

User Retrieval of Trademarks System Using Conceptual Similarity Approach

^{*1}Sonali Khamkar, ²Sushma Nandgaonkar

¹P.G. Student, Department of Computer Engineering, VPKBIET, Baramati, Maharashtra, India

²Assistant Professor, Department of Computer Engineering VPKBIET, Baramati, Maharashtra, India

**Email: khamkarsonali111@gmail.com*

Received: 09th July 2018, Accepted: 14th August 2018, Published: 31st August 2018

Abstract

The trademark is a form of intellectual property. There is a need to protect them. Trademarks are images, texts, slogan, and any domain name. It uniquely distinguishes the services or goods of businesses. There are a number of issues of trademark infringement. The conceptual similarity is also one issue regarding trademark infringement litigation for different trademarks. This paper retrieves the trademark suggestion on the basis of conceptual similarity. Also, the system will be used for logo comparison by using histogram algorithm.

Keywords: Conceptual Similarity, Infringement, Trademark, Similarity of Trademark.

Introduction

Trademarks are important for businesses. The trademarks are in different forms. These are images, slogans, characters or any domain name. The trademarks uniquely distinguish goods or services. When anyone wants to register trademark, it requires to apply for it in a trademark register office. It takes more time to process trademarks while checking for availability. They also do not calculate the logo similarity. So the numbers of trademark infringement cases are due in court. Researchers studied that cases and it developed a system to retrieve similar trademarks [1].

The system proposed deciphers the hypothetical similarities among the trademarks for this purpose. The system is based on text retrieval. This approach uses the trademark retrieval algorithm. It removes stopword and stemming and analyze trademark query which user has entered. It measures the conceptual similarity of trademark. But the system is not given the trademark suggestion. So this system contributes the trademark suggestion to the given trademark with the help of retrieval algorithm. Also the system can be used for finding the logo similarity. The existing system only finds query image similarity partially. This approach solves problem by using histogram algorithm. Because of histogram algorithm calculated logo similarity effectively.

Literature Review

F. M. Anuar, et al. proposed to retrieve text trademark on the basis of conceptual similarity. The system provides simple and easy to understand trademark retrieval system. The system motivates conceptual similarity to make them more accurate and more secure. It uses feature extraction, hash indexing and distance computation with the help of retrieval algorithm.

K. Lai, et al. proposed to introduce substance point of an exclusive gear and it is used to search nook pixel for image [8]. The system used precision recall method for the performance evaluation of retrieval of image. The experimental result concludes that the trademark image retrieval based on shape feature performs better and gives satisfactory results.

Qi, et al. introduced substance point of an exclusive figure and the point used to search nook pixel for it. The experimental result concludes that the trademark image retrieval is based on shape features. It performs satisfactory and better results. But the system cannot be used for calculating rotation invariant image features accurately [9].

M. -Y. Pai, et al. have proposed the various data reflow systems use information as user input data, but it seems hard and complicated semantic mechanism. The system overcome that problem by using a semantic based content mechanism for an information retrieval system [11].

F. M. Anuar, et al. have proposed trademark retrieval system. The trademark is a form of intellectual property. The system focus on this significant feature of text trademark [12]. The proposed system used natural language processing. Also it used semantic technology to compute the conceptual similarity between trademarks.

R. Setchi, et al. proposed, the semantic based image retrieval system tool. It produces a set of numbers for each text, which are used to retrieve information in the process called semantic expansion. The content based image retrieval and keyword based search complements traditional methods by providing images with high inspirational value [13].

J. J. Jiang et al. have proposed a new approach for measuring semantic similarity between concepts and words [14]. It gathers the corpus statistical information

with lexical taxonomy. So that the semantic distance between nodes in the semantic space constructed by the taxonomy can be contended with the computational semantic retrieval by data similarity. Trademark evidences are derived from distributional analysis of corpus data, which is enhanced by the node based approach of information content calculation.

L. Sbattella and R. Tedesco author proposed a fact for substance and listing information from main data. It uses conceptual level and lexical level for describing the main information. It provides better precision, which compares to regular search engine. It is simple and powerful system [15]. A comparison with classic search engine shows that the proposed approach is effective in providing better accuracy. The ordinary goal of such methodologies is to extract ordered information from natural language documents.

Thus, this system covers all the above useful aspects and overcome limitation of existing systems. This paper contributes with text trademark suggestion. Also logo comparisons trademark with correlation value with the help of histogram algorithm.

Problem Specification and Methodology

Problem Specification

The above review of literature observed that the number of researcher work on trademark retrieval based system. It uses conceptual and visual aspects of the trademark. But in recent years conceptual and visual infringement cases are increasing. So some researchers worked on that real time trademark infringement case. Therefore, this system works on user retrieval trademark system which is linked visual and conceptual aspects of trademarks which retrieved trademark similar to the user. It's problem specification as follows:

- As deep Trademark violation is an aspect of IP delinquency that hegemony to serious financial issue. The concept of trademark similarity is used in trademark retrieval system. The logo similarity is computed with the help of histogram algorithm.

The goals and objectives are followed:

- The logo comparison ranking result can be improved using histogram algorithm. The correlation value shows the similarity of the logo.
- To search and retrieve similar trademark on the basis of conceptual and visual aspects.
- Extract features of text trademark with the help of a semantic retrieval algorithm and extract logo features with texture and color basis by using histogram algorithm.
- To make an improvement of trademark conceptual similarity and to make them more accurate and more secure against the trademark infringement.

Feature Extraction

Feature extraction is the method of extracting features of text and image by using different techniques. In text retrieval trademark system used natural language processing technique. This technique is used for analysis of speech and natural language. This system uses NLP for natural text. The system first removes stopword from text trademark. For example, when a user enters “a amazon” trademark removes ‘a’ from it. Stemming is also used to take a root word like “playing”, “played” replaced by “play”. Extract the features in token and synonyms form. Extract the tokens by using tokenization. It separates the text into token (ex. Red Bull then its token is split by Red, Bull). Then system uses a wordnet dictionary for calculating synonyms. The wordnet dictionary is freely available on the internet.

Work of Trademark System

System Framework

The system framework is based on conceptual similarity [1]. The proposed system is used to retrieve of trademark hypothetical similarity, which should be more accurate and more secure against the trademark infringement. In this system first the trademark dataset is loaded. The dataset has different trademarks with different dimension. Then resize this trademark with fixed dimension. Apply feature extraction of text trademark. The feature extraction defines the token and synonym sets. The Natural Language Processing (NLP) techniques are used in text trademark. The WordNet dictionary is used for generating synonym sets. The hash indexing accepts the token key and the synonymous key for pre-processing. It uses the indexing in that key and creates a new trademark for the user, which is similar to the user requirement trademark. To remove the extra required time is the need of a procedure. The factors are listed using a hashing technique. The same procedure follows on entering trademark. Then the computation performed on with the help of retrieval algorithm. If trademark already exists in the system, then it sent to trademark matching and return the similar documents to the user. If the trademark doesn't exist in the system, then give the message you can use this trademark.

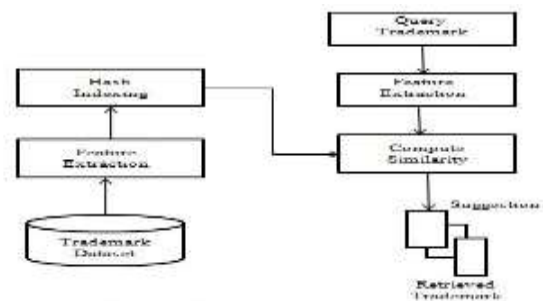


Figure1: Text Trademark Retrieval System

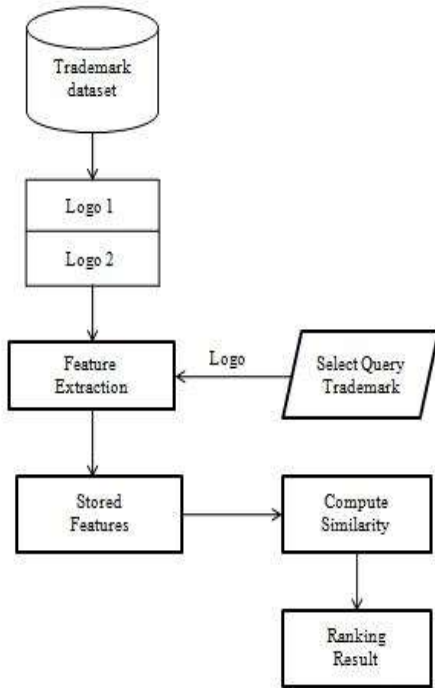


Figure 2: Logo Trademark Retrieval System

The logo comparison is the challenging task of the proposed system. So, it uses a histogram algorithm for logo comparison. In this system first select the query trademark logo and compare with whatever logo retrieved from the dataset. The feature extraction is applied to the query image with the help of histogram algorithm. In this system extract feature on the basis of texture and color. Then these features are stored in the data structure. The same steps are followed by retrieving images from datasets. Finally, compute similarity between a selected query trademark logo and retrieval logos from the dataset. These computations show the ranking results of logos with correlation value. That shows the similarity between the logos.

Mathematical Formulation

For displaying the distance computation of text trademark similarity is based on [1].

$$\text{sim}(Q,T) = \frac{|Q_{ft} \cap T_{ft}|}{|Q_{ft} \cup T_{ft}|} + \frac{|Q_{fs} \cap T_{fs}|}{D} + \frac{\sum_{i=1}^I \sum_{j=1}^J \max(\text{wordsim}(x_i, y_j))}{|Q_{ft} \setminus T_{ft}| \cdot |T_{ft} \setminus Q_{ft}|} \dots (1)$$

Q: Query trademark.

T: Trademark.

Q_{ft} & Q_{fs} : Token set and synonyms set of trademark query.

T_{ft} : Token set of one trademark.

$D = \max(|Q_{ft}|, |T_{ft}|)$;

i, j : set elements

Algorithm

1. Logo similarity algorithm

Semantic Retrieval of Trademark is based on rotation invariant local features.

Input: Query Image/ Logo, Image Database.

Output: Ranked Image.

Step 1: Select or browse the query image.

Step 2: Preprocessing technique is used on selected query image and database logos.

Step 3: Extract the rotation invariant features

```

for (int x=0; x<image. width; x++)
{
  for (int y=0; y<image. height; y++)
  {

```

$$R_{mn}(x, y, I) = Q_{mn}(x, y, I) Q_{mn}(x, y, I)^* \dots (2)$$

Step 4: Store all features as multidimensional vector.

Vector multi=new Vector ();

multi. Add (R_m);

Step 5: Histogram is computed and normalized for each trademark image.

$$H(k) = \frac{h(k)}{\sum_{m=1}^K h(m)} \dots (3)$$

Step 6: Extract query image features with histogram computation.

Step 7: Calculate the similarity of query image and i -th image of dataset.

$$S = \sum_{k=1}^K |H_q(k) - H_i(k)| \dots (4)$$

Step 8: Display ranked targeted images according to the output of step 7.

Output: Bubble_Sort (S).

Results and Discussion



Figure 3: Load the Trademark Database

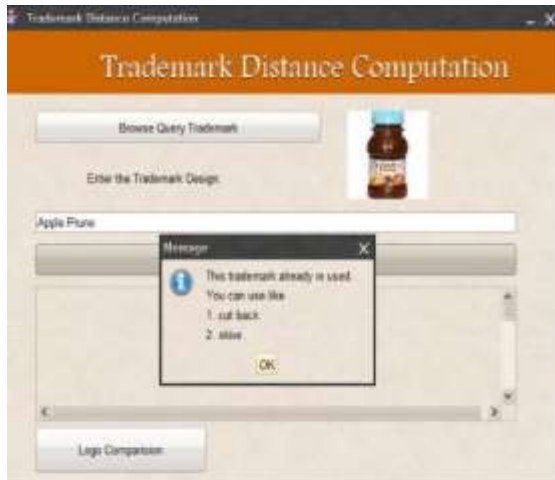


Figure 4: Input Query Trademark Suggestion



Figure 5: Analysis graph for Apple Prune Query

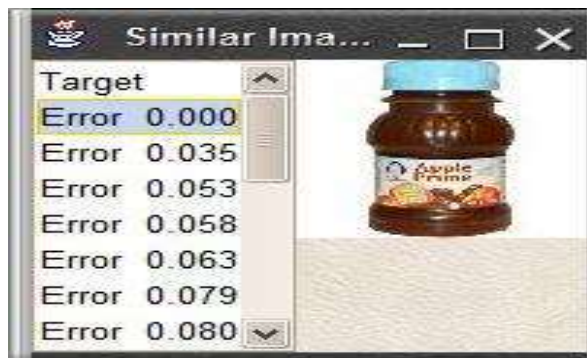


Figure 6: Retrieved Query Image Similarity

The results are generated on the basis of the Conceptual similarity approach. The goal of this system is to retrieve similarity of trademark. The system load the trademark dataset which is shown in

figure 3. Extract features and Hash indexing used. Apply trademark distance computation. In figure 4, the system retrieves suggestion when entered query text trademark is already available. The system also retrieved ranking result of a query image with histogram algorithm which is shown in figure 5, 6.

Conclusion

The work is motivated by data similarities increasing of fraud cases, where information retrieval system does not handle this particular issue and trademark similarity. The advantages and limitations of each data similarity of retrieval algorithm are described. The system worked conceptual similarities among trademarks like equal or relevant semantic implant. The natural language processing technology, lexical resources are used to calculate hypothetical similarity between different trademarks. The system gives some suggestion to the user if the entered trademark query is similar to in trademark database and also gives logo comparison ranking results.

References

- [1] Fatahiyah Mohd Anur, "Semantic Retrieval of Trademarks Based on Conceptual Similarity," IEEE Transaction on Systems, vol 46, no. 2, pp. 220-223, 2015.
- [2] Office for Harmonization in the Internal Market, Annual report 2012. [Online]. Available: <https://oami.europa.eu> (accessed 10.12.2013).
- [3] L. Dodell. (2013). The Trademark Problem: Casualty Insurance's Dirty Little Secret. [Online]. Available: <http://www.carriermanagement.com>, accessed Dec. 2013.
- [4] U.S. Congress Joint Economic Committee. (2012). The Impact of Intellectual Property Theft on the Economy. [Online]. Available: <http://www.jec.senate.gov/>, accessed Dec. 1, 2013.
- [5] C. D. Scott, "Trademark strategy in the Internet age: Customer hijacking and the doctrine of initial interest confusion," J. Retail., vol. 89, no. 2, pp. 176–189, Jun. 2013.
- [6] U.S. International Trade Commission. (2011). China: Effects of Intellectual Property Infringement and Indigenous Innovation Policies on the U.S. Economy. [Online]. Available: <http://www.usitc.gov/>, accessed Dec. 20, 2013.
- [7] Office for Harmonization in the Internal Market. (2014). Guidelines for Examination in the Office for Harmonization in the Internal Market on Community Trade Marks, Part C Opposition, Section 2 Identity and Likelihood of Confusion, Chapter 3 Comparison of Signs.

- [Online]. Available: <https://oami.europa.eu>, accessed Feb. 1, 2014.
- [8] F. M. Anuar, R. Setchi and Y. K. Lai, "Trademark image retrieval using an integrated shape descriptor," *Expert Syst. Appl.*, vol. 40, no. 1, pp. 105–121, 2013.
 - [9] H. Qi, K. Q. Li, Y. M. Shen, and W. Y. Qu, "An effective solution for trademark image retrieval by combining shape description and feature matching," *Pattern Recognit.*, vol. 43, no. 6, pp. 2017–2027, 2010.
 - [10] C. H. Wei, Y. Li, W. Y. Chau, and C. T. Li, "Trademark image retrieval using synthetic features for describing global shape and interior structure," *Pattern Recognit.*, vol. 42, no. 3, pp. 386–394, 2009.
 - [11] M.-Y. Pai, M. -Y. Chen, H. -C. Chu, and Y.-M. Chen, "Development of a semantic based content mapping mechanism for information retrieval," *Expert Syst. Appl.*, vol. 40, no. 7, pp. 2447-2461, 2013.
 - [12] F. M. Anuar, R. Setchi, and Y. K. Lai, "A conceptual model of trademark retrieval based on conceptual similarity," in *Proc. 17th Int. Conf. Knowl. Based Intell. Inf. Eng. Syst., Kitakyushu, Japan, 2013*, pp. 450-459.
 - [13] R. Setchi, Q. Tang, and I. Stankov, "Semantic-based information retrieval in support of concept design,, " *Adv. Eng. Inf.*, vol. 25, no. 2, pp. 131-146, 2011
 - [14] J. J. Jiang and D. W. Conrath. Stankov, "Semantic similarity based on corpus statistics and lexical taxonomy," in *Proc. Int. Conf. Res. Comput. Linguist., Taipei, Taiwan, 1997*, pp. 19-33.
 - [15] L. Sbattella and R. Tedesco, "A novel semantic information retrieval system based on a three-level domain model, " *J. Syst. Softw.*, vol. 86, no. 5, pp. 1426-1452, 2013.