will interfere with the subsequent search of skeleton joints and joint angles. Therefore, burrs need to be eliminated according to the characteristics of burrs. The process of skeleton deburring is also the process of mathematical morphology operation on binary image. After obtaining the complete skeleton, pose feature can be extracted. The pose feature information is mainly aimed at the angle of skeleton joint. In this paper, we propose a method to calculate the angle of skeleton joints. Firstly, we obtain each point on the skeleton and the set of intersection points on the skeleton. Traversing the intersections on the skeleton with a square radius. Complete the traversal of each intersection in turn. In the ergodic process, the points on the skeleton are connected to the intersection points, and then the angles formed by the connecting points are calculated to obtain the angles of each joint of the skeleton.

### 3.5 Feature fusion and similarity detection

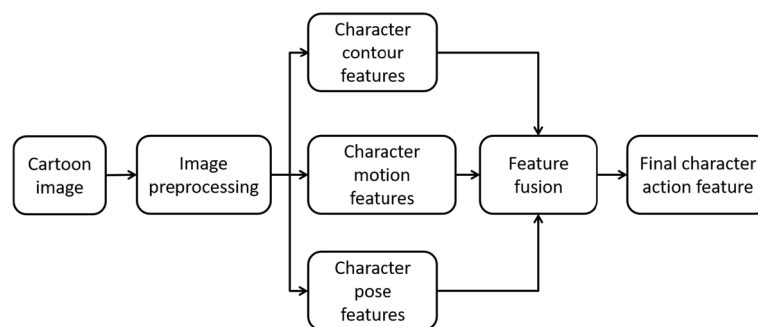
Feature fusion refers to the method of selecting the most distinguishing and relevant features and combining them to form new feature vectors. The specific method is to fuse the multi-type features extracted from the image into a new feature value, and this new feature value has higher discrimination ability than the single-type feature, and can also produce better action recognition results. The features of character personality actions extracted only by a single algorithm generally reflects only a certain part of the feature information, which is one-sided. Moreover, the process of feature extraction can be affected by external noises, which has a certain impact on the accuracy of the final recognition results. Multi-feature extraction can not only ensure the comprehensiveness of feature extraction, but also improve the accuracy and

robustness of feature recognition. Therefore, it is necessary to effectively fuse the features extracted in different ways, and use the fused features as the standard to judge whether the actions of a character are pirated or plagiarized. Figure 8 is the flowchart of feature fusion.

When detecting the similarity between the original action and the pirated action, the feature vectors established by the character contour features are compared first. The features extracted from the character contour include perimeter, area, minimum rectangle aspect ratio and tilt angle. There are altogether four feature data in the contour feature vector. The ranges of the four feature values are very different. The values of perimeter and area are much larger than those of minimum rectangle aspect ratio and tilt angle. Considering that these four data have great differences in values, standardization is employed to the feature values to keep balance between the values. Finally, Pearson correlation coefficient is calculated for the contour feature vectors between original cartoons and pirated cartoons. The calculation formula is shown in formula (6), where  $\bar{X}$  and  $\bar{Y}$ , respectively, represent the mean values of the two feature vectors:

$$R = \frac{\sum_{i=1}^n (x_i - \bar{X})(y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{Y})^2}} \quad (6)$$

For the character motion feature, the values of motion features are also standardized. Then calculate the Pearson correlation coefficient of the motion features of original and pirated cartoons, and finally obtain the motion similarity of the cartoon characters. According to the character pose feature, the obtained peripheral angles of each intersection can be set by referring to the maximum number of peripheral angles based on the standard of the same number of peripheral angles. The angles at each intersection point are stored in ascending order. Zero is added after the vector if the number of angles is insufficient. Finally, the Pearson correlation coefficients of the pose features are calculated for the original and pirated cartoon characters to obtain the similarity of character poses. In order to maximize the advantages of different types of features, multi-type of features are weighted and fused according to the ratio of 1:1:1. The fused features can contain more information conducive to identifying



**Fig. 8** Flowchart of the feature fusion

the action features of cartoon characters. After the synthesis of the three types of features and the calculation of the Pearson correlation, we propose following discriminant criteria to determine whether exist infringement:

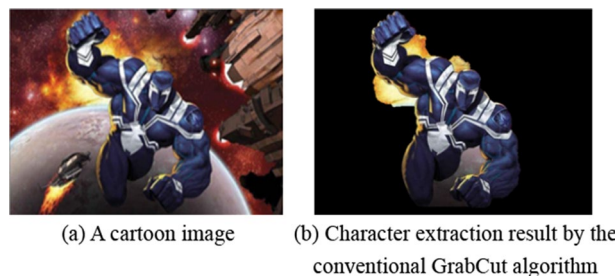
- (1) If  $0.8 \leq |r| \leq 1$ , it indicates obvious infringement.
- (2) If  $0.5 \leq |r| < 0.8$ , it indicates moderate infringement.
- (3) If  $0.3 \leq |r| < 0.5$ , it indicates low-level infringement.
- (4) If  $|r| < 0.3$ , it indicates no infringement.

## 4 Experiments

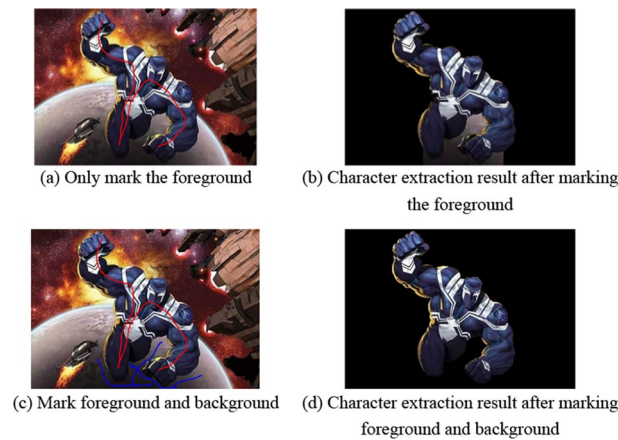
In this section, we conduct experimental verification for the cartoon character copyright recognition method of personality action feature proposed above, including character extraction experiment, image morphology processing experiment, contour feature extraction experiment, motion feature extraction experiment, pose feature extraction experiment and feature fusion experiment. The experiment in this paper is carried out under the environment of 64-bit Win8.1 operating system with Intel(R) Core(TM) I5-5257 U CPU @ 2.70 GHz 2.70 GHz processor. The programming language is Python.

### 4.1 Experiment of character extraction

In the proposed copyright recognition framework, the character extraction is the first step which may affect the performances of the later procedures. To analyze the performance of the character extraction, we compared the conventional GrabCut algorithm with its improved version by segmenting the cartoon character. Experimental results show that the GrabCut algorithm has a good extraction effect on the foreground objects, but there are some defects in the extraction process. Figure 9 shows the foreground extraction effect of the conventional GrabCut algorithm, in which the background pixels clearly exist. The improved interactive GrabCut algorithm can segment the target character more complete, greatly reduce the mixing of redundant background, and improve the stability, accuracy and robustness. Figure 10 shows the foreground extraction result of the improved GrabCut algorithm. It can be seen from the figure that the optimized target character is more accurate than the previously extracted result, and the extracted foreground is more controllable. The extracted target character basically has no large area of background interference, and it has a good recognition foundation for the subsequent feature extraction.



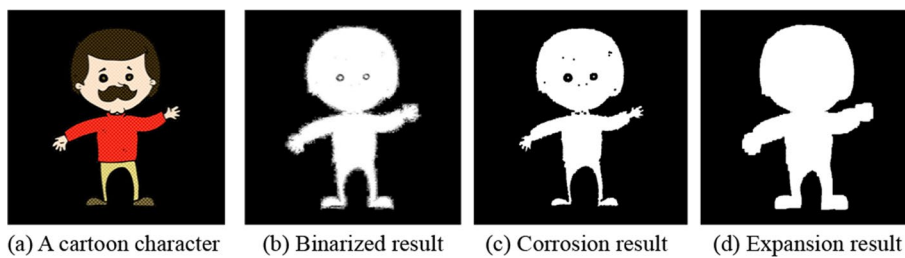
**Fig. 9** A cartoon image and its corresponding character extraction result



**Fig. 10** Character extraction results with the improved interactive GrabCut algorithm

#### 4.2 Experiment of image morphology processing

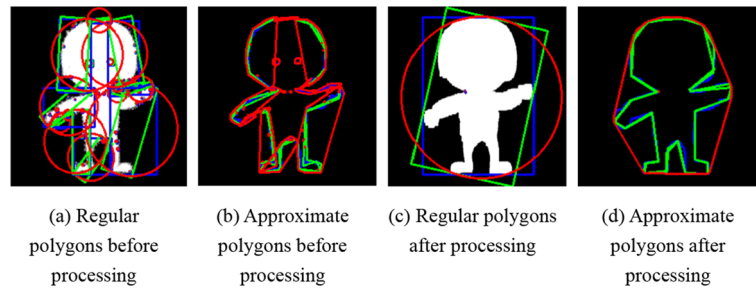
After the character is extracted, the residual tiny interference points and lines are further removed and flattened by morphological processing. Figure 11 shows the illustration of morphological corrosion and expansion. It can be seen that the edge of the target character is smoother through morphological corrosion, and the cavity can be filled through morphological expansion. Figure 12 shows the result of obtaining the contour of the character before and after the morphological processing. It can be clearly seen from the figure that the contour edges are not smooth and neat before the morphological processing, while the contour obtained after the morphological processing is smooth and tidy. Figure 13 shows the results of box selection of characters using peripheral circles, minimum rectangles, horizontal rectangles and approximate polygons before and after morphological processing. Because of the existence of multi-connected regions in the image before morphological processing, the subsequent feature extraction is greatly interfered. The experimental results show that the morphological operation of corrosion



**Fig. 11** Experiment of morphology processing with corrosion and expansion



**Fig. 12** Extracted contour before and after morphological processing



**Fig. 13** Experiment result of external polygon generation before and after morphological processing

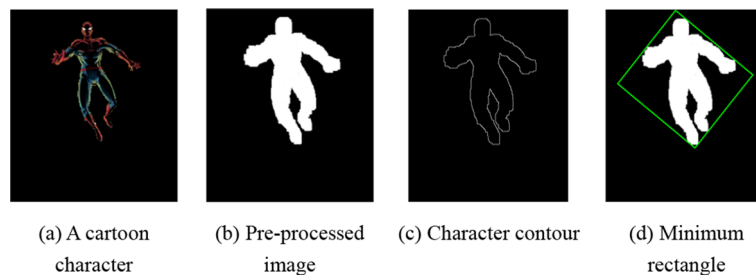
and expansion can remove the interferences except for the character and obtain the contour of the character and the polygon surrounding the character, which make preparation for the subsequent feature extraction.

#### 4.3 Experiment of contour feature extraction

The feature vector of the cartoon character contour is established by calculating the perimeter and area of the preprocessed character contour and by calculating the aspect ratio and tilt angle of the enclosing minimum rectangle after box selection. Figure 14 shows the experiment for obtaining the minimum rectangle attached to the contour of the character. It shows the experimental effect drawing of obtaining the cartoon character from the original drawing, corroding and expanding after acquiring the character, obtaining the contour of the character, and obtaining the minimum rectangle attached to the contour. The experimental results show that the preprocessed image can better obtain the contour of the character and the surrounding minimum rectangle. Table 1 shows the extracted data of character contour features, including contour perimeter, area, minimum external rectangle aspect ratio and tilt angle.

#### 4.4 Experiment of motion feature extraction

For the extraction of motion features in 3D coordinates, different actions of the same character in a group of four-panel cartoon images are captured in the experiment. In



**Fig. 14** Image processing for contour feature extraction

**Table 1** Feature data extracted from character contour

Perimeter (number of pixels)	Area (number of pixels)	Aspect ratio of the enclosing rectangle	Tilt angle of the outer rectangle (degree)
1529	31,982	1.08	39.08

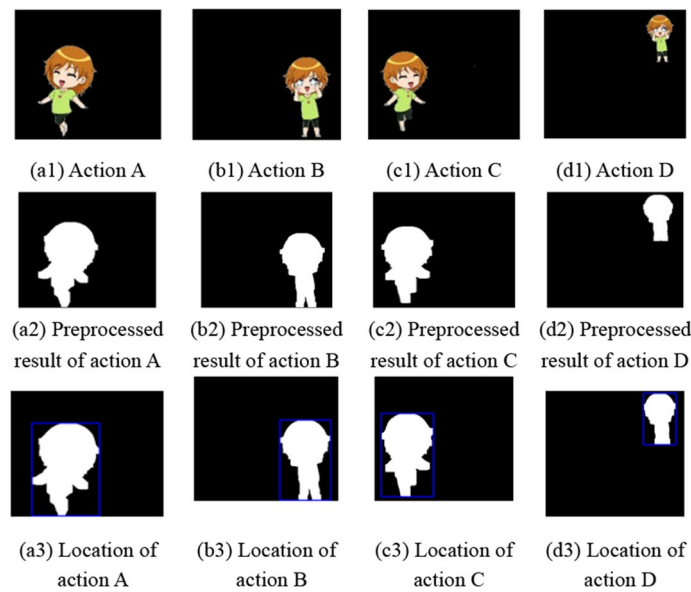


Fig. 15 Location analysis of character action

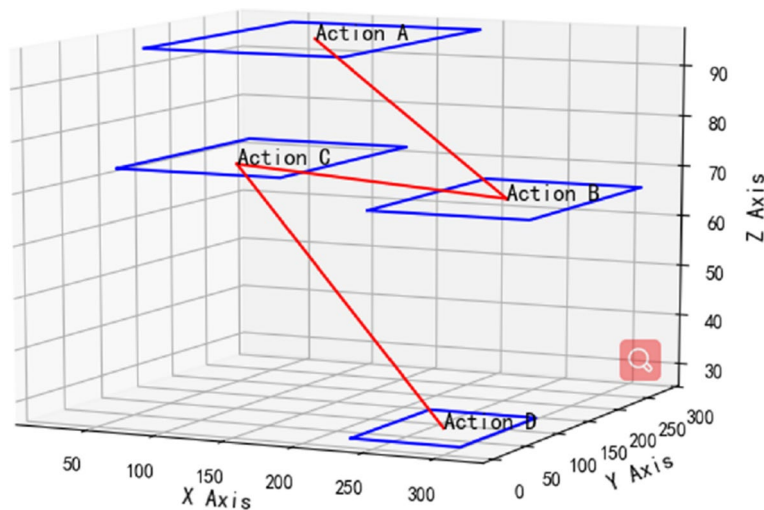


Fig. 16 Visualization of character motion angles in 3D coordinate system

order to better see the experimental effect, action B was reduced as action D to better reflect the changes of the character in 3D coordinates.

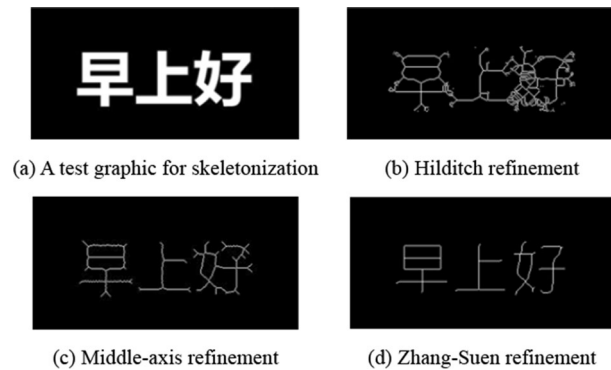
Figure 15 shows the analysis process of character action, by locating the outer horizontal rectangle for the characters in different action images. Figure 16 shows the visualization of the character motion angle in 3D coordinate system. The blue rectangles are the outer horizontal rectangles of the character action A, B, C and D, respectively, and the red lines are the motion track of the character in 3D coordinate system. Table 2 is the data extracted during the establishment of 3D coordinates, including the 2D coordinates of the horizontal rectangle center, the proportion of the rectangle to the image, the z-axis coordinate value and the 3D coordinates of the character

**Table 2** Establishment of 3D coordinates

Action number	2D coordinates of the center point of a horizontal rectangle	Proportion of rectangle to image (%)	Z-axis coordinates	The 3D coordinates of the center point of the character
Action A	(119,180)	32	96	(119,180,96)
Action B	(255,193)	22	66	(255,193,66)
Action C	(80,181)	24	72	(80,141,72)
Action D	(270,64)	9	27	(270,64,27)

**Table 3** Feature data of motion angle extracted from moving character

Point number	Cosine of the angle	Motion angle (degree)
Motion angle 1	0.93	22.17
Motion angle 2	0.90	25.31

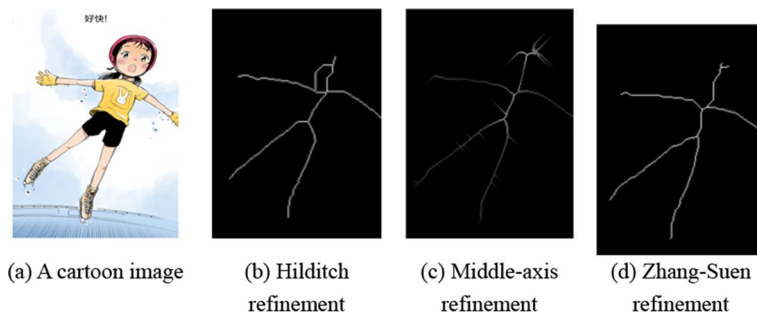
**Fig. 17** Comparison results of different skeletonization algorithms

center point. Table 3 shows that the motion vector established according to the 3D coordinates of the central point of the character, and the motion angle and its cosine value are calculated according to the motion vector.

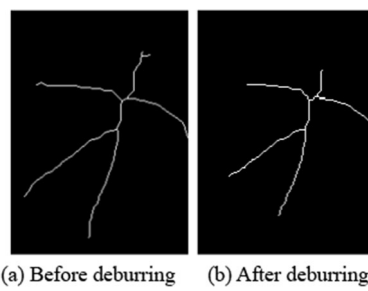
#### 4.5 Experiment of pose feature extraction

Before extracting the character skeleton, experiments are carried out on Hilditch refinement, middle-axis transformation refinement and Zhang-Suen refinement for comparing their refinement performances. According to the skeleton extraction effect, Zhang-Suen refinement is finally employed to extract the cartoon character skeleton. Figure 17 shows the skeletonization results of a simple graphic using the three refinement algorithms. Figure 18 shows the skeletonization results of a cartoon character. The thinning result of the Zhang-Suen refinement is more clear and accurate in skeleton extraction. Figure 19 shows the experimental comparison before and after deburring the skeleton obtained by the thinning algorithm. The deburring operation can remove the redundant skeleton points more effectively and ensure the accuracy and continuity of the extracted skeleton. Table 4 shows the feature data extracted from character pose, including the coordinates of the crossing points and the joint angles around the skeleton crossing points.





**Fig. 18** Extracted skeletons with different algorithms



**Fig. 19** Skeletons before and after deburring

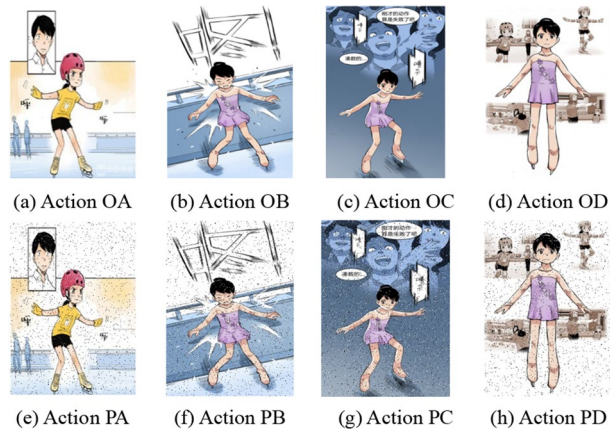
**Table 4** Feature data extracted from character pose

The joints	Cross-point coordinates	Angle (degree)
InterSect. 1	(99,72)	90, 127, 143
InterSect. 2	(92,74)	48, 121, 135
InterSect. 3	(89,99)	90, 104, 166

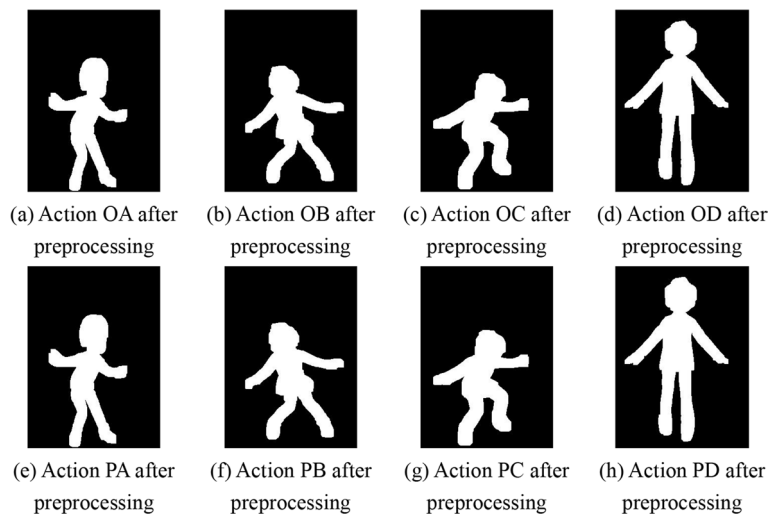
#### 4.6 Experiment of character action recognition

In order to validate the robustness, feasibility and effectiveness of the proposed framework, we conduct the character action recognition experiments with a simulated plagiarism cartoon by adding salt-and-pepper noises. The original cartoon images are chosen from a part of the cartoon "the ice flowers blossom". The following are the specific experimental results and analysis.

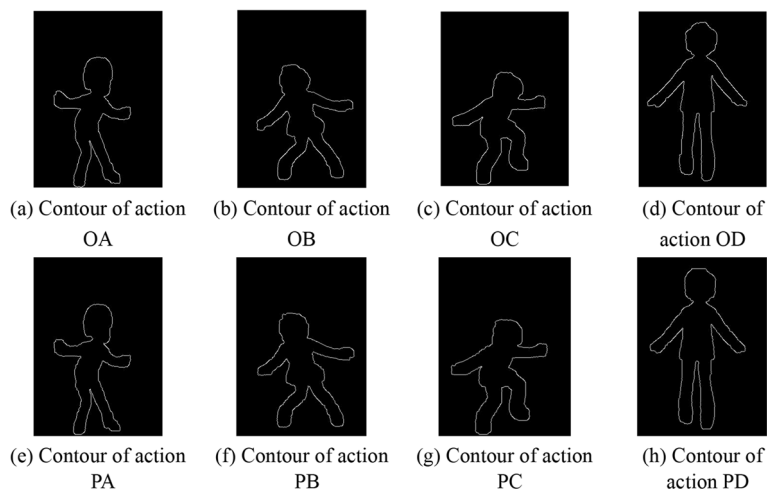
Figure 20 is the original and simulated plagiarism cartoon character images of four different actions. The four actions of the original cartoon are represented as OA, OB, OC and OD, and the actions of the simulated plagiarism cartoon after adding noises are represented as PA, PB, PC and PD. The resolution of the cartoon image is  $265 \times 363$ . Figure 21 shows the experimental results of image preprocessing for the original and pirated images. Figures 22 and 23 show the experimental results of character contour extraction and those of locating character with bounding box attached. Figure 24 shows the visualization results of the motion angles of the original and pirated characters in the 3D coordinate system. Figure 25 shows the experimental results of skeletonization for the original and pirated characters. Figure 26 shows



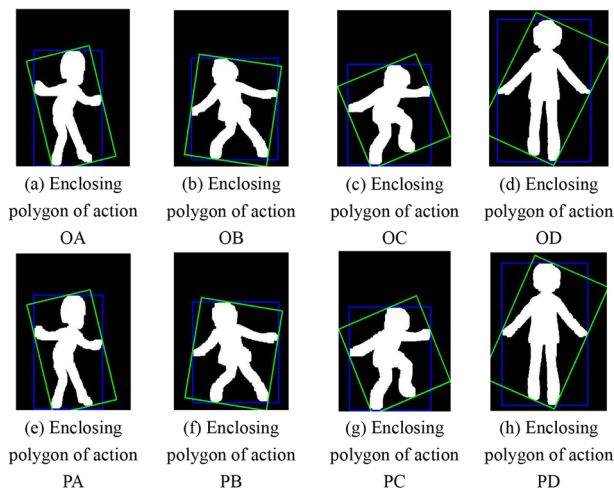
**Fig. 20** Original and pirated character action images with four actions



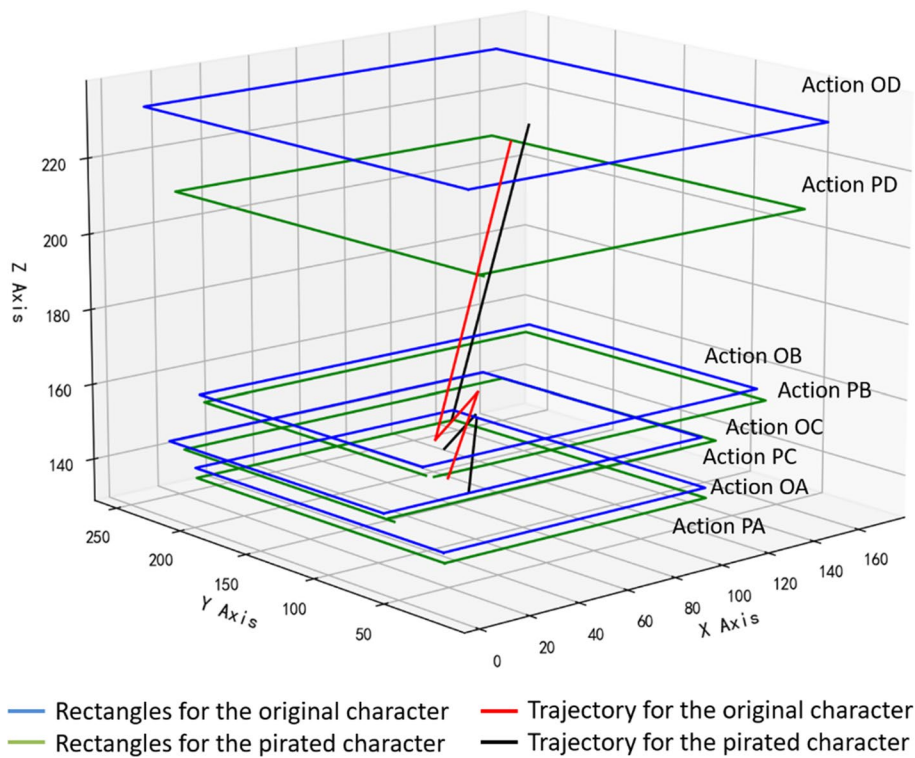
**Fig. 21** Experimental results of image preprocessing



**Fig. 22** Experimental results of contour extraction

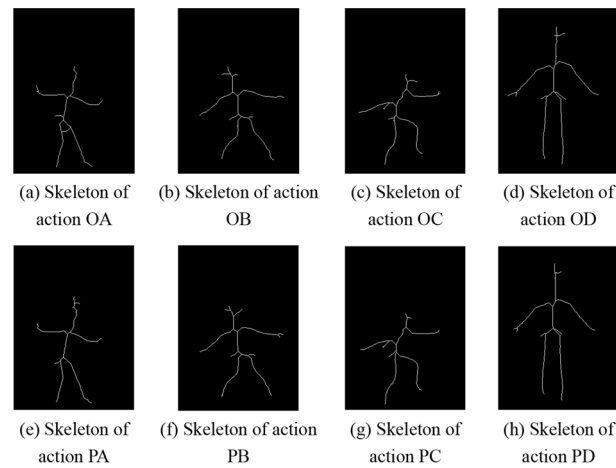


**Fig. 23** Experimental results of locating characters with bounding boxes

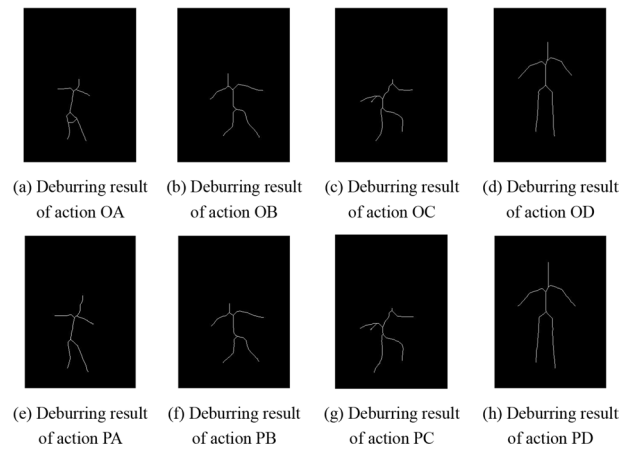


**Fig. 24** Visualization results of motion angles

the skeleton deburring results. It shows that the skeleton with deburring can remove redundant branches, which made it easier to extract effective skeleton joint angles. From Figs. 22, 23, 24, 25, 26, it can be seen that a certain amount of salt-and-pepper noise will affect the coefficient of corrosion and expansion, the result of character contour, character peripheral polygon, motion angle and character skeleton.



**Fig. 25** Experimental results of character skeletonization



**Fig. 26** Experimental results of skeleton deburring

Tables 5, 6, 7, 8, 9 show the extracted multi-type feature data, including contour feature data, motion feature data and pose feature data. It can be seen from Tables 5, 6, 7, 8, 9 that the addition of a certain amount of salt-and-pepper noise will not only change the perimeter and area of the character contour, as well as the aspect ratio and tilt angle of the smallest surrounding rectangle, but also change the motion angle of the character. In the process of skeletonization, due to the difference of the preprocessing results, the calculation of the skeleton will also be different because of the different angle. The closed area of the skeleton makes the calculated joint angle become different. Table 10 shows the feature similarities between the original and pirated cartoon images. The experimental results show that under the interference of some noise, the similarity of the contour feature of the original and pirated cartoon characters represents a result of a moderate infringement. The similarity is improved after combining the motion and pose features, which means that the motion and pose features are important in recognizing character actions. Finally, based on the similarity after feature fusion of the three types of features, an obvious infringement is obtained, which is the correct result, demonstrating that the

**Table 5** Feature data of character contour

Action number	Perimeter	Area	Minimum aspect ratio of enclosing rectangle	Minimum tilt of the enclosing rectangle
Action OA	1165	15,444	1.75	76.21
Action OB	1189	16,850	1.23	8.37
Action OC	1067	16,620	1.05	67.45
Action OD	1506	21,407	1.62	25.46
Action PA	1169	15,126	1.74	76.56
Action PB	1197	17,002	1.23	9.19
Action PC	1089	17,077	1.06	67.59
Action PD	1477	21,686	1.68	23.79

**Table 6** Generated 3D coordinate information by motion data

Action number	2D coordinates of the center point of a horizontal rectangle	Proportion of the rectangle to the image (%)	Z-axis coordinates	3D coordinates of the center point of the character
Action OA	(120, 228)	44	132	(120, 228, 132)
Action OB	(139, 229)	48	144	(139, 229, 144)
Action OC	(120, 242)	46	138	(120, 242, 138)
Action OD	(121, 186)	71	213	(121, 186, 213)
Action PA	(128, 227)	43	129	(128, 227, 129)
Action PB	(142, 229)	47	141	(142, 229, 141)
Action PC	(121, 243)	46	138	(121, 243, 138)
Action PD	(123, 186)	66	198	(123, 186, 198)

**Table 7** Feature data of character motion

Motion angle number	Cosine of the angle	Motion angle
Motion angle O1	0.99	3.30
Motion angle O2	0.99	4.49
Motion angle P1	0.99	2.48
Motion angle P2	0.99	4.79

proposed framework has a good performance for recognizing the copyright of cartoon character action.

## 5 Conclusion

Based on the characteristics of cartoon images, this paper studies the cartoon character action recognition methods and designs a copyright recognition method based on the personality actions of cartoon characters. Firstly, an image preprocessing scheme is designed before feature extraction. The interactive GrabCut algorithm is used to obtain the cartoon character, and the character is binarized, corroded and expanded to remove interference points and lines. Secondly, a feature extraction method is designed for character contour, motion and pose. A feature fusion mechanism is employed to combine the extracted feature data in a proportional weighted way, and similarity detection of original and pirated character actions is carried out by using Pearson correlation coefficient,

**Table 8** Feature data of the original character pose

The joints	Angle
Action A crossing point 1	90, 117, 153
Action A crossing point 2	98, 127, 135
Action A crossing point 3	106, 127, 127
Action A crossing point 4	80, 125, 156
Action A crossing point 5	90, 117, 153
Action B crossing point 1	104, 127, 129
Action B crossing point 2	106, 113, 141
Action B crossing point 3	90, 135, 135
Action C crossing point 1	90, 135, 135
Action C crossing point 2	53, 127, 180
Action C crossing point 3	80, 129, 153
Action C crossing point 4	90, 135, 135
Action D crossing point 1	56, 143, 161
Action D crossing point 2	80, 127, 153
Action D crossing point 3	98, 127, 135

**Table 9** Feature data of the pirated character pose

The joints	Angle
Action AA crossing point 1	90, 131, 139
Action AA crossing point 2	90, 121, 149
Action AA crossing point 3	106, 127, 127
Action BB crossing point 1	106, 127, 127
Action BB crossing point 2	108, 117, 135
Action BB crossing point 3	90, 135, 135
Action CC crossing point 1	108, 117, 135
Action CC crossing point 2	53, 127, 180
Action CC crossing point 3	80, 127, 153
Action CC crossing point 4	82, 132, 146
Action DD crossing point 1	63, 131, 166
Action DD crossing point 2	72, 135, 153
Action DD crossing point 3	90, 113, 157

**Table 10** Similarity detection results

Features used for similarity detection	Similarity
Similarity of contour feature	0.76
Similarity of contour and motion features	0.81
Similarity of motion and pose features	0.82
Similarity after feature fusion of the three features	0.85

so as to achieve the purpose of copyright protection. In the future work, we will apply the proposed framework to cartoon videos and 3D cartoon mesh models, and study the cartoon portrait feature learning in 3D space, and make a contribution for further research on copyright protection of 3D cartoon characters.

### Abbreviations

SUSAN	Smallest univalue segment assimilating nucleus
FAST	Features from accelerated segment test
ORB	Oriented brief
DOG	Difference-of-Gaussian
RANSC	Random sample consensus
SIFT	Scale invariant feature transform
ASIFT	Affine scale invariant feature transform
SURF	Speeded up robust feature
DCT	Dual clutch transmission
C3D	Converse3D
3D	Three-dimensional

### Acknowledgements

This research project was supported by the National Natural Science Foundation of China (Grant No. 62062064).

### Author contributions

All authors contributed to the research work. The authors designed the new method and planned the experiments. De Li led and reviewed the research work. Lingyu Wang and Xun Jin performed the experiments and wrote the paper.

### Funding

This research project was supported by the National Natural Science Foundation of China (Grant No. 62062064).

### Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

### Declarations

#### Competing interests

The authors declare that they have no competing interests.

Received: 22 December 2022 Accepted: 19 May 2024

Published online: 24 May 2024

### References

1. A.V. Malviya, S.A. Ladhake, Pixel based image forensic technique for copy-move forgery detection using auto color correlogram. *Procedia Comput. Sci.* **79**, 383–390 (2016)
2. S. Aouat, I. Ait-hammi, I. Hamouchene, A new approach for texture segmentation based on the Gray level co-occurrence matrix. *Multimed. Tools Appl.* **80**, 24027–24052 (2021)
3. F. Zhu, M. Dai, C. Xie, Y. Song, L. Luo, Fractal descriptors based on quaternion Fourier transform for color texture analysis. *J. Electron. Imaging* **24**(4), 043004 (2015)
4. M.S. Al-Ani, A.M. Darwesh, Target identification using a moment invariant approach. *IEIE Trans. Smart Process. Comput.* **8**(5), 335–346 (2019)
5. H. Zhang, S. Li, X. Liu, Research on Gesture Recognition Based on Improved Canny & K-Means Algorithm and CNN. *IOP Conference Series: Earth and Environmental Science* **440**(4), 1–8 (2020)
6. K. Rathi, P. Singh, Blind image forgery detection by using DCT and SURF based algorithm. *Int. J. Recent Technol. Eng. (IJRTE)* **8**(5), 2984–2987 (2020)
7. J. Shotton, A. Blake, R. Cipolla, Multi-scale categorical object recognition using contour fragments. *IEEE Trans. Pattern Anal. Mach. Intell.* **30**(7), 1270–1281 (2008)
8. A.F. Bobick, J.W. Davis, The recognition of human movement using temporal templates. *IEEE Trans. Pattern Anal. Mach. Intell.* **23**(3), 257–267 (2001)
9. C. Yan, Y. Sun, H. Zhong, C. Zhu, Z. Zhu, B. Zheng, X. Zhou, Review of omnimedia content quality evaluation. *J. Signal Process.* **38**(6), 1111–1143 (2022)
10. L. Shi, Y. Zhang, J. Cheng, H. Lu, Two-stream adaptive graph convolutional networks for skeleton-based action recognition. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 12018–12027, 2019.
11. D. Tran, L. Bourdev, R. Fergus, L. Torresani, M. Paluri, Learning spatiotemporal features with 3D convolutional networks. 2015 IEEE International Conference on Computer Vision (ICCV), pp. 4489–4497, 2014.
12. C. Yan, Z. Li, Y. Zhang, Y. Liu, X. Ji, Y. Zhang, Depth image denoising using nuclear norm and learning graph model. *ACM Trans. Multimed. Comput. Commun. Appl.* **16**(4), 1–17 (2020)
13. C. Yan, B. Gong, Y. Wei, Y. Gao, Deep multi-view enhancement hashing for image retrieval. *IEEE Trans. Pattern Anal. Mach. Intell.* **43**(4), 1445–1451 (2021)
14. C. Yan, Y. Hao, L. Li, J. Yin, A. Liu, Z. Mao, Z. Chen, X. Gao, Task-adaptive attention for image captioning. *IEEE Trans. Circuits Syst. Video Technol.* **32**(1), 43–51 (2022)
15. C. Yan, T. Teng, Y. Liu, Y. Zhang, H. Wang, X. Ji, Precise no-reference image quality evaluation based on distortion identification. *ACM Trans. Multimed. Comput. Commun. Appl.* **17**(3), 1–21 (2021)

16. C. Yan, L. Meng, L. Li, J. Zhang, Z. Wang, J. Yin, J. Zhang, Y. Sun, B. Zheng, Age-invariant face recognition by multi-feature fusion and decomposition with self-attention. *ACM Trans. Multimed. Comput. Commun. Appl.* **18**(1), 1–18 (2022)
17. M.F. Tsai, C.H. Chen, Spatial temporal variation graph convolutional networks (STV-GCN) for skeleton-based emotional action recognition. *IEEE Access* **9**, 13870–13877 (2021)
18. L. Gorelick, M. Blank, E. Shechtman, M. Irani, R. Basri, Actions as space-time shapes. *IEEE Trans. Pattern Anal. Mach. Intell.* **29**(12), 2247–2253 (2007)
19. B. Ni, G. Wang, P. Moulin, RGBD-HuDaAct: A color-depth video database for human daily activity recognition. 2011 IEEE International Conference on Computer Vision Workshops (ICCV workshops), pp. 1147–1153, 2011.
20. B. Chakraborty, M.B. Holte, T.B. Moeslund, J. Gonzalez, Selective spatio-temporal interest points. *Comput. Vis. Image Underst.* **116**(3), 396–410 (2012)
21. N. Dalal, B. Triggs, Histograms of oriented gradients for human detection. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), pp. 886–893, 2005.
22. H. Wang, A. Klaser, C. Schmid, C.L. Liu, Action recognition by dense trajectories. 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3169–3176, 2011.
23. L. Wang, C. Ma, D. Li, Research on character action recognition of digital comics. *Procedia Comput. Sci.* **208**, 286–292 (2022)
24. N.J. Naccache, R. Shinghal, An investigation into the skeletonization approach of hilditch. *Pattern Recogn.* **17**(3), 279–284 (1984)
25. T.Y. Zhang, C.Y. Suen, A fast parallel algorithm for thinning digital patterns. *Commun. ACM* **27**(3), 236–239 (1984)

### **Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.