

Currency Recognition using SIFT

S. A. Bhavani

Assistant Professor

Department of Computer Science and Engineering
Anil Neerukonda Institute of Technology & Sciences
Bheemunipatnam, Sangivalasa-53116
Visakhapatnam (District), AP

ABSTRACT

Currency recognition is an important area of pattern recognition. A system for the recognition of currency is one kind of intelligent system which is a very important need of the current automation systems in the modern world of today. Currency Recognition and is implemented to reduce human power to automatically recognize the amount monetary value of currency and without human supervision. The software interface that we are proposing here could be used for various currencies (we are using four in my paper). Many a times, currency notes are blurry or damaged; many of them have complex designs to enhance security. This makes the task of currency recognition very difficult. So it becomes very important to select the right features and proper algorithm for this purpose. The basic requirements for an algorithm to be considered as practically implementable are simplicity, less complexity, high speed and efficiency. Our main aim is to design an easy but efficient algorithm that would be useful for maximum number of currencies, because all currencies have different security features, making it a tough job to design one algorithm that could be used for recognition of all available currencies. Writing different programs for all is also a tedious job. The algorithm used in my paper is SIFT (**Scale Invariant Feature Transform**) algorithm. The aim of the paper is to recognize the currencies.

Keywords

Currency Recognition, SIFT, Digital Image Processing.

1. INTRODUCTION

There are many Currencies all over the world, with each of them looking totally different. For instance the size of the paper is different, the same as the color and pattern. The staffs who work for the money exchanging (e.g. Forex Bank) have to distinguish different types of currencies and that is not an easy job. They have to remember the symbol of each currency. This may cause some problems (e.g. wrong recognition), so they need an efficient and exact system to help their work. As we mentioned before, the aim of our system is to help people who need to recognize different currencies, and work with convenience and efficiency. For bank staffs, there is a "Currency Sorting Machine" helps them to recognize different kinds of currencies. The main working processes of "Currency Sorting Machine" are image acquisition and recognitions. It is a technique named "optical, mechanical and electronic integration", integrated with calculation, pattern recognition (high speed image processing), currency anti-fake technology, and lots of multidisciplinary techniques. It is accurate and highly-efficient. But for most staffs, they have to keep a lot of different characteristics and anti-fakes label for different commonly-used currencies in their mind. However, each of them has a handbook that about the characteristics and anti-

fakes labels of some less commonly-used currencies. Even for that, no one can ever be 100 per cent confident about the manual recognition. Otherwise our system is based on image processing, techniques which include filtering, edge detection, segmentation, etc. In order to make the system more comprehensive, we need to create a small database for storing the characteristics of the currency.

In our system, we scan the images of each Currency note and then detect the some of the important features or key features of a Currency note using SIFT (Scale Invariant Feature Extraction) Algorithm and we store the images in the data base and we train our system with the respective key features and the correlation values of each note. Now the testing sample image comes into picture and the comparison between the original images stored in the database and sample scanned image is done. The comparison is done between the features and correlation values of trained set and new sample images. This is done using Radial Basis Networks which consists of three layers. Input layer, Hidden Layer and Output layer.

Then the results are displayed accordingly. If the note is matched, it displays the matched note else no match of the note is displayed. The classification of images is done only on JPEG images only.

2. RELATED WORK

2.1 Definition of image

Data representing a two-dimensional scene. A digital image is composed of pixels arranged in a rectangular array with a certain height and width. Each pixel may consist of one or more bits of information, representing the brightness of the image at that point and possibly including color information encoded as RGB triples.

Picture a visual representation (of an object or scene or person or abstraction) produced on a surface, "they showed us the pictures of their wedding" a movie is a series of images projected so rapidly that the eye integrates them.

Effective indexing and retrieving desired image in large image database in the basis of features such as color, text and shape that can be automatically extracted from the images themselves.

2.2 Image Processing

The analysis of a picture using techniques that can identify shades, colors and relationships that cannot be perceived by the human eye. Image processing is used to solve identification problems, such as in forensic medicine or in creating weather maps from satellite pictures. It deals with images in bitmapped graphics format that have been scanned in or captured with digital cameras. Any image improvement,

such as refining a picture in a paint program that has been scanned or entered from a video source.

2.3 Image Retrieval and Information Retrieval

Since the 1970s Image Retrieval has become a very active research topic, with two major research communities, database management and computer vision. One is text-based and another is visual-based. Text-based image retrieval has become very popular since 1970s, which involves annotating the image with keywords, and use text-based database management systems (DBMS) to retrieve the images. In text-based image retrieval system, keywords of semantic information are attached to the images.

They can be typed manually or by extracting the captions of the images. It is very efficient for simple and small image databases, since the whole database can be described by just few hundreds of keywords. But in the 1990s, several large leaps in development of processor, memory and storage made the size of image databases grow dramatically. As the image database and image size grow, there will be more images having different contents and the images having rich contents cannot be described by only several Semantic keywords. The demand of labour on annotating the images also rises dramatically.

Retrieval image providing effective and efficient tool querying large image database. Information retrieval provides the textual representation of images. It requires the text descriptions to the respective images.

Recent technology development in various fields has made large digital image databases practical. Well organized database and efficient browsing, storing, and retrieval algorithms are very important in such systems. Image retrieval techniques were developed to aid these components.

Image Retrieval was originated from Information Retrieval, which has been very active research topic since 1940s. "We have huge amounts of information to which accurate and speedy access is becoming ever more difficult." In principle, Information Retrieval is simple. It can be illustrated by a scene of a store of documents and a person (user of the store). He formulates a question to which the answer is a set of documents satisfying his question. He can obtain the set by reading all the documents in the store, retaining the relevant documents and discarding all the others. In this scene, it is a 'perfect' retrieval. But in practice, we need to model the "read" process in both syntactic and semantic to extract useful information. The target of Information Retrieval is not only "how to extract useful information", but also "how to measure relevance among documents". These challenges also exist in Image Retrieval.

Also the keywords are very dependent on the observer's interest and they are subjective. Captions are not always precisely describing the picture. Indexing and searching a large image database via keywords are time-consuming and inefficient. Content Based Image Retrieval (CBIR) researches attempt to automate such complex process of retrieving images that are similar to the reference image or descriptions given.

2.4 Content Based Image Retrieval

Content-based image retrieval also known as query by image content and content-based visual information retrieval problem of searching for digital images in large database. Content-based means that the search will analyze the actual

contents of image. The term content in this context might refer to colours, shapes, textures or any other information that can be derived from the image itself.

The earliest use of the term Content Based Image Retrieval in the literature seems to be by Kato, was to describe his experiments in automatic retrieval of images from a database by color and shape features. The term has since been widely used to describe the process of retrieving desired images from a large collection on the basis of features (such as colour, Texture and shape) that can be automatically extracted from the images themselves. The features used for retrieval can be either primitive or semantic, but the extraction process must be predominantly automatic.

The ideal approach of querying an image database is using content semantics, which applies the human understanding about image. Unfortunately, extracting the semantic information in an image efficiently and accurately is still a question. Even with the most advanced implementation of computer vision, it is still not easy to identify an image of horses on a road. So, using low level features instead of semantics is still a more practical way. Until semantic extraction can be done automatically and accurately, image retrieval systems cannot be expected to find all correct images. They should select the most similar images to let the user choose the desired images. The number of images of retrieved set can be reduced by applying similarity measure that measures the perceptual similarity.

A typical CBIR system consists of three major components and the variations of them depend on features used.

- i. **Feature extraction** – Analyze raw image data to extract feature specific Information.
- ii. **Feature storage** – Provide efficient storage for the extracted information, also help to improve searching speed.
- iii. **Similarity measure** – Measure the difference between images for determining the relevance between images.

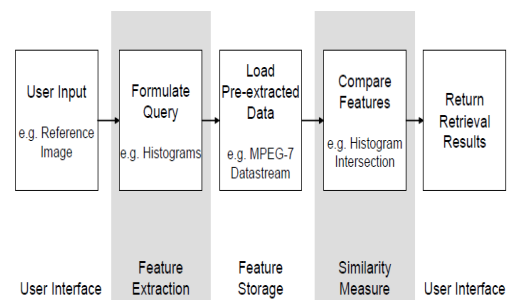


Fig 1: CBIR Function

2.5 Content Based Image Retrieval using colour

Retrieving image based on colour similarity is achieved by computing a colour histogram for each image that identifies the proportion of pixels within an image holding specific values. Current research is attempting to segment colour proportion by region and by spatial relationship among several colour region.

Among different types of low level features, colour is the most straightforward information which can be easily retrieved from digital images with simple and compact description, while others require more pre-processing and computational tasks such as pattern recognition or texture analysis.

While comparing image by colour feature, three properties are usually considered:

- i. **Area of matching** – Count the area or number of pixels having same or similar colours. Larger matched area means more similar.
- ii. **Colour distance** – Distance between colours, usually in a perceptually uniform color space. Closer between matched colours means more similar.
- iii. **Spatial distribution** – Usually used while combining colour with other features such as texture and shape.

In a typical colour similarity measure, area of matching is usually counted as the similarity colour distance is used to control the matching between colours and to adjust the similarity. In conventional colour image retrieval system, the most straight forward approach is using colour histogram. Histograms of each colour, for example, images of 256 colours, will be generated. Similarities between such images are then performed and measured by **Histogram Intersection Method** (HIM). This is the basic approach and can give simple and efficient representation of colour distribution. Histogram approach is not limited by taking the number of pixels of each colour in the image or using HIM similarity measure. Indexes of histograms can represent many types of features such as colours in different colour space, coefficients in transformed domain or spatial-related information. There are also many variations in comparing histograms. But histograms have a limitation that the feature space is fixed, compactness of the description is limited, because histograms will not skip non-existed colours. One argument that we can use a lower resolution histogram to improve the compactness but it is a trade off between compactness and accuracy.

2.6 JPEG visual descriptors

That shape often carries semantic information follows from the fact that many characteristic objects can be visually recognized solely from their shapes. This distinguishes shape from other elementary visual features such as colour, or texture. But the notion of object shape has many meanings. To deal with 3D real-world objects, JPEG standard has a 3D shape descriptor.

JPEG has defined a set of standard descriptors for description and storage of the most commonly used features. This makes the extracted features more accessible. Since the required storage size is much smaller than compressed images files. Moreover, the format of the data is fixed, so the data can be used in any JPEG compatible systems. Thus comparison between algorithms can be done easily if the implementations of the target Algorithms are JPEG compatible.

In JPEG visual standard, some colour descriptors are defined, including several histogram based descriptors representing different colour features, and a Dominant Colour Descriptor (DCD). DCD describes colour feature by a set of representative colours with their percentage and each colour have at least a certain distance away in CIE colour space controlled by a threshold T_d . It is very compact since there is no redundant information for non-existed colours, and similar colours are grouped into a palette colour.

2.7 Relevance Feedback

Although JPEG defined efficient and most commonly used CBIR methods, content based methods still have limitations that they may not be able to find the images that exactly match user's expectation. One reason is that a precise query cannot be formulated Although DCD can describe colour features in

a compact and effective way, Drawbacks of its default similarity measure method pull down the performance of DCD.

By just giving an image as query, Interactive searching may be used for improving the retrieval result by refining the query by user's feedback. JPEG did not handle interactive searching directly. They use content management approach to describe multimedia contents in a structural format. This uses textual semantics. This can improve the efficiency of browsing and text based searching, but not for content based searching. Also it is not interactive and the improvement may not show instantly. For content based interactive searching, Relevance Feedback (RF) is a commonly used technique which use user selected relevance information to refine the query.

It can be used as an extension of similarity measure and will not affect JPEG standard. Since the representations depend on the features, relevance feedback algorithms may be different for each descriptor. In this research, a merged palette histogram approach is proposed to improve Dominant Colour Descriptor searching with use of RF.

But the effect of this coefficient is still not very clear. It is unable to balance the effect of "area of matching" and "distance between colours".

It causes ambiguous retrieval results. These problems will be described in chapter 4, and a new Merged Palette Histogram Similarity Measure (MPHSM) will be proposed to tackle these problems.

3. PROBLEM ANALYSIS

3.1 Existing System

In earlier days, image retrieving from large image database can be done by following ways. We will discuss briefly about the image retrieving of various steps

- Automatic Image Annotation and Retrieval using Cross Media Relevance Models
- Concept Based Query Expansion
- Query System Bridging The Semantic Gap For Large Image Databases
- Ontology-Based Query Expansion Widget for information Retrieval
- Detecting image purpose in World-Wide Web documents.

3.2 Proposed System

Relevance feedback is an interactive process that starts with normal CBIR. The user input a query, and then the system extracts the image feature and measure the distance with images in the database. An initial retrieval list is then generated.

User can choose the relevant image to further refine the query, and this process can be iterated many times until the user find the desired images.

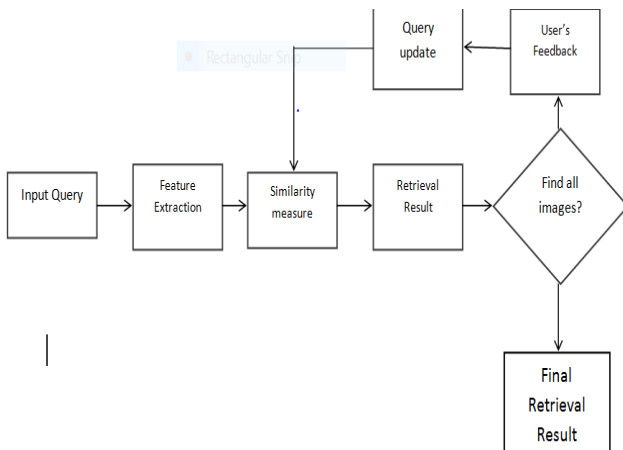


Fig 2. Block Diagram of proposed system

3.3 Structure of the project:

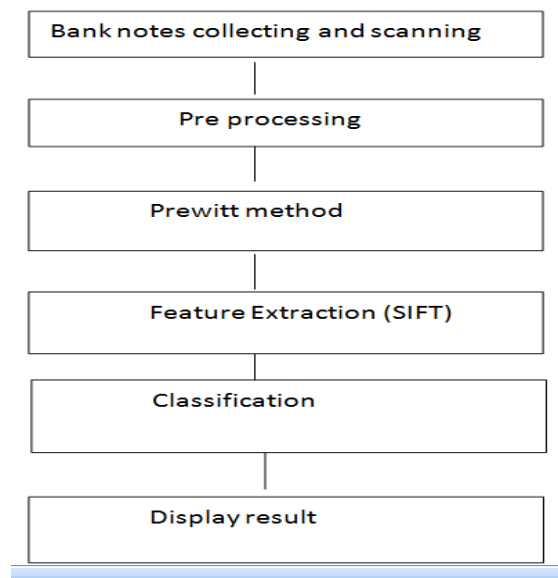


Fig 3: Architecture of project

Aim of the proposed algorithm is to develop an algorithm which can be easily applied to number of different currencies and has good efficiency and high speed. We consider 500,100,50,20,10 rupee of Indian Currency.

- Obtaining the Input Image through a scanner or a camera.
- Preprocessing Operations makes extraction of features easier. In this particular case, pre processing operations involve gray scale conversion, detecting edges by prewitt filter and Canny's edge detection method.
- The features of the image are extracted using **Scale Invariant Feature Transform** algorithm (SIFT).
- Classification of the currencies is by Radial Based Function.
- Displaying results.

4. IMPLEMENTATION

4.1 Module 1:

4.1.1 Pre Processing of image:

- The task of pre-processing is achieved by converting colored currency images into gray scale, which facilitates further pre-processing.
- After that, the edge of the image is filtered using Prewitt method and Canny's edge detection method.

A) RGB TO GRAY SCALE CONVERSION:

RGB Colour image:

In RGB colour model, each colour appears in its primary spectral components of red, green and blue. The colour of a pixel is made up of three components; red, green, and blue (RGB), described by their corresponding intensities.

$$I_{RGB} = (fR, fG, fB)$$

Where $fR(x,y)$ is the intensity of the pixel (x,y) in the red channel, $fG(x,y)$ is the intensity of pixel (x,y) in the green channel, and $fB(x,y)$ is the intensity of pixel (x,y) in the blue channel.

Gray Scale image: Gray scale is an image carries only intensity information.

It has range of shades of gray without apparent colour. The darkest possible shade is black and the lightest possible shade is white.

Black is represented by $R = G = B = 00000000$, and white is represented by $R = G = B = 255$ or 11111111 . Because there are 8 bits in the binary representation of the gray level, this imaging method is called 8-bit gray scale.

Conversion of RGB Color image to Gray Scale image:

$$color = (Fr + Fg + Fb) / 3$$

$$I_y = 0.299Fr + 0.587Fg + 0.114Fb$$

B) PREWITT OPERATOR

Prewitt operator is used for edge detection in an image. It provides two masks. One for detecting edges in horizontal direction and the other for detecting edges in vertical direction.

Vertical direction mask: When this mask is convolved in an image, it gives the vertical edges in the image. It simply works like a first order derivative and calculates the difference of pixel intensities in an edge region.

-1	0	1
-1	0	1
-1	0	1

Horizontal direction mask: When this mask is convolved in an image, it gives the horizontal edges in the image. It calculates difference among the pixel intensities of a particular edge.

-1	-1	-1
0	0	0
1	1	1

4.2 Module 2:

4.2.1 Feature Extraction

One very important area of application is image processing, in which algorithms are used to detect and isolate various desired portion or shapes of digitized image.

SCALE INVARIANT FEATURE TRANSFORM is one of the low level feature extraction techniques which we use.

SIFT is used to detect and describe local features in images and can help in object recognition. We begin by detecting points of interest, which are termed key points. The image is convolved with Gaussian filters at different scales and then the difference of successive Gaussian- blurred images are taken(DOG).The extraction of these features the SIFT algorithm applies a 3 stage filtering approach.

Scale-Space Extreme Detection

This stage of the filtering attempts to identify those locations and scales that are identifiable from different views of the same object. This can be efficiently achieved using a "scale space" function. Hence this algorithm is scale invariant.

- Difference-of-Gaussian scale-space function, $D(x,y,\sigma)$ with key point as the origin. This Taylor expansion is given by Taylor series which defines $D(X)$.

Key point Localization

This stage attempts to eliminate more points from the list of key points by finding those that have low contrast or are poorly localized on an edge. If $D(X) < 0.03$, then they are considered as the bad contrast. Also reject the points of strong edge response.

Orientation Assignment

This step aims to assign a consistent orientation to the key points based on local image properties. This Remove effects of scale and rotation.

4.3 Module 3:

4.3.1 Classification of Images

The correlation coefficient of two different image matrices gives very important information about the similarity matching between them. It can be evaluated between two matrices.

Where A is the first $m \times n$ matrix, B is the second $m \times n$ matrix, A is the mean of A , B is the mean of B , and r is the correlation coefficient between A and B . The greater the values of r , the more the two matrices are similar. It is worth mentioning that since the correlation values cannot be evaluated before hand, they are only evaluated after the input image is provided. This cause some delay as the input image is not directly compared to pre-stored values. However, correlation values are important and provide a very robust classification power.

4.3.2 Radial Basis Function

One of the mostly used types of Radial Based Neural Networks is the one based on Gaussian radial basis.

$$S_j(X) = \exp(-1/2 (X-U_j)^{-1} (X-U_j))$$

X is the input vector with elements X_i , and U_j is the center of the radial basis function.

The used Radial Basis Function Network Classifier contains 25 neurons in the hidden layer. Since the correlations are not available before hand, when the image is input, the system calculates all the correlations, then it builds and trains the network and finally classifies the images.

5. RESULTS

RGB to Gray Scale conversion:



Edge detection using Prewitt operator:



Input image for testing:



Final output:



6. CONCLUSIN

A new perceptual model based on a set of computational measures corresponding to perceptual textural features, namely coarseness, directionality, contrast, and busyness, was introduced in this paper. Computational measures are based on two different representations (viewpoints): original images and the autocorrelation function associated with images. Coarseness was estimated as an average of the number of extreme. Contrast was estimated as a combination of the average amplitude of the gradient, the percentage of pixels having the amplitude superior to a certain threshold and coarseness itself. Directionality was estimated as the average number of pixels having the dominant orientation(s). Busyness was estimated based on coarseness. The computational measures proposed for each perceptual textural feature were evaluated, based on a psychometric method, by conducting a set of experimentations taking into account human judgments. The psychometric method used is based on the sum of rank values and the Spearman coefficient of rank-correlation. Experimental results show an appreciable correspondence between the proposed computational measures and human judgments. Compared to related works, our results are better. In order to validate the proposed set of computational measures, we applied them in a content-based image retrieval experimentation using a large image database, the well-known Brodatz database, which contains 112 classes of 9 images each class for a total of 1008 images. Further research related to this work concerns mainly possible derivation of semantically-meaningful features based on the perceptual features used in this work as well as the use of additional features, such as randomness, in order eventually to further improve representation and retrieval effectiveness.

7. REFERENCES

- [1] N. Abbadeni, "Information retrieval from visual databases using multiple representations and multiple queries," in *Proc. ACM Symp. Appl. Comput.*, 2009, pp. 1523–1527.
- [2] N. Abbadeni, "Perceptual image retrieval," in *Proc. Int. Conf. Vis. Inf. Syst.*, Amsterdam, Netherlands, 2005, pp. 259–268.
- [3] N. Abbadeni, "Multiple representations, similarity matching, and results fusion for content-based image retrieval," *Multimedia Syst. J.*, vol. 10, no. 5, pp. 444–456, 2005.
- [4] N. Abbadeni, "Content representation and similarity matching for texture-based image retrieval," in *Proc. 5th ACM Int. Workshop Multimedia Inf. Retrieval*, Berkeley, CA, 2003, pp. 63–70.
- [5] N. Abbadeni, "A new similarity matching measure: Application to texture-based image retrieval," in *Proc. 3rd Int. Workshop Texture Anal. Synth.*, Nice, France, 2003, pp. 1–6.
- [6] N. Abbadeni, D. Ziou, and S. Wang, "Computational measures corresponding to perceptual textural features," in *Proc. 7th IEEE Int. Conf. Image Process.*, Vancouver, Canada, 2000, vol. 3, pp. 897–900.
- [7] N. Abbadeni, D. Ziou, and S. Wang, "Autocovariance-based perceptual textural features corresponding to human visual perception," in *Proc. 15th IAPR/IEEE Int. Conf. Pattern Recognit.*, Barcelona, Spain, Sep. 3–8, 2000, vol. 3, pp. 901–904.
- [8] M. Amadasun and R. King, "Textural features corresponding to textural properties," *IEEE Trans. Syst., Man Cybern.*, vol. 19, no. 5, pp. 1264–1274, Sep.-Oct. 1989.
- [9] J. Ashley, R. Barber, M. Flickner, J. Hafner, D. Lee, W. Niblack, and D. Petkovic, "Automatic and semi-automatic methods for image annotation and retrieval in QBIC," in *Proc. SPIE Conf. Storage Retrieval for Image and Video Databases*, 1995, vol. 2420, pp. 24–35.
- [10] J. R. Bergen and E. H. Adelson, "Early vision and texture perception," *Nature*, vol. 333, no. 6171, pp. 363–364, May 1988.
- [11] P. Brodatz, *Textures: A Photographic Album for Artists and Designers*. New York: Dover, 1966.
- [12] R. Datta, D. Joshi, J. Li, and J. Z. Wang, "Image retrieval: Ideas, influences, and trends of the new age," *ACM Trans. Comput. Surv.*, vol. 40, no. 2, p. 60, 2008.
- [13] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, and B. Dom *et al.*, "Query by image and video content: The QBIC system," *IEEE Computer*, vol. 28, no. 9, pp. 23–32, Sep. 1995.
- [14] J. C. French, A. C. Chapin, and W. N. Martin, "An application of multiple viewpoints to content-based image retrieval," in *Proc. ACM/IEEE Joint Conf. Digital Libraries*, 2003, pp. 128–130.
- [15] J. C. Gower, "A general coefficient of similarity and some of its properties," *Biometrics J.*, vol. 27, pp. 857–874, 1971.