A shape and texture features fusion to retrieve similar Trademark Image Material

R K Tripathi, S C Agrawal
Dept. of CEA, IET, GLA University, Mathura, 281406, UP, INDIA

Abstract—Trademark images or materials such as symbols, text, logos, image, design or phrase are used to unique representation of any organization. Retrieval of trademark material images are important to protect the new trademark image that is to be registered. Therefore, retrieval of similar trademark images is required. In this paper, an approach is presented to extract more similar trademark images so that a unique trademark image can be registered. In this paper, Zernike moment of the query image and dataset images are computed, then most similar images from the dataset are retrieved at the first layer refinement. In the second layer, texture features are extracted of query image and refined dataset images to retrieve most appropriate similar images. Zernike moments is applied to extract global shape features and Scale Invariant Feature Transform (SIFT) and Speeded-Up Robust Feature (SURF) are applied to extract texture features on the basis of a few key-points of the trademark images. A weighted average of both the key-points feature vectors is computed for retrieving the rank1, rank5, rank10, rank15 and rank20 most similar images using Euclidean distance. Experiments have been performed on a proposed dataset to perform the analysis and found that proposed work perform better and improves the accuracy.

Keywords: Trademark image material, rank, SIFT, SURF, Zernike moments

1. Introduction
Trademark image material retrieval is a challenging research area in the field of image retrieval. Trademark is unique identification of any commercial organization, individual or other authorized entity to identify the services or products to the clients by a symbol or indication. In other words, a trademark image material is used as an intellectual property or marketing materials. The number of trademarks growing rapidly in different countries. A trademark material can be a symbol, logo, image, design, phrase, text, or a fusion of these. Trademark materials are mainly useful in making distinction of one commercial organization from other and their services or goods as well [1].

In the recent scenario, trademarks are increasing day by day rapidly. Trademark image retrieval helps in unique trademark registration to any organization after confirming by matching proposed trademark with the retrieved trademark registered images. To protect the uniqueness of registered trademarks, it is important to check the uniqueness of the trademark by finding the similarity with the existing trademarks and trademarks that are to be registered.

Trademark image retrieval (TIR) can be performed by two methods: trademark image retrieval using description and image retrieval using content. Description based approach uses the text in which query is fired in textual form [2]. In content based image retrieval [3], a query is fired through an image and its feature vector is generated to describe its content and retrieve similar trademark images. Content may offer both accuracy and inexpensive solution. Based image retrieval is very challenging because it requires a strong discriminating features such as texture feature, shape feature and colour feature. To compute the similarity score between query and gallery trademark images, first feature vectors are computed for query trademark image as well as all the gallery trademark images stored in database.
We have proposed a robust approach in two steps to extract the Zernike moment in first step and then SIFT and SURF in second step. We have applied Zernike moments as global feature extractor which characterize to the shape feature, and SIFT and SURF as local texture orientation and description extraction based on key points for better performance. Proposed approach is robust to trademark image retrieval with the following features:

- Zernike moments as global shape feature and SURF and SIFT as local feature extractor, which are robust to illumination, rotation and scale invariant as well as better feature extractor. Proposed approach included Zernike feature at a first layer filter of interclass similar images and then SURF and SIFT at second layer as descriptor to retrieve most similar intra-class images.
- SIFT and SURF descriptors are key-points feature extractor having scale and rotation invariant feature to retrieve the trademark images after finding weighted average.
- The proposed method performed well to retrieve the similar images on proposed dataset and outperformed to Agrawal et al. [7].

Remaining part of the paper is structured as follows: Section II discusses the TIR approaches, Section III has process of proposed approach, Section IV shows the result analysis to measure the performance of the presented TIR system, and conclusion is presented in the section V.

2. Related Work

This section presents the progress of the research in the field of image retrieval of trademark images. In the recent decade, most of the researchers of this field have developed many TIR systems to improve the performance. Trademark [8], Star [9], and Artisan [10] are three important image retrieval systems of trademark image materials. There are different methods have been presented in the area of trademark image material retrieval. The graphical feature vectors have utilized by a TIR system to interpret the trademark image automatically and computes the similarity score based on human perception [8]. The Star system uses content-based TIR methods that include the moment invariants, grey level projection and Fourier descriptors. The Star system used spatial layout and shape components of an image [9]. However, it is difficult to recognize some significant components by an automated process [11]. Therefore, manual operation is required to segment some abstract trademark images to a certain extent. The Artisan system introduced a novel approach that adds principles derived from Gestalt psychology [10].

In 2001, Hsieh et al. [12] presented multiple classifier color image retrieval using region growing method. Wei et al. [13] proposed synthetic features to describe interior structure and global shape for the trademark image retrieval. Qi et al. [14] combined shape description and feature matching to retrieve trademark images. Anuar et al. [15] presented an integration of global-local descriptors to extract the features through Zernike’s moments coefficients and edge gradient co-occurrence matrix. Wu et al. [16] presented regional and boundary feature fusion for trademark image retrieval. Agrawal et al. [7] presented an architecture and combined texture and shape features to retrieve trademark images and performed well in case of Rank-1 identification while could not performed well in case of rank-5, rank-10 and rank-15. To improve the performance of TIR, we have used entropy as refinement layer to remove the dissimilar trademark images and then applied Zernike moments and SURF feature descriptor to retrieve most similar image. Ahmed et al. [17] used local and global features for image retrieval. In 2018, Perez et al. [18] employed combination of deep CNN models for trademark image retrievals. In 2020, Wei et al. [19] presented an intensity variation descriptor for image retrieval.

3. Proposed Work

In this section, we described the process of proposed approach. Fig. 1 presents the well-defined sequence of proposed approach- Zernike moment is computed as a first layer to remove interclass similar
images based on texture description and then SURF as a local feature computes scale and orientation invariant description in second layer helps to retrieve most similar rank-1, 5, 10, 15, 20 trademark images of intra-class trademark images.

**Fig. 1:** shows framework of trademark images retrieval- consists of first layer to refine the similar from dissimilar images using Zernike moments and then second layer retrieves correct trademark images using SIFT and SURF.

### 3.1 Zernike Moments

In this subsection, we applied widely used Zernike Moment which is highly utilized in CBIR.

We have following advantages of Zernike moments:

- Magnitudes of Zernike moments are rotational invariant which is helpful to retrieve similar trademarks existing in different rotation.
- Most important advantage of this shape descriptor that it is robust to noise as well as able to handle minor variations in shape.
- It has minimum information redundancy due to its orthogonal property.

For digital trademark images, if Pxy is the reference pixel then Zernike moment is computed with following formula:

$$S_{pq} = \frac{p+1}{\Pi} \sum_{i} \sum_{j} p_0 \left[V_{pq}(i, j)\right]^2, i^2 + j^2 \geq 1 \quad (2)$$

where $p = 0, 1, 2, \ldots \infty$ and $q$ represents to angular dependence. $V_{pq}(r, \theta) = R_{pq}(r) \exp^{in\theta}$ are the Zernike moments in polar coordinate system with some specified conditions, $i^2 + j^2 = 1$, $p \geq |q|$, $(p - |q|)$ is even and $s^2_{pq} = s_{p-q} \cdot R_{pq}(b)$ is orthogonal polynomial(radial) defined as:

$$R_{p,q}(b) = \sum_{s=0}^{s_{pq}} (-1)^s \frac{s! \left(p+q\right)^2}{s! \left(p+|q| - s\right) \left(2+|q|-s\right) \left(2-|q|-s\right) !} b^{p-2s} \quad (3)$$

To estimate Zernike moment of a trademark image, interested region is extracted and then mapping of that region is done to the origin of unit circle in the polar coordinate system. In estimation process, internal pixels of the unit circle are considered and rest pixels are excluded, then rotation invariance is obtained through this mapping.
3.2. SURF

SURF feature descriptor is known as Speed Up Robust Features which is scale and orientation invariant point detector. There are two basic steps for feature extraction:

3.2.1. Orientation Computation

Firstly, we identify reproducible orientation of interest points to make feature extraction rotation invariant. First, responses of the Haar wavelet are obtained in both x-y directions. Integral images are used for fast filtering. Once the responses are computed, weighted using a Gaussian which is centered at the point of interest, the responses are presented as vectors with the vertical and horizontal response strength along the ordinate and the abscissa respectively. The orientation which is dominant is computed by finding the addition of all the responses in a sliding window orientation which covers an angle of \( \pi/3 \). The both vertical and horizontal responses are summed within the window. The sum of both responses provides a new feature vector. These longest vector has its orientation to the interest point.

3.2.2. Descriptor Components Computation

To extract the feature description, firstly a region of square dimension is constructed centred around the point of interest and along the selected orientation. The region is divided into smaller sub-regions of dimension \( w \times w \) regularly which has important spatial information. We compute features at regularly spaced sample points for all sub-region. For simplification, we call \( h_{wx} \) as response of Haar-wavelet in horizontal direction and \( h_{wy} \) as response of wavelet in vertical direction. Responses \( h_{wx} \) and \( h_{wy} \) are weighted using a Gaussian that is centred at the point of interest to improve the robustness towards localization errors and geometric deformations.

After this, the \( h_{wx} \) and \( h_{wy} \) wavelet responses are added of each sub-region and used a first entries set to form the feature vector. In order to extract information about the polarity of the changes in intensity, we also compute the addition of the absolute values of the x and y direction responses, \( |h_{wx}| \) and \( |h_{wy}| \). Hence, a 4-dimensional feature vector \( v \) is calculated for each sub-region. This yields a descriptor vector of length 64 for all sub-regions of dimension. The responses of the wavelet are invariant to illumination. The descriptor is converted into unit vector to achieve invariance to contrast.

3.3. SIFT

SIFT has a scale and rotation invariant property which helps to retrieve the images of different rotation and scales. SIFT works in four steps:

- Scale space peak selection-key-points are detected to find the features.
- Correctly finding the location of key-points
- Assigning orientation
- Key-point descriptor describes the key-points in high dimensional space. Then, features are stored in a feature vector for finding the weighted similarity.

3.4 Similarity Measure

To generate the similarity score between query trademark image and stored trademark gallery images, Euclidean distance measure has been used. Euclidean distance measure yields better results in image retrieval scenario to retrieve similar images against the query images. The Euclidean distance is shown as: This module includes following sub-modules:

\[
\text{Dist}(q, D) = \left( \frac{1}{n} \sum_{i=1}^{n} (q_i - D_i)^2 \right)^{1/2}
\]

In the above formula, is the distance computed between the feature descriptor vector of query image \( q \) and trademark database images \( D \). In the proposed approach, we have computed the weighted similarity between SIFT feature and SURF feature. Consider \( D_{sift} \) is the value of shape similarity and \( D_{surf} \) is the SURF similarity value. We computed weighted similarity \( D_{ws} \) as:
\[ D_{ws} = \frac{W_{surf} \times D_{sift} + W_{sift} \times D_{sift}}{W_{surf} + W_{sift}} \] (5)

4. Results and Discussion

To evaluate the performance of proposed approach, we have extended dataset [7] from 300 images to 750 trademark images. To validate the accuracy and effectiveness of the proposed TIR, we have tested on own dataset and existing dataset [7]. We have implemented TIR system on Intel Core i7 8th Gen CPU, 8GB DDR 3 RAM. Matlab 2016a is used for the implementation and feature extraction and rank-1, 5, 10, 15, 20 retrievals. In the Graphical User Interface (shown in Figure 2), brown button has two operation-first is to select the image that is to be select as input and second is to display the retrieved result. The extended trademark database of [7] contains images with shape, color, illumination, size variations etc. The performance of the proposed approach is evaluated on the basis of precision. The precision [17] is defined as:

\[ \text{Precision} = \frac{N_A(q)}{N_R(q)} \] (6)

Here, \( N_A(q) \) is number of relevant trademark images retrieved and \( N_R(q) \) is total retrieved trademark images as a result of query images. Table I shows the rank-1, 5, 10, 15, 20 accuracy of the proposed approach and its components - Zernike moments and SURF descriptors at second layer responsible for retrieved result.

Figure 2 shows a GUI which has query image in first row and retrieved the results in second and third rows. Most of the images (9) are correctly retrieved while 1 image retrieved incorrectly in rank-10 retrieved images. This trademark image is text image and proposed approach works better with 90% accuracy.

![Fig. 2. shows nine correct retrieval results of trademark Cheetos from proposed dataset. First row is the query image and second and third row is the retrieved rank-10 matching results.](image-url)
Fig. 3. shows all correctly retrieval results of Trademark Make in India from own dataset. First row is the query image, second and third row is the retrieved rank-10 results.

Figure 3 shows the retrieved results of Trademark Make in India. Retrieved results have 10 correct retrieved trademark images and no incorrect retrieved images while images of this trademark has a lot of variation. Proposed approach is robust to scale and orientation. Figure 4 shows the retrieval results of Trademark Salesforce. Seven images of this trademark have been successfully retrieved while four images retrieved incorrectly.

Table 1 presents the performance of Shape Descriptors-SIFT, SURF, and proposed approach on own dataset. Table 2 presents the performance of Shape Descriptors SIFT, SURF, and proposed approach which fuse both after Zernike layer.

Fig. 4. Seven correctly retrieved results of Trademark Salesforce. First row is the query image, second and third row is the retrieved rank-10 matching results.
**Fig. 5.** Eight correctly retrieved results of Trademark PayPal. First row is the query image, second and third row is the retrieved rank-10 matching results.

**Fig. 6** Shows SanDisk images. First row is the query image, second and third row is the retrieved rank-10 matching.

**Table 1:** Performance of proposed approach on Own Dataset

| Approaches    | Rank-1 | Rank-5 | Rank-10 | Rank-15 | Rank-20 |
|---------------|--------|--------|---------|---------|---------|
| SIFT          | 100    | 84.60  | 70.33   | 68.41   | 62.16   |
| SURF          | 100    | 78.75  | 71.23   | 68.25   | 61.75   |
| SIFT & SURF   | 99.25  | 92.15  | 83.17   | 76.15   | 73.35   |
| Agrawal et al.| 100    | 91.55  | 79.92   | 73.70   | 71.34   |
| Proposed Approach | **100** | **94.15** | **85.25** | **79.56** | **77.95** |
**Table 2: Performance Analysis of proposed approach on existing dataset [7]**

| Approaches       | Rank-1  | Rank-5  | Rank-10 | Rank-15 | Rank-20 |
|------------------|---------|---------|---------|---------|---------|
| SIFT             | 95.5    | 85.75   | 75.25   | 73.50   | 69.97   |
| SURF             | 91.5    | 82.67   | 74.13   | 72.35   | 66.89   |
| SIFT & SURF      | 95.75   | 92.55   | 85.52   | 78.15   | 74.56   |
| Proposed Approach| **96.25** | **93.95** | **87.15** | **80.25** | **77.57** |

then compared with Agrawal et al. [7] on its trademark dataset.

Figure 5 shows 80% images retrieved correctly. Only two images retrieved falsely. Figure 6 also presents 6 SanDisk images retrieved correctly which are textual trademarks and 4 incorrectly retrieved images.

5. Conclusion

In this paper, a robust approach has been proposed to retrieve trademark image materials from a database by firing a query image. Zernike layer refines to the dissimilar images and retrieves similar 40 images and then SIFT and SURF descriptor retrieved most relevant images to the query images.

This proposed approach performed well and provided a better result in comparison to existing approach. This approach handles orientation, scale, shape features to retrieve more similar trademark images from the database. The proposed approach and its feature extractor performed well and yielded 100% accuracy in Rank-1 and more than 92% in Rank-5. Rank-1 and Rank-5 accuracy is better in comparison to Rank-10, Rank-15 and Rank-20.

In future, improvement is required to improve the recognition and retrieval accuracy of Rank 10, 15, 20 and more.

References

[1] Schietse J, Eakins J P, and Veltkamp R C 2007 Practice and challenges in trademark image retrieval ACM , pp. 518–24.
[2] Chalechale A, Mertins A, and Naghdy G 2004 Edge image description using angular radial partitioning.
[3] Rajam L F, Valli S, and Nagar J 2013 A survey on content based image retrieval
[4] Saryanidi E, Dal V, Tek S C, Tun B, and Gkmen M 2012 Local zernike moments: A new representation for face recognition 19th IEEE International Conference on Image Processing pp. 585–88.
[5] Teague M R 1980 Image analysis via the general theory of moments J. Opt. Soc. Am., vol. 70 no. 8 pp. 920–930.
[6] Kim W Y and Kim Y S 2000 A region-based shape descriptor using zernike moments Sig. Proc.: Image Comm. 16 pp. 95–102.
[7] Agrawal D, Jalal A S, and Tripathi R K 2013 Trademark image retrieval by integrating shape with texture feature International Conference on Information Systems and Computer Networks, pp. 30–33.
[8] Kato T 1992 Database architecture for content-based image retrieval in Electronic Imaging
[9] Wu J, Lam C, Mehfre B, Gao Y, and Desai N 1996 A Content based retrieval for trademark registration Multimedia Tools and Applications 3
[10] Eakins J P, Graham M E, and Boardman J M 1997 Evaluation of a trademark image retrieval system Proceedings of the 19th Annual BCS-IRSG Conference on Information Retrieval Research, ser. IRSG’97. Swindon, UK: BCS Learning & Development Ltd., pp. 1–1

[11] Alwis S and Austin J 1998 A novel architecture for trademark image retrieval systems Proceedings of the Challenge of Image Retrieval British Computer Society.

[12] Hsieh I S and Fan K C 2001 Multiple classifiers for color flag and trademark image retrieval Trans. Img. Proc. 10 6 pp. 938–950.

[13] Wei C H, Li Y, Chau W Y, and Li C T 2009 Trademark image retrieval using synthetic features for describing global shape and interior structure Pattern Recogn. 42 3 pp. 386–394

[14] Qi H, Li K, Shen Y, and Qu W 2010 An effective solution for trademark image retrieval by combining shape description and feature matching Pattern Recogn. 43 6 pp. 2017–2027

[15] Anuar F Mohd, Setchi R, and Lai Y K 2013 Trademark image retrieval using an integrated shape descriptor Expert Syst. Appl., 40 1 pp. 105–121

[16] Wu M, Xiao W, and Hong Z 2018 Similar image retrieval in large-scale trademark databases based on regional and boundary fusion feature PLOS ONE 13, p. e0205002

[17] Ahmed K T, Irtaza A, and Iqbal M A 2017 Fusion of local and global features for effective image extraction Applied Intelligence 47 pp. 526–543

[18] Perez C A, Estévez P A, Galdames F J, Schulz D A, Perez P J, Bastías D, Vilar D R 2018 Trademark Image Retrieval Using a Combination of Deep Convolutional Neural Networks International Joint Conference on Neural Networks (IJCNN)

[19] Wei Z and Liu G H 2020 Image Retrieval Using the Intensity Variation Descriptor Mathematical Problems in Engineering, Hindawi