

Currency Value Detection and Counting using Feature Extraction and Classification Algorithms

Sindhuja Reddy Kamidi¹, Janardhan Reddy Kamidi²

¹U.G. Student, Department of Information Technology, Chaitanya Bharathi Institute of Technology, Telangana, India

²Graduate from Department of Computer Science Engineering, Keshav Memorial Institute of Technology, Hyderabad, Telangana, India

Abstract: *Image processing is the area in which currently many predictions are performed from the captured images. Currency Value detection aims at searching the features of the training dataset with the query image feature and obtains the respective value of the feature that has the match with query image. In present system the images are compared and matched according to pixel values and few features extracted, which consumes a lot of time for execution instead we can extract the desirable features that are used to identify and classify the type of image. The proposed technique aims to reduce the searching time of image retrieval and hence it improves the performance of image retrieval system using best suitable feature extraction algorithms offered by Matlab and instead of sticking to one feature extraction algorithm it is better to apply multiple feature extraction methods that improves the accuracy. To build an effective classifier it is better to add new components and new changes to the existing algorithms that increase the good impact on the classifier.*

Keywords: Classification, Feature Extraction, Image Processing, Machine learning algorithms

1. Introduction

Currency value detection is very useful and necessary in counting the notes and obtaining the total of the amount. It also helps the blind people to recognize the value of the note that they are holding just by scanning. If we configure the output with the text to voice converter the results can be heard by the blind.

There has been lot of research and implementations for the Currency value detection. But the features extraction methods in image processing is also gaining its momentum. Efficient image database retrieval can be done only if we have a system that is able to automatically extract relevant features directly from the images stored in the database.

But before implementing some classification algorithm it is very important to collect the currency of different nations and its values. It is also important to collect the relative values conversions from one type of currency to other to make the calculations more appropriate algorithm. To implement these algorithms we can use the R TOOL or MATLAB or JAVA or any programming language which does the basic functions, but if we use R TOOL or MATLAB the pre-processing gets easier as it has some pre-defined functions which can be directly implemented.

Once we get the dataset for training further we can proceed with feature extraction and classification algorithms. MATLAB provides many predefined packages for feature extraction like Local Feature Extraction, Feature Matching, Image Registration, Geometric Transformations. For currency value detection the best feature extraction method is to use Local Feature Extraction as it deals with the extraction of internal and interesting points to extract feature descriptors. Using such feature extraction methods will increase in accuracy and efficiency.

For each feature extracted the corresponding class label is given and stored into the database. A classification

algorithm is implemented to train the system with the available data so that it can classify the query image based on the features extracted.

The implemented algorithm will detect the value of currency and applies the basic mathematical operations to find the total amount.

2. Methods and Procedures

In Local Feature Extraction there are many methods of detection of interesting points and extracting the features. SURF Feature extraction is one among them. The SURF algorithm alone may not be much efficient to make it more accurate we need to implement more than one feature extraction methods. As the currency detection is done basically by the number represented on the currency. So we can also implement the OCR (Optical character recognition) to recognize the numbers. To make it more efficient we can also take mean and standard deviation of neighbor pixel values to include a feature which deals with color of objects on currency.

2.1 Feature Extraction Methods

SURF Feature Extraction method: speeded up robust features (SURF) is a patented local feature detector and descriptor. It can be used for tasks such as object recognition, image registration, classification or 3D reconstruction. To detect interest points, SURF uses an integer approximation of the determinant of Hessian blob detector, which can be computed with 3 integer operations using a precomputed integral image. Its feature descriptor is based on the sum of the Haar wavelet response around the point of interest. These can also be computed with the aid of the integral image. SURF descriptors have been used to locate and recognize objects, people or faces, to reconstruct 3D scenes, to track objects and to extract points of interest.

The algorithm has three main parts: interest point detection, local neighborhood description and matching.. for interest point detection before applying the algorithm the SURF uses gaussian filter to smooth the image and get the effective and interesting points more accurately. The sum of the original image within a rectangle can be evaluated quickly using the integral image, requiring evaluations at the rectangle's four corners.

2.1.1 Surf Feature Extraction Method:

SURF uses a blob detector based on the Hessian matrix to find points of interest. The determinant of the Hessian matrix is used as a measure of local change around the point and points are chosen where this determinant is maximal. SURF also uses the determinant of the Hessian for selecting the scale. Given a point $p=(x, y)$ in an image I , the Hessian matrix $H(p, \sigma)$ at point p and scale σ , is shown in fig 2



Figure 1: SURF features extraction of interesting points

2.1.2 Optical Character Recognition Algorithm (Ocr)

OCR is another feature extraction method to extract the text and numbers on the currency. Optical character recognition is used to detect the characters of different form present in the image. So we can extract few features depending on the values on the currency. MATLAB provides a predefined implementation of OCR. We can use it to extract the features that we want.



Figure 2: OCR applied for a currency image

In order to recognize many variations of the same character, features that are invariant to certain transformations on the character need to be used. Invariants are features which have approximately the same values for samples of the same character that are, for example, translated, scaled, rotated, stretched, skewed, or mirrored

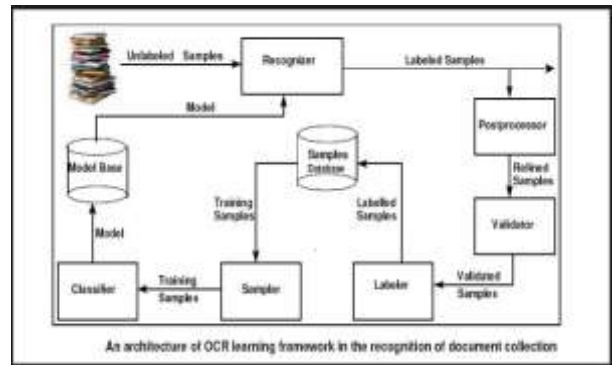


Figure 3: Flow of execution of algorithm

2.1.3 Mean and Standard Deviation

Mean and standard deviation of successive pixel values is also to be calculated and considered as the feature. Mean and standard deviation of pixel values obtains the feature depending on the color of the objects and their positions. As on any currency the arrangement of elements and color, font of the elements remains same.

Thus the features extracted using SURF, OCR and mean standard deviation we collect them and store it in a csv file format to apply classification algorithm and train the system.

$$\sigma = \sqrt{\frac{\sum [x - \bar{x}]^2}{n}}$$

Where X is attribute value, n is total number of examples

2.2 Dataset Format

Dataset holds the features that are extracted using feature extraction methods. The dataset may be in the format of CSV file. Where it contains many columns for concentrated point's average, OCR recognized numeric value, mean of pixels, standard deviation of pixels, actual class label. This dataset is used to train the classifier.

SURF points mean value	SURF points deviation value	OCR number recognized	Mean value of all the pixel values	Standard deviation of all the pixel values	Class label

2.3 Classification Methods

As the dataset we obtained is completely numeric the K-nearest neighbor algorithm can be used to train the classifier. KNN algorithm is best to use while training the classifier where there is a small dataset and it has numeric values. Basically KNN falls under supervised learning algorithms. KNN algorithm is a Lazy learner which starts comparing with the each and every tuple in the dataset and finds the class under which it falls. it is a non-parametric method used for classification and regression. In both cases, the input consists of the k closest training examples in the feature space. The distance to the kth nearest neighbor can also be seen as a local density estimate and thus is also a popular outlier score in anomaly detection. The larger the distance to the k-NN, the lower the local density, the more likely and the query point is an outlier.

Distance finding in KNN can be done in four ways:

1. Euclidean distance computes the root of square difference between attribute values of objects.

$$Dist_{XY} = \sqrt{\sum_{k=1}^m (X_{ik} - X_{jk})^2}$$

2. Manhattan distance computes the absolute differences between attributes of pair of objects

$$Dist_{XY} = |X_{ik} - X_{jk}|$$

3. Chebychev Distance is also known as maximum value distance and is computed as the absolute magnitude of the differences between attribute values of a pair of objects.

$$Dist_{XY} = \max_k |X_{ik} - X_{jk}|$$

4. Minkowski Distance is the generalized metric distance. Note that when p=2, the distance becomes the Euclidean distance. When p=1 it becomes city block distance. Chebyshev distance is a variant of Minkowski distance where p=∞ (infinity taking a limit). This distance can be used for both ordinal and quantitative variables.

$$Dist_{XY} = \left(\sum_{k=1}^d |X_{ik} - X_{jk}|^{\frac{1}{p}} \right)^p$$

2.4 Implementation Of Algorithms

Step 1: collect the currency pictures that are to be used to train the system.

Step 2: implement a feature extraction method and extract the features and save it into a dataset.

Step 3: train the classification system using KNN algorithm. (Note: give more coefficient value that have more impact on the output)

Step 4: Test the system by giving unknown image to system and find the accuracy of the system.

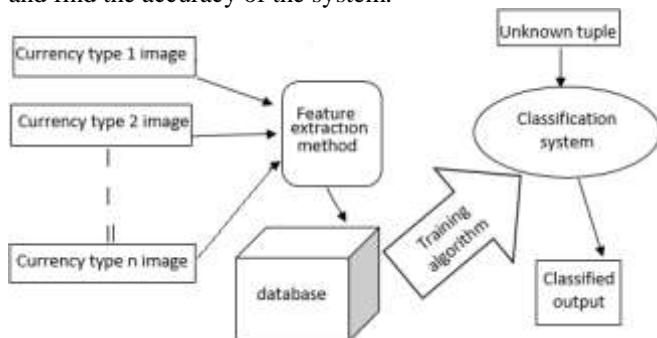


Figure 1: Flow Chart representing the process

Features that differentiates from currently existing system: As the three feature extraction algorithms are used the results are more appropriate. The system can identify the value of the currency even if the input is mirror image to the actual image. It detects the value in whatever direction the image may take, as OCR feature extraction algorithm is used. Once the features are extracted there is no need of holding the images in the dataset. The only values that we store in our dataset is numeric values. So it results in reduced size of database and easy to store and retrieve the data. As the coefficients are given according to the weightage of the feature the classification algorithm is more efficient.

3. A Real Life Example

Consider the currency counting machine which has to scan all the mixed currencies of different countries and find the total according to the selected country currency value and displays it. To do this the machine has to be intelligent enough to convert all the currencies accordingly to a single form and find the total. Here is the situation where the currency value detection comes into consideration with efficient and effective calculations. It can also used in ATM machines. It is very much helpful for blind people, when the currency is scanned and given to the system it automatically detects and if connected to text to voice conversion system it speaks out the result and the person can identify the currency value that he scanned.

4. Conclusion

Here by we conclude that the total amount of the currency can be calculated using currency converting formulas and the detected value by classifier by summing up the results after conversion. Using multiple feature extraction methods makes the system more efficient and accurate. Here, SURF is used to find the concentrated points and positions, OCR is used to find the characters and numbers on the currency, the color and texture is also considered as the feature by obtaining mean and standard deviation. While training the classifier it is important to prioritize the features that impact more on the output and add a high coefficient value such that the results become more sensitive to the prioritized feature.

References

- [1] mark Nixon & Alberto Aguado, "Feature Extraction and image processing"
- [2] Saravanan Thirumuruganathan, "A Detailed Introduction to K-Nearest Neighbor (KNN) Algorithm".
- [3] Data Mining Case Studies by Yachang Zhao.
- [4] Han and Kamber- Data Mining Concepts and Techniques 3rd Edition.
- [5] Richard Szeliski "Computer Vision: Algorithms and Applications".
- [6] Bahram Javidi, "image processing and classification"

Author Profile



Kamidi Sindhuja Reddy is presently Studying Final year B.E in Information Technology, CBIT, Hyderabad. She did few mini projects in the area of Machine learning, Image processing in previous years. She is currently doing project in Data mining Big data and Artificial Intelligence domain.



Kamidi Janardhan Reddy is B.Tech graduate in Computer Science and Engineering, KMIT, Hyderabad. He did few mini projects in the area of web applications, Machine learning, Image processing Data mining, Big data and Artificial Intelligence in previous years. He is currently doing project in machine learning in healthcare domain.