Statistical Analysis of Binarized SIFT Descriptors

M. Diephuis, S. Voloshynovskiy, O. Koval and F. Beekhof
Stochastic Information Processing Group
Université de Genève
1227 Carouge, Switzerland
Email: {Maurits.Diephuis, svolos*, Oleksiy.Koval, Fokko.Beekhof}@unige.ch

***Abstract*—SIFT descriptors are broadly used in various emerging applications. In recent years, these descriptors were deployed in compressed and binarized forms due to the computational complexity, storage, security and privacy cost incurred by working on real data. At the same time, the theoretical analysis of SIFT feature performance in different applications remains an open issue due to the lack of accurate statistics of binarized SIFT descriptors. We address this problem and statistically analyse projected binarized SIFT descriptors in this paper. The methodology is based on dimensionality reduction using random projections with binarization. Furthermore, we investigate the statistical models of intra- and inter-descriptor dependencies for various distortions. Finally, we demonstrate a simple heuristic to distinguish between descriptors from identical but distorted images and descriptors from non identical images.**

I. Introduction

With the rise of massive centralized multi-media repositories such as Flickr and Youtube comes an equally large demand for services such as copyright verification, automatic tagging, image-content identification and authentication. At the same time, the available computational storage and cost impose significant limits on the rendered services, given the shear size of the repositories. Additionally, search issues require efficient means for accurate comparisons of large scale multi-media repositories even when the data is degraded via various signal processing operations or geometrical transforms.

To address these challenging problems, the computer vision community developed a number of global and local descriptors which allow for matching distorted but identical images. In recent years, many researches have sought efficient matching strategies based on binarized SIFT descriptors, leveraging recent advances in fast hardware computation of Hamming distances [1]–[3]. Work from [4], [5] proposes to use coarsely quantized random projections to build a descriptor hash. An information-theoretic analysis of content identification based on similarly constructed hashes is performed in [6]. At the same time, the current state-of-the-art does not include a statistical analysis of the robustness, invariance, and information content of binary fingerprints extracted from randomly projected SIFT descriptors that define their discriminative power. The modelling of possible attacks is also an open issue.

Therefore, the main goal of this work is to provide an analysis of the statistical properties, and thus the discriminative power, of binarized SIFT descriptors. This work will primarily investigate the SIFT fingerprint statistics on real image datasets, and the statistics of the fingerprints under various distortions. Contrary to the state-of-the-art, we consider the relationship between fingerprints extracted from inter- and intra-image SIFT descriptors. We demonstrate that the binarized SIFT descriptors are characterized by a certain level of similarity, even for perceptually dissimilar images due to particularities, inherent to the SIFT design and the impact this phenomenon has on their discriminative power.

This work is organized as follows. Section 2 introduces the problems and issues that will be examined. It also presents the modelling framework that is deployed for our experiments, most notably the random projections and binarization procedure. Section 3 describes all statistical tests, the experimental results and analysis. Finally, the conclusions are presented in Section 4.

Notation Scalar random variables are denoted by capital letters X, and bold capitals $\mathbf{X}$ denote vector random variables. Corresponding small letters x and $\mathbf{x}$ denote their respective realizations, where $\mathbf{x} = \{x[1], x[2], \ldots, x[N]\}$. The binarized version of $\mathbf{x}$ is denoted by $\mathbf{b}_\mathbf{x}$. The projected SIFT descriptor k from image m is described by $\tilde{\mathbf{x}}^k(m)$, its individual elements are $\tilde{x}^k(m)[i]$. The binarized projected SIFT descriptor k from image m then becomes $\mathbf{b}^k_{\mathbf{x}(m)}$. Individual elements of a binarized projected SIFT descriptor are expressed by $b^k_{\mathbf{x}(m)}[i]$. $X \sim p(x)$ denotes that random variable follows $p_X(x)$. $\mathcal{N}$, $\mathcal{B}$ indicate the Gaussian and Binomial distributions. Finally, the dimensionality reducing transform matrix is designated by $\mathbf{W}^{L\times N}$ where L and N are numbers of its rows and columns.

II. Problem formulation and Methodology

This work will perform a statistical analysis of SIFT based image descriptors as a fundamental unit for image identification and retrieval systems [2], [6]. First, we model the statistics of the SIFT descriptors in order to justify their discriminative power. Secondly, we analyze SIFT descriptor resilience to various geometrical and signal processing distortions that are commonly found in image identification and retrieval. We then compare the performance of a commonly used matching heuristic in the original domain to the binarized projected domain. Finally, we analyze the matching of descriptors extracted from identical but distorted images versus the matching between descriptors originating from different images.

The data originates from the Airplane subset of the Caltech 101 dataset [7] from which about 36000 SIFT points were generated using the VLFeat toolbox [8]. In order to satisfy complexity and storage requirements, the original raw SIFT

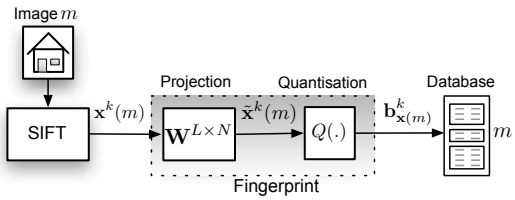


Figure 1: Binarized SIFT generating framework.

descriptors are converted to the binary domain. This transformation is done with random projections and sign based binarization.

Random Projection and Binarization

The binary fingerprints are obtained via a two staged process: a dimensionality reduction $\mathbf{W}^{L\times N}$ and binarization $Q(.)$ [9] shown in Figure 1.

The dimensionality reduction of the k-th SIFT descriptor from image m, $\mathbf{x}^k(m)$, is done as follows:

$$\tilde{\mathbf{x}}^k(m) = \mathbf{W}^{L\times N}\mathbf{x}^k(m). \quad (1)$$

where $\mathbf{W}^{L\times N} \in \mathcal{R}^{L\times N}$. L is the number of dimensions $\mathbf{W}^{L\times N}$ will map to, N is the length of the input column vector, which for SIFT vectors is 128. Random matrix $\mathbf{W}^{L\times N} = (\mathbf{W}_1, \mathbf{W}_2, \ldots, \mathbf{W}_N)^T$ consists of a set of approximately orthonormal basis vectors, where all elements are generated as $W_i[j] \sim \mathcal{N}(0, \frac{1}{N})$, $1 \leq i \leq N, 1 \leq j \leq L$, and as such behaves as an approximate *orthoprojector*. It guarantees that the projected vectors $\tilde{\mathbf{x}}^k(m)$ could be considered as realisations of a Gaussian source with a covariance matrix that converges to diagonal in probability [10]. Also, as stated in the introduction, one can reasonably expect raw SIFT descriptors to be correlated amongst each other. It is proven that random projections decorrelate data vectors amongst each other under certain assumptions [6]. Furthermore, such a dimensionality reduction is extremely fast and can be applied over time to an ever changing dataset. Finally, binarization is simply done by extracting and storing the sign of all individual elements of all projected SIFT vectors:

$$\begin{aligned}\mathbf{b}_{\mathbf{x}^k(m)} =& \{sign(\tilde{x}^k(m)[1]), sign(\tilde{x}^k(m)[2]),\\ & \ldots, sign(\tilde{x}^k(m)[L])\}, \end{aligned} \quad (2)$$

where for $i \in \{1 \ldots L\}$, $b^k_{\mathbf{x}(m)}[i] \in \{0, 1\}$ and $\forall a, sign(a) = 1$ if $a \geq 0$ and 0 otherwise. The goal of using such a threshold is justified by maximizing the entropy of the extracted binary data which should converge to 1 bit per sample. The binarized projected descriptors are then stored in the database in the bin corresponding to image index m.

Since our dimensionality reduction and binarization are inherently information lossy, a three-fold analysis of the binarized projected descriptors and the original raw descriptors will be performed in order to establish the trade-off between complexity and storage versus the performance limits.

To accurately model channel distortions for real and projected descriptors, one needs to annotate the ground truth, i.e., one must assure that the matching of descriptors between two identical but distorted images is correct. We distinguish the following situation: if a match between point correspondences $\{\mathbf{x}^k(m) \leftrightarrow \mathbf{y}^k(m)\}$ is correct, the distance between the matching descriptors $\mathbf{x}^k(m)$ and $\mathbf{y}^k(m)$ is kept in set. As no matching metric is flawless, and a manual annotation is not feasible, the following procedure is followed. For all points, we check, using a weak geometry based rule, if the matching point $\mathbf{y}^k(m)$ is in the expected region. If not, this distance between $\{\mathbf{x}^k(m) \leftrightarrow \mathbf{y}^k(m)\}$ is removed from the set. The same rule is also used to reassign matches. This method of assessment is possible as we know the channel distortions that will be applied in the tests. Note that this is a different procedure then the matching rule, as proposed by [11].

III. Performance Analysis

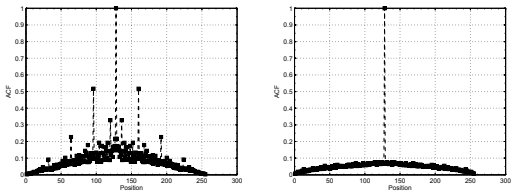
Descriptor statistics

In order to justify the discriminative power of both the original and binarized projected descriptors, we used the joint empirical entropy [1]. Estimating such a function of a joint distribution is extremely complex in a high dimensional space and was thus performed on the assumption that descriptor elements from a single originating vector are independent. This approach, while lacking a certain accuracy, provides an upper bound on the sought discriminative power according to the chain rule for entropy [1].

In order to evaluate the discriminative power, the 36000 raw SIFT descriptors were processed following the above advocated approach.

Transformation of the original SIFT descriptors into the binary domain according to (1) and (2) converges to i.i.d. binary strings, which, with high probability, exhibit $p(1) = p(0) = 0.5$. To demonstrate this, the auto-correlation function from the binarized projected fingerprints was calculated (Figure 2), which shows that there is asymptotically no correlation between bits of individual binarized projected SIFT descriptors contrarily to the binarization applied in the original domain. Under this assumption, one would be able to distinguish 2^{128} different images, if all fingerprints were independent. However, our experiments show that there is a certain amount of correlation between binarized projected descriptors (Figure 2b). The measured Hamming distance between all binarized projected descriptors from our dataset can be modelled with the Bernoulli distribution $\mathcal{B}(L, 0.36)$. Had they been independent, this would have been $\mathcal{B}(L, 0.5)$.

Additionally, in realistic image identification and content based image retrieval (CBIR) scenarios various types of distortions are present that reduce the discriminative power of such descriptors. The next Section is dedicated to analysis of the impact various channel distortions have on the discriminative performance of the binarized projected SIFT descriptors.



(a) (b)

Figure 2: The average correlation per image between descriptors before and after projection and binarization. The average correlation between original SIFT descriptors per image is shown via the normalized Auto Correlation Function in 2a. After projection and binarization this can be seen in 2b.

Channel Distortion Statistics

In order to examine how binarized projected descriptors behave under various image distortions and how their discriminative power is influenced, a digital communications approach has been deployed. Assuming that the influence of various distortions in the original raw SIFT domain will be reflected in bit flips in the binarized projected domain, the probability of bit error P_b [1] will be used to measure performance (Figure 3b). Then, the resulting discriminative power of the binarized projected SIFT descriptors for a particular channel distortion will be evaluated as $2^{N(1-H_2(P_b))} = 2^{128(1-H_2(P_b))}$, where $H_2(P_b) = -P_b \log_2 P_b - (1-P_b)\log_2(1-P_b)$ denotes the binary entropy [1]. Evidently, if $P_b = 0.5$, $H_2(P_b) = 1$ and $2^{128(1-H_2(P_b))} = 1$. Therefore, the discriminative power is determined by the distortions present in the identification or CBIR protocol. In the case the number of images $M > 2^{N(1-H_2(P_b)}$, one needs to enlarge N, i.e. increase the number of used SIFT descriptors.

Images in our dataset were exposed to the following distortions (Figure 3b, Table I). They were rotated with parameter θ in degrees and scaled with parameters $s = s_x = s_y$. The Similarity transformation was deployed with parameters s, θ and $t_x, t_y = 1$. Additive White Gaussian Noise was added to images with variance σ^2. And finally, images where subjected to JPEG compression with *quality factor* q.

The channel distortion is estimated by taking each individual image and its SIFT descriptors, making a distorted image copy and comparing the descriptors from both using the annotated ground truth as detailed in Section II. Two class distances are measured, the distance between matching descriptors, or intra-class distances, and the distance between non matching descriptors, or inter-class distances (Figure 3b, 3c).

Our experimental results (Figure 4) demonstrate that both original and binarized projected SIFT descriptors have a reasonable resilience against these distortions. The maximal $P_b = 0.08$ was induced by the similarity transform with $s = 0.2, \theta = 10$. Therefore, discriminative power in this identification protocol is bounded by $2^{128 \cdot 0.60}$ images. Note that this bound is obtained under the independence assumption, and can therefore not be considered tight due to the significant correlation between binarized projected SIFT descriptors (Figure 3a).

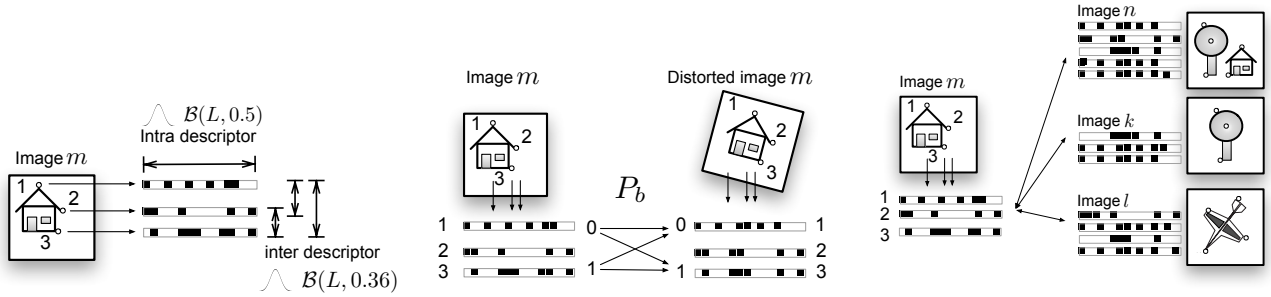
The second and third column in Figure 4 show the attained intra- en inter-class distances for these channel distortions and parameter that resulted in the largest probability of P_b. As the class distances overlap, it shows that perfect identification is not possible, neither in the original domain, nor in the binarized projected domain for these distortions.

Matching Heuristics

It is a common practice, to enhance the matching of descriptors, to discard matches that do not fulfill some kind of heuristic requirement. The commonly used method, proposed by [11], stipulates that a match between descriptors is only considered if the second best match is at least a certain percentage worse, usually 60 percent (Algorithm 1). To ascertain how such a rule behaves in the binarized projected domain, we executed the following experiment. For all images from our dataset, and all distortions as outlined in Section III we matched an image against its distorted counterpart using the non ambiguous matching rule [11] and annotated all attained matches between descriptors. We then ran the same test, based on the binarized projected descriptors while stipulating that a match is only valid if its Hamming distance is 60 percent smaller than the second best match. On comparing the attained point matches based on the original SIFT descriptors and binarized projected descriptors we found that the latter attained around 90 percent of the matches that were attained based on the original descriptor. All results can be seen in Table I.

Algorithm 1: Heuristically match binarized projected SIFT points [12].

Data: K Binarized SIFT descriptors $\mathbf{b}^k_{\mathbf{x}(m)}, k \in \{1, 2, \ldots, K\}$
Data: L Binarized SIFT descriptors $\mathbf{b}^l_{\mathbf{y}(m)}, l \in \{1, 2, \ldots, L\}$
Result: Vector $match$ with attained matches where $match(i) = j$ denotes a valid match between $\mathbf{b}^i_{\mathbf{x}(m)}$ and $\mathbf{b}^j_{\mathbf{y}(m)}$.
for $i \leftarrow 1$ **to** K **do**
  for $j \leftarrow 1$ **to** L **do**
    $distance(i) \longleftarrow$ Hamming($\mathbf{b}^i_{\mathbf{x}(m)}$, $\mathbf{b}^j_{\mathbf{y}(m)}$)
  end
  index is an array with the original index of the data prior to sorting
  $(distance(), index()) \longleftarrow$ sortAscend($distance()$)
  if $distance(1) \geq 0.6 * distance(2)$ **then**
    $match(i) \longleftarrow index(1)$
    else
      $match(i) \longleftarrow -1$
    end
  end
end



(a) (b) (c)

Figure 3: Schematic frameworks for inter and intra binarized projected descriptor Hamming distances (Figure 3a), determination of the probability of bit error P_b for channel distortions (Figure 3b) and inter image matching (Figure 3c).

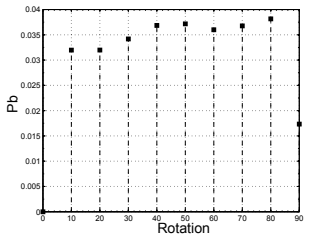
Similarity Transform, parameters s, θ (degrees).									
Distortion parameters:	(0.2, 10)	(0.4, 20)	(0.6, 30)	(0.8, 40)	(1, 10)	(1.2, 10)	(1.4, 20)	(1.6, 30)	(1.8, 40)
Relative heuristic matches between real and binarized descriptors:	**33 %**	85%	86%	90%	100%	93%	92%	93%	93%
Scale transform, parameter s.									
Distortion parameters:	0.25	0.5	0.75	1	1.25	1.5	2		
Relative heuristic matches between real and binarized descriptors:	**87%**	92%	90%	100%	93%	93%	96%		
JPEG, parameter is *quality factor* q.									
Distortion parameters:	10	20	30	30	50	60	70	80	90
Relative heuristic matches between real and binarized descriptors:	**95%**	96%	97%	98%	99%	99%	99%	99%	100%
Rotation transform, parameter θ (degrees).									
Distortion parameters:	80	70	60	50	40	30	20	10	0
Relative heuristic matches between real and binarized descriptors:	**89%**	90%	92%	90%	92%	92%	93%	94%	100%
AWGN, parameter σ.									
Distortion parameters:	0.1	0.08	0.06	0.04	0.02	0			
Relative heuristic matches between real and binarized descriptors:	**90%**	93%	92%	92%	92%	100%			

Table I: Impact of signal processing and geometrical distortions on fingerprint bits matching. This Table shows the used channel distortions and their parameters. Furthermore it presents the percentage of attained matches using the matching heuristic from Section III between binarized projected descriptors that are identical to the matches attained by the original SIFT descriptors for various distortions. The latter attained matches are treated as the baseline and set to a 100 percent.

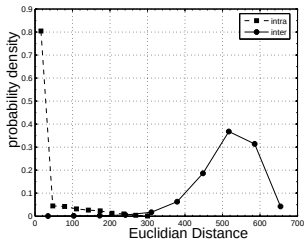
Inter Image Statistics

Finally, to test how binary projected descriptors behave between non identical images, descriptors from the 800 images from the Airplane category of the Caltech 101 dataset were again extracted, projected and binarized. The deployed matching rule is the heuristic Hamming distance rule as mentioned in Section III (Figure 3c). The result can be seen in Figure 5a. In this graph the Hamming distances between identical but distorted images are shown. Again, all distortions from Section III, Table I were tested. The inter-class distances are all Hamming distances between matching descriptors originating from different and undistorted images. It is beyond doubt that it is not possible to fully distinguish the distorted, yet identical images from a different image within the tested set based on this heuristic and these descriptors. Finally, one should note that the bounds in Figure 5a are optimistic as the Airplane set does contain a number of identical airplanes.

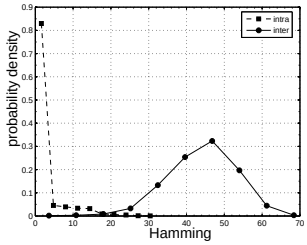
This means that to identify the fact that two sets of descriptors were extracted from non identical images, one needs to add side information. A simple heuristic, as outlined in Algoritm 2, uses the image coordinates around where the descriptors were extracted. These coordinates are sampled, and the samples are checked for weak geometric similarity. If the similarity transform needed to project points to each other deviates considerably, the heuristic drams the originating images as non identical and sets the Hamming distance between all descriptors to $L/2$ where L is the dimensionality of the binarized projected descriptor. The results of this algorithm on the tested dataset are shown in Figure 5b. The figure contains a small tail, which is caused by the earlier mentioned phenomena that the Airplane dataset contains identical airplanes and it contains airplanes that are just visually similar to the SIFT algorithm.



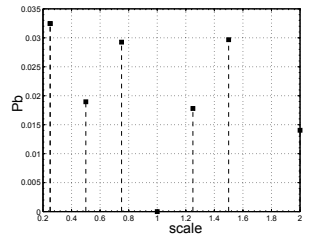
(a) P_b for rotation distortion.



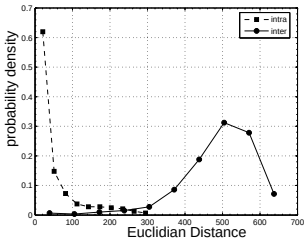
(b) Euclidean distance between $\mathbf{x}$ and $\mathbf{y}$ for $\theta = 50$ degrees rotation.



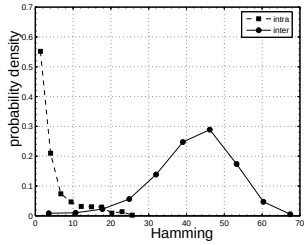
(c) Hamming distance between $\mathbf{b_x}$ and $\mathbf{b_y}$ for $\theta = 50$ degrees rotation.



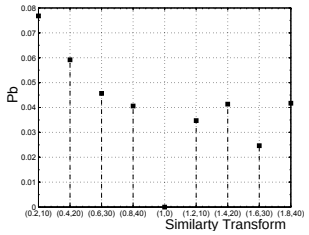
(d) P_b for scale distortion.



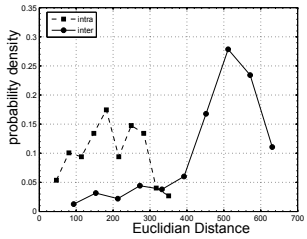
(e) Euclidean distance between $\mathbf{x}$ and $\mathbf{y}$ for scale distortion of $s = 0.2$.



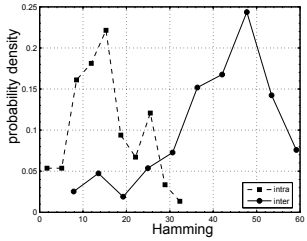
(f) Hamming distance between $\mathbf{b_x}$ and $\mathbf{b_y}$ for scale distortion of $s = 0.2$.



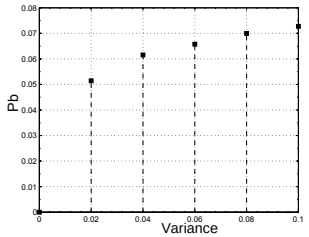
(g) P_b for the similarity transform distortion.



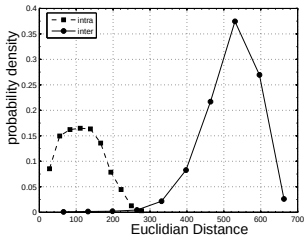
(h) Euclidean distance between $\mathbf{x}$ and $\mathbf{y}$ for the similarity transform distortion with $(s = 0.2, \theta = 10)$.



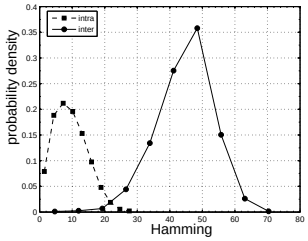
(i) Hamming distance between $\mathbf{b_x}$ and $\mathbf{b_y}$ for the similarity transform distortion with $(s = 0.2, \theta = 10)$.



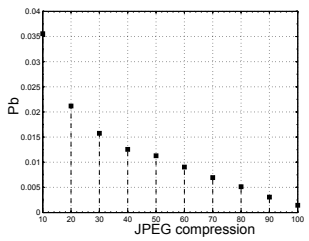
(j) P_b for AWGN distortion.



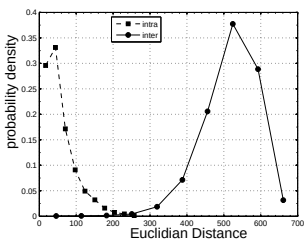
(k) Euclidean distance between $\mathbf{x}$ and $\mathbf{y}$ for AWGN with $\sigma^2 = 0.1$.



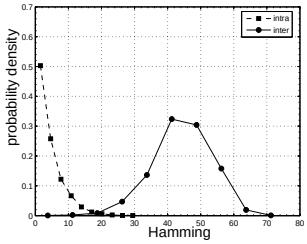
(l) Hamming distance between $\mathbf{b_x}$ and $\mathbf{b_y}$ for AWGN with $\sigma^2 = 0.1$.



(m) P_b for JPEG compression distortion.

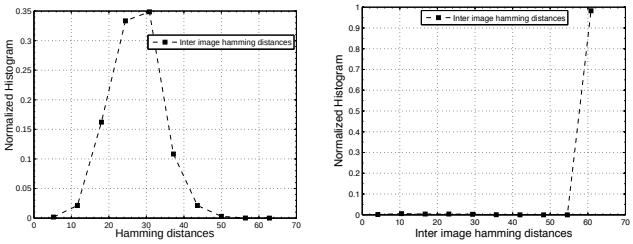


(n) Euclidean distance between $\mathbf{x}$ and $\mathbf{y}$ for JPEG compression distortion with $q = 10$.



(o) Hamming distance between $\mathbf{b_x}$ and $\mathbf{b_y}$ for JPEG compression distortion $q = 10$.

Figure 4: Channel statistics for various distortions for a set of original and projected SIFT descriptors. $\mathbf{x}$ and $\mathbf{y}$ denote the codebook of original and the distorted sift descriptors. The projected counterparts are denoted by: $\mathbf{b_x}$ and $\mathbf{b_y}$. Descriptors from identical indices are matching descriptors.



(a) (b)

Figure 5: 5a shows the Hamming distances between the matching binarized projected descriptors of unequal undistorted images, 5b shows the effect of deploying Algorithm 2.

Algorithm 2: Filtering out descriptors by checking weak geometric consistency.

Data: K Binarized Projected SIFT descriptors $\mathbf{b}^k_{\mathbf{x}(m)}, k \in \{1, 2, \ldots, K\}$
Data: Corresponding descriptor row column coordinates $(r^k_{\mathbf{x}(m)}, c^k_{\mathbf{x}(m)})$.
Data: L Binarized Projected SIFT descriptors $\mathbf{b}^l_{\mathbf{y}(m)}, l \in \{1, 2, \ldots, L\}$.
Data: Corresponding descriptor row column coordinates $(r^l_{\mathbf{x}(m)}, c^l_{\mathbf{y}(m)})$.
Result: Hamming distances $distance()$
for $i \leftarrow 1$ **to** K **do**
  for $j \leftarrow 1$ **to** L **do**
    Sample 6 point pairs $(\mathbf{a}, \mathbf{b})$ randomly from $(r^k_{\mathbf{x}(m)}, c^k_{\mathbf{x}(m)})$ and $(r^l_{\mathbf{x}(m)}, c^l_{\mathbf{y}(m)})$
    for *all combinations* $\binom{3}{6}$ **do**
      $(\textbf{scale}, \textbf{theta}) \longleftarrow$ similarityTransform$(\mathbf{a}, \mathbf{b})$
    end
    if $std(\textbf{\textit{scale}}) > 0.1 || std(\textbf{\textit{theta}}) > 0.1$ **then**
      $distance(i, j) \longleftarrow L/2$
    end
  end
end

IV. Conclusions

In this paper we examined the statistical properties of original and random binarized projected SIFT descriptors. Dimensionality reduction is achieved by deploying random projections followed by sign based binarization. We examine the basic statistical properties of such binarized projected descriptors and show that although not perfect, the random projections are a fast and sufficient method for decorrelation. Furthermore, the compressed binarized descriptors show similar resilience to a range of channel distortions in comparison to the original SIFT descriptors. We demonstrate that implementing a commonly used matching metric for raw SIFT descriptors in the projected binarized domain is possible without a significant loss in performance. Finally, images from the used image repository can not be identified errorless based on the binarized projected descriptors alone. This is inherent to their design which favors robustness. All Matlab code is available on: `http://sip.unige.ch/research/binarized-sift`

V. Acknowledgment

This work is partly funded by SNF-grant 20021-132337.

References

[1] T. M. Cover and J. A. Thomas, *Elements of information theory*. New York, NY, USA: Wiley-Interscience, 1991.

[2] H. Jégou, M. Douze, and C. Schmid, "Hamming embedding and weak geometric consistency for large scale image search," in *European Conference on Computer Vision*, ser. LNCS, A. Z. David Forsyth, Philip Torr, Ed., vol. I. Springer, oct 2008, pp. 304–317. [Online]. Available: http://lear.inrialpes.fr/pubs/2008/JDS08

[3] K. Mikolajczyk and C. Schmid, "Scale and affine invariant interest point detectors," *International Journal of Computer Vision*, vol. 60, no. 1, pp. 63–86, 2004. [Online]. Available: http://lear.inrialpes.fr/pubs/2004/MS04

[4] C. Yeo, P. Ahammad, and K. Ramchandran, "Coding of image feature descriptors for distributed rate-efficient visual correspondences," *International Journal of Computer Vision*, pp. 1–15, 2011, 10.1007/s11263-011-0427-1. [Online]. Available: http://dx.doi.org/10.1007/s11263-011-0427-1

[5] S. Roy and Q. Sun, "Robust hash for detecting and localizing image tampering," in *Image Processing, 2007. ICIP 2007. IEEE International Conference on*, vol. 6, 16 2007-oct. 19 2007, pp. VI –117 –VI –120.

[6] S. Voloshynovskiy, O. Koval, F. Beekhof, F. Farhadzadeh, and T. Holotyak, "Information-theoretical analysis of private content identification," in *IEEE Information Theory Workshop, ITW2010*, Dublin, Ireland, Aug.30-Sep.3 2010.

[7] L. Fei-Fei, R. Fergus, and P. Perona, "Learning Generative Visual Models from Few Training Examples: An Incremental Bayesian Approach Tested on 101 Object Categories," 2004, p. 178. [Online]. Available: http://visionlab.ece.uiuc.edu/documents/Fei-FeiFergusPerona_GMBV04.pdf

[8] A. Vedaldi and B. Fulkerson, "Vlfeat: An open and portable library of computer vision algorithms," 2008. [Online]. Available: http://www.vlfeat.org

[9] S. Voloshynovskiy, O. Koval, F. Beekhof, and T. Pun, "Random projections based item authentication," in *Proceedings of SPIE Photonics West, Electronic Imaging / Media Forensics and Security XI*, San Jose, USA, 2009. [Online]. Available: http://cvml.unige.ch/publications/postscript/2009/2009.SPIE.authentication.rp.pdf

[10] F. Farhadzadeh, S. Voloshynovskiy, O. Koval, T. Holotyak, and F. Beekhof, "Statistical analysis of digital image fingerprinting based on random projections," in *Proceedings of ISPA Image and Signal Processing and Analysis*. ISPA, 2011.

[11] D. Lowe, "Distinctive image features from scale-invariant keypoints," in *International Journal of Computer Vision*, vol. 20, 2003, pp. 91–110. [Online]. Available: citeseer.ist.psu.edu/lowe04distinctive.html

[12] M. Brown and D. Lowe, "Invariant features from interest point groups," 2002. [Online]. Available: citeseer.ist.psu.edu/brown02invariant.html