
Time-efficient Logo Spotting using Text/Non-text Separation as Preprocessing and Approximate Nearest Neighbor Search

Viet Phuong Le* — **Nibal Nayef*** — **Muriel Visani*** — **Jean-Marc Ogier*** — **Cao De Tran****

* *Laboratory L3I, Faculty of Science and Technology, La Rochelle University, France*

** *College of Information and Communication Technology, Can Tho University, Vietnam*

{viet_phuong.le, nibal.nayef, muriel.visani, jean-marc.ogier}@univ-lr.fr, tcde@cit.ctu.edu.vn

RÉSUMÉ. Dans les systèmes de vision par ordinateur et plus particulièrement les systèmes de recherche de documents, la recherche des similarités entre les vecteurs de descripteurs de grande dimension est la partie la plus coûteuse en termes de temps de calcul. Dans cet article, nous proposons un système de recherche de documents basé sur la détection de logos. Ce système est efficace en termes de temps de calcul. En effet, notre approche de détection consiste à appairer et grouper les différents descripteurs. Afin de réduire le nombre de descripteurs de type point-clé à appairer, nous proposons d'utiliser la méthode de séparation texte/non-texte pour éliminer les descripteurs de la couche texte qui sont non-pertinents pour l'appariement des logos. Cette méthode de séparation est utilisée comme une étape de prétraitement rapide et efficace. L'étape d'appariement des descripteurs de type point-clé est optimisée davantage en utilisant les algorithmes de recherche approximative de plus proches voisins. Notre approche de recherche globale de documents basée sur la recherche de logos est évaluée sur la base de données standard « Tobacco-800 » d'une part et notre propre base de données des publicités des magazines privés d'une autre part. Les résultats obtenus démontrent que les deux étapes d'optimisation proposées, plus précisément la séparation de texte, réduisent considérablement le temps de calcul du système de 75% à 47% respectivement dans les deux bases, tandis que la précision reste stable.

ABSTRACT. Searching for the most similar matches to high dimensional feature vectors is the most computationally expensive part of many computer vision and document retrieval systems. This work proposes a time-efficient document retrieval system based on logo spotting. The spotting approach is based on feature matching and grouping. In order to reduce the number of key-point features to be matched, we propose to utilize a text/non-text separation method to get rid

of text layer features which are irrelevant to logo matching. The separation method is used as a fast and effective preprocessing step. We further optimize the key-point feature matching step by using an approximate nearest neighbor search algorithms. The overall document retrieval with focused logo retrieval is evaluated on the standard Tobacco-800 database and also our private advertisement magazine database. The results show that the two proposed speed up steps – specially the text separation – reduce the computation time of the system sharply by 75% and 47% on the two databases respectively, while its precision remains unaffected.

MOTS-CLÉS : Détection et recherche de logos₁, Séparation texte/non-texte₂, Système efficace de détection₃.

KEYWORDS: Logo spotting and retrieval₁, Text/not-text separation₂, Efficient spotting systems₃.

1. Introduction

Logos are commonly used in documents, especially in business and administrative documents. They allow us to determine the source of the documents quickly and accurately, without any textual transcription and at a low cost. This brings about different interesting issues, such as logo spotting, recognition and also document retrieval and indexation based on spotting logos.

Different methods for logo spotting have been developed in the literature, such as key-point-based approaches (Bagdanov *et al.*, 2007; Rusinol et Lladós, 2009a; Rusinol *et al.*, 2013; Revaud *et al.*, 2012), index-based approaches and learning-based approaches. In the key-point-based approaches, pairs of key-points between sets of key-points of two images : a query image and a document image are computed. The key-points are typically described by high dimensional feature vectors. For each key-point of the query image, its best candidate match in the document image is found.

Bagdanov et al. (Bagdanov *et al.*, 2007) used this approach for detection and retrieval of trademarks appearing in sports videos, while Rusinol et al. (Rusinol et Lladós, 2009a; Rusinol *et al.*, 2013) and Revaud et al. (Revaud *et al.*, 2012) applied them to find the matches within a bag-of-words model. These works used a linear searching method to find matches which is the most efficient for nearest neighbor searching ; however, it is often too costly to deal with high dimensional feature spaces. This brings interest to many researchers to provide algorithms that perform approximate but more efficient nearest neighbor searching (Friedman *et al.*, 1977; Liu *et al.*, 2004; Nister et Stewenius, 2006; Mikolajczyk et Matas, 2007), in which the results are sometimes the best neighbors rather than the closest neighbors. Such approximate algorithms can be orders of magnitude faster than exact search, while still providing near-optimal accuracy. In the literature, the multiple randomized kd-tree is an algorithm which gives the best performance (Muja et Lowe, 2009; Muja et Lowe, 2014).

The classical kd-tree algorithm, presented by Freidman et al. (Friedman *et al.*, 1977), is the most widely used algorithm for nearest neighbor searching. This algorithm is reported with N build time and $\log(N)$ search time for N points. Despite being efficient in low dimensions, its performance degrades quickly in high dimensions. An improved version of kd-tree algorithm using multiple randomized kd-trees is presented by Silpa-Anan and Hartley (Silpa-Anan et Hartley, 2008). The randomized trees are built by choosing the split dimension randomly from the first D dimensions on which data have the greatest variance. The implementation of this algorithm is publicly available in the Fast Library for Approximated Nearest Neighbor (FLANN)¹.

Exploring another direction for speeding up logo spotting, we propose to reduce the input size, namely the number of features to be matched. As logos belong to the graphics – non-text – layer of document images, we investigated separating text and non-text in documents as a step which precedes the feature matching step. There, we remove almost all text content which could produce many key-points. In the field of

1. <http://www.cs.ubc.ca/research/flann>

document analysis, separating text and non-text has been used as a layout analysis step (Wong *et al.*, 1982 ; Bloomberg, 1991 ; Lin *et al.*, 2006 ; Le *et al.*, 2015) to improve the accuracy as well as the running time of OCR systems. We extend the application of text/non-text separation to information spotting tasks in documents, because it is easier and faster to perform spotting if the text and non-text are well segmented.

In this work, we show how to make logo retrieval systems more efficient. As case studies, we consider the logo spotting methods presented in (Le *et al.*, 2012 ; Le *et al.*, 2013), and we consider specifically the application of document retrieval presented in (Le *et al.*, 2014) by the same authors. We enhance this retrieval system by : firstly, using the text/non-text separation method presented in (Le *et al.*, 2015) as a preprocessing step. Secondly, using an Approximate Nearest Neighbor (ANN) search in order to make the overall logo spotting approach more time-efficient. The paper is structured as follows : the proposed approach is presented in Section 2. The proposed evaluation protocols of the different approach steps are discussed in Section 3. Section 4 presents the experimental evaluation, and we present conclusions in the last section.

2. The Proposed Approach

Our goal is to perform document retrieval based on a logo spotting approach which mainly uses feature matching and nearest neighbor search. Such techniques – although widely used – are time consuming and could be costly in the context of retrieval applications. In order to optimize and increase the efficiency of our logo focused retrieval system, we propose two main steps : Firstly, a text segmentation step to greatly reduce the content of documents in the database, which in turn reduces greatly the number of features to be matched, and hence increasing the speed of feature matching and grouping. Secondly, the use of approximate nearest neighbor (ANN) techniques for speeding up the feature matching process.

Figure 1 shows the block diagram of our proposed approach for document retrieval based on logo spotting. As a main preprocessing step, we apply a text/non-text separation method to segment a document into a text layer and a non-text layer. The latter layer should contain graphs, logos, separators or any content which is considered as non-text. The images of the non-text layers are given as input images to the document retrieval system. This system is mainly based on a matching and grouping approach of local visual features, where an ANN search is applied to speed up the key-point matching step. In the following subsections, we explain the proposed optimization steps of the framework in more detail.

2.1. Text/non-text Separation

As mentioned above, the main goal of this step is to segment a document image into a text layer and a non-text layer. By getting rid of all the text in the database documents, the number of extracted local features from a document is greatly reduced.

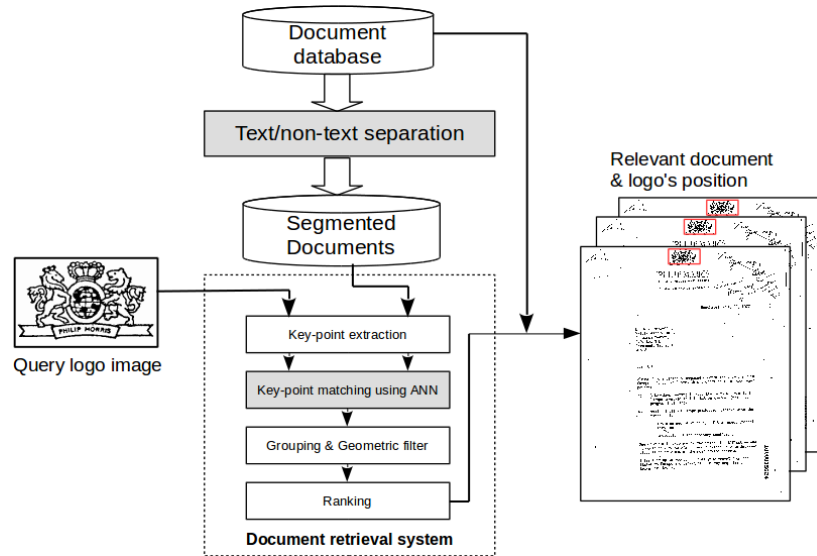


Figure 1. The block diagram of our proposed approach. Text/non-text separation is used as a preprocessing step and an ANN search is applied in the key-point matching step of the document retrieval system.

The feature extraction and matching is considered the bottleneck step of spotting and retrieval systems. Our system is based on spotting logos which belong to non-text or image content, hence, keeping only non-text content ensures that the necessary information remains in the database documents. We use a learning-based approach for text and non-text separation in document images (Le *et al.*, 2015). It has been shown in (Le *et al.*, 2015) that this approach is efficient, fast and can be generalized to different databases. It is worth mentioning that any suitable text and non-text segmentation method could be used. In Subsection 4.2, we also experimentally test text separation using an adapted version of the classical RLSA method (Wong *et al.*, 1982).

The training features are extracted at the level of connected components, a mid-level between the slow noise-sensitive pixel level, and the segmentation-dependent zone level. Connected components are classified through their characteristics represented by features. Therefore, it is important to find features which have the ability to well discriminate the different visual characteristics of connected components. Simple geometrical features such as elongation and solidity; moment features such as Hu moments; stroke width features are used in our method. In addition, the surrounding context of a connected component can play an important role in classifying the text and the non-text components. We extract context features based on the nearest neighbors of the connected component. Those features include shape, size and stroke width information. Finally, log-normal distribution is used to normalize the features.

For labeling connected components, Adaboosting with Decision trees is used. Decision trees is a simple learning method for our set of features, and it provides fast and good results. Finally, the classification of connected components into text and non-text is corrected based on classification probabilities and size as well as stroke width analysis of the nearest neighbors of a connected component.

2.2. Document Retrieval based on Logo Spotting

The document retrieval system is based on the approach presented in (Le *et al.*, 2014), where it has been shown that it performs better than state-of-the-art methods. The retrieval is based on logo spotting which in turn is based on matching pairs of key-points between the query logo image and each segmented document in the database. The key-point matching step has a two-stage filtering step with SIFT descriptors and BRIEF descriptors. After that, the matches are grouped by a density-clustering method, and then are filtered by a geometric verification step. Finally, each document is ranked by the number of matched key-points.

In (Le *et al.*, 2014), first, the interest points of the query logo images and the segmented document images are extracted and described by SIFT and BRIEF descriptors. In the second step – the key-point matching –, key-points are matched in the SIFT feature space using the nearest neighbor matching rule. If the ratio of the first and second nearest neighbors is lower than a given threshold ϕ ($\phi = 0.6$ in practice), then the key-point is representative enough to be considered. On the other hand, if the ratio is greater than ϕ , then it means that the matching is not reliable, as there is a possible ambiguity between the two nearest neighbors, and this key-point is moved to the second stage. In the second stage, the key-point and k its neighbors are re-filtered in BRIEF descriptor space to select the best match thanks to a given threshold θ . Segmentation by density-clustering and post-filtering by homography are used in the third step. Finally, each document image is ranked based on the number of matched key-points after normalization.

In order to speed up the system, we enhance the key-point matching step using an ANN search to find matches between sets of key-points of the query logo images and sets of key-points of the segmented document images (see Figure 1).

3. Evaluation Protocols

In order to better analyze the performance of our proposed approach, we evaluate each of the main parts of the approach separately as well as the overall retrieval performance. This shows the effect of the proposed optimization steps.

3.1. Text/non-text separation

In our experiments, the well-known Tobacco-800 database and our private magazine database are used to evaluate the performance. Unfortunately, both databases were not created for task of text/graphics separation. The ground-truth of Tobacco-800 provides only the logo regions and the signature regions (Zhu et Doermann, 2007 ; Zhu et al., 2007), and our magazine database provides logo regions and their classes as ground-truth. Hence, we consider only two types of regions : logo and non-logo regions. Therefore, a text/graphics separation is evaluated based on its ability to reject non-logo regions and keeping logo regions. We define the first evaluation factor based on the remaining area of logo region after the segmentation as follows :

$$\epsilon = \frac{Area(\text{logo_region_after_segmentation})}{Area(\text{original_logo_region})} \quad [1]$$

where the function $Area()$ returns the area of a region, $\text{logo_region_after_segmentation}$ is the region of a logo on a segmented document, and $\text{original_logo_region}$ is the logo region on the original document which is available in the ground-truth. The value of ϵ has to be chosen realistically. We count a correct segmentation if $\epsilon \geq 0.75$. Otherwise, it is a false one. Figure 2 shows examples of correct and false segmentation cases.

In addition, we define a second factor to compute the reduced area of all documents after the segmentation.

$$\text{reduction} = 1 - \frac{\sum Area(\text{document})}{\sum Area(\text{original_document})} \quad [2]$$

where $\sum Area(\text{document})$, $\sum Area(\text{original_document})$ are the sum of regions areas (except logo regions) of all segmented documents and original documents, respectively. In our context, a separation algorithm is better if it has more correct segmentation cases and a lower *reduction* factor.

3.2. Logo Spotting

The measure of performance of logo spotting is quite different from measures of logo recognition methods, since both recognition and localization have to be evaluated together. Similar to the evaluation protocol proposed for spotting systems in (Rusiñol et Lladós, 2009b), we evaluate at logo level. In this case, a binary concept of retrieval whether a logo is found or not is used.

If the logo area in the ground-truth can be overlapped with at least a certain percentage of the resulting spotted logo area, then the logo is considered as found correctly. In our experiments, a logo is recognized as *true positive* if it overlaps with at least a 75% its ground-truth area. Otherwise, it is considered as *false negative*. On the other hand, if the resulting area does not overlap with any ground-truth logos, it is considered as *false positive*.

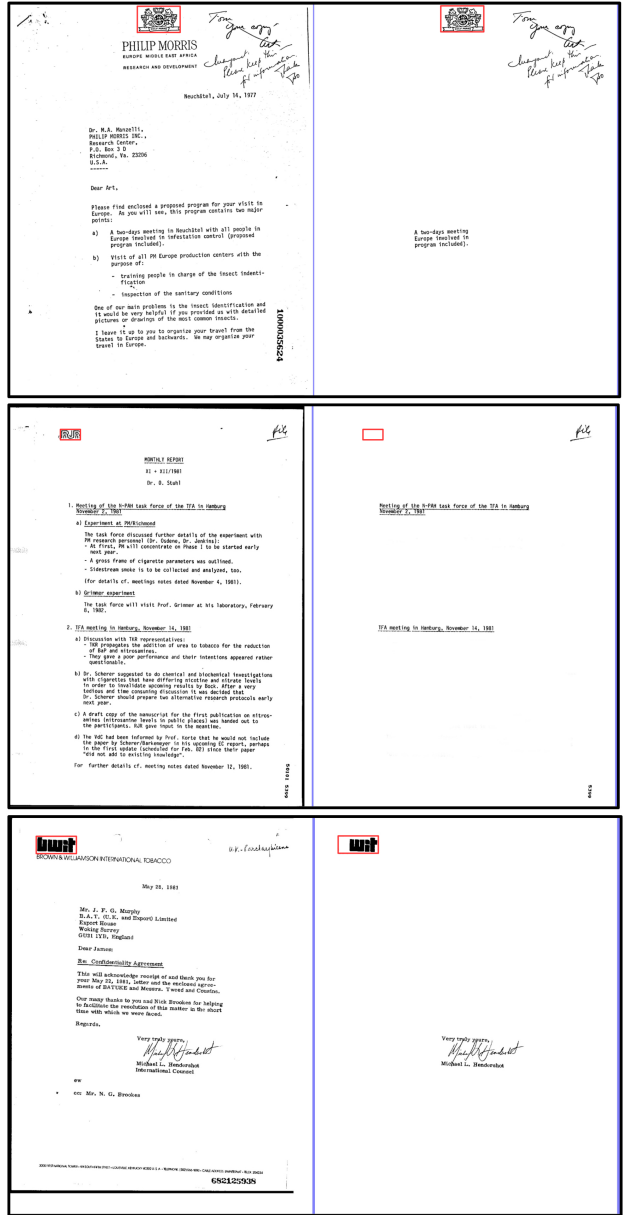


Figure 2. Examples of : correct segmentation (first row), false segmentation (second and third rows) where the remaining area of the logo is lower than 75% compared to the logo in the original document.

3.3. Evaluation protocol for document retrieval

To evaluate the performance of document retrieval system, we also use two most commonly used measures : Mean Average Precision (MAP) and Mean R-Precision (MRP). The overall system performance across all queries and all documents is measured.

$$MAP = \frac{\sum_{q=1}^Q AP(q)}{Q} \quad [3]$$

$$MRP = \frac{\sum_{q=1}^Q RP(q)}{Q} \quad [4]$$

where :

$$AP = \frac{\sum_{k=1}^n \{P(k) \times rel(k)\}}{\# \text{ relevant documents}} \quad [5]$$

$RP(q)$ is the R-precision of query q ; n is the number of retrieved documents; $P(k)$ is the precision at rank k , $rel(k)$ is 1 if the result at rank k is a relevant document, and otherwise 0. Q is the number of queries.

4. Experimental Evaluation

4.1. Databases

The Tobacco-800 dataset is a well-known public dataset for document analysis. It is a realistic dataset for document retrieval (Zhu et Doermann, 2007; Zhu *et al.*, 2007). The Tobacco-800 dataset has 1290 document images including 412 document images containing logos and 878 document images containing no logos. We test our system using a set of queries consisting of 432 logo images among 35 classes extracted from the dataset thanks to the ground-truth. Figures 3 shows examples of documents from the Tobacco-800 database.

The second database has been collected from advertisement magazines. It contains 100 images scanned at 400dpi. There are 113 logos with 6 different logo classes which appear in the database images. Building this database aims to test our proposed logo spotting method in contemporary real colored documents which have complex layouts. We also create its ground-truth information as bounding box around logo regions. Figure 4 shows examples of documents of this database.

4.2. Experiment 1 : text/non-text separation

In the first experiment, we compare RLSA (Wong *et al.*, 1982) to the method in (Le *et al.*, 2015). The steps and parameters of the RLSA method are adapted for best performance on the databases. For training the method in (Le *et al.*, 2015), we randomly select a set of 50 documents of Tobacco-800 database. We keep all non-text

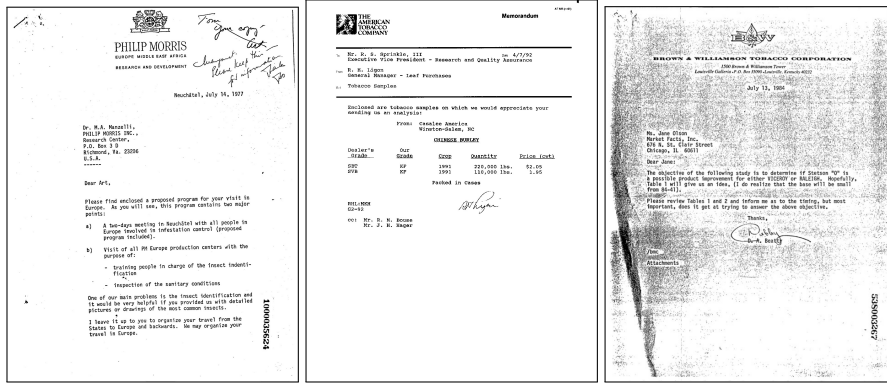


Figure 3. Example documents of the Tobacco-800 database.

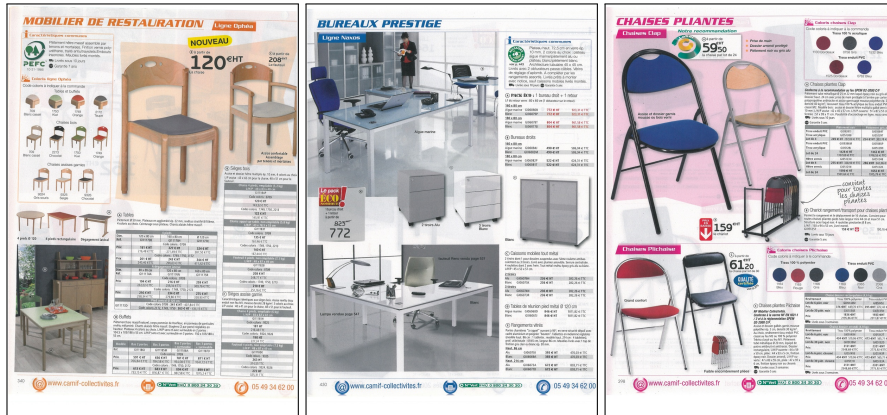


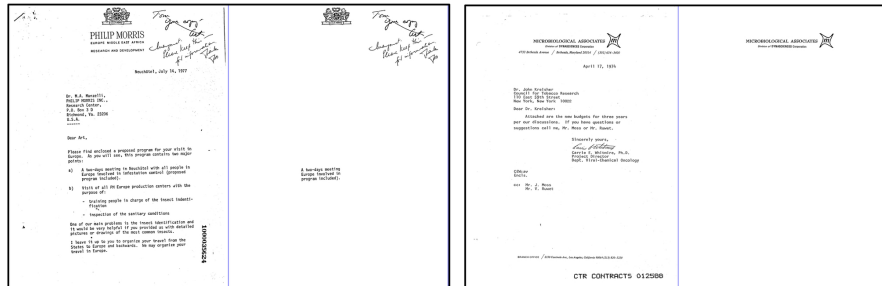
Figure 4. Example documents of our advertisement magazine dataset.

connected components and select randomly the number of text connected components to be twice as many as that of non-text connected components (because the number of text connected components is much higher than non-text components, which causes imbalanced training). For testing, we use the rest of the 1240 documents of Tobacco-800 database.

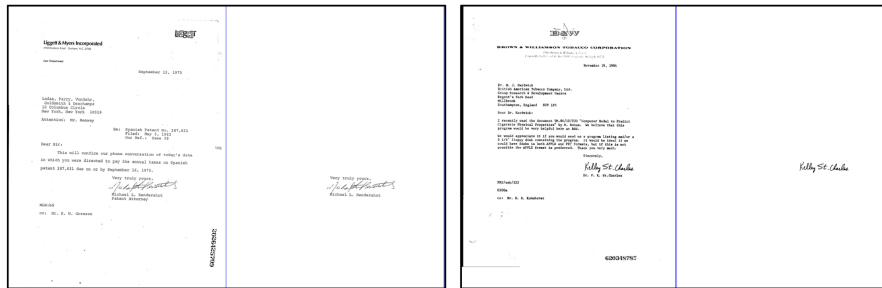
Table 1 shows the results. The method in (Le *et al.*, 2015) gives better results with only 1 missing case over 412 documents containing logos, and can reduce more irrelevant areas than RLSA. Example results are shown in Figure 5. This justifies our choice to adapt the method in (Le *et al.*, 2015).

Tableau 1. Comparison of performance between RLSA and Le et al. (Le et al., 2015).

	RLSA (Wong et al., 1982)	Le et al. (Le et al., 2015)
Number of missing cases	8	1
Reduction	52.82%	79.92%
Avg. computation time	0.97s	0.68s



(a) Correctly segmented : relevant graphical areas are kept along with some irrelevant text.



(b) Falsely segmented : the remaining area of a logo is lower than 75% compared to the logos in the document before segmentation.

Figure 5. Text/non-text separation results. Precise segmentation (removal of all text) is not required in our system as long as all the relevant graphical areas are kept. Those graphical areas are the only needed parts for the subsequent spotting step.

It is important to mention that adding this separation step to the retrieval system does not negatively affect the retrieval time compared to the system's retrieval time without this step. In retrieval applications, text separation could be performed offline when analyzing a database of documents. This step could also be online, it takes 0.68 second per document on average without code optimization.

4.3. Experiment 2 : document retrieval system

In the second experiment, the document retrieval performance is applied to compare between systems using : (1) original document database, (2) segmented document database with linear search and (3) segmented document database with ANN search using kd-tree algorithm. We evaluate the systems based on the number of extracted key-points, average matching time, and two retrieval measures (MAP and MRP). Table 2 presents the results. There is a small decline in precision by nearly 0.5% compared to segmented document database with linear search and segmented document database with ANN using kd-tree algorithm, respectively. However, the matching time decreased sharply over 66% for the system using the segmented document database and over 75% for the system using segmented document database with ANN using kd-tree algorithm.

Tableau 2. Comparison of performance on Tobacco-800 database between systems using original documents, segmented documents with linear search (LS) and segmented documents with ANN using kd-tree algorithm.

	Original documents & LS	Segmented documents & LS	Segmented documents & ANN
Number of key-points	17.9M	3.7M	3.7M
Avg. matching time	1.29s	0.43s	0.31s
MAP	88.31%	88.27%	88.12%
MRP	85.87%	85.51%	85.39%

4.4. Experiment 3 : logo spotting

For testing logo spotting on documents, we build a system similar the document retrieval system above. The ranking step is replaced by a decision step. In the decision step, a given threshold ω based on the percentage of the matches is used to estimate if the document contains the query logo or not. If the percentage of the matches is greater than ω , the document contains the logo, otherwise, the document does not contain this logo ($\omega = 0.025$ in practice). In the text/non-text separation step, to be able to use all 100 documents of advertisement database for testing, we use a set of 136 documents containing graphics from UW-III dataset to train the text/non-text separation model. We also keep all non-text connected components and select text connected components randomly, where their number is twice as many as that of non-text connected components.

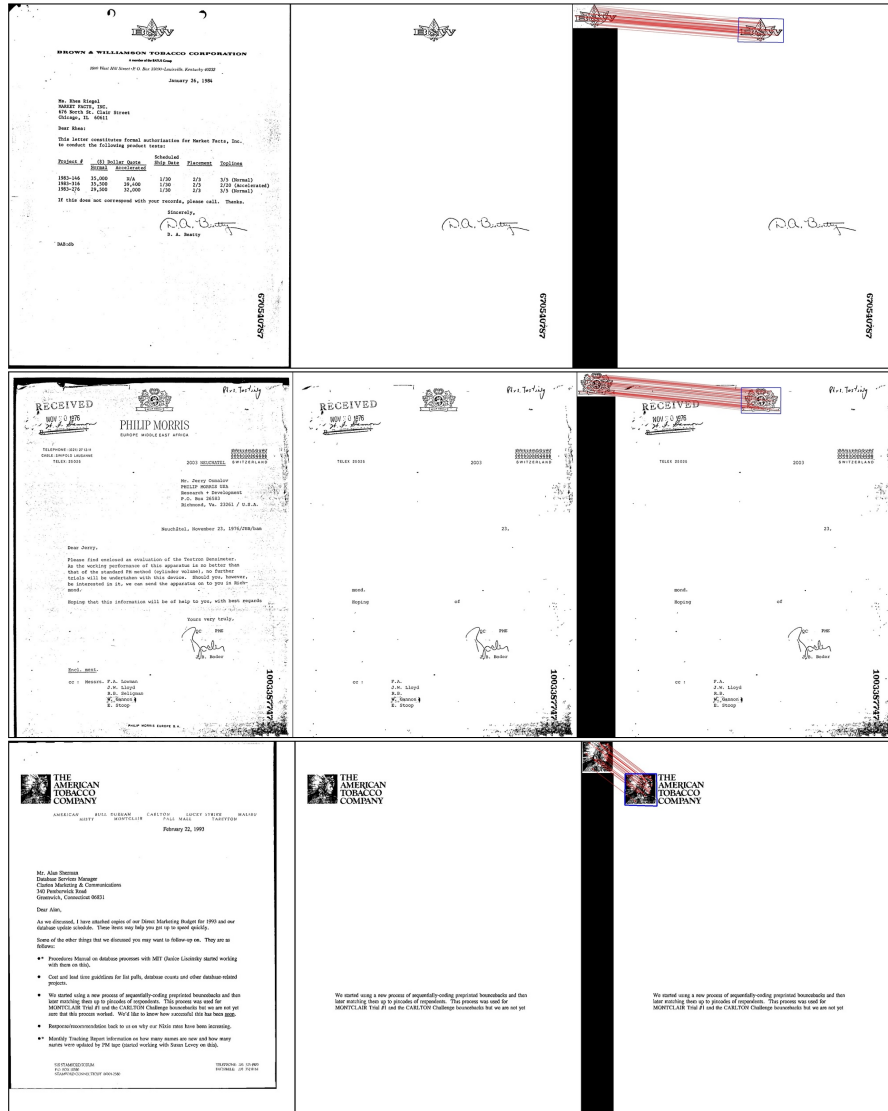


Figure 6. Some examples of logo matching on segmented documents and ANN algorithm on Tobacco-800 database. Column 1 : original documents, Column 2 : documents after text/non-text separation, Column 3 : spotting results.

In the third experiment, the logo spotting performance is also applied to compare between systems using : (1) original advertisement document database, (2) segmented advertisement document database with linear search and (3) segmented advertisement document database with ANN search using kd-tree algorithm. We evaluate the sys-

tems based on the number of extracted key-points, average matching time, and two measures : true positive rate (TPR) and false positive rate (FPR). The results in Table 3 show that the average matching time decreased over 47% while there is a slight increase in the false positive rate by nearly 0.32%. This is because the advertisement document database contains mostly graphics. In addition, most of logos in this database have parts which are text. Figure 7 shows example results of this experiment.

Tableau 3. *Comparison of performance on advertising magazine database between systems using original document database, segmented document database with linear search (LS) and segmented document database with ANN using kd-tree algorithm.*

	Original documents & LS	Segmented documents & LS	Segmented documents & ANN
Number of key-points	3.91M	2.53M	2.53M
Avg. matching time	2.25s	1.35s	1.18s
TPR	99.77%	99.12%	99.80%
FPR	0.08%	0.14%	0.40%

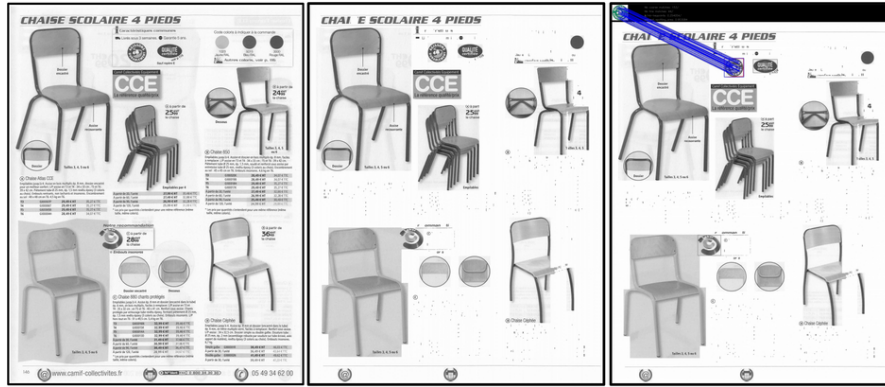
5. Conclusion

For the goal of increasing the efficiency of logo spotting and retrieval systems – which are based on feature matching –, we have integrated text/non-text separation and ANN search into such systems. The text/non-text separation method is fast and efficient in removing all irrelevant features. It uses a learning-based approach on a powerful feature set extracted from connected component and their context. In addition, the ANN with kd-tree algorithm is used to speed up key-point matching step. This work shows the efficacy of analyzing document contents when used as a preprocessing and preparation step for document retrieval systems.

For future work, we aim to enhance the non-text separation method to provide even more accurate segmentation. We would like also to investigate the use of our methods for other information spotting approaches.

6. Bibliographie

Bagdanov A. D., Ballan L., Bertini M., Del Bimbo A., « Trademark matching and retrieval in sports video databases », *Proceedings of the international workshop on Workshop on multimedia information retrieval*, ACM, p. 79-86, 2007.



(a)



(b)

Figure 7. Logo matching on segmented documents and ANN algorithm on advertising magazine database. Column 1 is original documents, Column 2 is the documents after text/non-text separation, and the spotting results are shown in Column 3.

Bloomberg D. S., « Multiresolution morphological approach to document image analysis », *Proc. of the International Conference on Document Analysis and Recognition, Saint-Malo, France, 1991.*

Friedman J. H., Bentley J. L., Finkel R. A., « An algorithm for finding best matches in logarithmic expected time », *ACM Transactions on Mathematical Software (TOMS)*, vol. 3, n° 3, p. 209-226, 1977.

Le V. P., Nayef N., Visani M., Ogier J.-M., De Tran C., « Document Retrieval Based on Logo Spotting Using Key-Point Matching », *Pattern Recognition (ICPR), 2014 22nd International Conference on*, IEEE, p. 3056-3061, 2014.

Le V. P., Nayef N., Visani M., Ogier J.-M., De Tran C., « Text and Non-text Segmentation based on Connected Component Features », *Document Analysis and Recognition (ICDAR), 2015 13th International Conference on*, IEEE, p. 1096-1100, 2015.

- Le V. P., Visani M., De Tran C., Ogier J., « Logo spotting for document categorization », *Pattern Recognition (ICPR), 2012 21st International Conference on*, IEEE, p. 3484-3487, 2012.
- Le V. P., Visani M., De Tran C., Ogier J.-M., « Improving logo spotting and matching for document categorization by a post-filter based on homography », *Document Analysis and Recognition (ICDAR), 2013 12th International Conference on*, IEEE, p. 270-274, 2013.
- Lin M.-W., Tapamo J.-R., Ndovie B., « A texture-based method for document segmentation and classification. », *South African Computer Journal*, vol. 36, p. 49-56, 2006.
- Liu T., Moore A. W., Yang K., Gray A. G., « An investigation of practical approximate nearest neighbor algorithms », *Advances in neural information processing systems*, p. 825-832, 2004.
- Mikolajczyk K., Matas J., « Improving descriptors for fast tree matching by optimal linear projection », *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, IEEE, p. 1-8, 2007.
- Muja M., Lowe D. G., « Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration. », *VISAPP (1)*, 2009.
- Muja M., Lowe D. G., « Scalable nearest neighbor algorithms for high dimensional data », *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 36, n^o 11, p. 2227-2240, 2014.
- Nister D., Stewenius H., « Scalable recognition with a vocabulary tree », *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, vol. 2, IEEE, p. 2161-2168, 2006.
- Revaud J., Douze M., Schmid C., « Correlation-based burstiness for logo retrieval », *Proceedings of the 20th ACM international conference on Multimedia*, ACM, p. 965-968, 2012.
- Rusinol M., D'Andecy V. P., Karatzas D., Lladós J., « Classification of administrative document images by logo identification », *Graphics Recognition. New Trends and Challenges*, Springer, p. 49-58, 2013.
- Rusinol M., Lladós J., « Logo spotting by a bag-of-words approach for document categorization », *Document Analysis and Recognition, 2009. ICDAR'09. 10th International Conference on*, IEEE, p. 111-115, 2009a.
- Rusiñol M., Lladós J., « A performance evaluation protocol for symbol spotting systems in terms of recognition and location indices », *International Journal on Document Analysis and Recognition (IJ DAR)*, vol. 12, n^o 2, p. 83-96, 2009b.
- Silpa-Anan C., Hartley R., « Optimised KD-trees for fast image descriptor matching », *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, IEEE, p. 1-8, 2008.
- Wong K. Y., Casey R. G., Wahl F. M., « Document analysis system », *IBM journal of research and development*, vol. 26, n^o 6, p. 647-656, 1982.
- Zhu G., Doermann D., « Automatic document logo detection », *Document Analysis and Recognition, 2007. ICDAR 2007. Ninth International Conference on*, vol. 2, IEEE, p. 864-868, 2007.
- Zhu G., Zheng Y., Doermann D., Jaeger S., « Multi-scale structural saliency for signature detection », *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, IEEE, p. 1-8, 2007.