Comparing the performance and accuracy of algorithms applied to tattoos images identification

Agnus Azevedo Horta
DCA, FEEC, University of Campinas
Av. Albert Einstein - 400
Brazil 13083-852, Campinas, SP
agnus@dca.fee.unicamp.br

Léo Pini Magalhães
DCA, FEEC, University of Campinas
Av. Albert Einstein - 400
Brazil 13083-852, Campinas, SP
leopini@fee.unicamp.br

ABSTRACT

This article presents results of the simulation of SIFT based algorithms in the context of the identification of tattoos. The algorithms studied are the SIFT - Scale Invariant Feature Transform, ASIFT - Affine SIFT, BOV - Bag of Visual Words and FV - Fisher Vector. The use of the OPF - Optimum-Path Forest and SVM - Support Vector Machine classifiers is exploited in conjunction with SIFT and ASIFT algorithms as well as BOV and FV. The present study uses the National Institute of Standards and Technology (NIST) Tatt-C dataset in a reduced and complete version. This work uses runtime and accuracy to compare the results of the simulations.

Keywords

Tattoo identification, SIFT, ASIFT, Bag of Visual Words, Fisher Vectors, Optimum-path forest classifier, Support Vector Machines, Soft biometric, Content-based image retrieval.

1 INTRODUCTION

The area of tattooing is a comprehensive area of activities that deal with the identification and detection of tattoos, search for similarities between tattoos, search for regions of interest in tattoos, treatment of tattoos on various materials/surfaces among their important applications. It is an area, for example, that supports important forensic activities related to law enforcement for offenders and victim support.

This paper is dedicated to the area of tattoo identification [1]. Although many important tattoo-related activities have been and are in progress in the scientific community, we believe that the comparison among algorithms we bring in this paper can be very useful in the choice of solutions for many applications in tattoo identification.

Most approaches that exist in the literature to solve the problem of identification of tattoos are strictly based on SIFT (Scale Invariant Feature Transform) and ASIFT (Affine SIFT) algorithms. This paper explores the idea to use both algorithms to provide reference results and to support other techniques - BOV (Bag of Visual words), FV (Fisher Vectors) and classifiers - to solve the tattoo identification problem. Thus we combined SIFT and ASIFT with BOV and FV approaches and additionally we applied these algorithms with the classifiers SVM (Support Vector Machine) and OPF (Optimum-Path Forest) to the tattoo identification problem. Our previous experience, in another area of application, showed a very encouraging performance of the OPF algorithm, compared to SVM, in that case with the descriptor BIC (Border/Interior pixel Classification) [DSFM11].

Section 2 gives a brief description of all simulated methods. These methods, as commented above, are the SIFT and ASIFT algorithms, the combination of both algorithms with the use of Bag of Visual words, Fisher Vectors, matching algorithms, and the use of the OPF and SVM classifiers. After that, Section 3 presents a comparison based on the accuracy and performance of the simulated approaches. For this comparison, the NIST Tatt-C dataset [NG15] is used in a reduced version and in its full version as explained in the section.

---

[1] "operational use-cases defined by the NIST challenge – 2015/2016: 1. Tattoo Similarity – matching visually similar or related tattoos from different subjects; 2. Tattoo Identification - matching different instances of the same tattoo image from the same subject over time; 3. Region of Interest - matching a small region of interest that is contained in a larger image; 4. Mixed Media - matching visually similar or related tattoos using different types of images (e.g. sketches, scanned print, computer graphics, or natural images); 5. Tattoo Detection - detecting whether an image contains a tattoo or not." [NG15]

The final section reviews the results of the simulations and discusses the next developments.

2 RELATED WORKS

Tattoo is already being used in identification processes due to its soft biometric characteristic. Although several studies have already been developed in this area, this is still a developing area. For example, there is a small number of available image databases for the development and evaluation of tattoos applications. The NIST National Institute of Standards and Technology has developed the Tatt-C dataset, an important and well-established example of such a set of tattoo images. [NG15]

Figure 1: Examples of tattoos images from the site The Commons (https://www.flickr.com/commons).

Figure 1 shows images that are similar to those used in this work. Due to copyright restrictions, we do not present images belonging to the NIST Tatt-C dataset used in this paper.

Another important point is related to the search for algorithms to identify tattoos, which should have a good relationship between runtime and accuracy. These algorithms should treat the biometric characteristics of tattoos associating the attributes of low level of abstraction that describe the tattoo. This is the point that we address in this article.

One of the first proposals to replace human beings in the assignment of labels for classes of images in the process of identification of tattoos is of 2007([JLJ07]). The proposal was based on the use of attributes of low-level of abstraction - color, shape, texture - to characterize tattoos and on the use of a simple relevance feedback strategy.

Currently, most approaches to tattoo treatment are based on the use of the SIFT descriptor [LOW99] and extensions such as affine SIFT (ASIFT) ([YM11]) exploring the invariance of these methods for scaling and rotation of the image, robustness to distortions related to changes in illumination and, for ASIFT, its property of complete affine invariance. In both approaches, the attributes of an image are represented by the key points of the image.

Lee et al. [LJJ08] describe a content based image retrieval (CBIR) system for combining and retrieving tattoo images. Subsequently, Lee et al. [LJJT12] still using SIFT key points proposes more robust similarity measures that used along with associated metadata improves performance in image retrieval.

The use of "pure" SIFT for the identification of tattoos can still be found in recent works such as Han et al. [HJ13], Marcetic et al. [MRSP14], and in the NIST challenge results Ngan et al.[NQG16].

The ASIFT algorithm is defined by considering a geometric interpretation of the finite decomposition of an image, where the angles corresponding to the latitude (ϕ) and longitude (θ) are gradually changed to generate a set of sample views of the initial image, in which SIFT is applied (see Figure 2). ASIFT can be used in the same applications of SIFT if time is not a strict restriction. The Section 3 will present the simulation results of both algorithms.
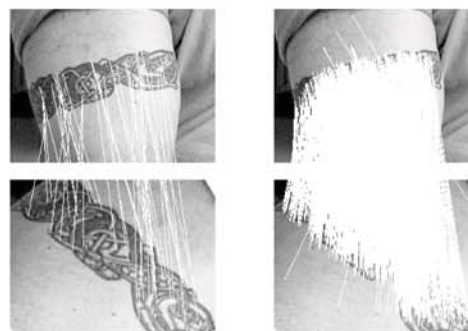

Figure 2: Images from the demo available at http://demo.ipol.im/demo/my_affine_sift/. Images showing the correspondence between the key points obtained by applying the SIFT (left) and ASIFT (right) in an original image and in its rotated image. The demo reports for the image at right 22938 ASIFT key points (top image) and 21577 ASIFT key points (bottom image); the key points computation was accomplished in 5 seconds, and 1482 matchings were identified. For the left image using SIFT 40 matchings were obtained and the key points matching was accomplished in 2 seconds.

Manger [Man12] also explored the categorization of tattoo images through an analogy with learning methods based on the categorization of text using the concept of a bag of words. In this analogy, one of the parameters is the size of a visual dictionary. SIFT is used to extract tattoo image characteristics that will be grouped using K-means. The model was inspired by the BOW model for text proposed by Sivic and Zisserman [SZ03] and extended to images by Csurka et al. [CDF04]. The original Manger modeling was discussed and improved in [MHW16], including an additional index based on Fisher vector coding. Fisher Vector (FV) derived from Fisher Kernel method [JH99]

encodes the visual vocabulary using Gaussian Mixture Models (GMM) and adds information related to local feature descriptors.

Regarding the classification process of images, the classifiers Support Vector Machines (SVM) and Optimum-Path Forest (OPF) are algorithms explored in this work. We use the libraries LIBSVM and LIBLINEAR for SVM and LIBOPF for OPF. LIBSVM Chang and Lin [CL11] implements algorithms for training and testing for SVM and supports several SVM formulations for classification, regression and distribution estimation. LIBLINEAR [FCH08] is an open-source library for large-scale linear classification. It supports logistic regression and linear SVM.

Perronnin and Dance [PD07] applied Kernels Fisher in visual vocabulary using nonlinear support machines (SVM) in the classification process. Schneider and Tuytelaars [ST14] successfully used FV together with the SVM classifier in sketches classification, surpassing the performance of existing techniques, a concept that can be applied to the recovery of tattoo images based on sketches. The present paper explores FV in tattoo identification.

OPF – Optimum-Path Forest classifier proposed by Papa et al. [PFS09] is a supervised classification method that represents each data class by its optimal path tree whose root nodes are called prototypes. The training samples correspond to the nodes of a complete graph whose arcs are weighted by the distances between its nodes.

The Section 3 will explore the combination of algorithms and classifiers using SIFT and ASIFT as base elements. The results will be compared considering accuracy and runtime.

3 RESULTS AND COMMENTS

The results of the simulations presented in this section were obtained by programs executed in a 4GB and 4-core virtual machine using a CPU-i7 with 8GB of memory, 4-core and 8 threads.

The simulations used the NIST Tatt-C dataset [NG15]; two datasets were generated, one for training processing and one for test processing. Each of the two sets was organized in five subsets - folds - to allow mean values across a five run-test, using the same procedure followed by Ngan et al. [NQG16].

In order to better understand and tunning the algorithms, the first set of simulations was performed using a "reduced" set of the Tatt-C dataset composed of, using the terminology of the NIST report, all #probes images (157 true images of interest), all #mates images (215 related images to probe-images) and none (zero) background images. In this way, the folds were defined for training containing approximately 302 images (157

minus 30 true images of interest added to 215 minus 40 images). For test processing, about 30 images were used (different true images for each fold). The results of these simulations are depicted in Table 1 and in Figures 3 to 5 always presenting mean and standard deviation of five rounds.

Considering the algorithms SIFT and ASIFT (lines 1 and 2 of Table 1) the process of searching for an image involves two basic steps, one denominated training which is the step in which the image database is processed and the image descriptor(s) is(are) generated, and the step denominated test that corresponds to the step where one or more images are displayed and the query is performed based on a match function. In the SIFT method, one aspect that influences in the results is the key point matching algorithm. In the experiments, we utilized the Knn Brute Force method of the OpenCV for both algorithms SIFT and ASIFT.

After the generation of the descriptors using SIFT or ASIFT we performed for the training phase the codebook generation for BOV based on KMeans and image coding, and for FV using the Gaussian Mixture Model (GMM) and image coding. Following, for the test phase, the image coding and rank generation (Figures 3 to 5) were performed for both methods. Numbers 75 to 1000 express the size of the codebook for BOV method and numbers 5 and 10 express the number of Gaussians for the FV method. Lines 3 to 12 depict the runtime for the different methods.

We have adopted in this paper the same definition of accuracy as NIST [NQG16]. Thus, accuracy is the amount of correctly identified images divided by the total number of probe images of the current fold.

The execution of the BOV for the reduced Tatt-C dataset with different codebook sizes showed us that for the BOV_SIFT the accuracy expressed by the Cumulative Matching Characteristic or CMC curve decreases with the increase of the codebook (Figure 3). In this way, the BOV_SIFT_75 has the best CMC curve compared to BOV_SIFT_500 and _1000. For ASIFT, a codebook of size 500 or 1000 resulted in almost the same CMC curve, overcoming codebooks with fewer elements (see Figure 4). Figure 5 showed us FV_SIFT_5 and _10 with similar behavior, and FV_ASIFT_5 with the best CMC for all FV.

In the next step, see Table 2, we applied classifiers SVM and OPF to the best algorithms - BOV_SIFT_75, BOV_ASIFT _500, FV_SIFT_5, FV_ASIFT_5 - obtained in the last step. The training and test phases show that the OPF classifier presents the best runtime and accuracy values compared to SVM. For training considering mean values, the best result for runtime was obtained for OPF + BOV_SIFT_75, followed by OPF + FV_ASIFT_5, OPF + FV_SIFT_5 and OPF + BOV_ASIFT_500. For test phase the

	Method	training (all images)		test (one image)	
		mean	std. deviation	mean	std. deviation
1	SIFT	207 *	– – –	25.56	1.63
2	ASIFT	200 *	– – –	536.51	38.87
3	BOV_SIFT_75	68.33	1.01	0.18	0.02
4	BOV_SIFT_500	104.24	2.64	0.30	0.03
5	BOV_SIFT_1000	165.15	3.56	0.41	0.04
6	BOV_ASIFT_75	339.27	5.48	0.78	0.05
7	BOV_ASIFT_500	659.64	13.17	1.34	0.06
8	BOV_ASIFT_1000	620.82	18.88	1.84	0.13
9	FV_SIFT_5	470.20	69.03	0.39	0.03
10	FV_SIFT_10	1,520.38	252.32	0.62	0.07
11	FV_ASIFT_5	2,855.31	768.71	1.85	0.30
12	FV_ASIFT_10	7,749.99	980.99	3.37	0.34

* single execution.

Table 1: Runtime in *seconds* for training and test phases using the reduced (no background images) Tatt-C dataset.

best result for runtime was obtained by OPF + BOV_SIFT_75, followed by OPF + FV_ASIFT_5, OPF + BOV_ASIFT_500 and OPF + FV_SIFT_5. Now, taking into account accuracy, again OPF presented better results than SVM. Considering accuracy mean values, OPF + BOV_ASIFT_500 presented the best result, followed by OPF + FV_ASIFT_5 and FV_SIFT_5, and BOV_SIFT_75.

Considering now the results provided by Table 1 in addition to the results discussed in the last paragraph (Table 2) we elected the algorithms BOV_SIFT_75 and FV_SIFT_5 to apply in the complete Tatt-C dataset in addition to the SIFT used as reference.
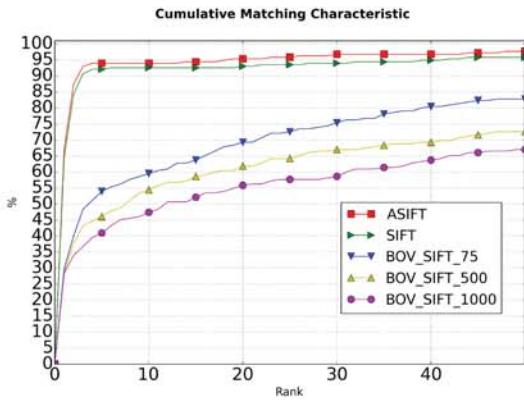

Figure 3: Cumulative Matching Characteristic for BOV_SIFT using the reduced Tatt-C dataset.

BOV_SIFT_75 and FV_SIFT_05 were chosen by the same reasoning of less runtime comparing with similar approaches (smaller dictionary and lower Gaussian number) (Tables 1 and 2) and comparable accuracy to the ASIFT approaches (Table 2). The use of the reduced dataset provided behavioral expectations of all algorithms addressed and, in this way, helped us to reduce the universe of simulation for the algorithms of interest to the complete dataset.

The second set of simulations was performed using the Tatt-C dataset composed by all #probes images (157
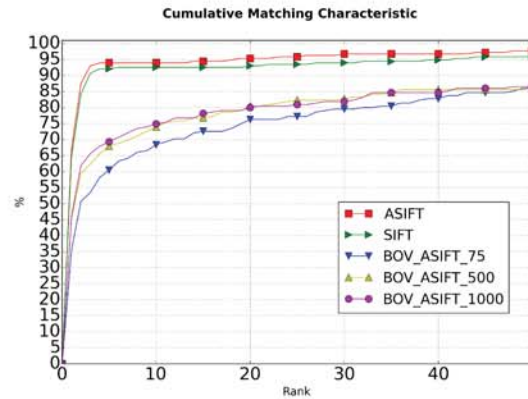

Figure 4: Cumulative Matching Characteristic for BOV_ASIFT using the reduced Tatt-C dataset.
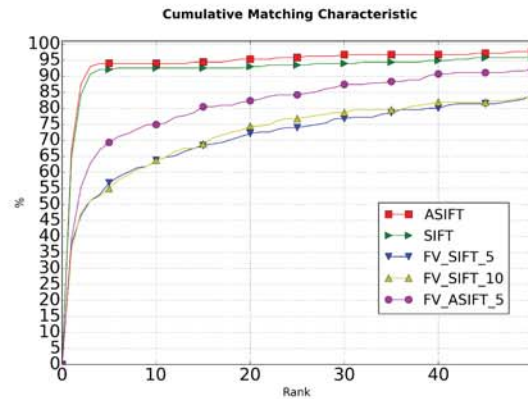

Figure 5: Cumulative Matching Characteristic for FV_SIFT and FV_ASIFT using the reduced Tatt-C dataset.

true images of interest), all #mates images (215 related images to probe-images) and background images (used to confuse the identification) as explained next. At first, a simulation was performed using all (100%) background images. In this way the folds were of approximate size of 4674 images (40 images of #mates images related to the 30 of #probes images, 215 minus 40 of #mates images, 157 minus 30 of #probes images and all 4332 background images) for training and

Method	OPF						SVM					
	Training time (s) (all images)		Testing time (s) (one image)		Accuracy (%)		Training time (s) (all images)		Testing time (s) (one image)		Accuracy (%)	
	mean	std dev	mean	std dev	mean	std dev	mean	std dev	mean	std dev	mean	std dev
BOV_SIFT_75	0.0293	0.0170	0.000054	0.000010	69.8517	2.3508	0.1066	0.0161	0.000347	0.000077	38.7097	5.2134
BOV_ASIFT_500	0.0668	0.0233	0.000199	0.000019	80.9683	2.6405	0.3181	0.0288	0.001290	0.000445	66.6667	4.1550
FV_SIFT_5	0.0490	0.0038	0.000225	0.000046	76.1766	2.3766	0.2979	0.0022	0.009501	0.000251	50.0000	3.2614
FV_ASIFT_5	0.0428	0.0089	0.000180	0.000056	76.3286	3.1973	0.2315	0.0619	0.000788	0.000203	59.3750	8.7565

Table 2: Runtime in *seconds* for the training and test phases for classifiers OPF and SVM with correspondent accuracy in % using the reduced Tatt-C dataset.

about 30 images (different for each fold) for test. After that, the simulation was repeated using now 25% of the background images (total of 1425 images) in the fold.

The use of a gradation in the number of background images corresponds to the fact that in a real world situation, non-tattoos exclusion algorithms can be applied before the tattoo identification process (with different degrees of failure, here represented by the notation 100% and 25% aggregated to the name of the algorithm).

It is important to note that in these two scenarios - amount of 100% and 25% of background images - we have reduced the amount of relevant images for training for BOV and FV. In the scenario of 100% the amount of relevant images was 0.9% (40/4674) and in the scenario 25% was 2.8% (40/1425). In the reduced Tatt-C dataset (first set of simulations) the ratio of relevant training image was 11.7% (40/342) in all cases.

Table 3 and Figure 6 presents runtime and accuracy values for SIFT, BOV_SIFT_75 and FV_SIFT_05 in both scenarios (100% and 25%). Table 3, as expected, shows that the use of the complete Tatt-C dataset has a significant impact on the performance of methods such as BOV and FV on the training phase (compare Tables 1 and 3). Figure 6 shows the effect in accuracy of the use of the background images. For FV_SIFT_5 (25%) and BOV_SIFT_75 (25%) the CMC curves were worse than in the case of the reduced dataset, for the case 100% of background the numbers were even worse (considering Figure 3 for BOV and Figure 5 for FV compared with Figure 6).

Next, we applied the above algorithms - BOV_SIFT_75, FV_SIFT_5 and SIFT for 100% and 25% of background with SVM and OPF classifiers; Table 4 presents runtime and accuracy results. The OPF classifier presents again better results in comparison to the SVM classifier in terms of runtime and accuracy. Considering accuracy and comparing reduced and complete dataset, see also Table 2, for BOV_SIFT_75 values were about 9% (100% background) and 3% (25% background) worse and for FV_SIFT_5 they were 10% and 6% worse.

Table 5 presents values of Rank 1 extracted from the values used for the construction of the CMCs. The best value for Rank 1 (72%) was obtained with the algorithm

OPF + FV_SIFT_5(25%), according to line 10 of the table.

The next section concludes this article by summarizing the results obtained and pointing to the use of variations of the Tatt-C dataset as a tool for selection of candidate algorithms before exhaustive simulations.

Method	training (all images)		test (one image)	
	mean	Std Dev	mean	Std Dev
SIFT (100%)	2848 *		331.32	27.34
SIFT (25%)	790 *		103.50	8.67
BOV_SIFT_75 (100%)	1632.59	2.43	0.16	0.02
BOV_SIFT_75 (25%)	321.95	3.93	0.18	0.01
FV_SIFT_5 (100%)	2381.33	67.37	0.39	0.04
FV_SIFT_5 (25%)	920.41	62.17	0.44	0.03

* single execution.

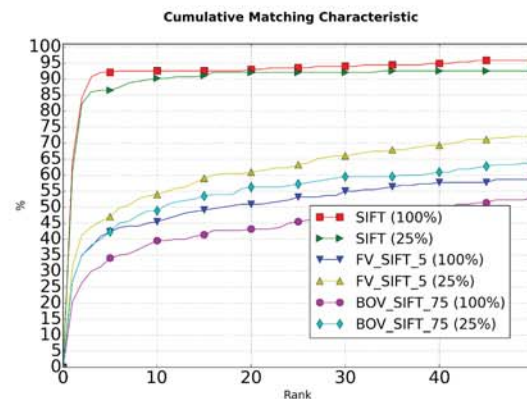Table 3: Runtime (in seconds) for the training and test phases.

Figure 6: Cumulative Matching Characteristic considering 25% and 100% of background.

4 CONCLUSION AND NEXT STEPS

In this work, we studied the use of SIFT, ASIFT, BOV and FV algorithms and the use of BOV and FV with the OPF and SVM classifiers applied to tattos identification. SIFT and ASIFT provided base values as well as the descriptors used in the BOV and FV methods.

To optimize the simulation process, we defined a strategy based on the application of an initial set of candidate algorithms to a reduced configuration (without the background images) of the Tatt-C dataset. Then, based on cost (time) / benefit (accuracy) analysis we identified

Method	OPF						SVM					
	Training time (s) (all images)		Testing time (s) (one image)		Accuracy		Training time (s) (all images)		Testing time (s) (one image)		Accuracy	
	mean	std dev	mean	std dev	mean	std dev	mean	std dev	mean	std dev	mean	std dev
BOV_SIFT_75 (25%)	0.1542	0.0147	0.000211	0.000040	67.8313	2.7196	2.9266	0.0748	0.018228	0.000939	38.7097	7.9573
BOV_SIFT_75 (100%)	1.3295	0.0196	0.000648	0.000051	64.1376	1.9630	34.8473	1.4027	0.313152	0.006051	32.2581	6.3311
FV_SIFT_5 (25%)	0.6112	0.0236	0.000904	0.000087	71.6280	2.1479	7.3247	0.2340	0.021415	0.001260	43.7500	5.5115
FV_SIFT_5 (100%)	6.7016	0.0208	0.003131	0.000104	68.6940	2.5677	81.6082	1.4602	0.376530	0.015259	40.0000	5.4561

Table 4: Runtime in *seconds* for the training and test phases for classifiers OPF and SVM with correspondent accuracy in % using the complete Tatt-C dataset.

	Method	Rank 1 (in %)
1	SIFT (100 %)	65
2	SIFT (25 %)	61
3	BOV _SIFT _75 (100 %)	20
4	BOV _SIFT _75 (25 %)	27
5	FV _SIFT _5 (100 %)	27
6	FV _SIFT _5 (25 %)	32
7	OPF + BOV _SIFT _75 (100 %)	64
8	OPF + BOV _SIFT _75 (25 %)	68
9	OPF + FV _SIFT _5 (100 %)	69
10	OPF + FV _SIFT _5 (25 %)	72
11	SVM + BOV _SIFT _75 (100 %)	32
12	SVM + BOV _SIFT _75 (25 %)	39
13	SVM + FV _SIFT _5 (100 %)	40
14	SVM + FV _SIFT _5 (25 %)	44

Table 5: Rank 1 list for the used algorithms.

the more suitable algorithms to be used in the simulations with the complete Tatt-C dataset.

As pointed in Section 3 we selected the SIFT (baseline), BOV_SIFT_75 and FV_SIFT_5 to be simulated with OPF and SVM algorithms using two Tatt-C dataset configurations, 100% and 25% of the background images. As already highlighted in Section 3, the idea of using two different amounts of background images was to simulate the quality of the image dataset (number of true tattoos).

The Rank 1 information provided by Table 5 in conjunction with the information provided by Table 3 and Table 4 allowed the completion of the experiment.

The training time column provided the runtime for the training phase of the different methods. So, for example to create the set of information used by OPF + FV_SIFT_5 (25%) we had to process the SIFT (runtime = 790s - Table 3), then the FV_SIFT_5 (25%) (runtime = 920s - Table 3) and then the OPF_FV_SIFT_5 (25%) (runtime = 0.61s - Table 4) totalizing 1710.6 seconds. The same goes for the other cases.

The test time columns are calculated in the same way. For example, for OPF + FV_SIFT_5 (25%), runtime is taken from Table 4, line 3 (0.000904s), while Table 5 points to this algorithm as the best for Rank 1 choice, presenting 72% accuracy.

Comparing (see Tables 3 and 4) the values obtained with the SIFT algorithm (baseline) and with SIFT combined with BOV or FV (without OPF) we observed a longer execution time and less precision compared to the same algorithms used in conjunction with OPF. In these cases, we obtain precision values that exceed the values obtained with SIFT (baseline). In the simulation results of Rank 1 presented in Table 5 using the Tatt-C data set with 100% of background images we obtained accuracy of 69% versus 65% (lines 9 and 1 respectively) and with a quantity of 25% of background images, accuracy of 72% versus 61% (lines 10 and 2 respectively).

In conclusion, it is important to bear in mind that the presented results were obtained with a relatively small image base and few cases of use per image. Thus, although we presented time values, for the test phase, by image, factors such as a considerable increase in the amount of images in the base will imply in the need to use memory management certainly impacting the access time of the images. On the other hand factors such as preprocessing in the images and optimization in the libraries, elements not used here, can contribute to the minimization of the times.

As a continuation of this work we see some scenarios. For example, the use of Deep Learning in the area of tattoos identification; the current literature presents studies in other use-cases - detection, similarity and de-identification [DP16], [HBRM16], [XGXHK16] - possibly due characteristics of current image databases (for example the NIST Tatt-C dataset presents a reduced number of tattoos for detection-training). A scenario we are already considering, taking advantage of a previous work [DSFM11] and also results related by Jain et al. [JLJ07], intend to couple the use of feedback by relevance to the identification process of tattoos leading the user to the decision loop.

5 ACKNOWLEDGMENTS

6 REFERENCES

[CDF04] Gabriella Csurka, Christopher R. Dance, Lixin Fan, Jutta Willamowski, and Cédric Bray. Visual categorization with bags of keypoints. In Workshop on Statistical Learning in Computer Vision, ECCV, pages 1–22, 2004.

[CL11] Chih-Chung Chang and Chih-Jen Lin. Lib-svm: A library for support vector machines. ACM Trans. Intell. Syst.Technol., 2(3):27:1–27:27, May 2011.

[DP16] Xing Di and Vishal M. Patel. Deep Tattoo Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) , Pages: 119 – 126, Year: 2016

[DSFM11] André Tavares da Silva, Alexandre Xavier Falcão, and Léo Pini Magalhães. Active learning paradigms for CBIR systems based on optimum path forest classification. Pattern Recognition, 44(12):2971 – 2978, 2011.

[FCH08] Rong-En Fan, Kai-Wei Chang, Cho-Jui Hsieh, Xiang-Rui Wang, and Chih-Jen Lin. Liblinear: A library for large linear classification. J. Mach. Learn. Res., 9:1871–1874, June 2008.

[HBRM16] Tomislav Hrkac, Karla Brkic, Slobodan Ribaric, Darijan Marcetc. Deep Lerning Architectures for Tattoo Detection and De-identification. 1St Int. Conf on Sensing, Processing and Learning for Intelligent Machines – SPLINE, July 6-8, 2016, Aalborg Univ, Denmark (pp. 45-49)

[HJ13] H. Han and A. K. Jain. Tattoo based identification: Sketch to image matching. In 2013 International Conference on Biometrics (ICB), pages 1–8, June 2013.

[JH99] Tommi S. Jaakkola and David Haussler. Exploiting generative models in discriminative classifiers. In Proceedings of the 1998 Conference on Advances in Neural Information Processing Systems II, pages 487–493, Cambridge, MA, USA, 1999. MIT Press.

[JLJ07] Anil K. Jain, Jung-Eun Lee, and Rong Jin. Tattoo-ID: Automatic Tattoo Image Retrieval for Suspect and Victim Identification, pages 256–265. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007.

[LJJ08] J.-E. Lee, A. K. Jain, and R. Jin, "Scars, marks and tattoos (SMT): Soft biometric for suspect and victim identification," Proceedings of the Biometrics Symposium, pp. 1–8, September 2008, Tampa, Florida, USA.

[LJJT12] J.-E. Lee, R. Jin, A. K. Jain, and W. Tong, "Image retrieval in forensics: Tattoo image database application," IEEE MultiMedia, vol. 19, no. 1, pp. 40–49, 2012.

[LOW99] D. G. Lowe. Object recognition from local scale-invariant features. In Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on, volume 2, pages 1150–1157 vol.2, 1999.

[Man12] D. Manger. Large-scale tattoo image retrieval. In Computer and Robot Vision (CRV), 2012 Ninth Conference on, pages 454–459, May 2012.

[MHW16] D. Manger, C. Herrmann, and D. Willersinn. Towards extending bag-of-words-models using context features for an 2d inverted index. In 2016 International Conference on Digital Image Computing: Techniques and Applications (DICTA), pages 1–5, Nov 2016.

[MRSP14] D. Marcetic, S. Ribaric, V. Struc, and N. Pavesic. An experimental tattoo de-identification system for privacy protection in still images. In 2014 37th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), pages 1288–1293, May 2014.

[NG15] M. Ngan and P. Grother. Tattoo recognition technology - challenge (tatt-c): an open tattoo database for developing tattoo recognition research. In Identity, Security and Behavior Analysis (ISBA), 2015 IEEE International Conference on, pages 1–6, March 2015.

[NQG16] Mei Ngan, George W. Quinn, and Patrick Grother. Tattoo recognition technology – challenge (tatt-c) outcomes and recommendations. revision 1.0. technical report nist 8078. Technical report, National Institute of Standards and Technology, sep 2016.

[PD07] F. Perronnin and C. Dance. Fisher kernels on visual vocabularies for image categorization. In 2007 IEEE Conference on Computer Vision and Pattern Recognition, pages 1–8, June 2007.

[PFS09] J. P. Papa, A. X. Falcão, and C. T. N. Suzuki. Supervised pattern classification based on optimum-path forest. Int. J. Imaging Syst. Technol., 19(2):120–131, June 2009.

[ST14] Rosália G. Schneider and Tinne Tuytelaars. Sketch classification and classification-driven analysis using fisher vectors. ACM Trans. Graph., 33(6):174:1–174:9, nov 2014.

[SZ03] J. Sivic and A. Zisserman. Video google: a text retrieval approach to object matching in videos. In Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on, pages 1470–1477 vol.2, Oct 2003.

[XGXHK16] Qingyong Xu, Soham Ghosh, Xingpeng Xu, Yi Huang, Adams Wai Kin Kong. Tattoo detection based on CNN and remarks on the NIST Database. 2016 International Conference on Biometrics (ICB) , Pages: 1 – 7, 2016

[YM11] Guoshen Yu and Jean-Michel Morel. ASIFT: An Algorithm for Fully Affine Invariant Comparison. Image Processing On Line, 1, 2011.