

# Naive Bayes Image Classification Based on Multiple Features

Kun Fang\*

China Jiliang University, Hangzhou 310018, China. E-mail: 354655821@qq.com

---

## ABSTRACT

In the field of computer science, in order to realize the classification of images, a Naive Bayes method based on multiple features has been proposed. This method is widely used and rich. It uses four features, such as extracting gray histogram features, SIFT features, SURF features, and reducing the dimensions of the dataset from data image sets. Naive Bayesian methods are used to obtain the accuracy, recall, and F1 value of the image for each feature. This paper analyzes the application and feature comparison of Naive Bayes method in images, and shows that image representation using SURF feature description can achieve better classification results.

**Keywords:** Naive Bayes Method; Naive Image Feature Classification; SURF Image Feature

---

## Introduction

As it is known to all, with the continuous development of modern scientific image technology, computer image technology has been widely developed and applied in many scientific fields of human society, bringing convenience to the research, design, development, production and use of various scientific things. The widespread application of computer science graphics and image processing technology has made image signal processing and mathematical computer graphics into two independent and closely related disciplines.

## 1. Concept and image application of Naive Bayes method

In order to effectively implement image classification for more accurate optimization of various image features, a Naive Bayes method based on multiple image features was first proposed and widely used in image processing in various computer 3D graphics. Naive generalized Bayesian method is actually an image classification type method based on Bayes' theorem and image feature generation conditions to generate hypotheses independently. It is an extremely simple

image classification type algorithm. Theorem's theorem is the theoretical basis, so it is generally referred to as the Bayesian classification in the industry. Naive Bayesian method has occupied an unparalleled important position in the field of image information processing, which has important guiding significance for the field of image information processing. The use of computer naive Bayesian processing technology can complete the set work goals in a short time, with high accuracy and good efficiency. When using naive Bayesian methods to process graphic images for industrial industry design and production, the design is higher, the accuracy is higher, and the efficiency is significantly improved, which is compared to manual calculation and processing. In addition, some specifications require particularly advanced industries. In the design, the level of image processing using artificial technology often cannot reach the same level as that using computer naive Bayesian method. Computer naive Bayesian processing technology can complete the set work goals in a short time, with high accuracy and good efficiency. In addition, computer naive Bayesian processing of graphic images can better meet the requirements of mass production, effectively reduce the cost of the project in the quality inspection process, and has profound significance for

industrial design graphic image processing.

The various characteristics and advantages and disadvantages of Naive Bayesian method

## **2.1 Features and advantages and disadvantages of gray histogram**

The gray histogram feature is also called the order statistic feature. Select a part of the target area in the entire image for systematic statistical calculation, obtain the corresponding statistical quantity, and then use the gray level to describe the image. The gray histogram feature has strong environmental adaptability. The gray histogram feature prescriptive feature can transform the histogram of the image into the shape we need, and purposefully increase the image of the gray area of a certain area, so that users get the required information points. On the other hand, the gray histogram shape operation is susceptible to noise interference. The image processing built on the robustness of the segmentation algorithm still has many shortcomings. The gray histogram features of the Naive Bayes method are only applicable to Gray image processing. If you want to perform order statistics on color images, you need to grayscale the color image first, and then proceed to the next image processing. There are still many things to be improved in the details of the image processing program.

## **2.2 SIFT characteristics and advantages and disadvantages**

The SIFT scale feature in the Naive Bayes method was first proposed as early as 1999. It is a feature algorithm based on computer 3D vision and is widely used to detect and analyze various localities in the 3D image space. The scale feature because it has the feature of finding an extreme value in the scale in three-dimensional space, and by extracting and finding out the position, scale, and rotation of the program different from the center of the variable, and the operation is rigorous, scientific, and precise. In image graphics processing, it is a very good local image feature description processing algorithm, which is widely used in various image graphics processing and various computer graphics. The characteristics of SIFT can also continue to maintain relative data stability in real-time processing of various visual signal transformation, noise

and anti-interference data, but the operating time will be highly real-time and inefficient, detecting different feature values. There will be more points than other detection feature points, and the complexity may be higher, and the real-time processing data efficiency of normal operating applications may need to be continuously improved.

## **2.3 SURF features and benefits**

SURF characterization features and technical advantages. When performing image processing on an image feature that has not been registered, we may easily encounter similar problems with different image feature sizes. The distance and size of each point in each feature of the image is different. If we modify the distance of the feature points of the image by registration, it may directly cause the matching of the feature strength of each picture. The SURF characterization technology is proposed to solve this similar problem, and the scale is fixed. Feature SURF characterization feature detection also plays an important guiding role in modern image picture processing and feature detection. SURF feature detection in the naive Albus method is an excellent and enhanced version based on SURF characterization features. SURF The representative feature detection algorithm uses the naive Hessian feature matrix method to determine each candidate point, which greatly reduces the complexity of image processing calculations. The application program design is simple and the efficiency is relatively higher than the two feature points with SIFT. It is also more convenient when it comes to graphic images, and realizes the scale feature Detection and image matching, is applied to the generalized mathematical computer graphics, graphic binocular vision and science of metrology systems.

## **2.4 Data reduction methods reduce the characteristics and advantages and disadvantages of the dimensional method**

The method of data set reduction is to reduce the dimensional method. The method of data set reduction is to reduce the dimensional method. First, the picture is gray-scaled, and the position of the target in the image is used to extract. The data set can be processed to generate a new target data set to achieve the purpose of reducing

dimensions. This method can efficiently process large batches of images. The disadvantage is that the time required to reduce the dimensional method features by using the data set reduction method will be longer, which cannot meet the urgent needs of large batches.

### 3. The role of each feature of Naive Bayes method in image processing

Through the analysis to understand the four methods of feature classification based on Naive Bayes image extraction based on multi-feature extraction. This classification method is used to apply the four special feature reduction methods mentioned in the above article to one thousand naive graphic images, that is, features in gray histograms, SIFT image features, SURF image features, and gray data set feature reduction calculation methods, as well as features in methods to reduce feature dimensions, accuracy in recalling a thousand plain graphic images, recall Rate, F1 value. One thousand square pictures are specially divided into ten groups, each group of pictures has one hundred square pictures, and an average of the precision, recall, and F1 values in each group of graphic images is obtained, and the precision, recall, and F1. The value calculation results can be gray histogram ellipse graphic features: 0.497 0.484 0.47. Sift histogram features: 0.540 0.550 0.539. Surf features: 0.587 0.591 0.575. The calculation method of data set gray cuts features: 0.473 0.466 0.460. The accuracy, recall, and F1 average result data show that the Naive Bayesian method SIFT features and SURF features are better in processing images.

### 4. Naive Bayesian feature SURF feature processing image optimum

The SIFT feature and the SURF feature are both excellent scale-invariant features. The SIFT feature is first operated using a non-maximal suppression procedure and then filtered using a Hessian matrix, while the SURF feature is first operated using a Hessian matrix program and then non-polar. The operation is greatly suppressed, and the order of the operation steps is changed. The effect is indeed very significant. When the SURF feature algorithm is used to process the image, the original image remains unchanged and only the filter size is changed. The speed of processing the image naturally

increases. Graphic image processing adopts the method of extracting SURF features. Two strategies are adopted: detection and description. First, the image target feature points are detected and then the image target feature points are described. The feature points are easy to detect and not easily affected by light. Efficiency will be more efficient. Although the stability of SIFT features is also high, the key point feature points it finds are limited to some very prominent points. Through the above analysis and the above image processing results using the Naive Bayes method, the accuracy, recall, and F1 values of the image are analyzed. The results show that the SURF feature is the best in processing images.

### 5. Conclusion

The purpose of early digital image processing was to improve the quality of the image. Mostly, people were the main object, and the purpose was to present better visual effects to the viewer. Naive Bayes image classification based on various features has been widely used. There are many fields, such as scientific research, industry, agriculture, military field, medical field, literature and culture, playing an increasing role in life, and now it has become a new type of graphics subject with strong development potential and broad market.

### References

1. Naive Bayes Classification (very simple and clear interpretation) (in Chinese) [Internet]. 2018. Available from: [https://blog.csdn.net/wyq\\_wyj/article/details/79485618](https://blog.csdn.net/wyq_wyj/article/details/79485618).
2. Image Grayscale Features-first-order Statistical Features / Grayscale Histogram Features (Matlab Implementation) (in Chinese) [Internet]. 2019. Available from: <https://blog.csdn.net/SongGu1996/article/details/98749834>.
3. Shen Y. Research on criminal image classification and retrieval based on machine learning. Xi'an Technological University 2018.
4. Zhang B. Naive bayesian classifier method based on the classification of the probability weighted. Journal of Chongqing University of Technology (Natural Science) 2012; 26(7): 81-83. doi: 10.3969/j.issn.1674-8425-B.2012.07.017.
5. Machine learning methods of low-dose CT imaging have achieved good results (in Chinese). Journal of Health and Science 2019.
6. Zhang X, Jiang J, Hong R, et al. Accelerated image

- classification algorithm based on naive Bayes K-nearest neighbor. Journal of Beijing University of Aeronautics and Astronautics 2015; 41(2): 302-310. doi: 10.13700/j.bh.1001-5965.2014.0471.
7. Liu C. Image recognition comparison based on naive bayes and semi-naive bayes. Information Technology and Network Security 2018; 37(12): 48-51. doi: 10.19358/j.issn.2096-5133.2018.12.010.
  8. Accuracy, Precision, Recall, F1 Value, ROC / AUC Collation Notes (in Chinese) [Internet]. 2018. Available from: <https://blog.csdn.net/u013063099/article/details/80964865>.