Proceedings of the OAGM&ARW Joint Workshop 2017 DOI: 10.3217/978-3-85125-524-9-17

Confusing Similarity between Visual Trademarks A Dataset Based on USTTAB Examinations*

Lukas Knoch[1] and Mathias Lux[2]

***Abstract*— Trademarks are an important visual clue for customers to identify brands, products and companies, and can influence the buying decision significantly. One major problem with visual trademarks is, that newly registered trademarks are required by law not to be visually similar to existing ones. Therefore, automatic detection of visually similar trademarks is an important use case for content based image retrieval. Confusing similarity between trademarks is defined by law, and numerous cases of the *United States Trademark Trial and Appeal Board* (USTTAB) handling trademark similarity are available. In this paper we present a novel and freely available data set for evaluation of trademark similarity algorithms based on real life data, ie. all registered trademarks in the USA as well as USTTAB decisions and expert opinions. The data set should serve as a basis for further investigations, ie. extension of the data set by crowd sourcing and consideration of the intuitive concept of visually confusing similarity.**

I. Introduction

Visual trademarks, or *logos*, often influence our buying decisions and are therefore valuable goods for the companies owning the visual trademark. A common and well known example is the Apple company logo (compare Figure I) present on iPhones, iPads and Apple computers. Apple Computers invests time and money to find out if other companies worldwide use similar logos on similar products. The same approach is also taken by many companies who define themselves through their brands, like Nike, Adidas, or Red Bull.

Fig. 1. Examples of well known logos and protected trademarks in many countries including the Apple logo, the Github logo and the Nike swoosh.

To avoid confusion between different trademarks, they must be *dissimilar enough* to each other. Some companies even try to trick customers by deliberately using trademarks that are similar to well known signs. To avoid fraud, trademarks can be protected by law. There are several offices in charge of managing trademark registrations for different regions including the *European Union Intellectual Property Office* (formerly Office for Harmonization in the Internal Market, short OHIM) or the *United States Patent and Trademark Office* (short USPTO). If a new trademark is registered, it has to be ensured that there is no confusing similarity to any other previously registered marks. This difficult job is executed by professional *trademark examiners* who compare the different trademarks to each other and decide about the similarity. While there are systems in place like the textual *Vienna Classification* [21], taxonomies which are intended to help the examiner, these systems are tedious and error prone as they rely on manual annotation.

Another way of assisting the examiners are visual trademark retrieval systems. These systems can take a specific trademark as an input and deliver a set of trademarks ranked by similarity to the query image, which is commonly referred to as query by example in content based image retrieval. While several systems have been proposed [28], [9], [15], their retrieval performance leaves a lot of room for improvement [25]. There are several papers suggesting new algorithms for visual trademark retrieval, but their evaluations are based on trademark datasets downloaded from the internet [22], [23], pure shape datasets like MPEG-7 [13], [1] or hand picked ground truth [27], [20], [5], [26]. Unfortunately, objective evaluation of these systems is currently hardly possible as there are no datasets available that (i) represent real world data, ie. the actual visual trademarks registered at the trademark offices, and (ii) that are based on expert opinions and court decisions.

To aid with the development of content based visual trademark retrieval systems, this paper introduces a realistic novel dataset based on real world trademark trials. Our dataset can provide the base for research on content based visual information retrieval systems. The dataset contains 1,859,218 visual trademarks registered at the United States Patent Office (USPTO) as well as three different sets of ground truths based on trials at the United States Trademark Trial and Appeal Board (USTTAB). The raw visual trademarks and trial data is provided by Google[1][2], the extracted meta data is available at a public website[3].

* This article is based on the master's thesis of Lukas Knoch and has been done with the support of the World Intellectual Property Organization, WIPO, partially as the result of an internship at the UN Headquarters in Geneva, CH

[1] Lukas Knoch was student at Alpen-Adria Universität Klagenfurt, Austria `lukas.knoch@aau.at`

[2] Mathias Lux is Associate Professor at the Institute for Information Technology at Alpen-Adria Universität Klagenfurt `mathias.lux@aau.at`

[1] `https://www.google.com/googlebooks/uspto-trademarks-usamark.html`, last visited 2016-01-19

[2] `https://www.google.com/googlebooks/uspto-trademarks-ttab.html`, last visited 2016-01-19

[3] www.rumpelcoders.at/usttabdataset

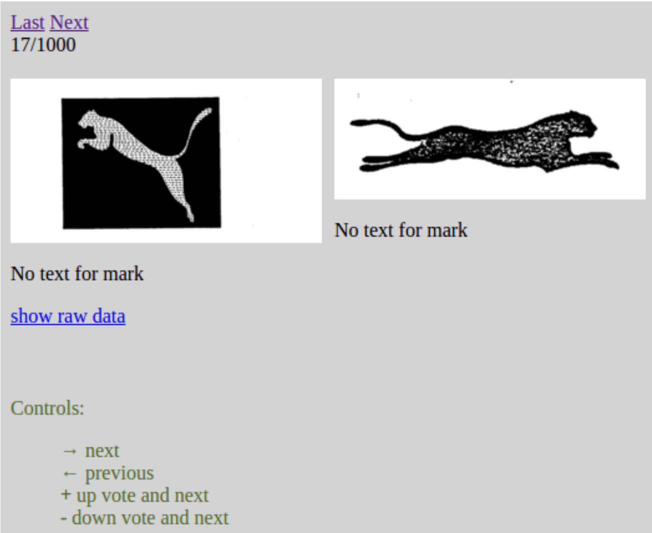

Fig. 2. The view on a pair of logos in the visual similarity evaluation application.

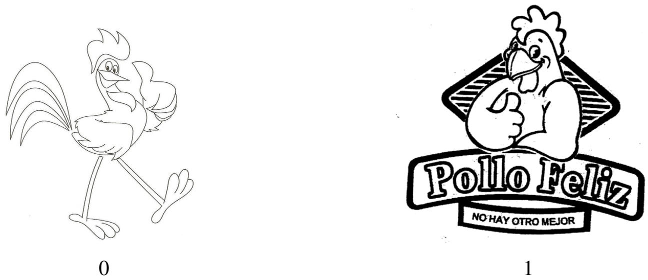

Fig. 3. One of the logo pairs in the USTTAB strict ground truth

II. Dataset

As already mentioned Google offers several trademark collections as free download[4] in cooperation with the USPTO. Note at this point that these downloads offer the actual USPTO data, ie. the actual image files filed for registration as well as the resulting metadata. On the Google site, daily trademark applications, images and the USTTAB trials data from 1955 until today are available. All of these can be downloaded in chronologically ordered ZIP-archives containing an XML file with describing all trials in the specific period of time.

A. Selection Criteria

For the creation of our new ground truth, the trials from 1955 until end of August 2015 were chosen, being all trials available at the time of extraction. Each trial entry in the retrieved data contains the *party-information*, a section that includes information about all parties involved in the trial. Each party has zero or more properties, which correspond to the trademarks associated with it. The properties are identified by a unique identification and a serial number.

A first filtering step was taken by selecting only those trials that do regard an opposition. For the dataset only oppositions are interesting, as those contain cases of confusing similarities, in contrast to obvious ones. In the next step, all entries with exactly two parties and exactly one associated property per party were selected. In all other cases it is not possible to distinguish the trademarks relevant for this claim. By joining this data with all available US trademark images, trials regarding non-visual trademarks could be removed.

As the presence of trademark images does not guarantee that the trial was filed because of visual similarity, the next step was to detect the *type of similarity*. Unfortunately, there is no formal classification contained in the data. To overcome this problem, a web-based application was developed, which allows experts to decide whether the trial was based on visual similarity or not. The experts were chosen from three different areas of expertise: One from the field of visual information retrieval at the University of Klagenfurt, one from the field of trademark retrieval at the World Intellectual Property Organisation and one with appropriate knowledge in both fields. To be able to create a sufficiently big ground truth in reasonable time, 1000 trials were randomly chosen from the previously selected. The application showed two trademark images next to each other and asked the expert to decide whether the claim was due to visual similarity or not. To assist the experts in their decision, the trademark name was presented beyond the image if one was present (compare Fig. I).

For the 1000 logo pairs, all experts agreed on visual similarity in 160 cases. At least two of the three experts agreed on visual similarity in 384 cases while there are 451 trials in which only one expert judged that the trial was due to visual similarity. The 1000 pairs included nine control pairs of obvious visual similarity, which were correctly answered by all experts.

B. Properties

The resulting dataset consists of 1.8 million visual trademarks. Those trademarks are either registered, pending or canceled in the USPTO registration data base. The set is composed by 1,587,248 verbal signs, 533,910 non-verbal signs and 4,867,626 combined trademarks. The signs are of varying image quality with different resolution, in color, gray scale or binary black & white format. As this data is directly form the USPTO's registration data base, its composition is realistic and, therefore, well suited for objective evaluations.

From the USTTAB trials and the expert's decisions, three blends of the data set were created. The first blend includes only logos on which all experts agreed. It is therefore referred to as *strict ground truth*. An example for this set can be seen in Fig. 3. The second blend consists of the logo pairs a majority of experts agreed on, the *majority ground truth* (cp. Fig. 4). Finally, the *minority ground truth* consists of all pairs with at least one expert voting for visual similarity (cp. Fig. 5).

[4] `https://www.google.com/googlebooks/uspto-trademarks.html`, last visited 2016-01-19

0 1

Fig. 4. One of the logo pairs in the USTTAB majority ground truth not being part of the strict set.

0 1

Fig. 5. One of the logo pairs in the USTTAB minority ground truth not being part of the strict set or the majority set.

C. Data Format

The dataset is defined in multiple text files. The first file, *data_full.txt*, contains the registration number of all trademarks used as diversifiers as well as all trademarks from the ground truth. Each line contains one number. The files *data_10.txt* and *data_1.txt* contains a 10% and the 1% random sample in the same format for test on smaller data sets, while still providing comparability. The ground truth is available in the folder *groundtruth*. This folder contains the files *gt_strict.txt*, *gt_majority.txt* and *gt_minority.txt*, which hold a comma separated list of trademark registration numbers identifying the visually similar logo pairs.

III. Retrieval Baseline

To provide a baseline for comparison, several state of the art algorithms were tested on the new dataset. The tests were executed with a benchmark software based on LIRE [19], which was presented in [16]. Note at this point that all descriptors used in the test as well as the benchmarking suite have been contributed to the LIRE open source project[5].

A. Tested Features

The following features were chosen to be tested on the new dataset because they not only cover a wide diversity of features types like color, shape, texture and combinations of them, but also because some of them were proposed as well suited in the trademark retrieval domain [2]. *Local Binary Patterns* (LBP) [11] represent the local texture of an image by encoding the threshold of each pixel's neighborhood in a binary number. A rotation invariant version can be achieved by restricting the observed patterns the so-called *uniform patterns*. For *Binary Patterns Pyramid (BPP)* a spatial pyramid was applied on the LBP. The *Shapeme Histogram Descriptor* (Shapeme) captures the global shape of an image by extracting the shape context and clustering with K-nearest neighbors. In this experiment, the shape contexts were calculated for 256 points chosen by Jitandta's algorithm with three time oversampling and 512 bins for the descriptor [3]. *Centrist* is a feature similar to LBP and also captures local texture. *Joint Composite Descriptor* (JCD) [29] combines the two fuzzy histogram features Color and Edge Directivity Descriptor [8] and Fuzzy Color and Texture Histogram [6]. *Adaptive Contours and Color Integration Descriptor (ACCID)* [12] captures visually salient shapes and combines them with a fuzzy color histogram. *Pyramid Histogram of Oriented Gradients* (PHOG) [4] extracts information about the local shape and the layout of the shape with a with a Spatial Pyramid Kernel. In this experiment, 15 orientation bins were used as that has been found effective in the context of trademark retrieval ($PHOG_{15}$, cp. [16]).

For the evaluation, the logos were resized to a maximum width and height of 512 pixel retaining aspect ratio. In an additional preprocessing step, a despeckle filter was applied and the white pixels were trimmed. Table III-A shows the result of the outlined features on the full USTTAB dataset utilizing the strict ground truth. As can be seen easily from Table III-A, $PHOG_{15}$ outperforms the other descriptors regarding recall and mean average precision. In terms of average and normalized rank, the Shapeme feature performs better than PHOG.

Fig. 6 shows the comparison of the mean average precision (MAP) for $PHOG_{15}$, Shapeme, ACCID, JCD, BBP, and Centrist on the three different ground truths. For Shapeme and $PHOG_{15}$, the MAP correlates to the agreement of the experts. The less agreement in the ground truth, the lower the MAP.

IV. Conclusion and Challenges

The data set as presented provides a hard challenge to researchers in visual information retrieval. While the data from the USTAB trials provides pairs of trademarks with confusing similarity, for both of the pairs it is very likely to find numerous visually similar other logos, which were not part of a trial. Moreover, companies often file trademarks in different version, re-register them or have multiple data records in the USTAB registration data base. Fig. 7 shows an example result list from searching for a visual trademark from the ground truth. At position 0 the query is shown and only on position 49 of the list the offending trademark is found. However, it can be easily seen that the logos in between are visually similar to the trial's logo pair.

While this is definitely a problem for a common use case like digital photo retrieval, in the visual trademark domain the experts doing inquiries certainly go beyond the first few results and finding the offending logo in the first 100 or even 500 results helps them with their work. Note also at this point that the data set is especially about confusing similarity, not near duplicate search, as the latter one has been subject to a lot of research already. Therefore, for future work we

[5] `http://www.lire-project.net/`, last visited 2016-01-19

Feature	Rank	$\widetilde{Rank}$	Recall@100	Recall@500	MAP
LBP	230,784.8	0.124	0.267	0.323	0.178
LBP (RotInv)	250,123.1	0.135	0.305	0.389	0.164
Shapeme	201,853.2	**0.10856828071845802**	0.488	0.513	0.378
Centrist	307,558.3	0.165	0.500	0.502	0.496
BPP	327,727.0	0.176	0.500	0.503	0.496
JCD	267,515.1	0.144	0.503	0.512	0.492
ACCID	227,305.3	0.122	0.505	0.510	0.499
PHOG$_{15}$	220,036.4	0.122	**0.5248344370860927**	**0.5364238410596026**	**0.5157031013278772**

TABLE I

THE RESULTS OF THE STRICT GROUND TRUTH (302 QUERIES) EVALUATED ON THE FULL USTTAB COLLECTION IN TERMS OF AVERAGE RANK, NORMALIZED RANK, RECALL AT 100, RECALL AT 500, AND MEAN AVERAGE PRECISION.

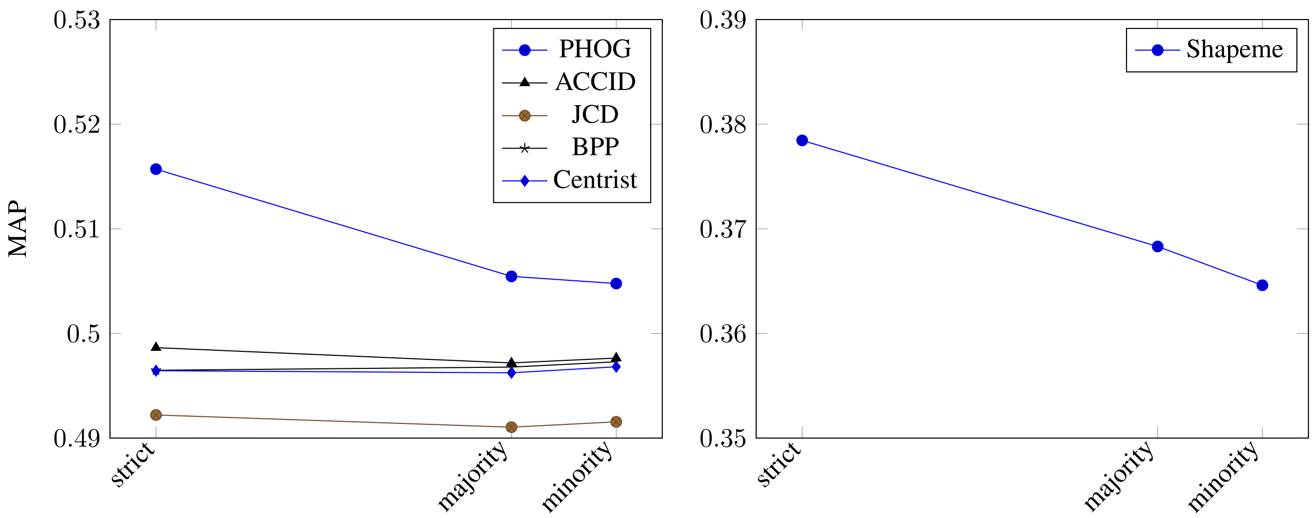

Fig. 6. Comparison of MAP results of different algorithms on the three USTTAB ground truths strict, majority and minority for the full collection. The x axis is scaled to represent the number of queries in each ground truth (302 for strict, 750 for majority and 882 for minority). While BPP, ACCID, JCD and CENTRIST hardly show any change in value, PHOG and Shapeme seem to mirror the human perception.

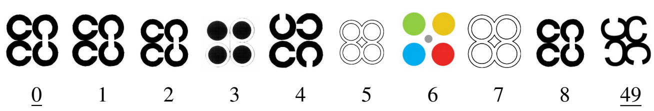

Fig. 7. Examples of retrieval results for a logo pair from the ground truth. At rank 0 the image shows the query, then the first eight results and only at rank 49 the logo from the corresponding USTAB trial.

aim to take a close look at the evaluation procedure, ie. by investigating the possibility of taking into account similar images that have not been in trials, as has been done for the pooling method in text information retrieval [18].

The data set has already been employed for testing different parameters of the PHOG and Shapeme features as well as extensive evaluations using other local and global features alike. The findings have already been integrated in the trademark search engine of the World Intellectual Property Organization (WIPO) of the United Nations[6].

However, there is a long way to go and there are several tasks, for which we propose crowd workers to be employed:

Identification of multiple instances. As noted before logos are submitted and re-submitted by the same company all around the world. These duplicate entries, which are often near duplicates in the visual domain, are visually similar, but should be considered separately. Crowd workers could identify and label the (near) duplicate entries.

Offending logos not investigated by the appeal board. As it is a lengthy and complicated process to file an appeal, there are a lot of visually confusing similarities that have no been investigated by the appeal board. In the current version of the data set these offending logos might show up as false positives in benchmarking. Crowd workers could label the offending logos to be treated separately.

Judging visually confusing similarity. While we had experts judge the offending logos upon visual vs. conceptual confusion, we think that the intuitive concept of visually confusing logos in the head of actual consumers is different to the concept adopted by legal experts. With the help of crowd workers we could paint a picture of how consumers see visual trademarks as well as the relevance and impact of offending logos and provide feedback to the legal experts.

[6] `http://www.wipo.int/branddb`, last visited 2016-08-30

Acknowledgements

We'd like to thank the Glenn MacStravic from the World Intellectual Property Organization for his ongoing support and critical reflection and discussion of our work.

References

[1] S. Agarwal, N. Chaturvedi, and P. K. Johari, "Content based trademark retrieval by integrating shape with colour and texture information," *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 7, no. 4, pp. 295–302, 2014.

[2] S. Belongie and J. Malik, "Matching with shape contexts," in *Content-based Access of Image and Video Libraries, 2000. Proceedings. IEEE Workshop on*, 2000, pp. 20–26.

[3] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 4, pp. 509–522, Apr 2002.

[4] A. Bosch, A. Zisserman, and X. Munoz, "Representing shape with a spatial pyramid kernel," in *Proceedings of the 6th ACM International Conference on Image and Video Retrieval*, ser. CIVR '07. New York, NY, USA: ACM, 2007, pp. 401–408. [Online]. Available: http://doi.acm.org/10.1145/1282280.1282340

[5] A. Cerri, M. Ferri, and D. Giorgi, "A new framework for trademark retrieval based on size functions," *Vision, Video, and Graphics*, 2005.

[6] S. Chatzichristofis and Y. Boutalis, "Fcth: Fuzzy color and texture histogram - a low level feature for accurate image retrieval," in *Image Analysis for Multimedia Interactive Services, 2008. WIAMIS '08. Ninth International Workshop on*, May 2008, pp. 191–196.

[7] B. D. Cullity, *Introduction to Magnetic Materials*. Reading, MA: Addison-Wesley, 1972.

[8] R. Datta, D. Joshi, J. Li, and J. Z. Wang, "Image retrieval: Ideas, influences, and trends of the new age," *ACM Comput. Surv.*, vol. 40, no. 2, pp. 5:1–5:60, May 2008. [Online]. Available: http://doi.acm.org/10.1145/1348246.1348248

[9] J. Eakins, J. Boardman, and K. Shields, "Retrieval of trade mark images by shape feature-the artisan project," in *Intelligent Image Databases, IEE Colloquium on*, May 1996, pp. 9/1–9/6.

[10] R. K. Gupta and S. D. Senturia, "Pull-in time dynamics as a measure of absolute pressure," in *Proc. IEEE International Workshop on Microelectromechanical Systems (MEMS'97)*, Nagoya, Japan, Jan. 1997, pp. 290–294.

[11] D.-C. He and L. Wang, "Texture unit, texture spectrum, and texture analysis," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 28, no. 4, pp. 509–512, Jul 1990.

[12] C. Iakovidou, "Development, implementation and evaluation of methods for the description and the retrieval of multimedia visual content using intelligent techniques," pp. 84–100.

[13] T. Iwanaga, H. Hama, T. Toriu, and T. T. Zin, "A modified histogram approach to trademark image retrieval," *International Journal of Computer Science and Network Security*, vol. 11, no. 4, April 2011.

[14] R. Jain, K. K. Ramakrishnan, and D. M. Chiu, "Congestion avoidance in computer networks with a connectionless network layer," Digital Equipment Corporation, MA, Tech. Rep. DEC-TR-506, Aug. 1987.

[15] T. Kato, "Database architecture for content-based image retrieval," pp. 112–123, 1992. [Online]. Available: http://dx.doi.org/10.1117/12.58497

[16] L. Knoch, "Content based search and retrieval in visual trademarks and logos," 2016.

[17] Q. Li, "Delay characterization and performance control of wide-area networks," Ph.D. dissertation, Univ. of Delaware, Newark, May 2000. [Online]. Available: http://www.ece.udel.edu/~qli

[18] A. Lipani, M. Lupu, and A. Hanbury, "Splitting water: Precision and anti-precision to reduce pool bias," in *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, 2015, pp. 103–112.

[19] M. Lux and S. A. Chatzichristofis, "Lire: Lucene image retrieval: An extensible java cbir library," in *Proceedings of the 16th ACM International Conference on Multimedia*, ser. MM '08. New York, NY, USA: ACM, 2008, pp. 1085–1088. [Online]. Available: http://doi.acm.org/10.1145/1459359.1459577

[20] A. Nigam, A. K. Garg, and R. Tripathi, "Content based trademark retrieval by integrating shape with colour and texture information," *International Journal of Computer Applications*, vol. 22, no. 7, May 2011.

[21] W. I. P. Organization, *International Classification of the Figurative Elements of Marks: Vienna Classification*, ser. WIPO publication. World Intellectual Property Organization, 1997. [Online]. Available: https://books.google.at/books?id=nw3QAAAACAAJ

[22] M. Rusiñol, D. Aldavert, D. Karatzas, R. Toledo, and J. Lladós, "Interactive trademark image retrieval by fusing semantic and visual content," in *Advances in Information Retrieval - 33rd European Conference on IR Research, ECIR 2011, Dublin, Ireland, April 18-21, 2011. Proceedings*, 2011, pp. 314–325. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-20161-5_32

[23] Z. Shaaban, "Trademark image retrieval system using neural networks," *International Journal of Computer Science and Network*, vol. 3, no. 1, February 2014.

[24] W. V. Sorin, "Optical reflectometry for component characterization," in *Fiber Optic Test and Measurement*, D. Derickson, Ed. Englewood Cliffs, NJ: Prentice-Hall, 1998.

[25] TM5, "Report on the TM5 image search project," 2015, Project Report.

[26] R. H. van Leuken, M. F. Demirci, V. J. Hodge, J. Austin, and R. C. Veltkamp, "Layout indexing of trademark images," in *Proceedings of the 6th ACM International Conference on Image and Video Retrieval*, ser. CIVR '07. New York, NY, USA: ACM, 2007, pp. 525–532. [Online]. Available: http://doi.acm.org/10.1145/1282280.1282356

[27] 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 Recognition*, vol. 42, no. 3, pp. 386–394, 2009.

[28] J. Wu, C. Lam, B. Mehtre, Y. Gao, and A. Narasimhalu, "Content-based retrieval for trademark registration," *Multimedia Tools and Applications*, vol. 3, no. 3, pp. 245–267, 1996. [Online]. Available: http://dx.doi.org/10.1007/BF00393940

[29] K. Zagoris, S. Chatzichristofis, N. Papamarkos, and Y. Boutalis, "Automatic image annotation and retrieval using the joint composite descriptor," in *Informatics (PCI), 2010 14th Panhellenic Conference on*, Sept 2010, pp. 143–147.

[30] S. Zhang, C. Zhu, J. K. O. Sin, and P. K. T. Mok, "A novel ultrathin elevated channel low-temperature poly-Si TFT," *IEEE Electron Device Lett.*, vol. 20, pp. 569–571, Nov. 1999.