Statistical Analysis of Binarized SIFT Descriptors

M. Diephuis, S. Voloshynovskiy, O. Koval, F. Beekhof

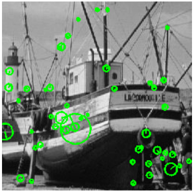
Stochastic Information Processing
Université de Genève

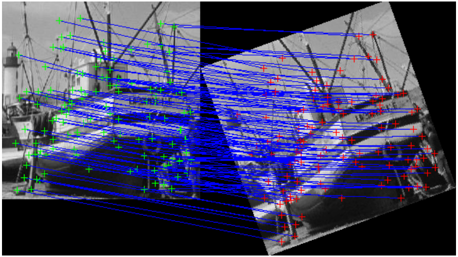
October 3, 2011

Outline

Introduction to SIFT

- SIFT[1] comprises a feature detector and a feature descriptor
- SIFT features are known to be robust to geometrical distortions

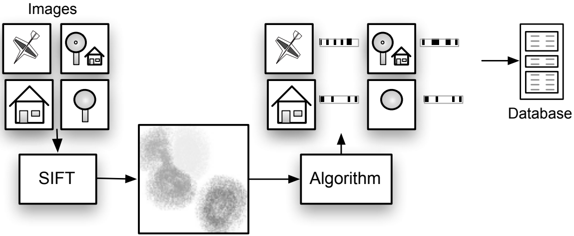




[1] *VLfeat toolbox*, by A. Vedaldi and B. Fulkerson, see `vlfeat.org`

SIFT applications in retrieval

- Aggregated SIFT descriptors are used increasingly in multimedia identification and retrieval tasks, e.g the bag of words approach [2].

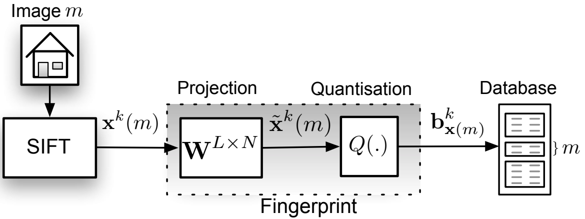


[2] *Hamming embedding and weak geometric consistency for large scale image search*, H. Jégou, M. Douze, C. Schmid

Architecture

Goal

Possibilities and limitations of using binarized projected SIFT descriptors for image identification



- $\mathbf{x}^k(m)$ is the k-th SIFT descriptor from image m
- $\mathbf{W}^{L\times N}$ is the Random Projection Matrix
- $\mathbf{x}^k(m)$ are the projected SIFT descriptors
- $\mathbf{b}^k_{\mathbf{x}(m)}$ is binarized descriptor

Testing setup

Testing binarized projected descriptors $\mathbf{b}_{\mathbf{x}^{k}(m)}$:

1. Correlations within an individual descriptor and between descriptors.

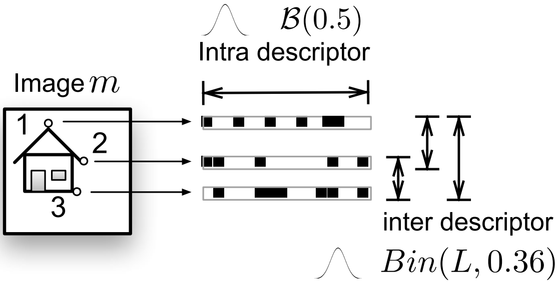


Figure: Test 1

Testing setup

Testing binarized projected descriptors $\mathbf{b}_{\mathbf{x}^{k}(m)}$:

1. Channel statistics for geometrical, noise and compression distortions.

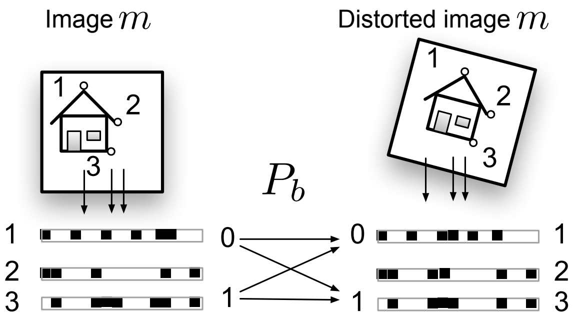


Figure: Test 2

Testing setup

Testing binarized projected descriptors $\mathbf{b}_{\mathbf{x}^k(m)}$:

1. Hamming distances between descriptors originating from different images.

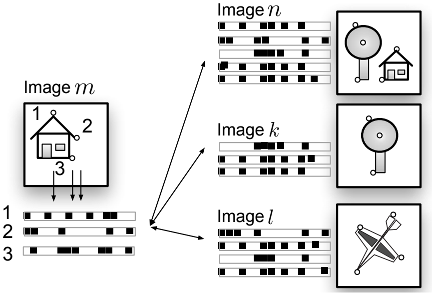


Figure: Test 3

Random Projections and other decorrelation methods

For security and storage reasons the method should provide:

- Fast data decorrelation
- Minimize privacy leakage

Method	Complexity	Privacy	Data Dependent
KLT	$\mathcal{O}(N^3)$	Basis vectors	Yes
DCT	$\mathcal{O}(N \log_2 N)$	Fixed basis vectors	No
DWT	$\mathcal{O}(N \log_2 N)$	Fixed basis vectors	No
RP	$\mathcal{O}(N^2)$	Basis vectors	No

Random Projection and Binarization

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

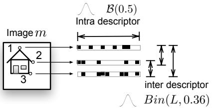
$$\tilde{\mathbf{x}}^k(m) = \mathbf{W}^{L \times N}\mathbf{x}^k(m). \quad (1)$$

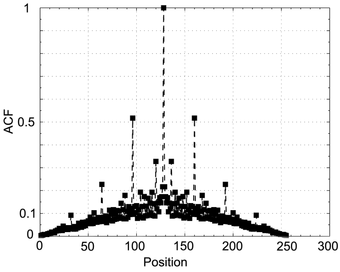
- L is the number of dimensions $\mathbf{W}^{L \times N}$ will map to, $L = 512$
- N is the length of the input column vector, which for SIFT vectors is 128.
- Random matrix $\mathbf{W}^{L \times N} = (\mathbf{W}_1, \mathbf{W}_2, \ldots, \mathbf{W}_N)^T$ consists of a set of approximately orthonormal basis vectors, where all elements are generated as $W_i[j] \sim \mathcal{N}(0, \frac{1}{N})$, $1 \leq i \leq N, 1 \leq j \leq L$

$\mathbf{W}^{L \times N}$ behaves as an approximate orthoprojector.

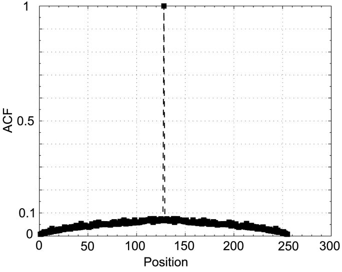
Test 1: Correlation after projection and binarization

- Correlation **within** descriptors is 0, e.g. $p(0) = p(1) = 0.5$, $\mathcal{B}(0.5)$
- Correlation between descriptors remains, distances adhere to $Bin(L, 0.36)$.





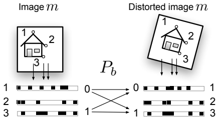
(a) ACF Original descriptors

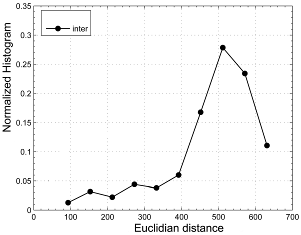


(b) ACF Projected Binarized descriptors

Test 2: Channel distortions

- Channel distortions between descriptors, originating from identical images.
- Distortions include scale, rotation, similarity transforms, JPEG compression and AWGN.





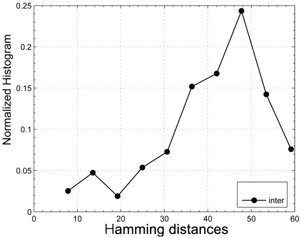
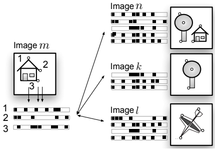


Figure: Channel distortion distances for a similarity transform $(s = 0.2, \theta = 10)$, resulting in a $P_b = 0.08$.

Test 3: Inter Image Matching

Match binarized projected descriptors from different images, based on the *non ambigu matching rule*[3].



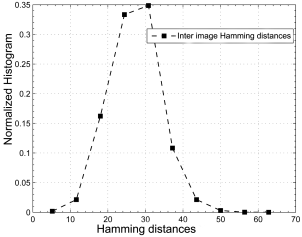


Figure: Hamming distances between matching descriptors between non-identical images.

[3] *Invariant features from interest point groups*, M. Brown and D. Lowe

Conclusions and Future Work

In short:

- We have shown statistical properties for projected binarized SIFT descriptors
- Random Projections are a simple and secure method for decorrelation
- In the Caltech 101 dataset, it is not possible to identify distorted images flawless based on the Hamming distance

Future

- Only retain the most valuable projections and their positions to achieve dimension reduction
- Testing on much larger dataset

Resources

Software, code, demo and paper available on
`http://sip.unige.ch/research/binarized-sift`