Morphological Pattern Based Approach for Retrieval of Similar Trademarks

Ranjeet Kumar
ranjeet@iiita.ac.in
Research and Development Department
Indian Institute of Information Technology, Allahabad

R.C.Tripathi
rctripathi@iiita.ac.in
Research and Development Department
Indian Institute of Information Technology, Allahabad

M.D.Tiwari
mdt@iiita.ac.in
Indian Institute of Information Technology, Allahabad

Abstract

In the current scenario of the economic developments in to capture the world market growing role of Trademarks is to attract customers and beat over the imitated/fake products and services. Trademark registration is becoming a critical issue. New trademarks need to be distinct from existing millions of trademarks. This demand searching their similarity out of all trademarks is a typical combination of characters, graphics, images, colors and texture. The number of company and brand is increasing day by day and they all demand a trademark registration for trade practices. In this paper we propose a novel approach to obtain noise free trademark, break it into the character/texture and shape of the trademark logo or design and their relation attributes to build an efficient trademark retrieval tool. We create the input image in a negative image format for ignoring the background and objects reverse intensities values. The paper illustrates the use of morphological opening approach for removing the very small objects which can act as a noise. The work is also implemented to extract the image from the trademark database will be matched further for similar input trademarks on the basis of trademark retrieval process in the database.

Keywords: *Image extraction, retrieval and matching, Trademark, Trademark extraction and retrieval, Morphological pattern.*

1. Introduction.

Image processing tools and techniques can be used to solve different problems related to image, text, graphics and color etc. A trademark can be a combination of text, graphics, image, and colored texture. Based on these one it can divide them in these components for finding the similarity among different trademarks and retrieval of same trademark from the database. The problem for finding the similar trademark has become a challenge because today's world is growing for global economic scenario caused by different trade related practices are coming closer to each other at international level. Millions of trademarks being submitted to various trademark offices world over for registration need to have distinctiveness from the existing trademarks as per definitions and trade practices in different countries. Trademark registration with manual searching is very arduous task for the officials and so they need an ICT based practical approach which can show them similar trademarks in the existing database to decide whether to grant the new query trademark a registration.

Most of the recent techniques used for the image retrieval have mainly utilized the features like color, texture, shape etc. They used existing CBIR technique, i.e. Content Based Image Retrieval Systems to retrieve the images based on visual features like texture, color, shape etc [1]. In this technique extraction of color feature using the color histogram technique is utilized. It also considers the shape feature because it is an important feature in CBIR applications. Many techniques or approaches have been utilized for the image retrieval, some of which based on improved pattern matching algorithms. Some others take a much broader approach like searching just from the text files. Some are based on shape and color feature and some have attempted morphological pattern based image matching and retrieval using a database. A shape based technique introduced for the logo retrieval exported in a paper [2] is also inadequate to solve the problem amicably. The main objective of our approach is object extraction to improve the results of trademark image retrieval in some predefined sense. There are different shapes and objects combined in a trademark image like characters, lines, image and color etc. For identify these objects we are breaking them in separate parts and matching each of them separately to arrive at the optimum solution.

2. Related Work.

Many research groups have been working and facing the challenges in the automatic trademark retrieval. Several techniques have already been discussed in the "prior art research" section above which are based on similarity matching of shape-multi component images, texture based matching, morphological based similarity matching and retrieval, contour and boundary based image retrieval techniques etc. trademark from databases.

2.1. Shape Based Trademark Retrieval : ARTISAN

The ARTISAN project depends on finding a solution of shape based similarity image matching of multi-component images [3]. The architecture of the ARTISAN system requires the basic steps as enumerated below [4].

- To accept images in an appropriate standard format
- To build up a database of stored image descriptions from these images
- To extract retrieval features from these descriptions
- To allow formulation of visual queries
- To provide efficient and effective matching of query and stored images
- To display query results in an appropriate format.

In the figure1 below shows the input window designed in the ARTISAN system for user's easy access. After given input by the user after process by the ARTISAN system is shown the output in the window given in fig2 below.

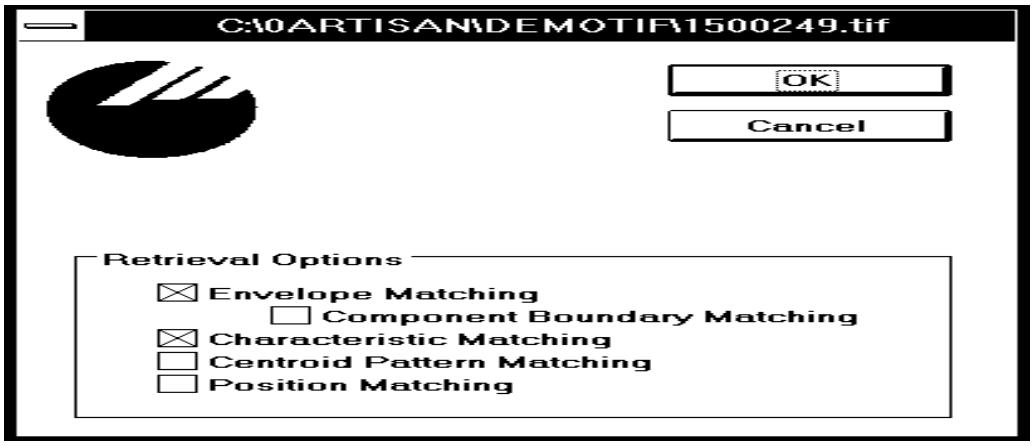

Fig1. Input query image for the retrieval in ARTISAN system [3]

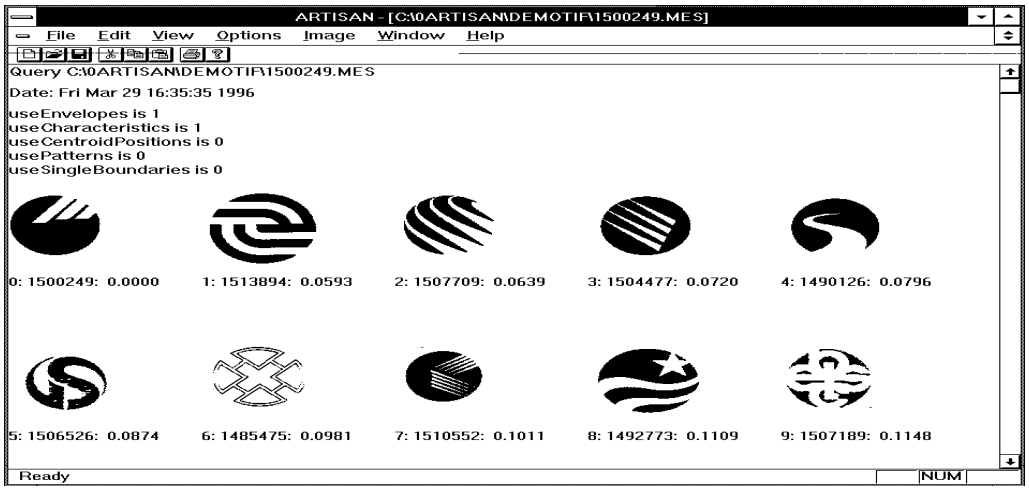

Fig2. Image retrieval results after input image given in fig1 in an ARTISAN system [3].

The system ARTISAN does not resolve the wider range of image type problems like recognition of the majority of implied shape features. Its use thereafter gets limited accordingly.

2.2. Contour and Region Based Trademark Retrieval

In this method, one considers two prime attributes- contour and region as similarity measure of two trademarks. The whole system works in two steps, in first steps, contour signature of an input trademark is extracted and used to filter out unlikely matched trademark from the database. After performing the first step, in the second step the region feature represented by MPEG-7 ART is used to search the best match from the database of trademarks [5]. In the above mentioned technique of trademark retrieval contains two main stages; database of trademarks is categorized into 10 classes by contour types. After completing the contouring process, region feature extraction and matching is performed. In this stage input trademark image is first normalized in size. After normalization of trademark symbols, it is then transformed into 35 moments by define it ART. Each moment is quantized

into 16 bins with a non-uniform quantizer [6]. After then it is used to form a 35-D feature vector. Now feature vector of the first stage and second stage are compared. After whole process is over, the output of the similar trademarks is retrieved from the database by the system.

2.3. Trademark Retrieval Based on Size Function

This technique of retrieval is based on size and has been proposed to be an effective system as Content based trademark retrieval. Basically it uses size functions. Three different classes of shape descriptors are combined for a total amount of 25 measuring functions. The evaluation of the system has been performed on a database of 1182 trademark images [7]. In this technique, trademark retrieval based on size functions has been derived with geometrical topological descriptors, conceived for formalizing qualitative aspects of shapes. At the outset, one set defines the size functions (SF's) and then describes the set of measuring functions. The similarity score concept is then introduced for the trademark retrieval [8].

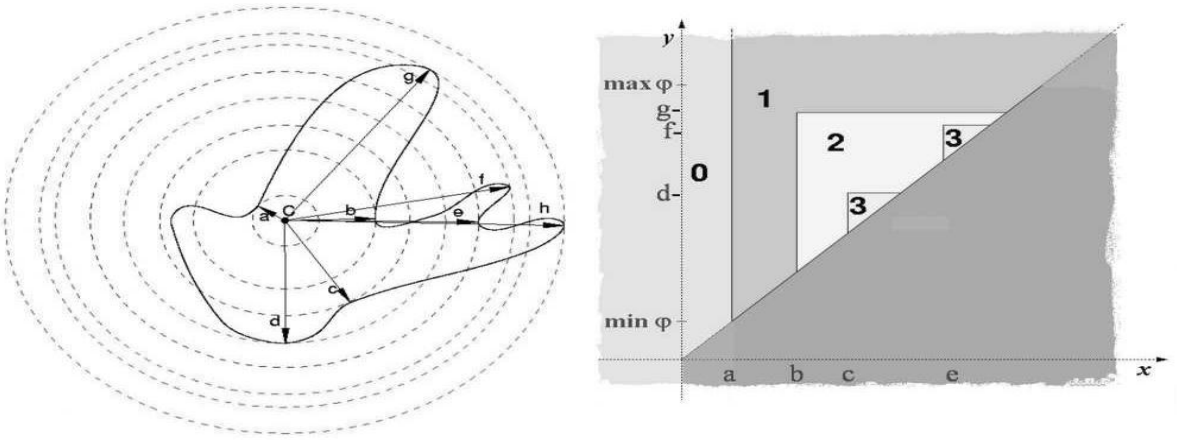

Fig3. Shows a simple example of SF. Here, the topological space M is a curve, while the measuring function is the distance from point c [7]

3. The Proposed Method.

Trademarks contain combination of text, graphics, image, color and figurative elements. Different approaches efficacy has been assessed. In our approach, we are trying to separate all the combinations of the text and images in the trademark and then arrange them in specific manner and then check the similarity between different trademarks available in the database. For this purpose, we define the entire process as outlined in fig 4 below.

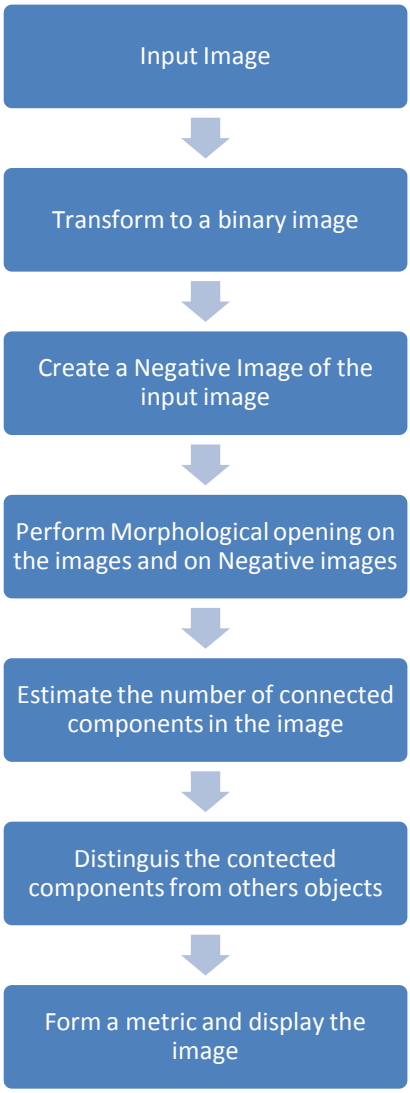

Fig 4. The process of the define approach
This has been implemented using MATLAB 7.0 as below.

Processes in our approach

- Input of a query image by the user
- Transform the image to a binary image using im2bw function.
- Create a negative image of the input image. The reason to create a negative image is, because background sometimes and objects may have reverse intensities values and it can lead to extraction of no objects.
- Apply morphological opening on the image as well as on its negative image using "bwareaopen" function. Morphological opening of an image helps to remove the very small objects present in the trademarks since these can be considered as noise. Scattered pixels are also be removed by this function.
- Estimate the number of connected components in the image using "bwlabel" function. It returns a metric of the input image in a format that consists of values 1 to N, where N is the total number of connected components in the image.

- Enclose each connected component in a rectangle to distinguish it from other objects using "regionprops" function.
- Decompose the image in various objects found in the image. After getting the metric of connected components, scan the metric in "for loop from 1to N", and construct a metric for that object using "imagen" function. After getting the metric, convert the metric back to the image format and display it.

4. Results and Discussion.

First we create a negative image from the input image using "im2bw" function. Figure 5 shows the negative image from the input image.

I = im2bw(Input_Image, level) converts the input grayscale image to a binary image. The output image will contain all pixels in the input image replaced with luminance greater than value specified by level with the value 1 as 'white' and replaces all other pixels with the value 0 'black'. The value of the level should be in the range (0,1)

Fig5 shows the creation of the negative image.

we use "bwlabel" function to find out the connected objects and define it in a suitable manner.

L = bwlabel(BW, n) returns a matrix L, of the same size as BW, containing labels for the connected objects in BW. The variable n can have a value of either 4 or 8, where 4 specifies 4-connected objects and 8 specifies 8-connected objects. If the argument is omitted, it gets default value as 8.

The elements of L are integer values greater than or equal to 0. The pixels labeled 0 are the background. The pixels labeled 1 make up one object; the pixels labeled 2 make up a second object; and so on. [L, num] = bwlabel (BW, n) returns the number of connected objects found in BW. Figure5 shows the connected objects in a image.

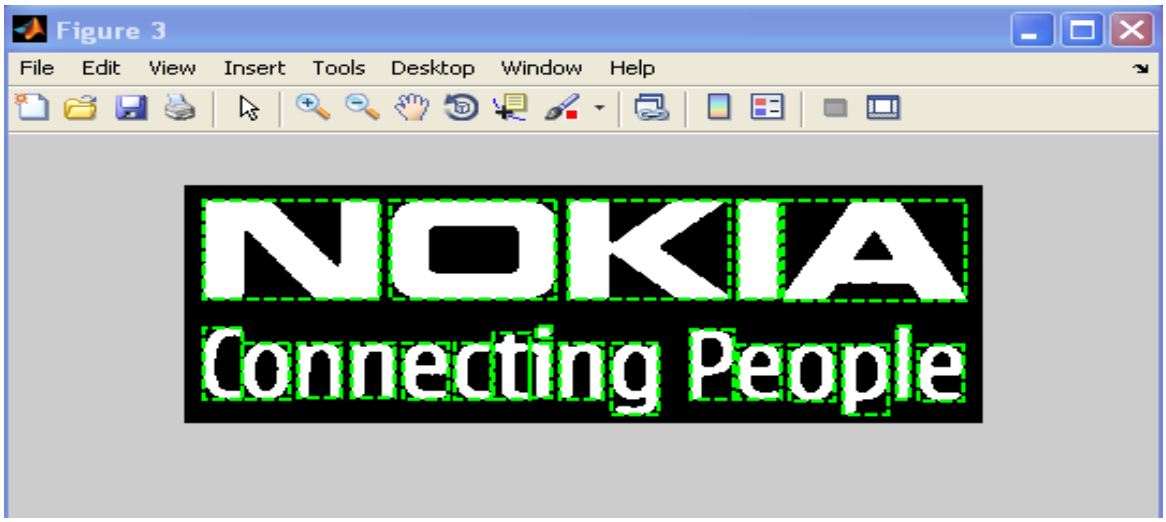

Fig6. Shows the connected objects in a Trademark image.

After getting the connected objects in the given trademark, input image is scanned. These are then decomposed in terms of all connected objects in the image. Fig7 shows the input image and its decomposition of the different connected objects present in the image.

Fig7. Showing the decomposition of all connected objects in the given input image.

5. Future Work and Conclusion.

Trademark image retrieval for similarity matching has been a complex problem and tried using different techniques and approaches. he present approach is based on decomposition of different connected objects. Removal of very small objects representing noise is done by way of morphological opening technique. This enables us to get all the objects which are connected in a given input image.

Our approach of matching of all the connected objects with a maximum number of objects matches using a database of different trademarks. The maximum number of objects and images which match with the trademark are outputted from which the user can find out the most similar trademark. In the case of a connected image, which is extracted from the input image, one can be match from the trademark database for similarity check also. After arranging the extracted objects which is connected in the input image, arrangement in such a way, it shows the objects which is maximum. For example in trademark of NOKIA, N is four times, P is two times, O is three times etc. These are used for matching of the maximum number of objects similar in the trademark in a database. After similarity check with extracted connected objects, system will generate the most similar trademarks in the output section.
We can search and retrieve the similar trademarks in the database using this technique which is found to be most efficient so far.

References.

[1] M.J.Swain and D.H.Ballard, "Color Indexing," *International Journal of Computer Vision*, vol. 7, no. 1, pp. 11–32, September 1991

[2] A. K. Jain and A. Vailaya, "Shape-Based Retrieval: A Case Study With Trademark Image Databases," *Pattern Recognition*, vol. 31, no. 9, pp. 1369–1390, 1998

[3] J P Eakins, J M Boardman and K Shields, Retrieval of Trademark Images by Shape feature- The ARTISAN Project, Department of Computing, University of Northumbria at Newcastle, Newcastle upon Tyne NE1 8ST, United Kingdom, 1996

[4] Eakins, J P "Retrieval of trade mark images by shape feature" Proc First International Conference on Electronic Library and Visual Information System Research, de Montfort University, Milton Keynes, 101-109, 1994

[5] Mao- Hsiung Hung, Chaur-Heh Hsieh, Chung-Ming Kuo, An Efficient Two- Stage Trademark Retrieval System, International Computer Symposium, Dec. 15-17, 2004, Taipei, Taiwan

[6] Mao- Hsiung Hung, Chaur-Heh Hsieh, Chung-Ming Kuo Similarity retrieval of shape images based on database classification, Journal of visual communication and image representation, volume 17, issue 5, October 2006, pages 970-985.

[7] A Cerri, m Ferri and D Giorgi, Vision, Video, and Graphics, 2005, The Eurographics Association

[8] P. Frosini, C. Landi, Size Theory as a TopologicalTool for Computer Vision. In: Pattern Recognition and Image Analysis, vol. 9(4), pp. 596-603, 1999

Author Profile

Biographical notes: Mr. Ranjeet Kumar has been working in Patent Referral Centre of IIIT Allahabad since 2007. Main contribution is in R&D department of the centre and doing research work in technological advancement in the field of IPR. The vetting of original expression and ideas emanating from BTech (IT) and MTech (IT) students of the Institute, Plagiarism check up and novelty assessments in the IT patents is his main activity.

Biographical notes: Prof.R.C.Tripathi is Incharge of the Patent Referral Centre in Indian Institute of Information Technology Allahabad. He is Dean (R&D) since 2007 and is promoting IPR culture in the Institute. Earlier he won the Government of Netherlands Fellowship in 1975 and obtained Diploma in Industrial Quantity Control at Bouwcentrum Netherlands. Earlier he worked for 32 years in different capacities to promote R&D in the country by support from Ministry of Communications and IT. He also headed the "Patent and IPR Division" for one Decade in the Ministry and looked after enhancing IPR portfolio of the 11 laboratories in the country setup by the Ministry of Communications and IT GoI.

Biographical notes: Dr. M.D.Tiwari a D.Phil from University at Allahabad and D.Sc from Garhwal University, He has been fellow of Alexander von Humboldt Foundation Germany. He is a founder Director of Indian Institute of Information Technology, Allahabad. He is President of Association of Indian Universities (AIU), New Delhi Since March 2009. Earlier he has worked as Vice Chancellor, MJP Rohilkhand University, Bareilly and First Chairman of UP State Council for Higher Education (State UGC) Lucknow. Earlier he was Adviser, All India Council for Technical Education, New Delhi and Joint Director, University Grants Commission, New Delhi.