IRSTI 28.23.15

D. Tursyn[1], U. Sadyk[2], A. Aitimov[3]
[1,2,3]Suleyman Demirel University, Kaskelen, Kazakhstan.

METHODS FOR HANDLING FRAMES OF HAND GESTURE RECOGNITION

Abstract. One of the foremost imperative errands within the field of computer vision is motion acknowledgment. This article deals with the problem of recognizing hand gestures from a static image. To solve this problem, the article suggests using the segmentation method based on the skin color model, and the recognition method based on the analysis of the shape of the hand and the selection of features that classify the gesture image. Most strategies of preparing outlines of hand gestures are considered: smoothing and binarization of pictures, their revolution, and changing the foundation. The preferences of utilizing these strategies in understanding the issue of hand motion acknowledgment are displayed.

Keywords: image processing, image smoothing, binarization, gesture recognition, dynamic gesture recognition, static gesture recognition, computer vision, binary image, threshold image, skeletonization.

***

Аннотация. Одной из важнейших императивных задач в области компьютерного зрения является распознавание рук. В данной статье рассматривается проблема распознавания жестов рук по статическому изображению. Для решения этой задачи в статье предлагается использовать метод сегментации, основанный на цветовой модели кожи, и метод распознавания, основанный на анализе формы руки и выделении признаков, классифицирующих жестовое изображение. Рассмотрено большинство стратегий подготовки контуров жестов рук: сглаживание и бинаризация рисунков, их переворот, изменение основы. Показаны предпочтения использования этих стратегий в понимании проблемы подтверждения движения руки.

Ключевые слова: обработка изображений, сглаживание изображений, бинаризация, распознавание жестов, динамическое распознавание жестов, статическое распознавание жестов, компьютерное зрение, бинарное изображение, пороговое изображение, скелетонизация.

***

Аңдатпа. Компьютерлік көру саласындағы маңызды императивті міндеттердің бірі қолды тану. Бұл мақалада статикалық кескіннен қол қимылдарын тану мәселесі қарастырылады. Бұл мәселені шешу үшін

мақалада қол терісінің түс моделіне негізделген сегментация әдісін және қолдың пішінін талдауға және қимыл кескінін жіктейтін белгілерді анықтауға негізделген тану әдісін қолдану ұсынылады. Қол қимылдарының контурларын дайындаудың көптеген стратегиялары қарастырылады, олар: суреттерді тегістеу және бинаризациялау, олардың төңкерілуі, фонын өзгерту. Сонымен қатар қолдың қозғалысын растау мәселесін түсінуде осы стратегияларды қолданудың артықшылықтары көрсетілген.

Түйін сөздер: суретті өңдеу, кескінді тегістеу, бинаризация, қимылды тану, динамикалық қимылды тану, статикалық қимылды тану, компьютерлік көру, екілік кескін, шекті сурет, қаңқалау.

I. Introduction

A high-quality strategy of signal acknowledgment will permit the advancement of numerous frameworks, such as cleverly motion interfacing, sign dialect interpretation frameworks, and administration for virtual and expanded reality frameworks.

The task of gesture recognition has various solutions in scientific works. Existing approaches can usually be divided into two classes:

1) wearable device-based recognition,

2) recognition based on computer vision.

One of the primary innovations for tackling the issue of hand recognition was uncommon gloves [1, 2]. This innovation is utilized for information collection and advanced transmission. The information is hand developments, such as speed and revolution points. This information is sent to the computer, and the computer performs acknowledgment employing an extraordinary calculation. For example, Takahashi and Kishino developed a data glove capable of recognizing 46 types of gestures. [3] Yangsheng et al. used Saibo gloves to accurately identify 14 different gestures and then control the robot. [4] The use of special gloves to solve the problem of gesture recognition has not become widespread due to the inconvenience of use.

Recently, methods and solutions that do not require special gloves are increasingly appearing to solve problems of hand gesture recognition. Next, we analyze the existing methods of gesture recognition using neural networks or infrared sensors.

To improve the quality of dynamic gesture recognition, you need to prepare a data set – a training sample. To do this, each frame of the hand gesture is processed. We considered the main methods of image processing used to solve the problem.

II. Types of hand gestures

Figure 1 shows some of the static gestures that can be diagnosed from video analysis. It can be seen that for the diagnosis of gestures, it is necessary to fix the position of the hand and fingers, while there are many similar situations that can be represented as a gesture, but not being such. When analyzing gestures, it is necessary to understand what positions and shapes the brush can take, relative to a standing / sitting person. The hand practically does not bend in the wrist in the direction of the thumb, in the direction of the little finger it bends about 70-90°. In the direction of the back and inner side of the palm, the brush bends (it can bend, and this value depends most often on the age) by 90°.

The elbow joint allows you to rotate the palm about 180° towards the thumb. The rotation of the hand relative to the line passing through the center of the palm and perpendicular to it is performed at 270°. If you turn the back side to yourself, then the rotation is carried out from 0 to 270° counterclockwise. If you turn the inner side to yourself, then the rotation is carried out from 90 to -180° clockwise. The brush has five fingers. This defines a set of gestures and the corresponding recognition algorithms.

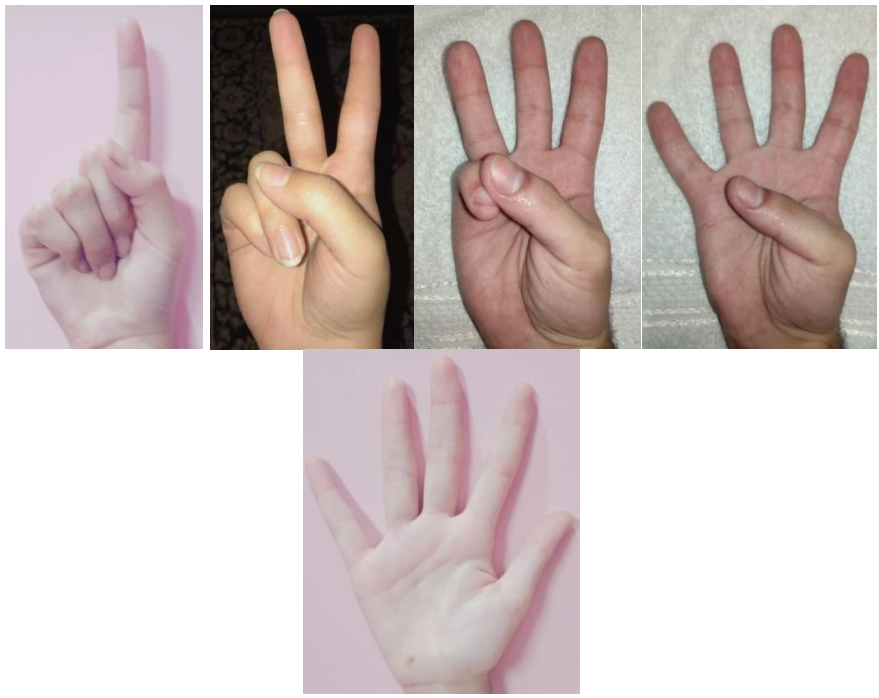

Fig. 1. Hand and finger positions during gesture display

Consider the possible combinations with the fingers (Fig. 1.). If the palm is clenched into a fist and each finger is straightened separately, you will get six options: fist, open thumb, index, middle, ring finger with little finger, little finger. Next, you can collect various combinations, with some combinations

straightening the third phalanx of the clenched fingers (if the wrist joints allow it). Of course, the sprain and rotation in the joints are individual, so here is the average data. For the fingers, there are 2 positions: compressed, unbent. Also, the rotation angles are selected as multiples of 90° (this is done to facilitate data processing). As a result, the gesture is determined by the position of the palm relative to the video device (rotation in different planes) and flexion/compression of the fingers.

III. Experiment and results

To improve the quality of hand gesture recognition, need to prepare a dataset – a training example. To do this, each frame of the dynamic gesture is processed. Consider the basic image processing methods used to solve the problem. Image smoothing - the transition locations in the image appear smoothest, eliminating irregularities in pixel structures. Smoothing eliminates noise and allows you to get an output suitable image for further study. In order to recognize dynamic hand movements, hand contours need to be most accurately identified, it is not enough to simply filter out the noise in the image [5]. Therefore, the bilateral filter (fig. 2) is the most suitable type of smoothing to solve the given task.

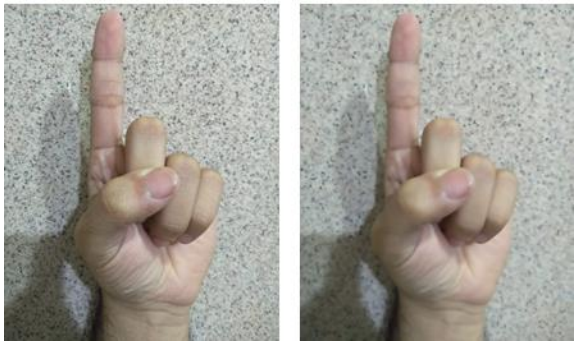

Fig.2. Example of a bilateral filtering

A binary image allows you to get a clear image of a recognizable object, but with a low level of contrast and brightness, binarization can give a less clear image of the object. In order for the result not to depend on these indicators, it is necessary to set a threshold level so that the number of white and black pixels is in a certain ratio, for example, from 50% to 50%. To get an image suitable for binary image, you need to convert it to grayscale (fig.3) [6].

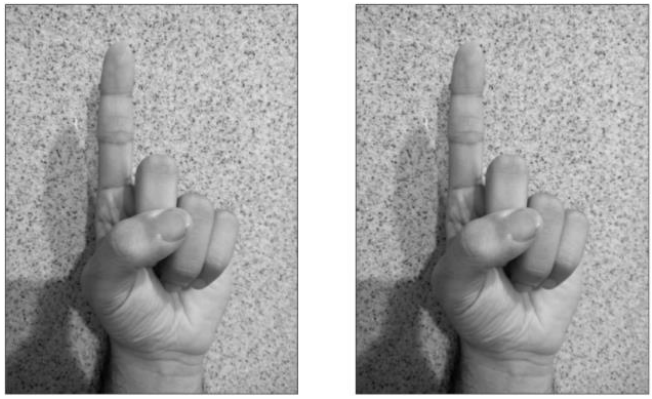

Fig.3. Converting an image to grayscale

Threshing is the process of converting a grayscale image into a binary image. Depending on a specific threshold value greater than the pixel is converted to a threshold of 255 (white), and a value lower than 0 (black) is converted to a threshold.

Fig.4. Threshold image

Skeletonization (fig.5) is the most appropriate way to represent an object in an image, since it can display both the general and detailed structure and shape of the object. The skeleton is a convenient and informative tool for analyzing the topological and metric properties of shapes.

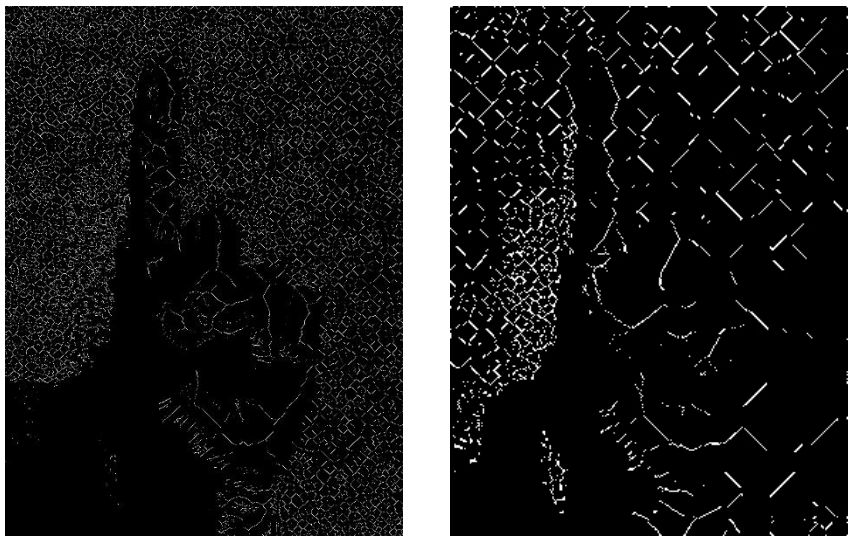

Fig.5. Skeletonization of the image using opencv

The dataset used to recognize the dynamic gesture frame must include several options for which background and position the hand is placed.

Change the background of the image binary is as follows: in the first stage, you need to remove the old background by subtracting the background image and the image of the hand; in the next step, the new background image and the resulting x, y axes is the final stage, the hand is to place a new background image.

Conclusion

As a result of the study, the main methods of processing video stream frames for solving the problem of recognizing hand gestures were identified. The advantages of these methods are to improve the quality of hand detail in the image, as well as the ability to generate a set of data for recognizing the frame of a hand gesture.

The experiment is based on a set of hand gesture recognition data based on the interval of numbers 1 and 5. Progress work designed in the Anaconda programming environment.

References

1 Takahashi, T., Kishino, F. Hand gesture coding based on experiments using a hand gesture interface device // ACM SIGCHI Bulletin. – 1991. – Vol. 23 (2). – P. 67–74.

2 Lee, C., Xu, Y. Online, interactive learning of gestures for human/robot inter-faces // Proceedings of the IEEE International Conference on Robotics and Automation. – 2002. – Vol. 4. – P. 2982–2987.

3 Smedt, Q. De, Wannous, H., Vandeborre, J.P. Heterogeneous hand gesture recognition using 3D dynamic skeletal data // Computer Vision and Image Understanding. – 2019. – Vol. 181. – P. 60–72.
4 Dynamic gesture recognition by directional pulse coupled neural networks for human-robot interaction in real time / J. Dong, Z. Xia, W. Yan, Q. Zhao // *Journal of Visual Communication and Image Representation.* – 2019. – Vol. 63. – P. 102583.
5 Oleari, C. Standard colorimetry: definitions, algorithms and software. Chichester: John Wiley & Sons, 2016. 512 p.
6 Клетте Р. Компьютерное зрение. Теория и алгоритмы / пер. с англ. под ред. А.А. Слинкин. М.: ДМК Пресс, 2019. - 506 с.