OPEN ACCESS INTERNATIONAL JOURNAL OF SCIENCE & ENGINEERING

SEMANTIC RETRIEVAL BY DATA SIMILARITY OF TRADEMARKS

Vishakha Chilpipre[1] , Prof. Varsha R. Dange[2]
Student, Department of Computer Science Dhole Patil College of Engineering, Pune, Maharashtra, India[1]
Assistant Professor, Department of Computer Science Dhole Patil College of Engineering, Pune, Maharashtra, India[2]

Abstract: **A trademarks is a sign that you can use to distinguish your business goods or services from those of other traders. Trademark can be defined expressly in the form of any symbol, logo, titles etc. so, they need to be secure. This project deciphers the hypothetic similarities among trademarks, which happens when more than two or more trademarks hail equal or relevant semantic implant. The state-of-the-art by offering a semantic algorithm to similitude trademarks in pre conditions of hypothetic parallelism. By using data similarity, it is derived that search and indexing technique developed similarity distance. The offered reflow algorithm is confirmed using two resources: a trademark database of conflicting cases and a databases company names. Extends the conceptual model by developing and evaluating a semantic algorithm for trademark retrieval based on conceptual similarity. The conceptual comparison of text documents that share similar domain, use similar concepts, or express similar ideas has been studied extensively. The underlying technology embedded in existing trademark search systems is primarily based on text-based retrieval. Use the different domains to measures the accuracy of the algorithm which gathered different data.**

Keywords: Conceptual similarity, trademark infringement, trademark retrieval, trademark similarity.

---∴∴∴---

I INTRODUCTION

A trademark generally refers to a brand or a logo. Trademark registration can also be obtained for a business name distinctive catch phrases, tag line or captions. Properly used and promoted a trademark may become the most valuable asset of a business. Trademarks such as Coca-Cola , HP, Canon , Nike, Adidas, puma and many more serve as an indication origin of the goods as well as an indication of quality. It is also essential to obtain trademark registration for the business name or trade name under the trademarks act. Registration of company or business name under the companies are does not itself give protection against others who might commence using identical or similar marks. A trademark is any unique expression related to a product or service that distinguishes it from others. A trademark can be a visual symbol which may be a word, a name, a device, label or numerals used by a business to distinguish it from other goods of different business. The system will be scalable across multiple platforms and will provide more reliability due to its efficiency and robustness.

II LITERATURE REVIEW

F. M. Anuar, R. Setchi, and Y. K. Lai, " Trademark image retrieval using an integrated shape descriptor ", 2013.

F. M. Anuar, R. Setchi, and Y. K. Lai [2] author proposed trademark image retrieval using an integrated shape descriptor.as the proposing innovatory trademark reflow technique to use the reform performance of expositor. Trademarks are distinctive visual symbols with high reputational value, due to the perception of quality and innovation associated with them. They are important reputational assets used as a marketing tool to convey a certain assurance of quality, innovation, and the standards, which the manufacturer seeks to maintain. This motivates the need for trademark protection by providing a solution to prevent infringement. This problem can be addressed by developing retrieval systems capable of comparing the visual similarity of trademarks. The visual similarity is checked during the trademark registration process where one of the steps involved is making sure that the trademark to be registered is not similar to any trademark which is already registered. This is important in order to avoid infringements as well as to protect the rights of the existing trademarks.

H. Qi, K. Q. Li, Y. M. Shen, andW. Y. Qu, " An effective solution for trademark image retrieval by combining shape description and feature matching ", 2010.

H. Qi, K. Q. Li, Y. M. Shen, and W. Y.Qu [3] uuthor proposed to introduced substance point of a exclusive figure and this the point used to search nook pixel from it. Number of image collections available has increased due to ease of capturing images by different acquisition systems. The storage format of image data is relatively standardized however the effective retrieval of images from such databases remains a significant challenges. For the performance evaluation of the system we use the most commonly used method namely precision-recall. From the experimental result we conclude that the Trademark Image Retrieval based on

shape feature perform better and gives satisfactory result. The restricted work to only handle the rotational challenges.

L. Sbattella and R. Tedesco, " A novel semantic information retrieval system based on a three-level domain model ", 2013.

L. Sbattella and R. Tedesco [4], author proposed a fact and ideal for substance and listing information from main data. Use the conceptual level and lexical level for describes the main information. The stochastic model is then used, during the document indexing phase, to disambiguate word meanings. The semantic information retrieval engine we developed supports simple keyword-based queries, as well as natural language-based queries. The engine is also able to extend the domain knowledge, discovering new and relevant concepts to add to the domain model. The validation tests indicate that the system is able to disambiguate and extract concepts with good accuracy. A comparison between our prototype and a classic search engine shows that the proposed approach is effective in providing better accuracy. The common goal of such methodologies is to automatically extract structured information from natural language documents. The used of model for knowledge extraction from natural language documents.

M. Y. Pai, M. Y.Chen, H. C. Chu, and Y. M. Chen, " Development of a semantic based content mapping mechanism for information retrieval ", 2013.

M.-Y. Pai, M.-Y.Chen, H.-C.Chu, and Y.-M. Chen [5] Author proposed the many data reflow systems use search information as user input data, but it is a mainly hard and complicated so use the semantic mechanism. To address this problem developed a semantic-based content mapping mechanism for an information retrieval system. This approach employs the semantic features and ontological structure of the content as the basis for constructing a content map, thus simplifying the search process and improving the accuracy of the returned results. Information retrieval systems include the searching technologies and functions that can help users find the information that they need based on criteria they are given. Existing IR systems mostly perform searches based on keywords entered by the user, although keywords cannot render a complete representation of the content semantics.

F. M. Anuar, R. Setchi, and Y. K. Lai, " A conceptual model of trademark retrieval based on conceptual similarity ", 2013.

F. M. Anuar, R. Setchi, and Y. K. Lai [6], author proposed to mainly focus on main fact by proposing a notation flow of the different procedure, to main at reflow the same trademarks. Trademarks are proprietary words and images with high reputational value; they are important assets, often used as a marketing tool, which require infringement protection. One of the issues considered during infringement litigation is the visual, conceptual and phonetic similarity of different trademarks. In particular, the conceptual similarity of trademarks is an area never previously studied in information retrieval. The focuses on this important aspect by proposing a conceptual model of the comparison process, aimed at retrieving conceptual Semantic Retrieval of Trademarks and logos by Data Similarity s actually similar trademarks. The proposed model employs natural language processing and semantic technology to compute the conceptual similarity between trademarks. The model employs natural language processing techniques, knowledge sources and a lexical ontology to compute conceptual similarity between textual trademarks.

J. Oliva, J. I. Serrano, M. D. del Castillo, and A. Iglesias, " SyMSS : A syntax-based measure for short-text semantic similarity ", 2011.

Jess Oliva , J Serrano, ngel Iglesias [7], author proposed, a new method for computing short-text and sentence semantic similarity. The method is based on the notion that the meaning of a sentence is made up of not only the meanings of its individual words, but also the structural way the words are combined. The captures and combines syntactic and semantic information to compute the semantic similarity of two sentences. Semantic information is obtained from a lexical database. To evaluate the improvements produced by a better lexical semantic similarity measure. Semantic information is obtained from a lexical knowledge base such as WordNet, which models common human knowledge about words in natural language and allows different types of measures of semantic similarity between concepts to be calculated.

R. Setchi, Q. Tang, and I. Stankov, " Semantic-based information retrieval in support of concept design ", 2011.

Rossitza Setchi , Qiao Tang, Ivan Stankov [8], author proposed, The semantic-based image retrieval tool tags images by first processing all significant words in the text around them, extracting all keywords and key phrases in it, ranking them according to their significance, and linking them to ontological concepts. It generates a set of concept numbers for each text, which is then used to retrieve information in a process called semantic expansion, where a keyword query is also processed semantically. The semantic-based image retrieval tool developed has demonstrated good performance and scalability, and has been integrated with keyword-based indexing and content retrieval algorithms in an industrial prototype. The concept-based search combined with content-based image retrieval and keyword-based search complements traditional methods by providing images with a degree of diversity and high inspirational value.

III SYSTEM ARCHITECTURE

The analysis of the trademarks is needed to comprehend the main of conceptual and logo similarities coming from different factors. The focuses on this main fact by proposing a hypothetical model of the comparison process, purposed at retrieving conceptually similar trademarks. The hash indexing accept the token key and synonym key to pre-processing and use the indexing in that key and create a new trademark for the user, its similar to that user requirement trademark. The feature extractions are defined the token and synonyms. A trademark reflow technique using the proposed retrieval algorithm is evolved, and the algorithm is tested on conceptual similarity. The retrieval trademark list is stored in database for next future trademark use in the next trademark retrieval concepts. To remove extra required time throughout the find procedure, the factors are listed using a hashing technique. The hash indexing is taken the trademark as the key index. Through

trademark retrieval process user can enter a text which he wants to trademark. If trademark is already exist in system then it sent to trademark matching and return the similar documents to the user. If trademark is not existed in system then trademark is stored in database. The return document is send to user is the use lexical resource and apply the hash indexing to that trademark for create new trademark to get the user.

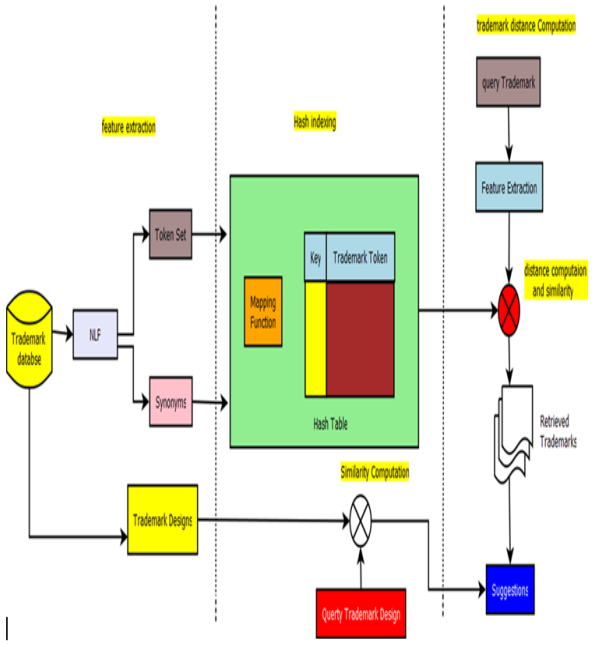

Figure 1 System Architecture

Modules:

The modules are as following:

Trademark text : User can check the availability in first module. The trademark is already use then provide to user token and synonym for new trademark. The synonyms is use in lexical resources.

Token : The token is provided to user for identification in process of developing new trademark for registration. The synonym set, as defined in the context

Trademark Indexing : The hash indexing takes the trademark as the key index. To reduce computational time during the search process, the features are indexed using a hashing technique. The final indexing table is merely a table that maps each trademark in the database to a set of trademarks from the same database for the trademark similarity computation.

Histogram :

1. Learning Stage : The histogram represents which kind of features is dominant and which is less dominant. We regard the histogram as the image descriptor. The histograms are calculated for learning images and memorized in the dictionary. The learning images are used as targeted images in the retrieval stage.

2. Retrieval Stage : In the retrieval stage, first, rotation invariant features are extracted from the query image. Then, the quantization and the histogram computation are performed in the same way as in the learning stage. Next, the similarity between the histogram obtained from the query image and each histogram in the dictionary is calculated. The similarity is defined as the absolute difference between two normalized histograms of the query image and the i^{th} histogram in the dictionary. Finally, the targeted images are ranked according to the similarity

IV ALGORITHM

This Histogram algorithm consists of three stages.

Step 1: In learning stage 1, extract the rotation invariant features

$$R_{mn}(x,y;I) = Q_{mn}(x,y;I)Q_{mn}(x,y;I)^*$$

Step 2: store all features as multidimensional vector.

Step 3: Apply vector quantization technique using K-means clustering

Step 4: store clustered result stored as dictionary 1 i.e. code book.

Step 5: In learning stage 2, step 1-2 are same.

Step 6: each feature is quantized based on based on code book generated at step4.

Step 7: Histogram is computed and normalized as

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

Step8: store this histogram as dictionary 2.

Step9: repeat step5-step7 for query image.

Step10: calculate the similarity of query image and i-th image from dictionary 2

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

Step11: Display ranked targeted images according to the output of step10.

V APPLICATION AND ADVANTAGES

1. Provide a simple and easy to understand trademark retrieval system.
2. Provided good precision compare to regular search engine that is a simple and well powerful system.

VI CONCLUSION AND FUTURE WORK

The work was motivated by increasing of fraud cases best an data similarities, where Information Retrieval system do not handle this particular issue and trademark similarity. The advantages and limitations of each data similarity of re-flow algorithm are described. The proposed work, hypothetical similarities among trademarks hail equal or relevant semantic implant. The model employs natural language processing techniques, knowledge sources and a lexical resource to compute conceptual similarity between trademarks. Also confirm that the comparison of trademarks in terms of conceptual similarity. The proposed model on existing trademark search models by providing a means of refining the search to conceptually related trademarks. In future work to improve the precision of the proposed

semantic algorithm should include a study comparing the use of various lexical resources.

REFERENCES

1. Fatahiyah M Anuar, Yu-Kun Lai, R. Setchi,"Semantic Retrieval of Trademarks Based on Conceptual Similarity." IEEE permission, 2015.
2. 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.
3. 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.
4. 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.
5. M.-Y. Pai, M.-Y.Chen, H.-C.Chu, and Y.-M. Chen, "Development of a semanticbased content mapping mechanism for information retrieval," Expert Syst. Appl., vol.40, no. 7, pp. 2447-2461, 2013.
6. 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.
7. J. Oliva, J. I. Serrano, M. D. del Castillo, and A. Iglesias, "SyMSS: A syntax-based measure for short-text semantic similarity," Data Knowl. Eng., vol. 70, no. 4, pp. 390-405, 2011.
8. 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.
9. J. J. Jiang and D. W. Conrath, "Semantic similarity based on corpus statistics and lexical taxonomy," in Proc. Int. Conf. Res. Comput. Linguist., Taipei, Taiwan, 1997, pp. 19-33.
10. Y. S. Kim and W. Y. Kim, "Content-based trademark retrieval system using a visually salient feature," Image Vis. Comput., vol. 16, nos. 12-13, pp. 931-939, 1998.
11. 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.
12. J. P. Eakins, J. M. Boardman, and K. Shields, "Retrieval of trademark images by shape feature-The ARTISAN project," in Proc. IEEE Colloq. Intell. Image Databases, London, U.K., 1996, pp. 9/1-9/6.
13. Company Names Database. [Online]. Available: http://uscompanydatabase.com, accessed Jul. 10, 2013.
14. Office for Harmonization in the Internal Market, Annual Report 2012. [Online]. Available: https://oami.europa.eu, accessed Dec. 10, 2013.
15. S. A. Fadzli and R. Setchi, "Concept-based indexing of annotated images using semantic DNA," Eng. Appl. Artif. Intell., vol. 25, no. 8, pp. 1644-1655, 2012.
16. J. Schietse, J. P. Eakins, and R. C. Veltkamp, "Practice and challenges in trademark image retrieval," in Proc. 6th ACM Int. Conf. Image Video Retrieval, Amsterdam, The Netherlands, 2007, pp. 518-524.
17. Trade Mark Search. [Online]. Available: http://www.ipo.gov.uk, accessed Dec. 10, 2013.
18. B. Furlan, V. Batanovic, and B. Nikolic, "Semantic similarity of short texts in languages with a deficient natural language processing support," Decis. Support Syst., vol. 55, no. 3, pp. 710-719, 2013.