Ministry of Education of the Republic of Belarus
Educational institution
Belarusian State University of
Informatics and Radioelectronics

UDC 621.391

Zhongmin
Liao

ALGORITHMS FOR RECOGNIZING OF GESTURES IN IMAGES

Abstract
for a Master's Degree
in the Specialty 1-45 80 01 Infocommunication Systems and Networks

_____

Supervisor
PhD,Head of the Department of
Infocommunication technologies
Tsviatkou Viktar

_____

Minsk 2023

INTRODUCTION

With the development of modern technology and more intelligent lifestyles, people are increasingly longing for a more natural way of interaction. Therefore, it is of great significance to research a natural and comfortable human-computer interaction mode for this phenomenon. At present, several commonly used ways of human-computer interaction are keyboard-based input, mouse-based input, voice, facial expression, and gesture recognition technology. The characteristics of natural, intuitive, concise, humanized, and flexible gestures can fully stimulate the potential of the hand, without the hand attached to the mouse, keyboard, and other external devices, to avoid limitations. And gestures can be reflected in the process of human-computer interaction in a natural and simple way. Gestures have nothing to do with language, and their meanings are common even in different cultures and customs. In addition, there will be communication barriers in both graphic interaction and natural language interaction. Gesture can play a bridging role to solve this problem well. For example, the user does not know the specific use process when using the computer keyboard and mouse, but for how to use and express gestures, it can quickly grasp. Human gestures have diversity and variability, and the efficiency of human-computer interaction can be significantly improved by endowing gestures with specific connotations and inputting them into the computer. Therefore, through the research of gesture recognition, we can better develop a natural and efficient human-computer interaction mode, which is in line with the current background needs.

SVM, ANN and KNN are typical algorithms used for gesture recognition. The application principle of Support Vector Machine (SVM) is based on the structural risk minimization principle and statistical learning theory. It has the characteristics of stable identification, less parameters required, and does not produce local minimum problem. Even when the sample information is not sufficient, it can find the optimum in the range of complexity and learning ability of the model, and then achieve better generalization performance by compromise. For samples with good discrimination ability (Support Vector Machine), SVM can find them automatically, and the optimal segmentation surface constructed from these samples can maximize the interval between classes, to achieve effective classification effect. Therefore, SVM can achieve good classification on small sample training sets.

Artificial Neural Network (ANN) is a mathematical tool abstracted from the brain, which is generally simulated as a system composed of many processing units (namely neurons) connected with each other. Each neuron has a specific output function, called the excitation function. Each connection between two neurons represents the coefficient weighting value of the signal passing through the

connection, which is called the weight. The output of the network is determined by the connection mode, the weight, and the excitation function of the network. Artificial neural network is widely used in image classification and recognition system because of its powerful learning ability. The learning and recognition process of neural network is the updating and adjusting process of each neuron weight.

Neural network has strong self-learning and self-organization ability and has parallel cooperative processing structure, recognition speed is relatively fast, strong robustness and fault tolerance, can integrate the training and recognition process. However, it takes a long time to train the network, requires a large amount of computation, has poor scalability, and is not strong in processing time series.

KNN algorithm is a commonly used data mining algorithm, which is widely used in data analysis, image processing, text classification and other fields. The basic idea of KNN algorithm is given a training dataset, a suitable distancing mechanism is selected for a new input instance, and *K* instances closest to the instance are found in the training dataset. Most of these *K* instances belong to a certain class, and then the input instance is classified into this class.

KNN faces two problems, the first is that all its classification samples need to be kept, which not only makes the computational space occupation high, but also computationally complex. The second problem is that the performance of KNN is very sensitive to the neighborhood size *K*. How to choose the value of the parameter *K* relies heavily on practice and experience. SVM can identify accurately even with small training samples, and the parameters of the core classification function of SVM are specific support vectors, the computational complexity of the algorithm is determined by the number of support vectors, not the overall dimensionality of the sample space, avoiding the "dimensional disaster" of many operations. Therefore, SVM can help KNN to solve the problem of small training sample size and the problem of choosing parameter *K*. The use of ANN can also be considered, resulting in a higher recognition accuracy of the whole system.

This paper combines the advantages of SVM, ANN and KNN to propose a combination of the three classification algorithms. The small training samples of SVM can overcome the optimization of KNN K-values, ANN improves the running speed by forward propagation, and in addition ANN has a higher accuracy rate. The dataset formed by extracting the eigenvalues of the gesture images is put through a hybrid model, and the respective predictions of SVM, KNN and neural network are voted on before the result is derived.

GENERAL DESCRIPTION OF WORK

Relevance of the subject

The work corresponds to paragraph 1 «Digital information and communication and interdisciplinary technologies, production based on them» of the State Program of innovative development of the Republic of Belarus for 2021–2025.

The work was carried out in the educational institution Belarusian State University of Informatics and Radioelectronics within the framework of research work 21-2033 "Processing, coding and transmission of information in network-centric systems".

The aim and tasks of the work

The aim of the work is designing an SVM-ANN-KNN hybrid model algorithm, obtain a dataset by extracting feature values, apply the dataset to the hybrid model and perform gesture recognition, and design a gesture recognition model with high recognition rate.

To achieve this aim, the following tasks were solved in the dissertation:

1 First is the collection of data sets, because the data set for each test result is different, so this paper will make your own data set, through the OpenCV library for image acquisition, and then the bilateral filtering processing make smooth the palm of your hand, will be skin color detection and processing of image binarization, but for skin color detection often at the different moment of the day, Image brightness changes will be very obvious. Normally, images use RGB gamut, which is sensitive to brightness, and the parameters for detection need to be adjusted according to different brightness. Therefore, YCrCb, namely YUV gamut, which is not affected by brightness, is used to detect *Cr* and *Cb*, excluding the influence of brightness *Y*.

2 After get standardized images, in order to simplify dataset, the need for feature selection and extraction, traditionally, it is the picture for cutting, there is the image of 1, 0 no images, finally become a matrix of only 1 and 0, but in another way, this paper find out the palm center point, and then calculate the distance to edge of the palm to palm, A circle of edges down, get a group of data, draw into the image, and then the data and image can be analyzed and feature extraction.

3 SVM, ANN and KNN classifiers were used to train and test the obtained dataset, and the recognition rate was obtained. Finally, it was compared with the SVM-ANN-KNN model proposed in this paper. To obtain a recognition model with both high accuracy recognition rate and generalization ability, 10-fold cross validation method was used. A part of the data is temporarily selected from the data set as the training set, and the rest is used as the test set to obtain an error rate. Then another part is selected as the training set, and the rest is used as the test set to obtain an error rate. In this way, the obtained error rate is averaged, which is the error rate of cross-validation.

Personal contribution of the author

In this paper, we designed a SVM-ANN-KNN hybrid model for gesture recognition, firstly, we designed SVM,ANN and KNN classifiers respectively to perform gesture recognition on the collected dataset, then we applied the dataset to the combined SVM-ANN-KNN model for gesture recognition, finally we got the error rate of each classifier, and after taking the average of each set of data for comparison, we found that the recognition rate using The gesture recognition rate using the SVM-ANN-KNN hybrid model was found to be 99.3%,which is higher than the traditional classifier.

Testing and implementation of results

The main provisions and results of the dissertation work were reported and discussed at: International scientific and technical seminar "Technologies of information transmission and processing" (Minsk, March 2022) and International scientific and technical seminar "Technologies of information transmission and processing" (Minsk, March 2022) and International scientific and technical seminar "Technologies of information transmission and processing" (Minsk, May 2023) and 59th scientific conference of postgraduates, undergraduates and students, (Minsk, May 2023)

Author's publications

According to the results of the research presented in the dissertation, 4 articles and abstracts in conference proceedings.

Structure and size of the work

The dissertation work consists of introduction, general description of the work, four chapters with conclusions for each chapter, conclusion, bibliography, eight appendixes.

The total amount of the thesis is 80 pages, of which 50 pages of text, 20 figures on 6 pages, 7 tables on 3 pages, a list of used bibliographic sources (33 titles on 3 pages), a list of the author's publications on the subject of the thesis (4 titles on 1 pages), 1 appendix on 10 pages, graphic material on 7 pages.

Plagiarism

An examination of the dissertation «Title of the master's thesis» by Author's Full Name was carried out for the correctness of the use of borrowed materials using the network resource «Antiplagiat» (access address: https://www.bigan.net/) in the on-line mode 26.05.2023. As a result of the verification, the correctness of the use of borrowed materials was established (the originality of the thesis is 87.2%)

SUMMARY OF WORK

The introduction addresses the problems of efficiency and recognition of human-computer interaction. The classifier formed by SVM, ANN and KNN algorithms alone has shortcomings in terms of recognition rate and parameter optimization, through extensive research and reading of literature it was found that a hybrid SVM-AKK-KNN model can be used to compensate for the shortcomings between the three classifiers.

The general description of work consists of the following three parts:

1 The collection of data sets, because the data set for each test result is different, so this paper will make your own data set, through the OpenCV library for image acquisition, and then the bilateral filtering processing make smooth the palm of your hand, will be skin color detection and processing of image binarization, but for skin color detection often at the different moment of the day, Image brightness changes will be very obvious. Normally, images use RGB gamut, which is sensitive to brightness, and the parameters for detection need to be adjusted according to different brightness. Therefore, YCrCb, namely YUV gamut, which is not affected by brightness, is used to detect *Cr* and *Cb*, excluding the influence of brightness *Y*. The detected image is shown in Figure 1.

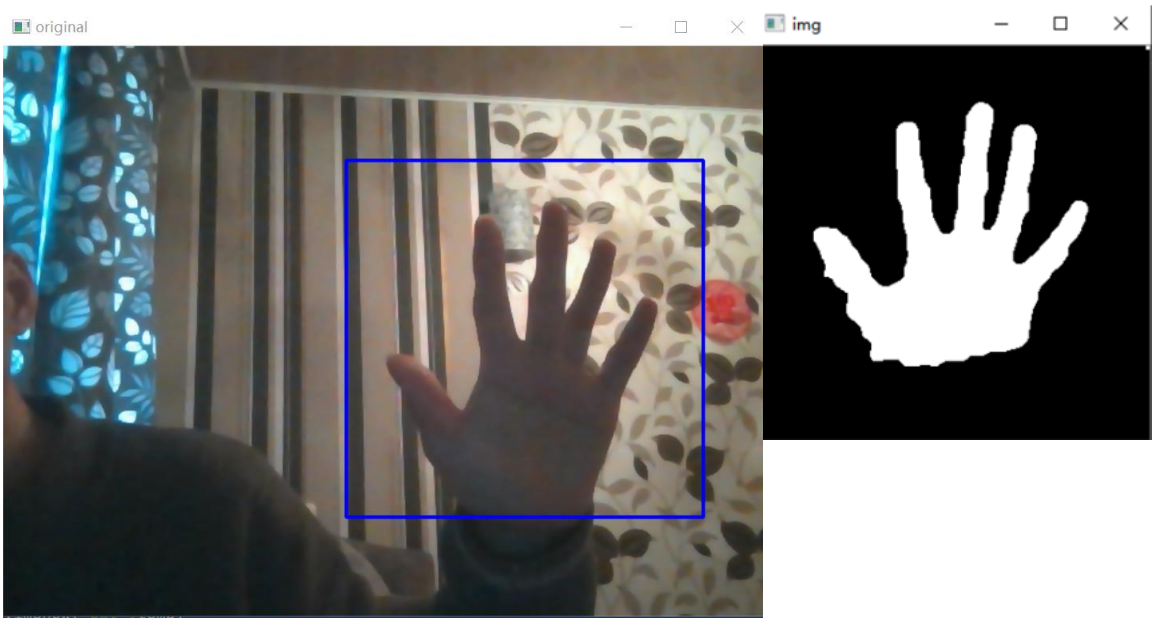

Figure 1 – Standardized image extraction process

2 The images were preprocessed and then binarized after image segmentation to obtain the images as shown in Figure 2. A total of 360 images, 9 gestures, and 40 images per gesture were collected. In order to simplify dataset, the need for feature

selection and extraction, traditionally, it is the picture for cutting, there is the image of 1, 0 no images, finally become a matrix of only 1 and 0, but in another way, this paper find out the palm center point, and then calculate the distance to edge of the palm to palm, A circle of edges down, get a group of data, draw into the image, and then the data and image can be analyzed and feature extraction.

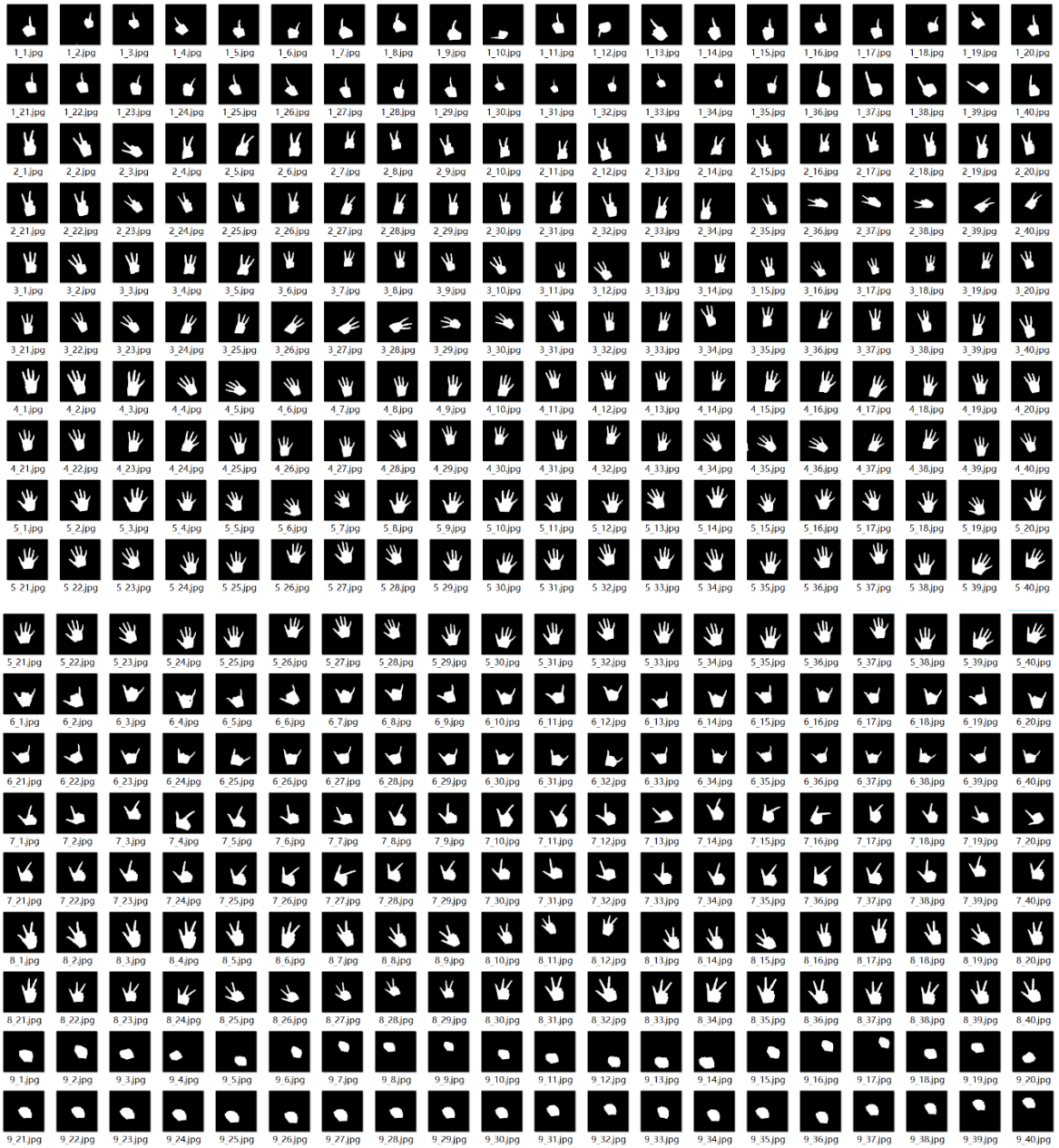

Figure 2 – Standardized picture

3 SVM, ANN and KNN classifiers were used to train and test the obtained dataset, and the recognition rate was obtained. Finally, it was compared with the SVM-ANN-KNN model proposed in this paper. To obtain a recognition model with

both high accuracy recognition rate and generalization ability, 10-fold cross validation method was used. A part of the data is temporarily selected from the data set as the training set, and the rest is used as the test set to obtain an error rate. Then another part is selected as the training set, and the rest is used as the test set to obtain an error rate. In this way, the obtained error rate is averaged, which is the error rate of cross-validation.

In the first chapter, before performing gesture recognition, the gesture in the image is segmented and the feature vector of the gesture is extracted. However, due to the influence of external environment and equipment during image acquisition, the input image needs to be processed appropriately before performing gesture segmentation to improve the accuracy of the results. This chapter introduces the common methods about image pre-processing, such as median filtering and mathematical morphology processing, and gives the results after median filtering and morphology processing.

In the second chapter introduces the traditional classification algorithms (KNN, ANN, Random Forest and SVM) and provides the theoretical basis for the hybrid SVM-ANN-KNN model proposed in the next section of this paper.

In the third chapter proposes a gesture recognition algorithm with a hybrid SVM-ANN-KNN model by comparing the advantages and disadvantages of traditional classification algorithms and introduces a cross-validation method that makes the dataset reusable and solves the problem of overfitting of small datasets, and finally introduces the design architecture of this proposed model.

In the fourth chapter, the algorithm proposed in this paper is validated. First, experiments on gesture segmentation and feature extraction are conducted, and then the extracted feature values are applied to SVM, ANN, KNN and the SVM-ANN-KNN model proposed in this paper. Finally, comparison tests are conducted on the four models to compare the accuracy of gesture recognition of the four models, and the results show that the accuracy of the SVM-ANN- KNN model proposed in this paper has an accuracy rate of 99.3%, which is higher than the other three classifiers.

CONCLUSION

In this paper, a gesture recognition method with a hybrid SVM-ANN-KNN model is proposed. Using this model to perform gesture recognition experiments on a small data set, the experiments show that this model is more accurate than some traditional models. The specific research results are as follows:

1 In this paper, a segmentation method based on YCbCr space was used to segment the images, and digital morphology processing was used to obtain the normalized images.

2 In this paper, four classes of gesture feature values are extracted, and the accuracy and generalization ability of recognition is improved by extracting four classes of feature values. The first category calculates the distance from the palm contour to the palm of the hand to obtain the number of peaks, the second category calculates the angle between the peaks, the third category is the average distance from the palm contour to the palm of the hand, and the fourth category is the sum of the peaks to the palm of the hand.

3 The algorithm proposed in this paper is verified. First, experiments on gesture segmentation and feature extraction are conducted, and then the extracted feature values are applied to SVM, ANN, KNN and the SVM-ANN-KNN model proposed in this paper. Finally, the four models are compared and tested to compare the gesture recognition accuracy of the four models, and the results show that the accuracy of the SVM-ANN-KNN model proposed in this paper is 99.3 %, which is higher than the other three classifiers.

LIST OF AUTHOR'S PUBLICATIONS

1–A. Liao, Z. M. Research on skeleton matching classification / Z. M. Liao, J. Ma, Y. P. Y. Qiu // Технологии передачи и обработки информации : материалы международного научно-технического семинара, Минск, март-апрель 2022 г. / Белорусский государственный университет информатики и радиоэлектроники. – Минск, 2022. – С. 93 – 95.

2–A. Liao, Z. M. A hybrid classication algorithm based on SVM, ANN AND KNN for gesture recognition / Z. M. Liao // Технологии передачи и обработки информации: материалы международного научно-технического семинара, Минск, март-апрель 2023 г. / Белорусский государственный университет информатики и радиоэлектроники. – Минск, 2023. – С. 159 –162.

3–A. Liao, Z. M. Research on gesture recognition based on support vector machine / Z. M. Liao // 59-я научная конференция аспирантов, магистрантов и студентов БГУИР. / Белорусский государственный университет информатики и радиоэлектроники. – Минск, 2023. – Р. 168 – 170.

4–A. Qiu, Y. P. Y. Object recognition based on skeleton / Y. P. Y. Qiu, J. Ma, Z. M. Liao // Технологии передачи и обработки информации : материалы международного научно-технического семинара, Минск, март-апрель 2022 г. / Белорусский государственный университет информатики и радиоэлектроники. – Минск, 2022. – С. 74 –76.