Bits of Images: Inverting Local Image Binary Descriptors

Emmanuel d'Angelo, Pierre Vandergheynst
Ecole Polytechnique Fédérale de Lausanne
Lausanne, Switzerland
Email: firstname.lastname@epfl.ch

Laurent Jacques
Université Catholique de Louvain
Louvain-la-Neuve, Belgium
Email: laurent.jacques@uclouvain.be

Alexandre Alahi
Stanford University
San Francisco, California 96678-2391
Email: alahi@stanford.edu

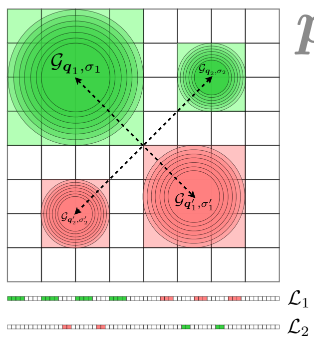


Fig. 1. Example of a local descriptor and the corresponding sensing matrix.

***Abstract*—Local Binary Descriptors (LBDs) are good at matching image parts, but what information is actually carried? This question is usually masked by a comparison of matching performances. In this work, we leverage an inverse problem approach to reconstruct the image content from LBDs. We show that this task is achievable with only the information in the descriptors, excluding the need of additional data. Furthermore, our reconstruction scheme reveals differences in the way different LBDs capture image information. The possible applications of our work are various, from privacy issues on mobile communication networks to the design of better descriptors and smart cameras.**

I. Introduction

Image keypoints are central to many image-based Pattern Recognition and Computer Vision algorithms, that rely on detecting and matching interest points. However, their qualities are usually assessed from their performance in image matching benchmarks, thus only implicitly measuring the embedded information.

We explored instead an explicit approach to evaluate the image content captured by a descriptor by reconstructing the geometry from the descriptor. We used two binarized descriptors: BRIEF [1] and FREAK [2], computed by retaining the sign of the difference between smaller cells contained in the source image patch and written $\mathcal{B}(\mathcal{L}(\boldsymbol{p}))$ for an image patch $\boldsymbol{p}$.

II. Problem statement

Unlike previous work based on SIFT keypoints [3], we use a regularized inverse problem, thus removing the dependance on a learning database. Given an observed descriptor $\bar{\boldsymbol{p}}$ and a wavelet analysis operator $\boldsymbol{W}$, we are interested by the solution of:

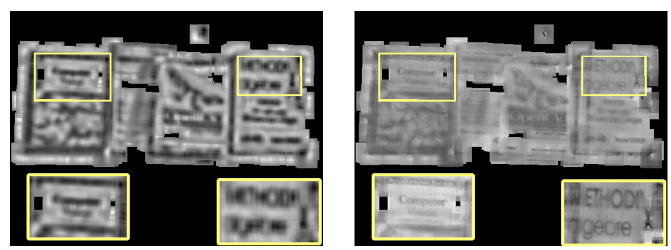
$$\hat{\boldsymbol{p}} = \operatorname*{arg\,min}_{\boldsymbol{x}\in\mathbb{R}^N} \mathcal{J}(\boldsymbol{x}) \text{ s.t. } \|\boldsymbol{W}\boldsymbol{x}\|_0 \leqslant k \text{ and } \boldsymbol{x} \in \mathcal{S}, \tag{1}$$

where the constraints enforces both the validity domain $\mathcal{S}$ of $\hat{\boldsymbol{p}}$ and its k-sparsity in the wavelet domain. Inspired by [4], our data term is

$$\mathcal{J}(\boldsymbol{x}) = \|[\bar{\boldsymbol{p}} \odot \mathcal{B}(\mathcal{L}\boldsymbol{x})]_-\|_1, \tag{2}$$

where $\boldsymbol{x} \odot \boldsymbol{y} = \sum_i x_i y_i$ is the Hadamard products of $\boldsymbol{x}$ and $\boldsymbol{y}$. Qualitatively, $\mathcal{J}$ measures the LBD consistency of $\boldsymbol{x}$ with $\boldsymbol{p}$, with $\mathcal{J}(\boldsymbol{x}) = 0$ iff $\mathcal{B}(\boldsymbol{\mathcal{L}}\boldsymbol{x}) = \bar{\boldsymbol{p}}$. $\mathcal{J}(\boldsymbol{x})$ counts the number of components where the binarized descriptors do not have the same sign. The problem (1) can then be solved with a slightly modified version of BIHT [4], where we also enforce the mean value of a patch (which defines the validity domain $\mathcal{S}$) that was lost by the operator $\mathcal{L}(.)$.

The results obtained using BRIEF or FREAK descriptors reveal the original image structure, but also emphasize the differences in the sensing strategies. Notice in Fig. 2 the blurred edges reconstructed from BRIEF, while the FREAK reconstruction has a poor contrast but fine details.



(a) FAST + binarized BRIEF (b) FAST + binarized FREAK

Fig. 2. Reconstruction of book covers. While BRIEF extracts salient shapes such as the auroch and the butterfly, FREAK is better at reconstructing the text and allows to read 3 titles out of 4

III. Conclusion

Our work shows that LBDs carry enough data to reconstruct the original image content. Beyond the privacy breach that could come from LBDs eavesdropping in cloud based image search applications, this also has the enthusiastic potential of designing better descriptors by the analysis of their content. Furthermore, this could serve as a basis for the development of disruptive smart cameras sensing binarized keypoint values instead of images.

References

[1] M. Calonder, V. Lepetit, C. Strecha, and P. Fua, "BRIEF: Binary Robust Independent Elementary Features," *Computer Vision–ECCV 2010*, pp. 778–792, 2010.

[2] A. Alahi, R. Ortiz, and P. Vandergheynst, "FREAK: Fast Retina Keypoint," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, 2012, pp. 510–517.

[3] P. Weinzaepfel, H. Jegou, and P. Pérez, "Reconstructing an image from its local descriptors," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, 2011, pp. 337–344.

[4] L. Jacques, J. N. Laska, P. T. Boufounos, and R. G. Baraniuk, "Robust 1-Bit Compressive Sensing via Binary Stable Embeddings of Sparse Vectors," *arXiv.org*, vol. cs.IT, Apr. 2011.