A contrario matching of SIFT-like descriptors

Julien Rabin, with Julie Delon, Yann Gousseau and Jean-Michel Morel

Master MVA, IPOL #4

October 21, 2011

Introduction

Most of computer vision problems are based on **local representation** of images:

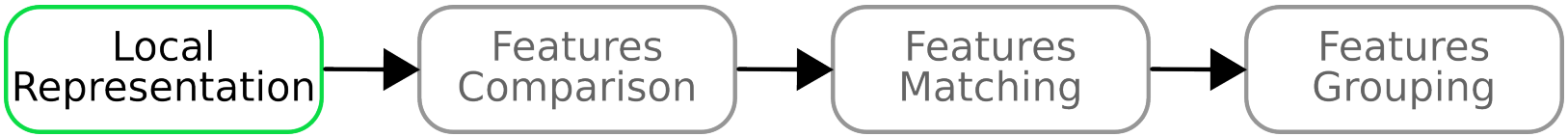
Image registration, indexation, classification, mosaicing, motion segmentation, object detection, object recognition, camera calibration, etc.

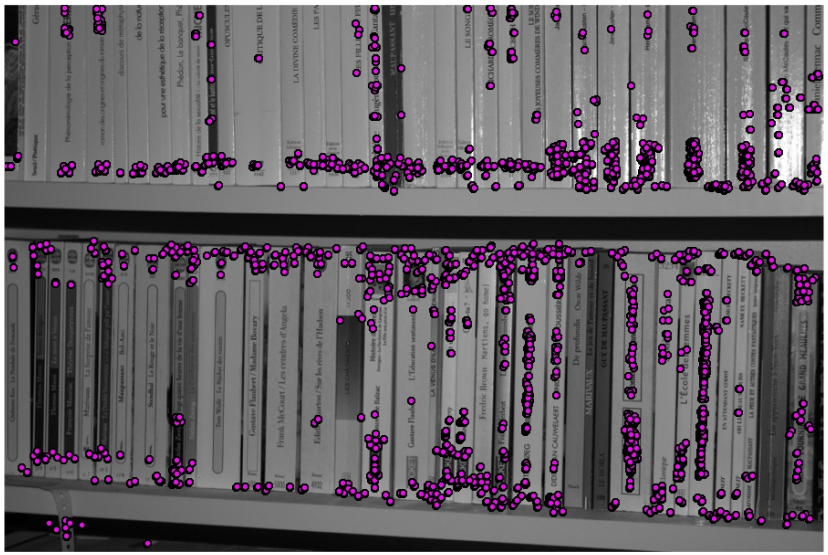
Dealing with local features requires some specific processing tasks.

Object Recognition example

Is there any similar **objects** between these images ?

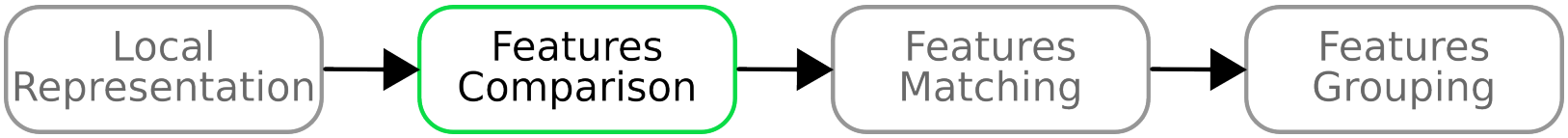
Preliminary step: Local representation of images

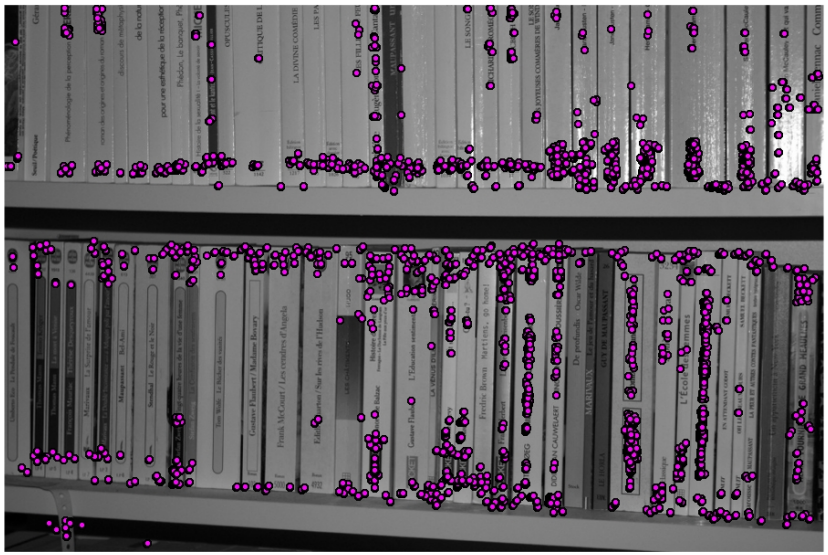




▷ Extract local and invariant descriptors, SIFT [Lowe 1999], Shape Context [Belongie and Malik, 2000], MSER [Matas *et al.* 2002], *etc.*

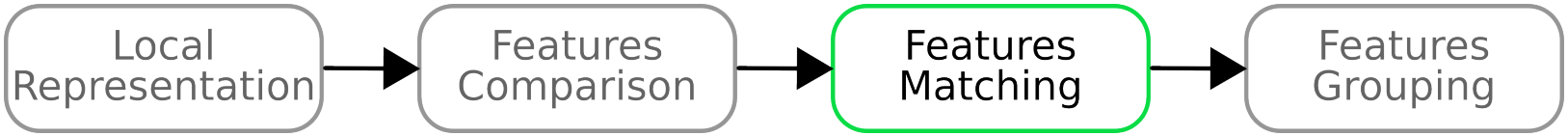
Step 1: Local features comparison

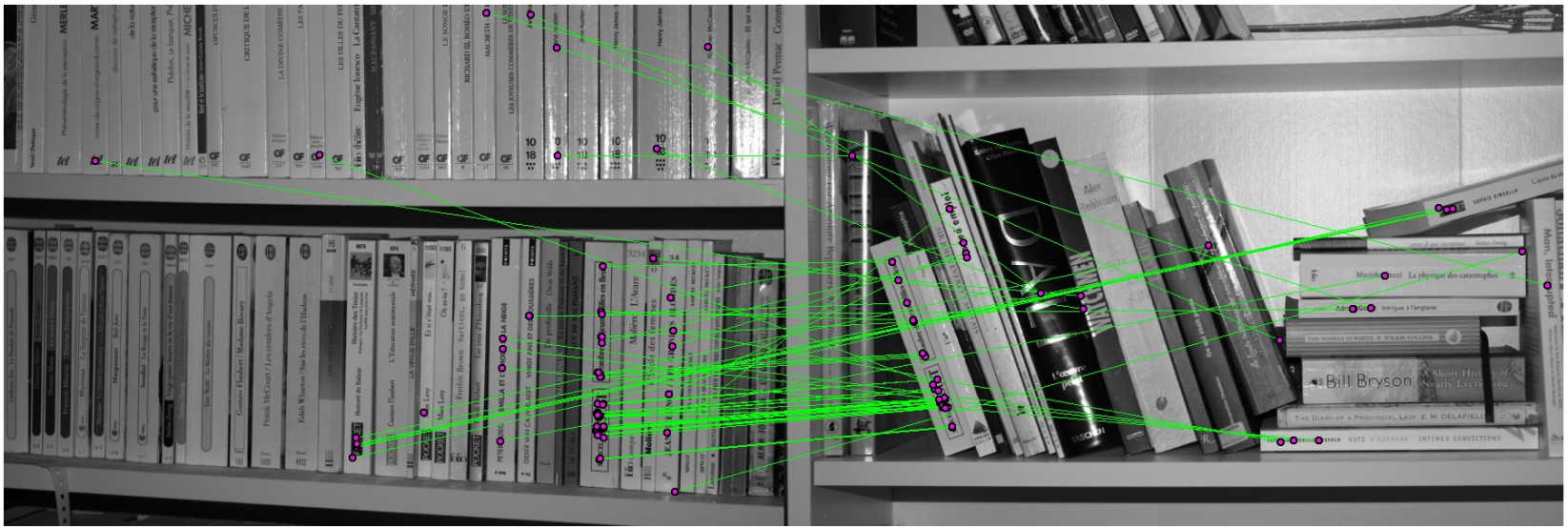




▷ Define a robust dissimilarity measure between local features

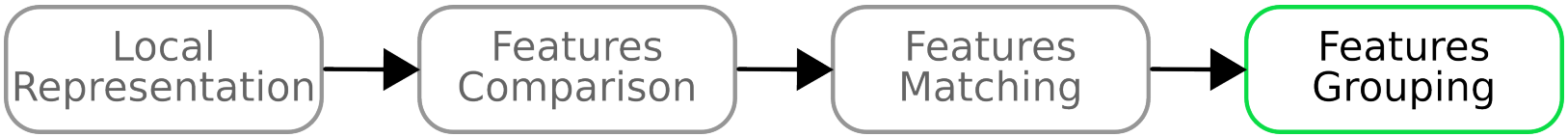
Step 2: Local features matching



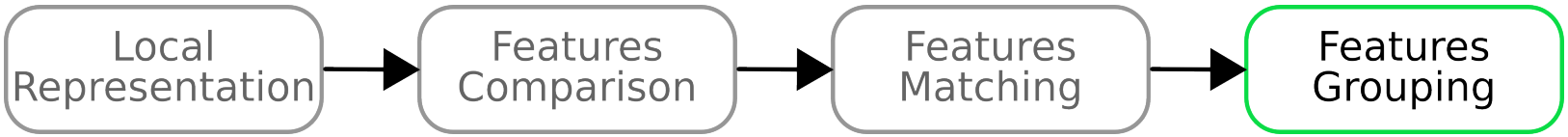


▷ Select reliable (multiple) correspondences between local features

Step 3: Local features grouping



▷ Detect a group and estimate the geometrical model's parameters related to the recognized object



▷ Detect a group and estimate the geometrical model's parameters related to the recognized object

Problem statement

Why the feature matching step is **critical**?

Testing a tentative match between two interest points has 4 different issues (hypothesis testing)

Ground truth \ Test	Positive	Negative
Correct	True-Positive (tp)	False-Negative (fn) [Type II error]
Incorrect	False-Positive (fp) [Type I error]	True-Negative (tn)

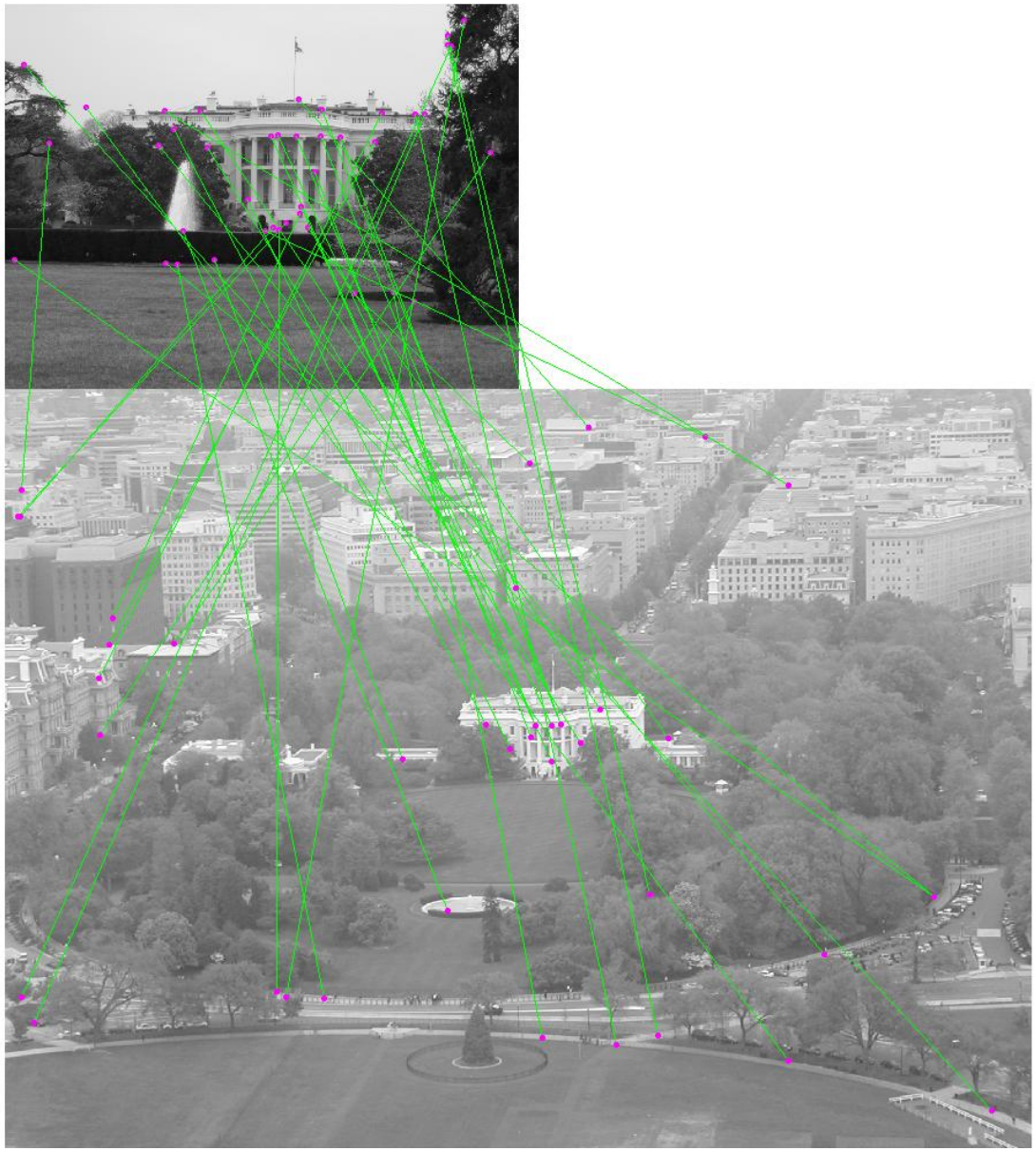
The grouping step in the processing chain is very sensitive to “outliers” (irrelevant correspondences).

$\Rightarrow$ need to control type I error

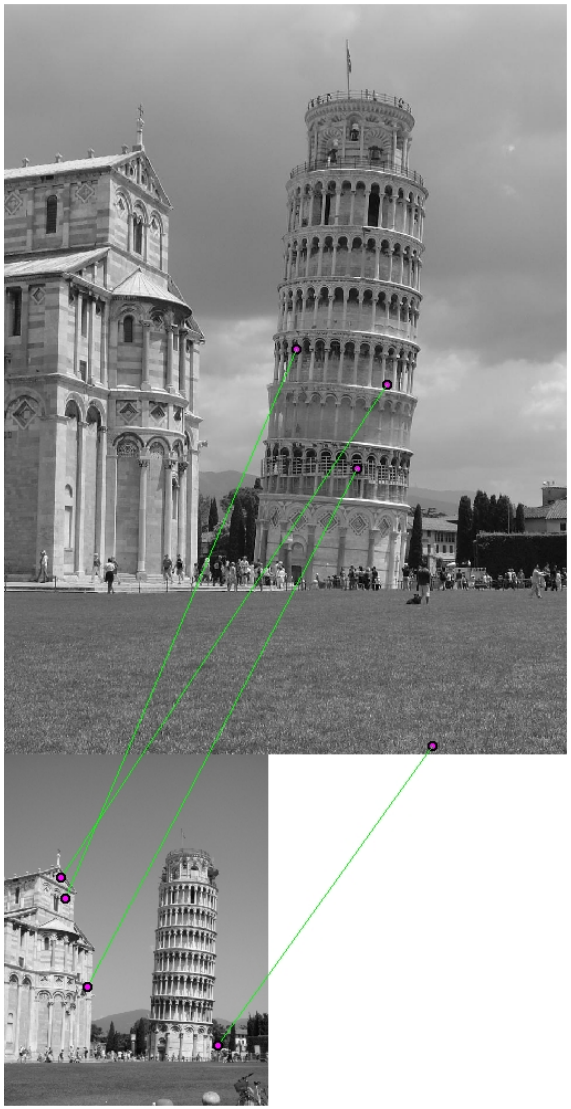
Examples using the original SIFT method

Examples using the original SIFT method

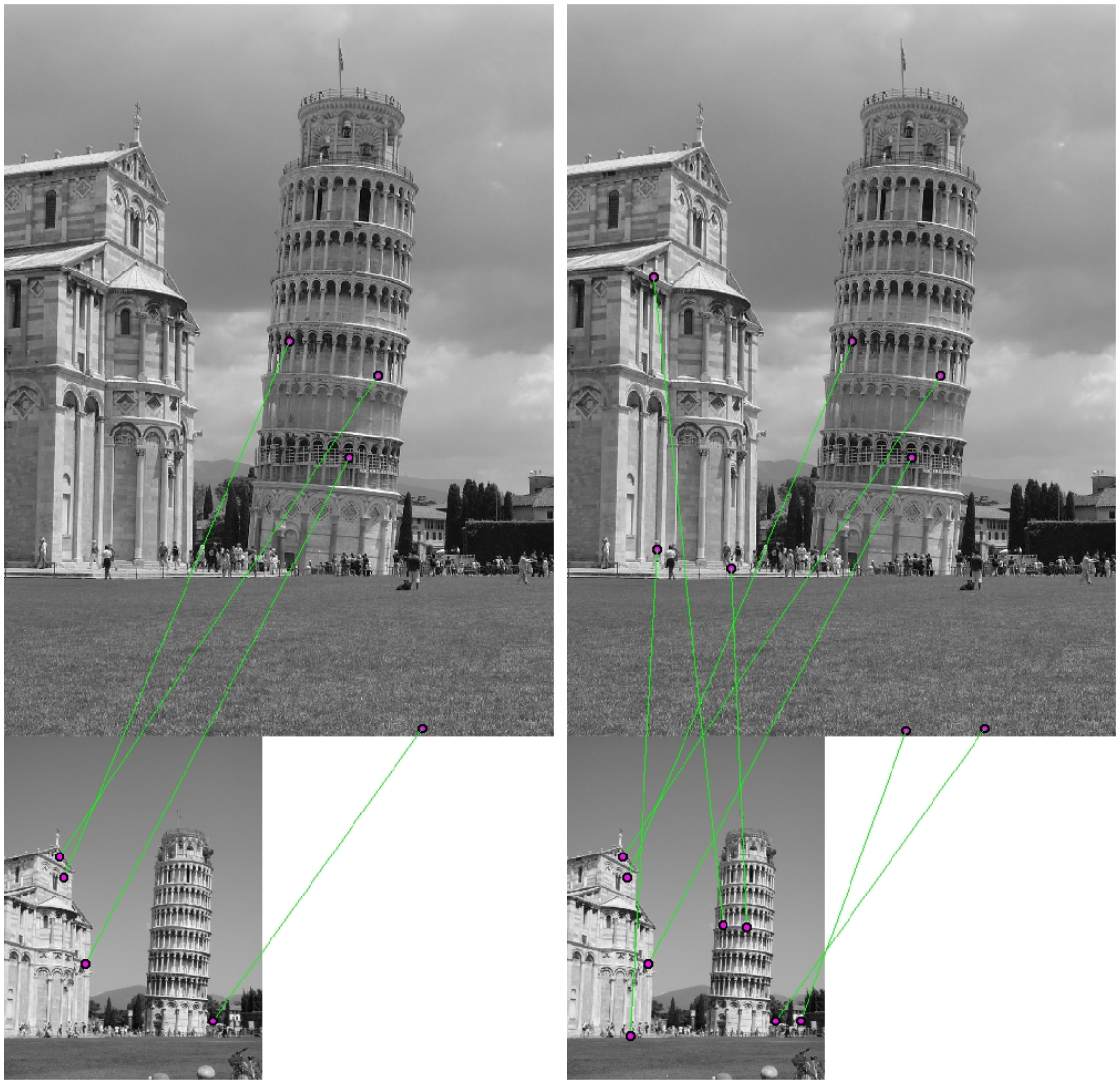
Examples using the original SIFT method



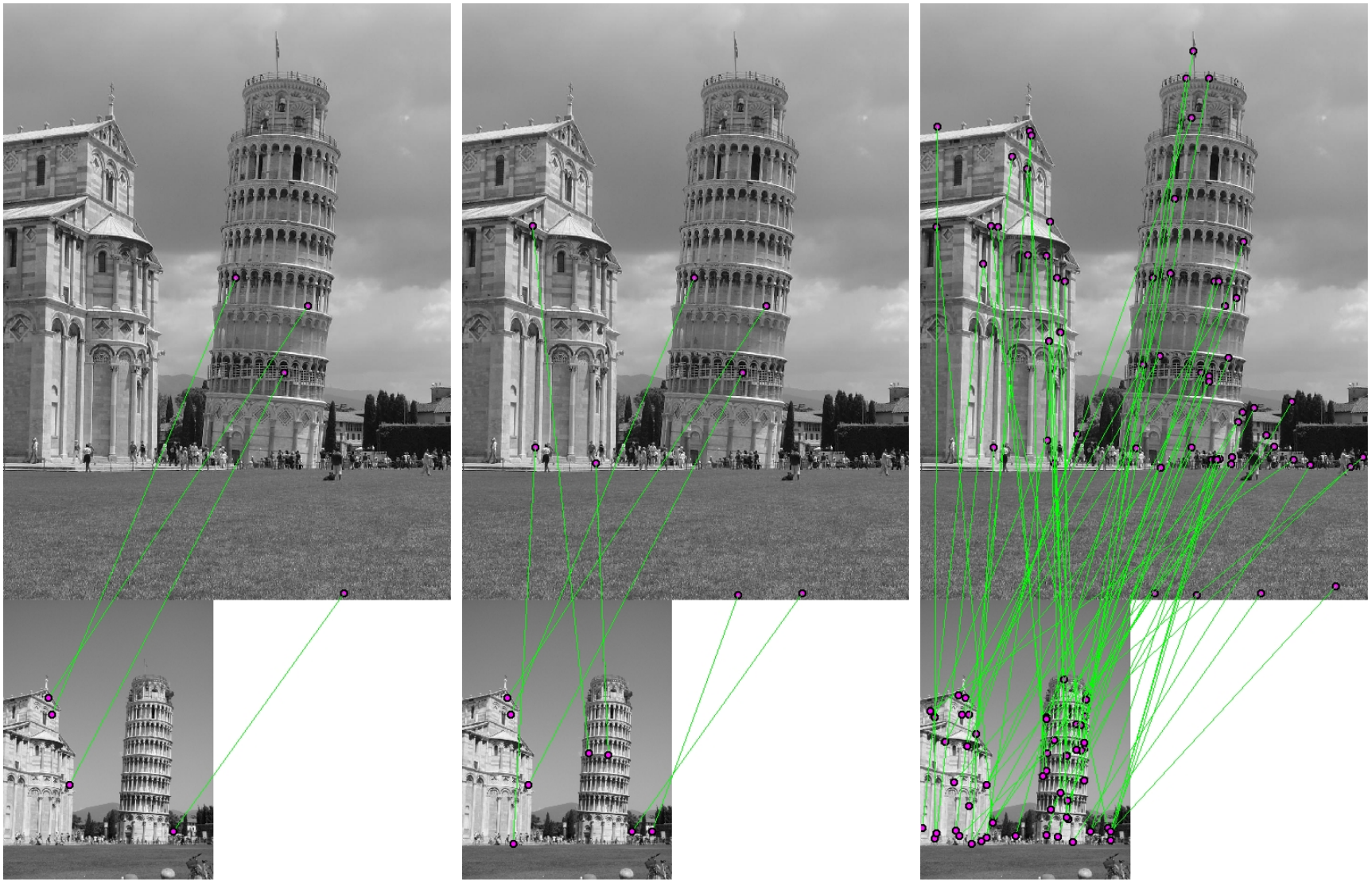
Examples using the original SIFT method



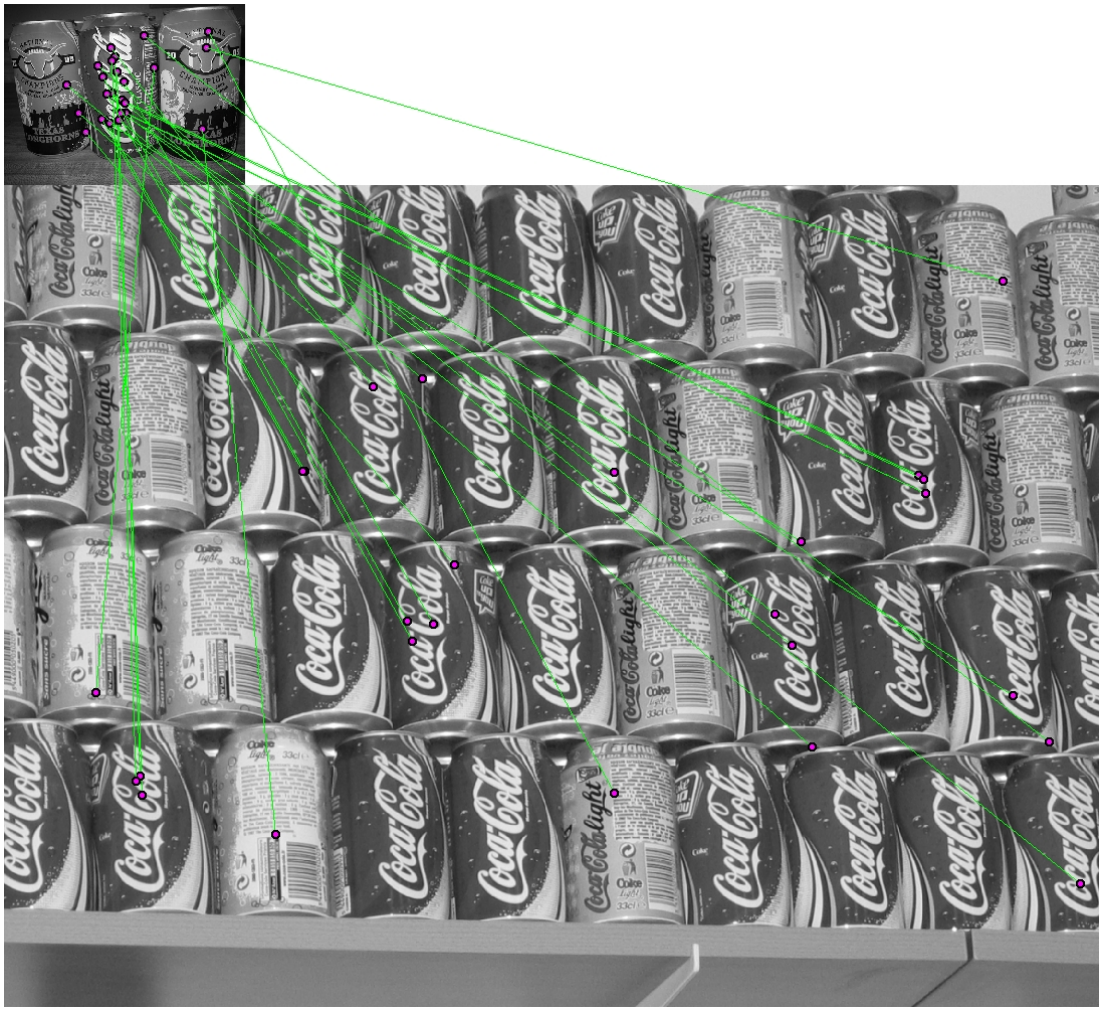
Examples using the original SIFT method



Examples using the original SIFT method



Examples using the original SIFT method



Objectives

Each of the previous tasks relies on **decision criteria**.

In this course, we will study decision criteria based on the ***a contrario*** methodology [Desolneux et al., 2000], applied to the two following tasks:

- SIFT orientation assignment,
- SIFT matching

Outline

Part I. Local features **detection**

Part II. *A contrario* methodology

Part III. Histogram **mode selection**

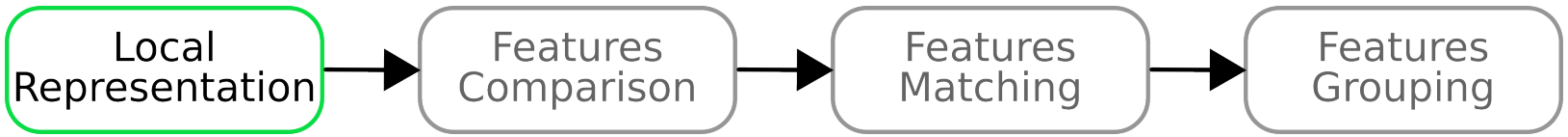
(break)

Part IV. SIFT **matching**

Part V. IPOL Projects

Part I

SIFT overview - local descriptors



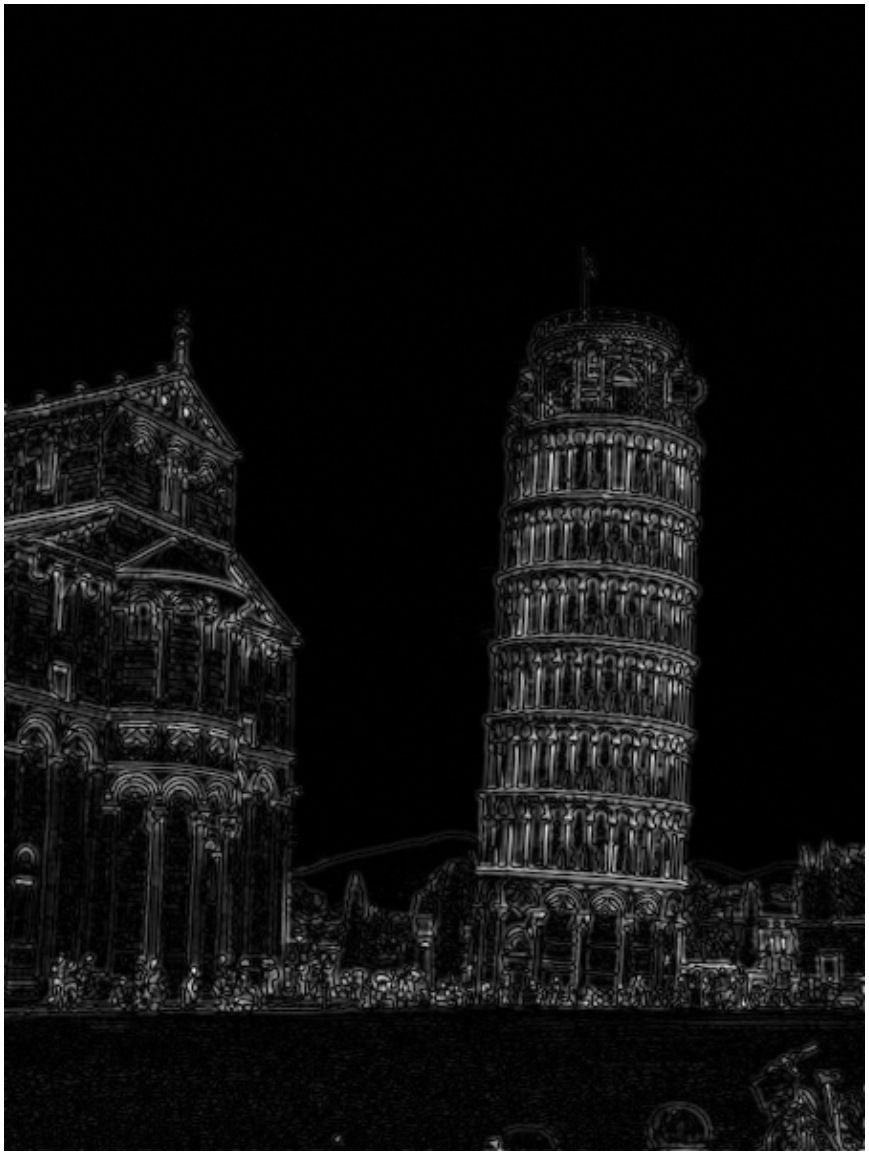
Local features

1. Discrete image u

Local features

2. Scale-space representation
   $\forall \sigma, \quad u_\sigma = g_\sigma * u$

Local features



3. Local extrema $(\vec{x}, \sigma)$ in space and scale of $\sigma^2 \Delta u_\sigma$

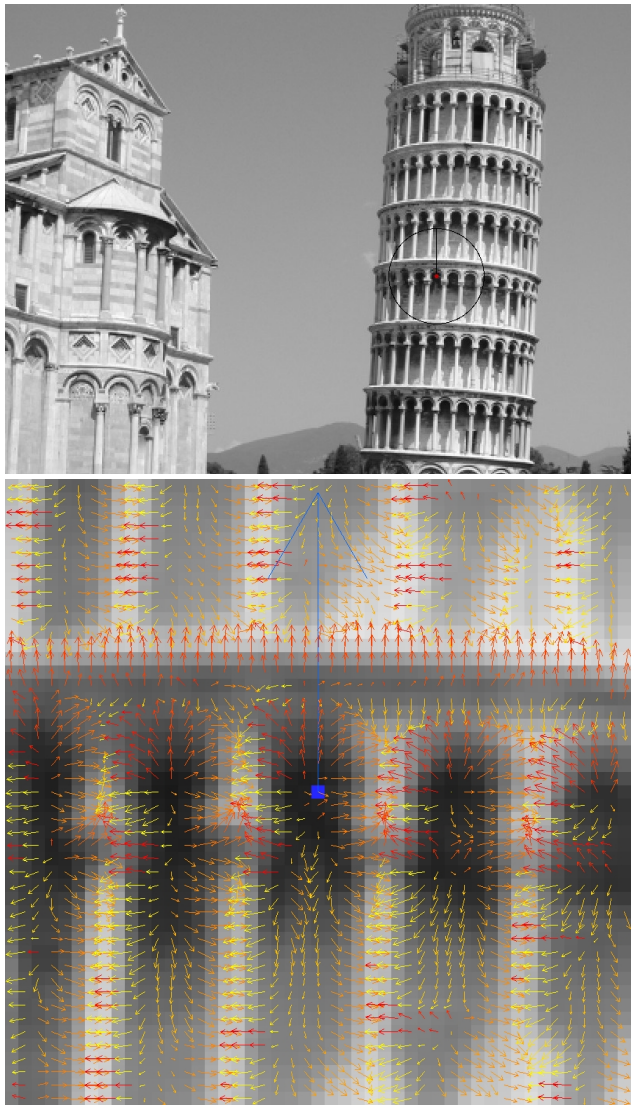
Local features

4. Harris (or Hessian) multiscale criterion to eliminate edge points $\rightarrow$ interest points $(\vec{x}, \sigma)$.

Local features

4. Harris (or Hessian) multiscale criterion to eliminate edge points $\rightarrow$ interest points $(\vec{x}, \sigma)$.

Local features



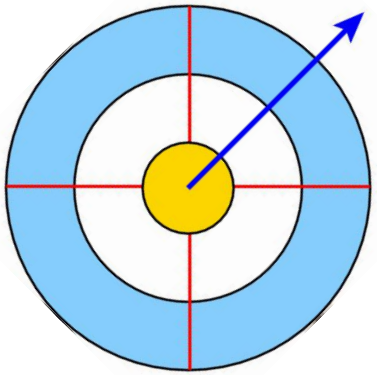
5. Main orientations (directions of ∇u_σ) assigned at each point $\rightarrow$ interest points $(\vec{x}, \sigma, \alpha)$.

Local features: example of SIFT descriptors [Lowe, 1999]

Construction of a **local descriptor a at each interest point** $(\vec{x}, \sigma, \alpha)$.

Mask (*e.g.* a square, a disk) around $\vec{x}$:

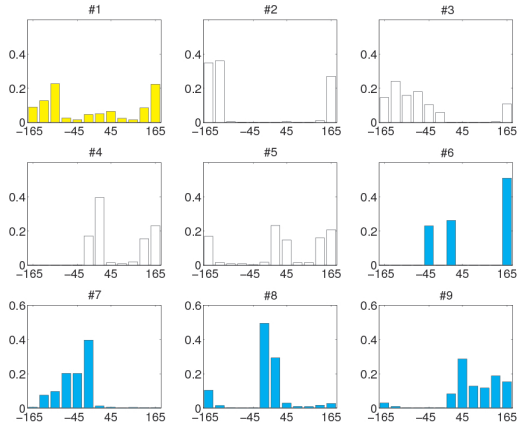
- M sectors,
- size proportional to σ.
- orientation given by α



Descriptor $a = (a_1, \dots a_M)$

a_m = normalized histogram of the gradient orientation (*), weighted by the gradient magnitude, in the m^{th} sector.

(*) Orientations defined with respect to the reference direction α.



Part II

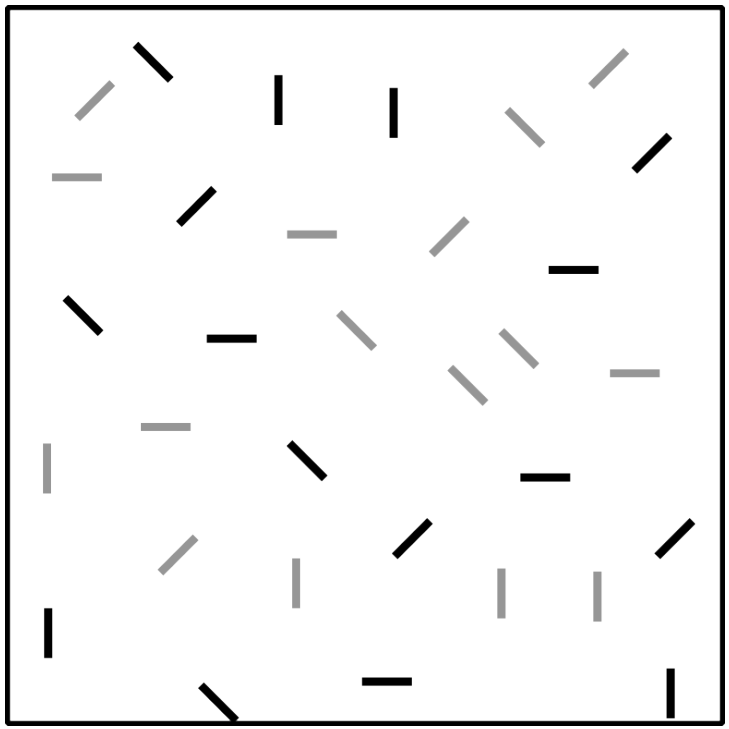
A Contrario detection theory

Gestalt theory and the Helmholtz Principle

«no structure should be detected in a noise image.»

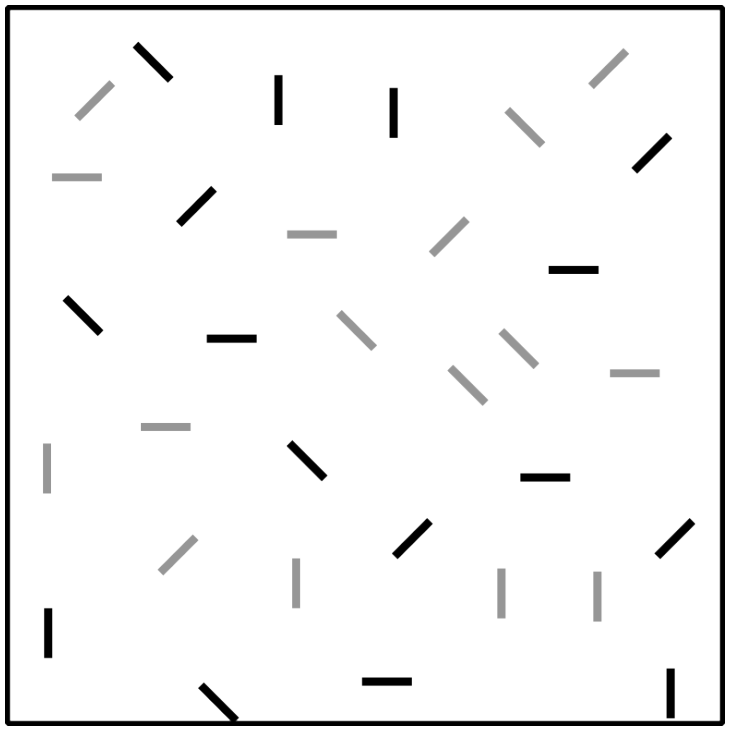
Gestalt theory and the Helmholtz Principle

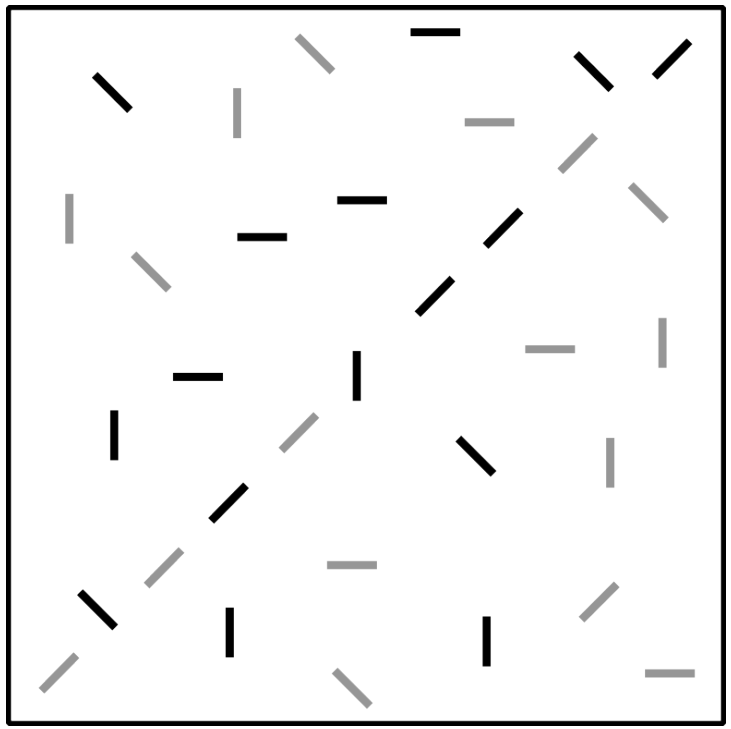
«no structure should be detected in a noise image.»



Gestalt theory and the Helmholtz Principle

«no structure should be detected in a noise image.»





A contrario methodology

Basic principles of *a contrario* methodology [Desolneux, Moisan and Morel, 00]:

Null hypothesis: The background model rejection in the *a contrario* approach rely on a **null hypothesis** testing, denoted $\mathcal{H}_0$, that states that the observed features are **mutually independent**.

Similarity measure: Now, given a group of features G, a similarity measure S has to be defined to evaluate the adequacy of this group with the type of structure we want to detect.

P-value: the probability of observing a random group of features $\mathbf{G}$ following the null hypothesis with a better similarity measure than G:

$$\mathbb{P}_{\mathcal{H}_0}(S(\mathbf{G}) \geqslant S(G)) \,. \tag{1}$$

Significativity measure with multiple tests: Let $\{G_i\}_i$ be the set of tested structures and $\mathcal{N} = |\{G_i\}_i|$ be the total number of tests. We call NFA, the significativity measure

$$\mathbf{NFA}(G_i, S(G_i)) := \mathcal{N} \times \mathbb{P}_{\mathcal{H}_0}(S(\mathbf{G}) \geqslant S(G_i)) \,. \tag{2}$$

A contrario methodology (II)

Group validation:

Definition (ε-meaningful group)

The group G_i is validated if

$$\mathrm{NFA}(G_i, S(G_i)) \leqslant \varepsilon$$

Fundamental property:

«The expected number of ε-meaningful groups following the null hypothesis is smaller than ε.»

Automatic thresholds on the similarity measure:

$$t_i(\varepsilon) = \min\{t, \mathrm{NFA}(G_i, t) \leqslant \varepsilon\} \tag{3}$$

Thus, a group is ε-meaningful *only if* $S(G_i) > t_i(\varepsilon)$.

Overview of the *a contrario* philosophy

To sum up:

- Detect groups of features that are very unlikely under the hypothesis that features are independent *(null hypothesis)*
- Unlikeliness ensured by controlling the expected number of false alarms.

Many applications since [Desolneux, Moisan, Morel, 00].

Part III

SIFT orientation assignment

Original SIFT approach

The key idea of D. Lowe in [Lowe, 1999] is to estimate the main orientation of a given keypoint by looking at the **statistical distribution of the gradient orientation** in the vicinity of the interest point. To do so, an histogram is first built in which **peaks** are then detected.

Circular histogram: Let consider a circular histogram H defined from M samples $\{\theta_1, \dots, \theta_M\}$ (orientations) quantized uniformly on L values for the interval $[-\pi, \pi[$:

$$\forall\, i \in \{1, \dots, L\}\; H[i] = \left| \left\{ \theta_m \in \left[\frac{2\pi}{L}(i-1) - \pi, \frac{2\pi}{L} i - \pi \right[, m = 1, \dots, M \right\} \right| . \quad (4)$$

Peack detection: In [Lowe, 2004], the histogram of the local gradient orientation is quantized on $L = 36$ accumulators and is built from gradient samples extracted in the neighborhood of the interest point $mi : (x_i, y_i, \sigma_i)$. Each entry in the histogram is weighted by the gradient norm and by a Gaussian kernel $G_{3\sigma_i}(. - x_i, . - y_i)$ with variance $9\sigma_i^2$ and mean (x_i, y_i).

The peaks in the histogram are then simply defined as local maxima, where local maxima lower than 80% to the global maximum are pruned, that is

$$P = \left\{ i \in \mathcal{I}\,,\; H[i-1] \leqslant H[i] \text{ and } H[i] \geqslant H[i+1] \text{ and } H[i] \geqslant 0.8 \max_{j \in I} H[j] \right\} \quad (5)$$

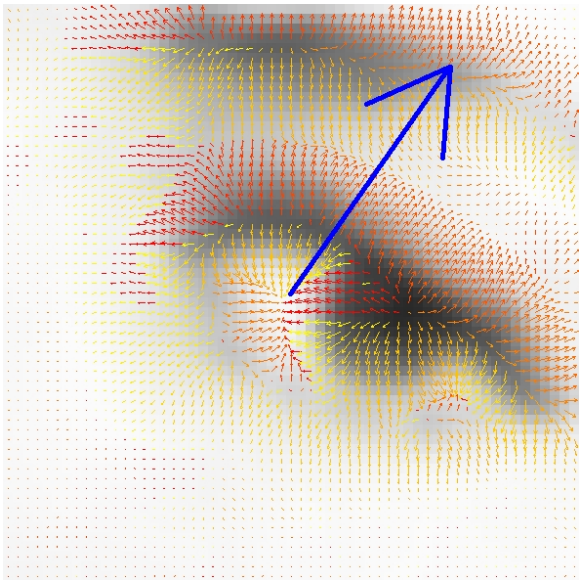
Original SIFT approach

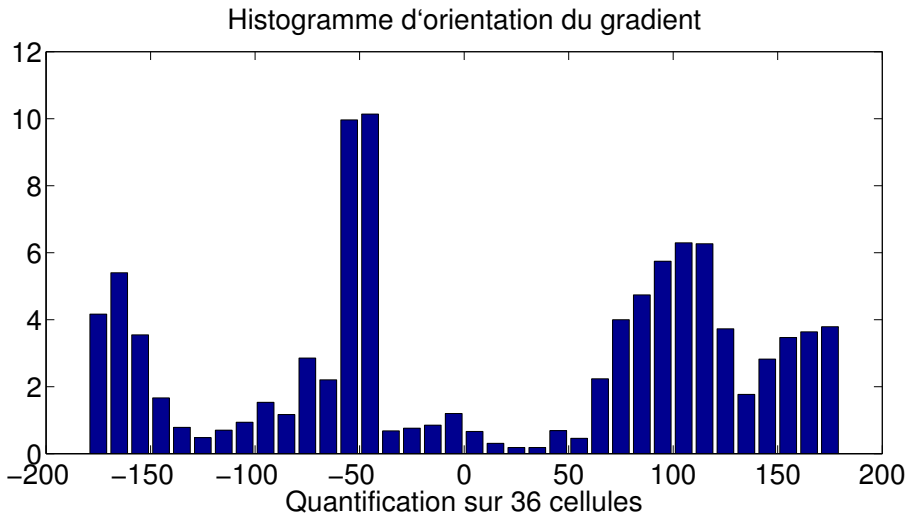
The location of each peak is then refined as the vertex α of the parabola fitted to the 3 histogram values around the peak, that is

$$\forall i \in P\,,\ \alpha_i = -\frac{\pi\, b}{L\, a}\,[2\pi] \text{ with } \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} (i-1)^2 & i-1 & 1 \\ i^2 & i & 1 \\ (i+1)^2 & i+1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} H[i-1] \\ H[i] \\ H[i+1] \end{bmatrix}\,, \quad (6)$$

where $H[0] = H[L]$ and $H[L+1] = H[1]$.
Finally, the orientation α_i is assign to the interest point m_i so that $m_i : (x_i, y_i, \sigma_i, \alpha_i)$.

Example




Histogramme d'orientation du gradient
12
10
8
6
4
2
0
−200
−150
−100
−50
0
50
100
150
200
Quantification sur 36 cellules

A contrario detection of meaningful intervals

Let first define a "meaningful interval" as an interval containing significantly more sample than the expected value.

Null hypothesis $\mathcal{H}_0$:

Definition (Null Hypothesis (background model))

the samples $\{\theta_m\}_{\{m=1,\dots,M\}}$ are mutually independent random variables, identically and uniformly distributed in $[-\pi, \pi($.

Observe that this null hypotheses corresponds to a background model of white noise image, since the orientation of the discrete gradient is uniformly distributed [Desolneux et al., 2002].

Then, the probability that one sample θ_m fall into $[a, b]$ under the null hypothesis $\mathcal{H}_0$ (*i.e.* assuming that the sample has been drawn using the background model) is hence:

$$p(a,b) = \frac{|[a,b]|}{L} = \frac{1}{L}\begin{cases} b-a+1 & \text{if } b \geqslant a \\ a-b+1+L & \text{otherwise} \end{cases} . \tag{7}$$

Meaningful intervals

P-value: Let denote by $k(a,b)$ the number of samples $\{\theta_m\}_m$ which belong to the discrete circular interval $[a,b]$. Then, the probability that at least $k(a,b)$ random samples of points among M fall into interval $[a,b]$ under the null hypothesis is given by the tail of the binomial distribution $\mathcal{B}(M,k(a,b),p(a,b))$ defined as:

$$\mathbb{P}_{\mathcal{H}_0}(\mathbf{k} \geqslant k(a,b)) = \mathcal{B}(M,k(a,b),p(a,b)) = \sum_{i=k(a,b)}^{M} \binom{M}{i} p(a,b)^i (1-p(a,b))^{M-i} . \tag{8}$$

Number of tests: In our setting, the number of tests is the total number of distinctive circular intervals

$$\mathcal{N} := |\{(a,b)\,,\ a \in \mathcal{I}, b \in \mathcal{I}\backslash\{a-1\}\}\cup\{[1,L]\}| = \sum_{i\in\mathcal{I}, j\in\mathcal{I}\backslash\{i-1\}} 1+1 = L(L-1)+1 . \tag{9}$$

Significativity measure: Then, we define the following quality measure, called NFA, for a given interval $[a,b]$ with $k(a,b)$ samples as:

$$\text{NFA}([a,b]) = \mathcal{N}\,\mathcal{B}(M,k(a,b),p(a,b)) . \tag{10}$$

Meaningful intervals

Definition (Meaningful interval)

An interval I is a $\varepsilon -$ *meaningful* interval when

$$\text{NFA}(I) = \mathcal{N}\,\mathcal{B}(M, k(a,b), p(a,b)) \leqslant \varepsilon\ .$$

Recall that a false alarm is a validated interval that follows the null-hypothesis.

Proposition

Let H be an histogram built from M random samples following the null hypothesis $\mathcal{H}_0$. The expected number of false alarms when validating ε-meaningful intervals is smaller than ε.

Meaningful gaps and modes

The previous definitions enables us to define a meaningful interval as an interval containing sufficiently more samples than it is expected from the background model. Nevertheless, it is insufficient in practice to select main modes of an histogram, since **a meaningful interval can contains some "gaps"**.
To take this into account, we first define meaningful gaps and then meaningful modes.

Gap: The definition of a meaningful gap is straight-forward: it is an interval that contains sufficiently *few* samples than expected from the background model. More precisely, the probability that at most $k(a,b)$ samples of points among M fall into interval $[a,b]$ under the null hypothesis is given by

$\mathbb{P}_{\mathcal{H}_0}(\mathbf{k} \leqslant k(a,b)) = \sum_{i=0}^{k(a,b)} \binom{M}{i} p(a,b)^i (1-p(a,b))^{M-i}$, which can be expressed from again from the tail of the binomial distribution

$$\mathbb{P}_{\mathcal{H}_0}(\mathbf{k} \leqslant k(a,b)) = \mathcal{B}(M, M-k(a,b), 1-p(a,b)) \ . \tag{11}$$

Definition (Meaningful gap)

An interval I is a $\varepsilon -$ *meaningful* gap when

$$\text{NFA'}(I) = \mathcal{N}\,\mathcal{B}(M, M-k(a,b), 1-p(a,b)) \leqslant \varepsilon \ . \tag{12}$$

Meaningful gaps and modes

Mode:

Definition (Meaningful mode)

An interval I is a ε-meaningful mode if it is a ε-meaningful interval that does not contain any ε-meaningful gap.

Maximal Meaningful Mode: several meaningful modes with overlaps could be detected (redundant detections). Therefore we need a maximality criterion to select only non-overlapping modes.

Definition (Maximal meaningful mode)

An interval I is a maximal-ε-meaningful mode if it is a ε-meaningful mode and if for all ε-meaningful mode $J \subset I$, $\text{NFA}(J) \geqslant \text{NFA}(I)$ and for all ε-meaningful mode $J \supsetneq I$, $\text{NFA}(J) > \text{NFA}(I)$.

Algorithm

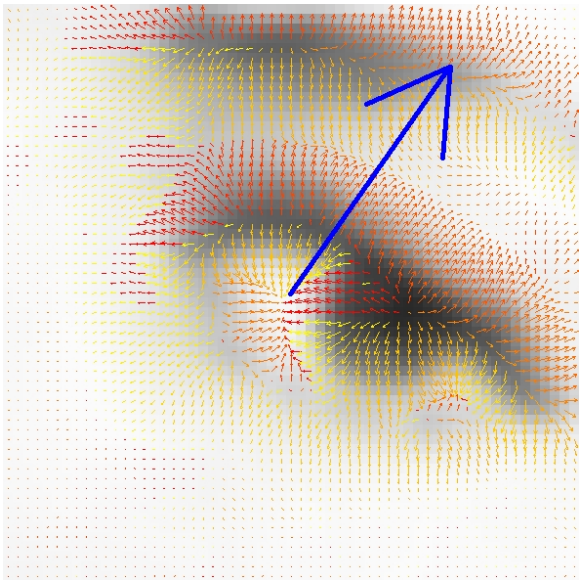
Algorithm: Automatic mode selection.

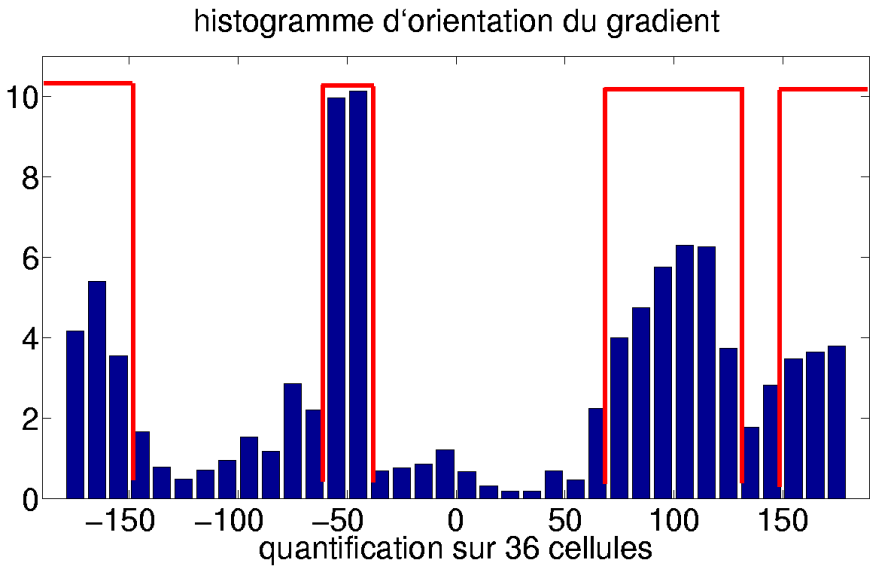
Input: Histogram H with M samples and L bins.
parameter $\varepsilon = 1$.

1) Find ε-meaningful intervals (definition 3);
2) Find ε-meaningful gaps (definition 4);
3) Find ε-meaningful modes (definition 5);
4) Find maximal ε-meaningful modes (definition 6).

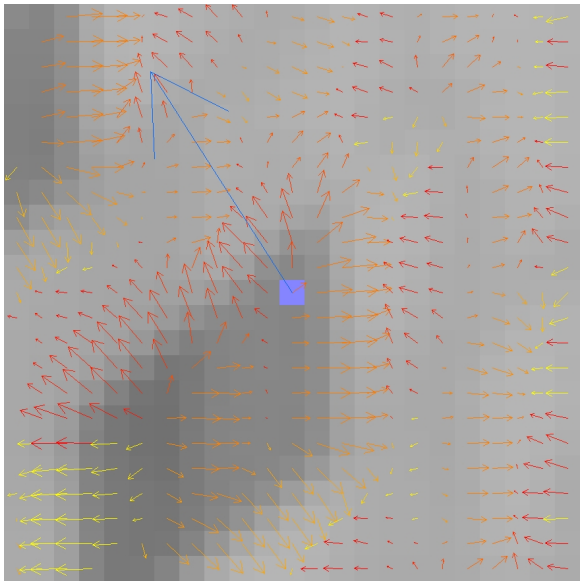
Output: List of intervals.

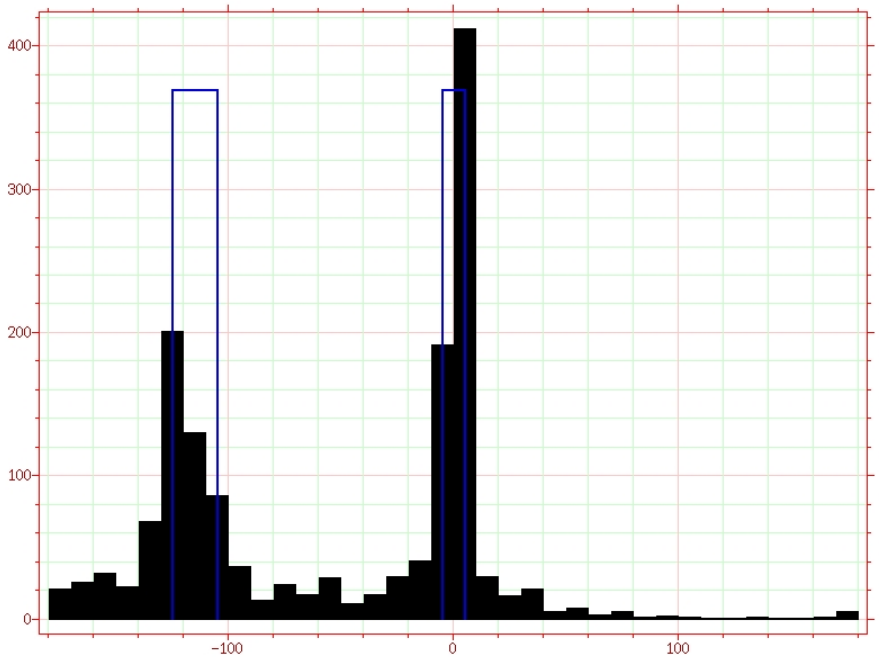
Example



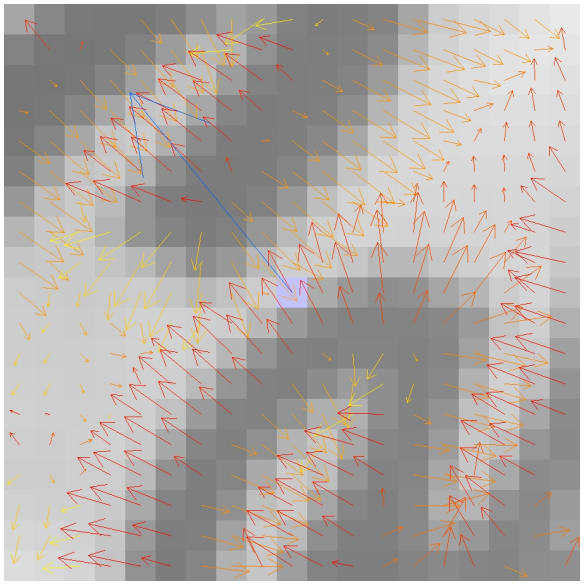

histogramme d'orientation du gradient
10
8
6
4
2
0
−150
−100
−50
0
50
100
150
quantification sur 36 cellules

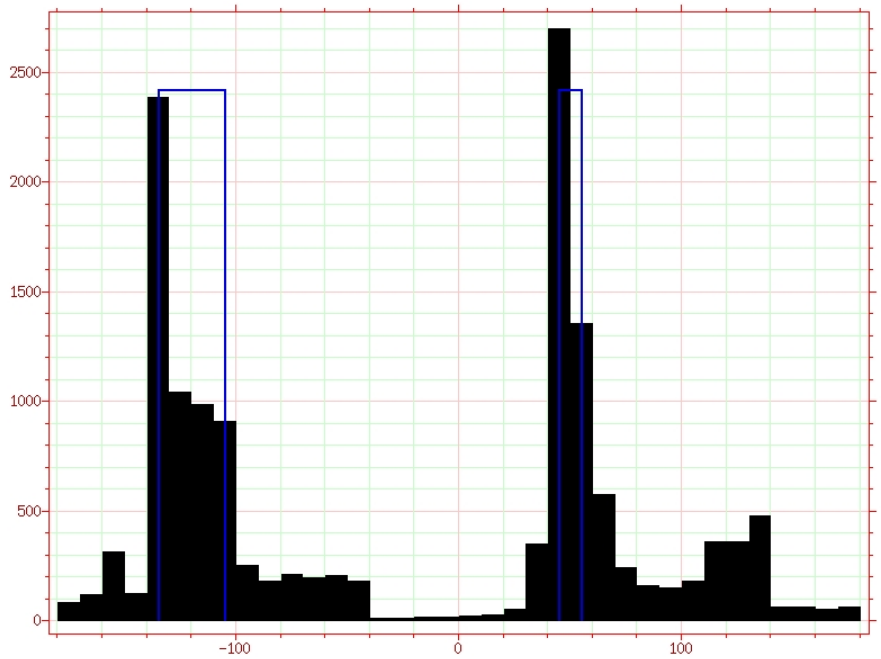
Example



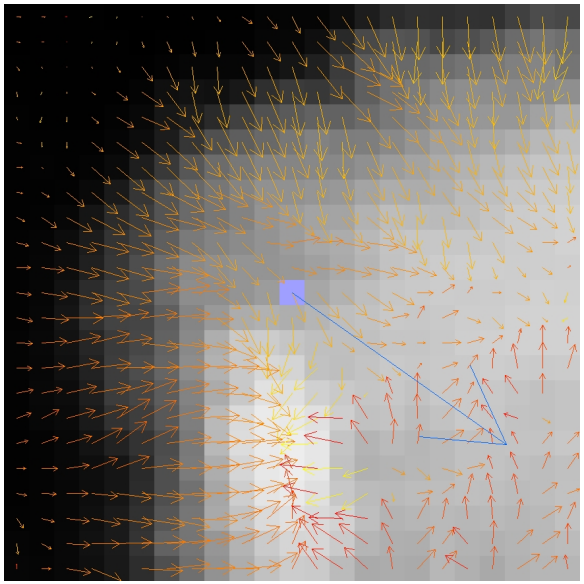


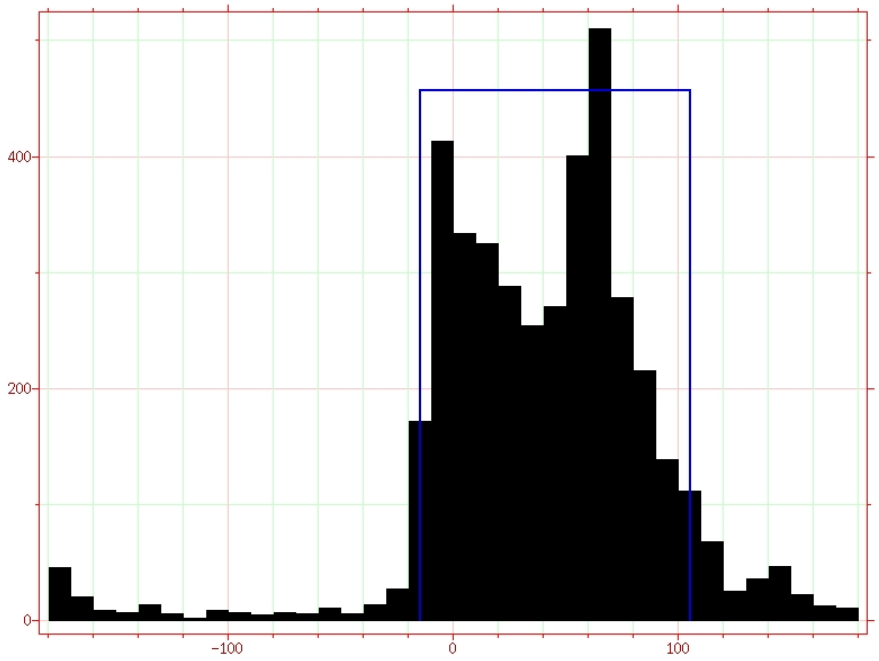
Example





Example





Interest point filtering and orientation assignment

For a given 1-meaningful mode $[a, b]$, the corresponding orientation $\alpha_{[a,b]}$ is simply defined as the *circular barycenter* of the histogram values in this interval, *i.e.*

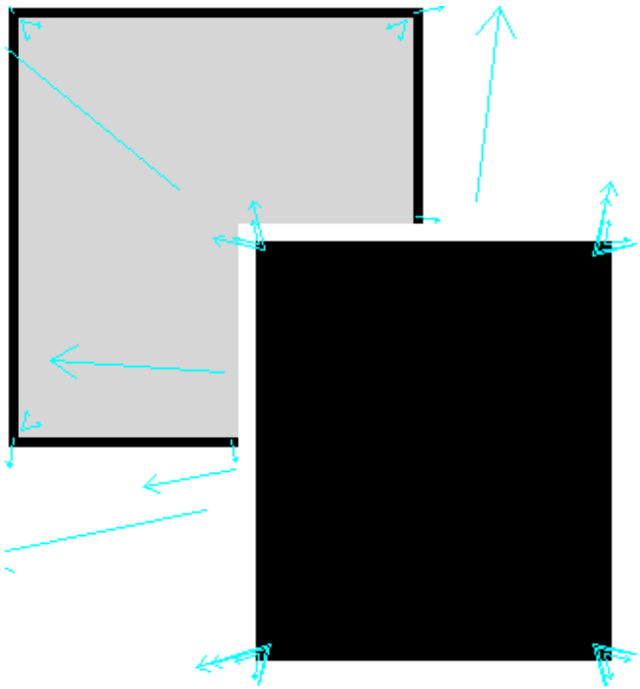
$$\alpha_{[a,b]} = \frac{2\pi}{L} \begin{cases} \sum_{i=a}^{b} i \cdot H[i] & \text{if } a \leqslant b \\ \sum_{i=a}^{L} i \cdot H[i] + \sum_{i=1}^{b} (i+L) \cdot H[i] & \text{if } a > b \end{cases} [2\pi] \,. \tag{13}$$

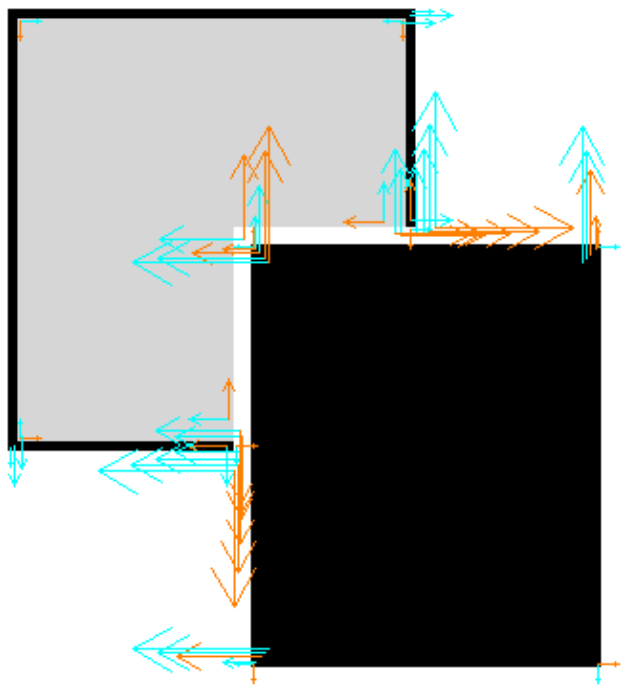
Eventually, observe that the *a contrario* mode selection makes it also possible to discard interest points that lies on edges structures. Indeed, only interest points with at least two orientation assignments should be kept.

Example

Example

Example

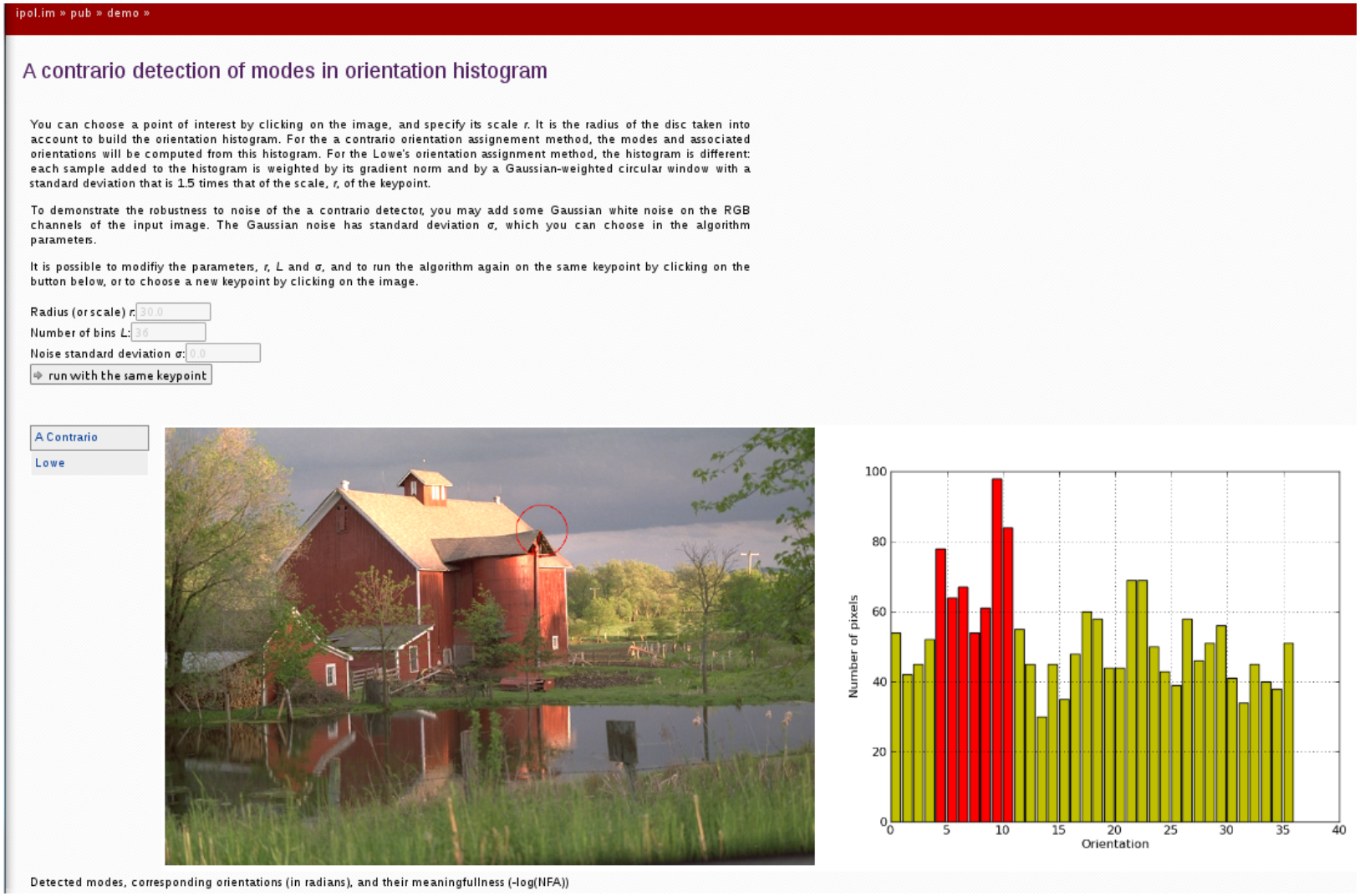




IPOL demo

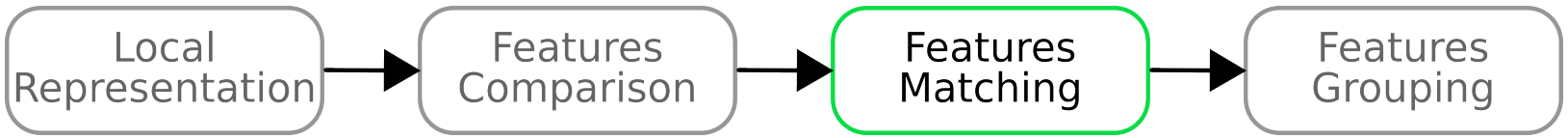
Last year IPOL project of **Carlo de Franchis**

http://dev.ipol.im/~carlo/ipol_demo/d_a_contrario_orientation_modes/



Part IV

SIFT matching



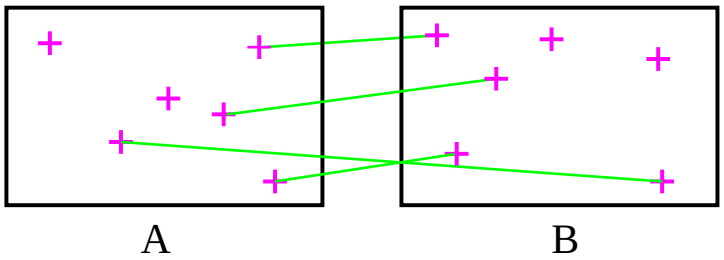
Matching framework with SIFT-like descriptors

Query descriptors: $a^1, \dots a^{N_Q}$.

- each descriptor a^i is made of M histograms, $a^i = (a^i_1, \dots, a^i_M)$

statement

Candidate descriptors (database): $b^1, \dots b^{N_C}$.



$\Leftrightarrow$

$$\begin{pmatrix} a^1 \\ \vdots \\ a^i \\ \vdots \\ a^{N_Q} \end{pmatrix} \longrightarrow \begin{pmatrix} b^1 \\ \vdots \\ b^j \\ \vdots \\ b^{N_C} \end{pmatrix}$$

Problem statement: Among all $N_Q \times N_C$ possible matches, which ones are relevant?

Matching framework with SIFT-like descriptors

Define a dissimilarity measure D between descriptors. For instance:

$$\mathrm{D}_{L^p}(a,b) := \left(\sum_{m=1}^{M}\sum_{i=1}^{N} \mid a_m[i] - b_m[i] \mid^p\right)^{\frac{1}{p}} \quad (14)$$

$$\mathrm{D}_{\chi^2}(a,b) := \sum_{m=1}^{M}\sum_{i=1}^{N} \frac{(a_m[i] - b_m[i])^2}{a_m[i] + b_m[i]} \quad (15)$$

$$\mathrm{D}_{M}(a,b) := \left(\sum_{m,n=1}^{M}\sum_{i,j=1}^{N} (a_m[i] - b_m[i]) \cdot w_{i,j,m,n} \cdot (a_n[j] - b_n[j])\right)^{\frac{1}{2}}, \quad (16)$$

Assumption: Dissimilarity measures $\{D(a^i, b^j)\}_{\{i,j\}}$ have been computed for all $N_Q \times N_C$ possible matches.

Need of a decision criterion $\leftrightarrow$ thresholds on D.

DT matching criterion

The simplest matching criterion, that we call DT (Distance Threshold), relies on a global threshold on distances. That is, each query a^i is simply matched with candidates $\{b^j\}$ that are at a distance $d(a^i, b^j)$ smaller than the threshold.

Definition (DT Criterion)

The set of matches is defined as:

$$\mathcal{C}_{\mathrm{DT}} := \left\{ (a^i, b^j)\,,\ i \in \{1, \dots, N_Q\} \text{ and } j \in \{1, \dots, N_C\} \ :\ D(a^i, b^j) \leqslant t \right\}$$

NN-DT matching criterion

Usually, matches are restricted to the nearest neighbor [Baumberg, 2000, Jia and Tang, 2008] for each query descriptor, in order to limit multiple false detections that often affect some query descriptors. We will refer to this criterion as NN-DT (Nearest Neighbor Distance Threshold).

Definition (NN-DT Criterion)

The set of matches is defined as:

$$\mathcal{C}_{\text{NN-DT}} := \left\{ (a^i, b^{J(i)}) \ , \ i \in \{1, \dots, N_Q\} \ : \ D(a^i, b^{J(i)}) \leqslant t \text{ s.t. } J(i) = \underset{j \in \{1,\dots,N_C\}}{\arg\min} \ D(a^i, b^j) \right\}$$

Two main drawbacks inherent to this approach still restrict its use in practice:

1. Impossible to define an optimal threshold for different queries and database ("universal" threshold does not exists!);
2. No multiple detections.

NN-DR matching criterion

In order to reduce the variability of the chosen threshold, Lowe [Lowe, 2004] introduces another criterion by comparing the distances between a^i and its closest and second-closest neighbors respectively. If the ratio between the two distances is below a threshold r, the match with the closest neighbor is validated.

Definition (NN-DR Criterion)

The set of matches is defined as:

$$\mathcal{C}_{\text{NN-DR}} := \left\{ (a^i, b^{J1(i)}) \ , \ i \in \{1, \dots, N_Q\} \ : \ \frac{D(a^i, b^{J1(i)})}{D(a^i, b^{J2(i)})} \leqslant r \ , \text{ s.t.} \right.$$
$$\left. J1(i) = \underset{j \in \{1,\dots,N_C\}}{\arg\min} \ D(a^i, b^j) \text{ et } J2(i) = \underset{j \in \{1,\dots,N_C\} \setminus J1(i)}{\arg\min} \ D(a^i, b^j) \right\}$$

This popular criterion, that we call NN-DR (Nearest Neighbor Distance Ratio), benefits from its simplicity and the fact that it is by far more robust than a simple threshold on distances. Nevertheless, it suffers from the following drawbacks:

1. Discard matches related to self similarity and multiple occurrences [Zhang and Kosecka 06], [Noury, Sur and Berger 10];
2. No multiple detections (NN restriction);
3. The significance of the statistic test based on the two nearest neighbors is strongly dependent to the picture size;

ADT matching criterion

Another interesting alternative matching criterion has been proposed by J-L. Lisani in [Cao et al., 2008] to adapt the distance ratio test in such a way that it **tolerates multiple detections**.
To do so, one consider an alternative database B' composed of N'_C descriptors in which we look for the nearest neighbor for each query descriptor. Now, a match is validated if the ratio between the distance from the query to a given candidate in B and the distance from the query to its most similar candidate in B' is below a fixed threshold r.

Definition (ADT Criterion)

The set of matches is defined as:

$$\mathcal{C}_{\text{ADT}} := \left\{ (a^i, b^j)\,,\ i \in \{1,\dots,N_Q\} \text{ and } j \in \{1,\dots,N_C\} \ :\ D(a^i, b^j) \leqslant r \cdot \min_{j \in \{1,\dots,N'_C\}} D(a^i, b'^j) \right\}$$

This method can be seen as an *a contrario* approach in which one estimates *adaptive thresholds* on the dissimilarity measure depending on the query and a background model (the database B').
Nevertheless, it requires some extra computations with the database B'.
Moreover, contrary to the *a contrario* methodology introduced in Part II, it does not make it possible to set the threshold r in such way that we control the expected number of false alarms.

The null hypothesis

We assume that

$$D(a^i, b^j) = \sum_{m=1}^{M} d(a^i_m, b^j_m)$$

Let **b** be a **random descriptor**, and a^i any query descriptor.

Definition (Null hypotheses)

Given a query descriptor a^i, a random descriptor b satisfies the hypothesis $\mathcal{H}_0^i$ ***:*** ***"the distances*** $d(a^i_m, \mathbf{b}_m)$ ***(***$m \in \{1, \dots M\}$***) are mutually independent random variables."***

Then

$$\mathbb{P}\left(D(a^i, \mathbf{b}) \leqslant \delta \,|\, \mathcal{H}_0^i\right) = \int_{-\infty}^{\delta} \underset{m=1}{\overset{M}{*}} p^i_m(x)\, dx,$$

where p^i_m is the pdf of the random variable $d(a^i_m, \mathbf{b}_m)$.

In practice : for every $m \in \{1, \dots M\}$, p^i_m is empirically estimated over the database $\{b^1, \dots b^{N_C}\}$,

$$p^i_m(x) = \frac{1}{N_C} \# \left\{ b^j ; d(a^i_m, b^j_m) = x \right\}.$$

$\Rightarrow$ for each i and each value of δ, one can estimate $\mathbb{P}\left(D(a^i, \mathbf{b}) \leqslant \delta \,|\, \mathcal{H}_0^i\right)$.

Meaningful matches

Definition:

$$\mathrm{NFA}(a^i, \delta) := N_Q\, N_C\, \mathbb{P}\left(D(a^i, \mathbf{b}) \leqslant \delta \,|\, \mathcal{H}_0^i\right)$$

Adaptive threshold computation:

$$\tilde{\delta}_i(\varepsilon) = \arg\max_{\delta} \left\{\mathrm{NFA}(a^i, \delta) \leqslant \varepsilon\right\}$$

A Contrario (AC) Matching criterion: A match between a^i and b^j is

- validated if $\mathrm{NFA}(a^i, D(a^i, b^j)) \leqslant \varepsilon \Leftrightarrow D(a^i, b^j) \leqslant \tilde{\delta}_i(\varepsilon)$,
- rejected if $\mathrm{NFA}(a^i, D(a^i, b^j)) > \varepsilon \Leftrightarrow D(a^i, b^j) > \tilde{\delta}_i(\varepsilon)$.

Control of false alarms:

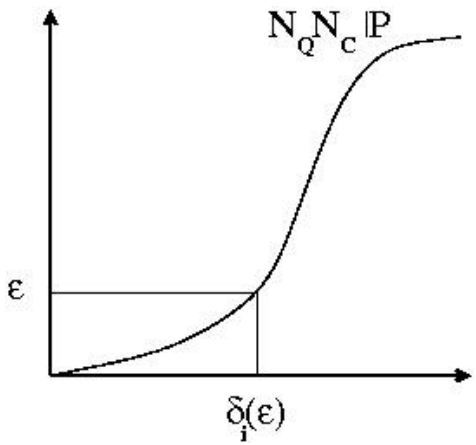
Proposition

Under the null hypothesis, the expected number of matches among the $N_C \times N_Q$ possible matches is smaller than ε.

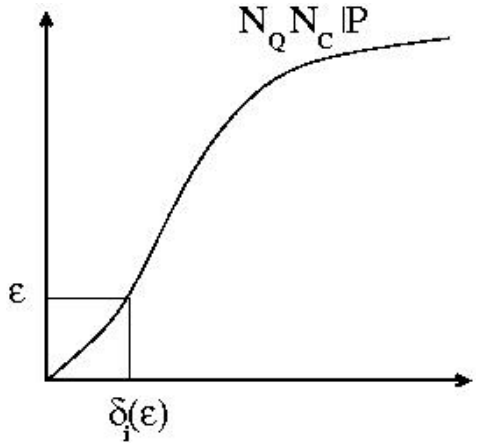
Adaptive distance thresholds

Advantages of **A Contrario criterion** (AC):

- only one threshold $\varepsilon \Rightarrow$ adaptive thresholds $\tilde{\delta}_i(\varepsilon)$ in function of a^i and of the database
- total number of matches not limited *a priori*



a^i « rare »



a^i « common »

A Contrario (AC) matching algorithm

Algorithm: Automatic distance threshold setting.

Input: N_Q query descriptors $\{a^i\}$ and N_C candidate descriptors $\{b^j\}$, parameter $\varepsilon > 0$.

For each query descriptor a^i, $i = 1, \ldots, N_Q$:

1) computation of distances $d_m(a^i, b^j)$ for all $m = 1, \ldots, M$ and $j = 1, \ldots, N_C$;
2) estimation of probability density functions: for each m, p_m^i computed as the empirical distribution of $d_m(a^i, b^j)$, when b^j spans the database;
3) computation of p-value $\delta \mapsto \mathbb{P}\left(D(a^i, \mathbf{b}) \leqslant \delta \,|\, \mathcal{H}_0^i\right)$;
4) computation of threshold $\delta_i(\varepsilon)$;
5) matching of a^i with each descriptor b^j such that $D(a^i, b^j) \leqslant \delta_i(\varepsilon)$;

Output: List of correspondences.

Can experiment

Small dataset illustration

Can experiment

NN-DT matching criterion

Can experiment

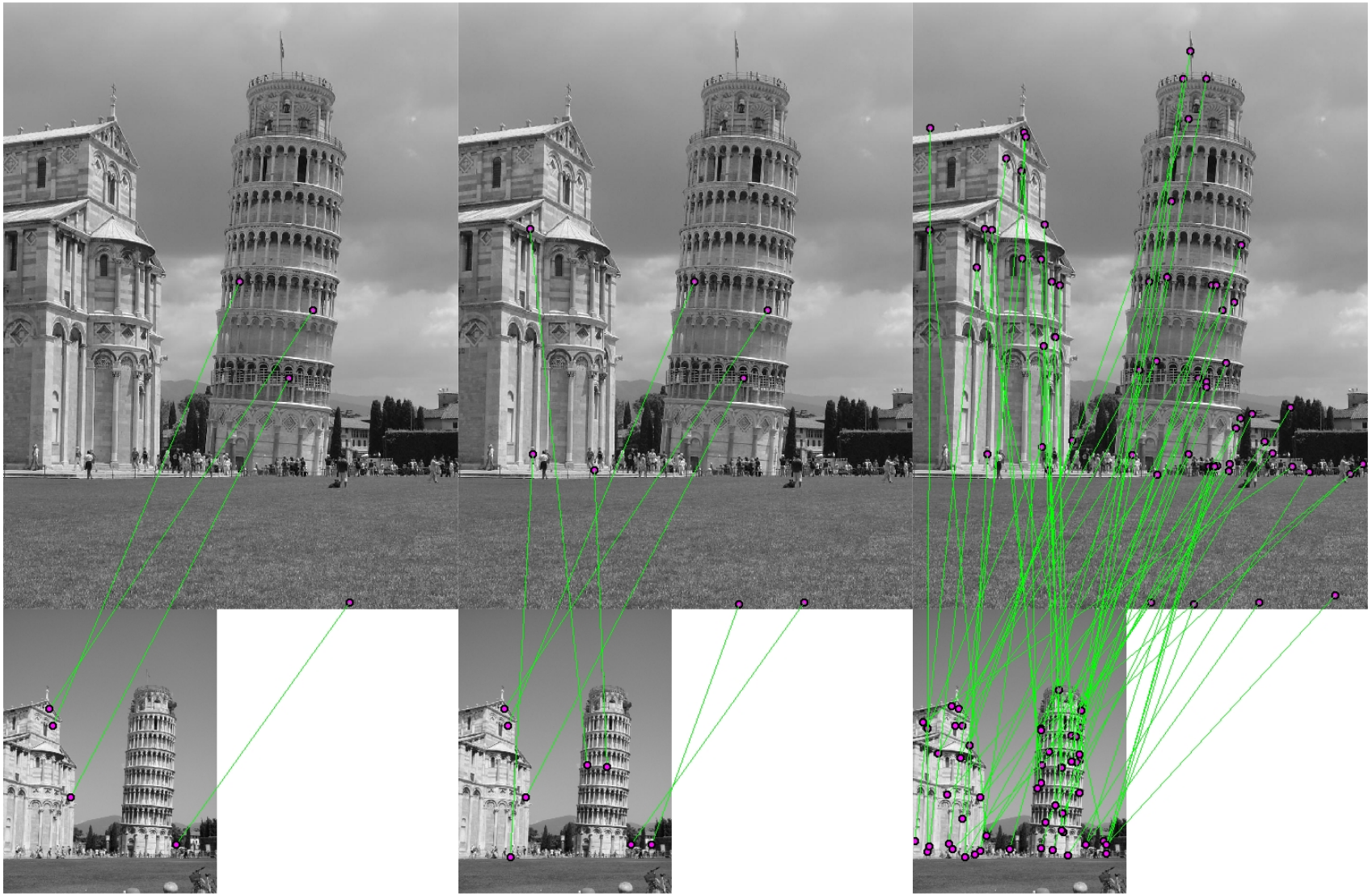
NN-DR matching criterion

Can experiment

AC matching criterion

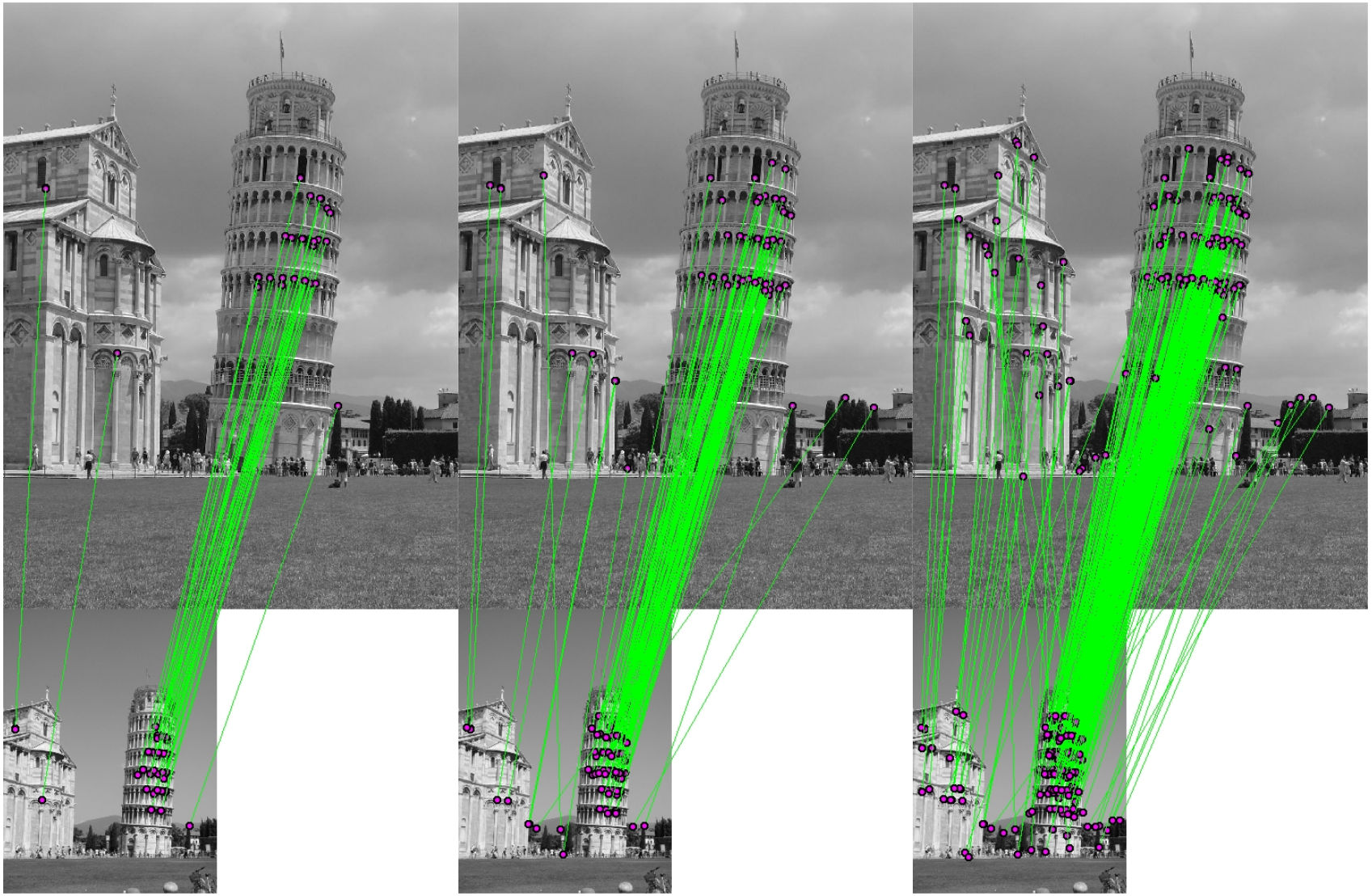
Pisa experiment

NN-DR with $r = 0.65$, $r = 0.7$, $r = 0.8$



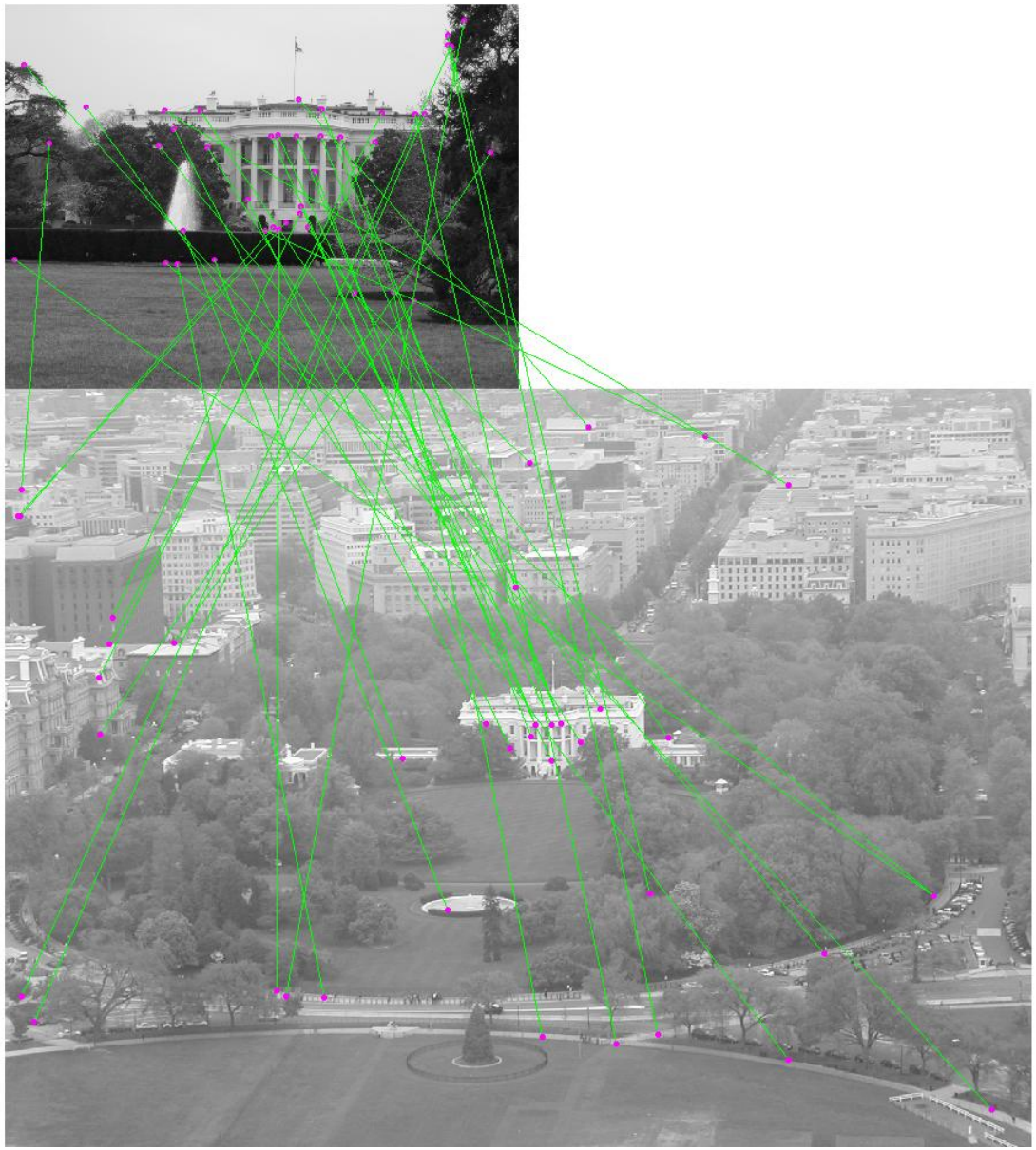
Pisa experiment

AC with $\varepsilon = 10^{-2}$, $\varepsilon = 10^{-1}$, $\varepsilon = 1$.



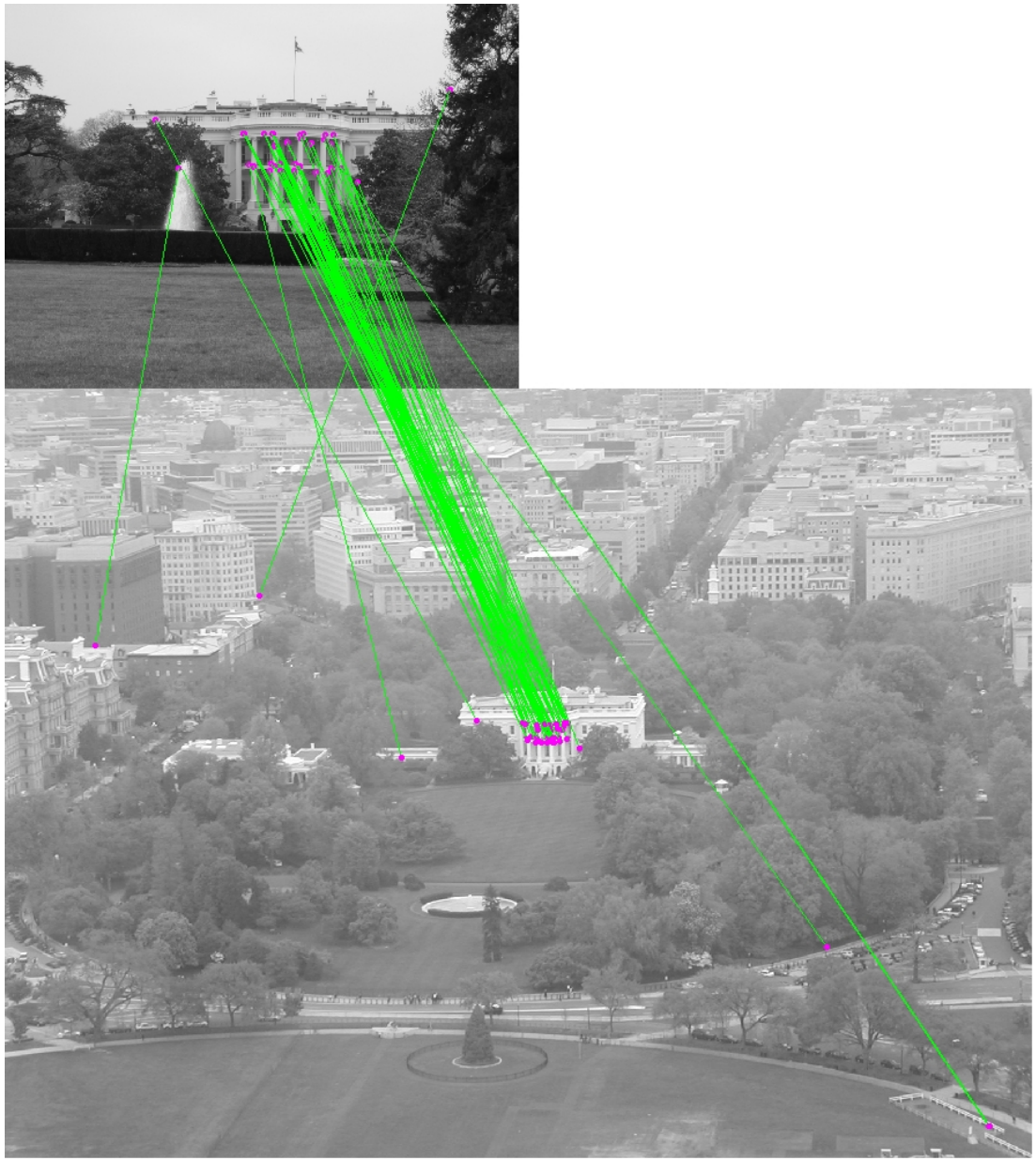
White House

Lowe's algorithm at $r = 0.8$.



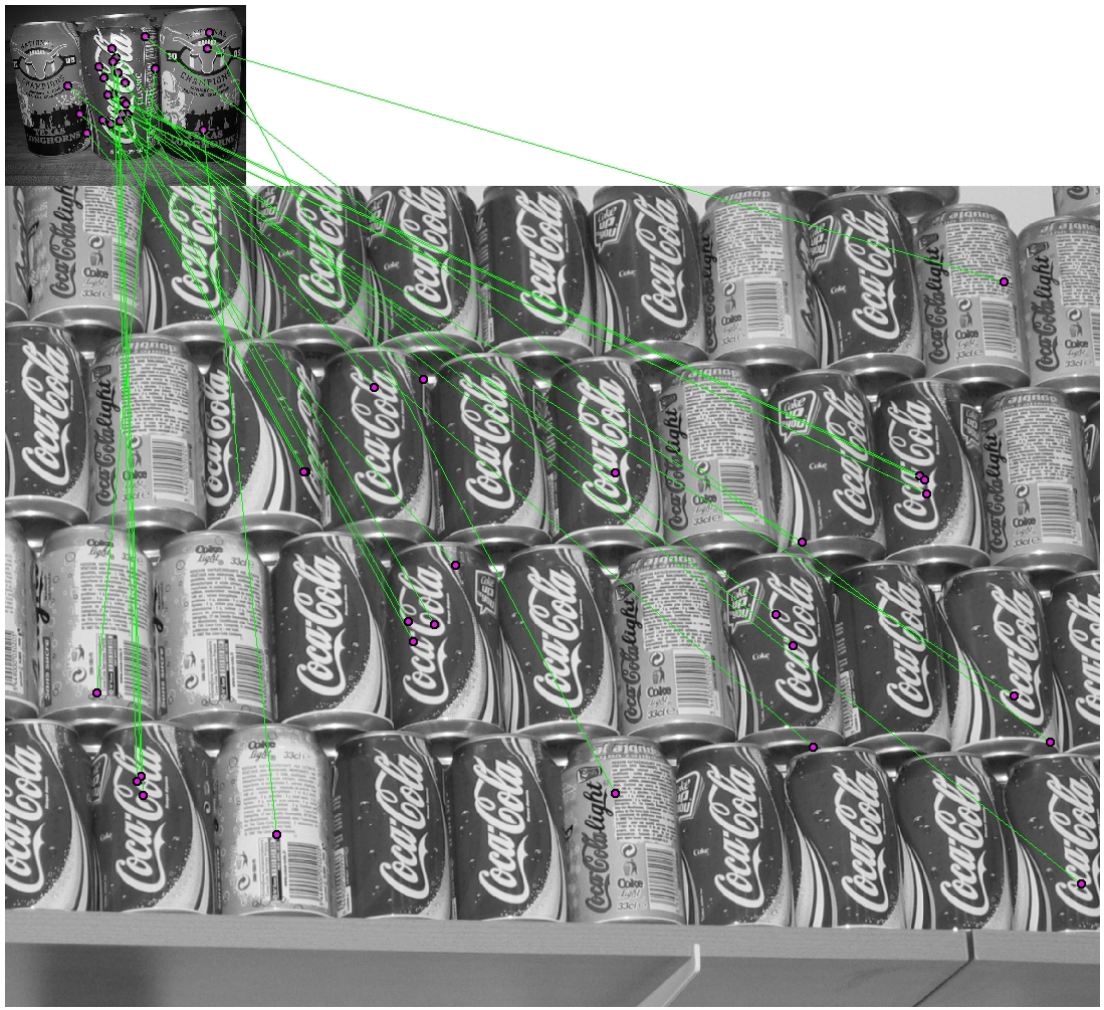
White House

AC at $\varepsilon = 10^{-1}$.



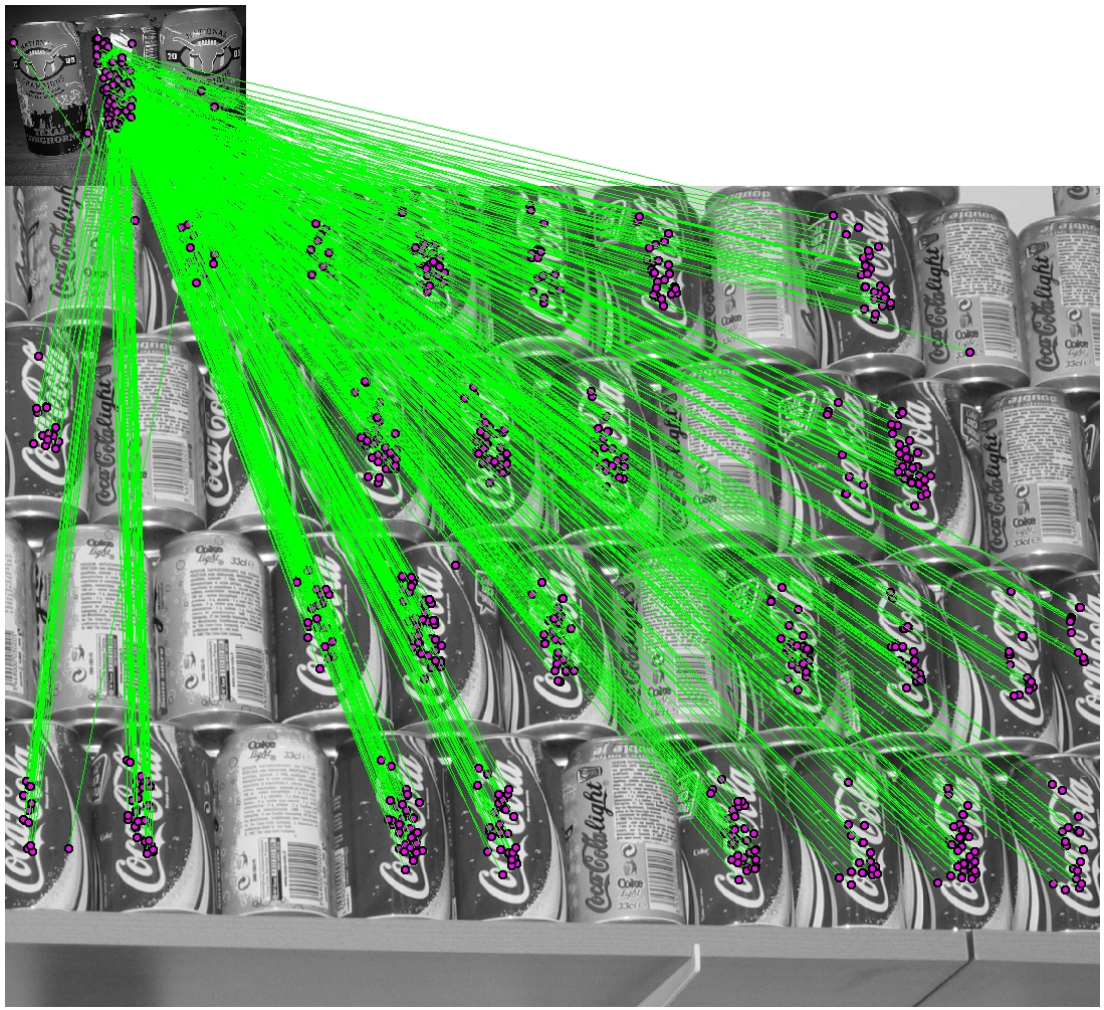
Cans

NN-DR with $r = 0.8$



multiple soda cans

AC with $\varepsilon = 10^{-1}$ (1115 matches between the logos out of 1120)



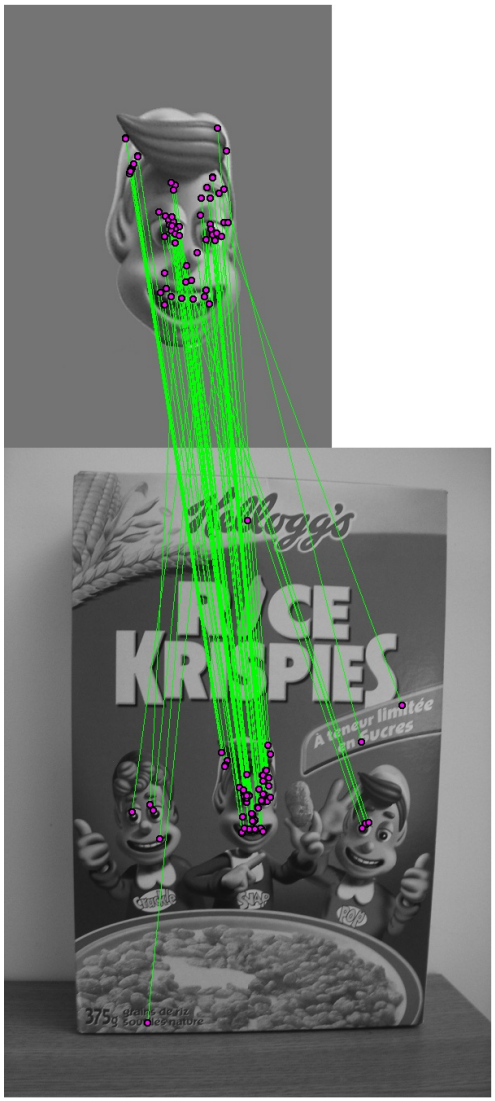
multiple soja cans

multiple soja cans

NN restriction for AC criterion

NN restriction for AC criterion

AC criterion with NN restriction



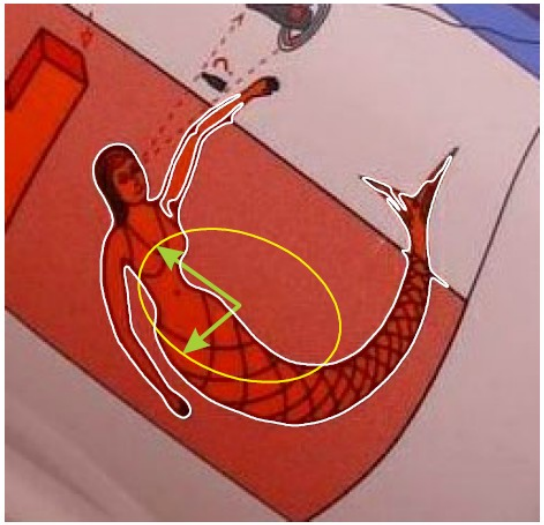
NN restriction for AC criterion

AC criterion without NN restriction

Part V

IPOL Projects

Projects #1: A SIFT descriptor for MSER

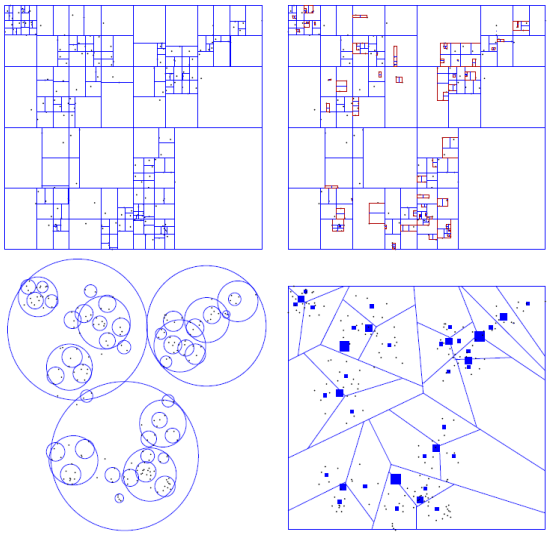


This project is devoted to the study of the SIFT description technique introduced in [Forssén and Lowe, 2007] for **Maximally Stable Extremal Regions** (MSER) detector proposed in [Matas et al., 2002] for wide baseline stereo imaging.

Goals:

1. study of the MSER algorithm which –similarly to SIFT– aims at detecting robust and repeatable local features for image comparison (codes are available for experiments).
2. study and implement the multi-scale variant of the MSER proposed by Forssen and Lowe.

Projects #2: Fast SIFT matching



This project is devoted to the study of the *M-SIFT* descriptors [Mikolajczyk and Matas, 2007] for fast similarity computation.

Motivation: For on-line image search with large database and for real-time applications, **exhaustive** are prohibitive. using such approaches are generally too much time consuming. In practice, **Approximate Nearest-Neighbor** (ANN) search approaches are often used to speed-up the matching step.

Goals:

1. analyze the SIFT variant proposed in [Mikolajczyk and Matas, 2007], which is claimed to improve SIFT matching accuracy when using ANN techniques;
2. IPOL demo for SIFT using ANN matching.

Projects #3: *A contrario* SIFT matching

This project is devoted to the study of the *a contrario* matching criterion for SIFT descriptors (co-supervised with **Pascal Monasse**)

Motivation: The major drawback AC technique is that it greatly depends on the accuracy of the null hypothesis to reject mismatches. For SIFT matching, the null hypothesis rely on the assumption that the distances between histograms are **independent**, which is not completely verified in practice for random pairs of descriptors.

Goals:

1. comparison of the AC matching criterion to other matching criteria (NN approaches, distance and ratio thresholding). Matlab and C/C++ codes are available for experimental study;
2. study the practical interest of **Principal Component Analysis** (PCA) to decorrelate histograms and to improve matching results;
3. IPOL demo for a contrario SIFT matching.

Bibliography

Baumberg, A. (2000).
Reliable feature matching across widely separated views.
In *Proc. CVPR*.

Cao, F., Lisani, J., Morel, J.-M., Musé, P., and Sur, F. (2008).
A theory of shape identification, volume 1948 of *Lecture Notes in Mathematics*.
Springer.

Desolneux, A., Ladjal, S., Moisan, L., and Morel, J.-M. (2002).
Dequantizing image orientation.
11(10):1129–1140.

Desolneux, A., Moisan, L., and Morel, J.-M. (2000).
Meaningful alignments.
Int. J. Comput. Vision, 40(1):7–23.

Forssén, P. and Lowe, D. (2007).
Shape descriptors for maximally stable extremal regions.
In *IEEE ICCV*.

Jia, J. and Tang, C.-K. (2008).
Image stitching using structure deformation.
IEEE Transactions on Pattern Analysis and Machine Intelligence, 30(4):617–631.

Lowe, D. G. (1999).
Object recognition from local scale-invariant features.
In *ICCV '99: Proceedings of the International Conference on Computer Vision-Volume 2*, page 1150, Washington, DC, USA. IEEE Computer Society.

Lowe, D. G. (2004).
Distinctive image features from scale-invariant keypoints.
Int. J. Comput. Vision, 60(2):91–110.

Matas, J., Chum, O., Urban, M., and Pajdla, T. (2002).
Robust wide baseline stereo from maximally stable extremal regions.
In *BMVC*, pages 384–393.

Mikolajczyk, K. and Matas, J. (2007).
Improving SIFT for fast tree matching by optimal linear projection.
page 8, Los Alamitos, CA, USA. IEEE Computer Society.

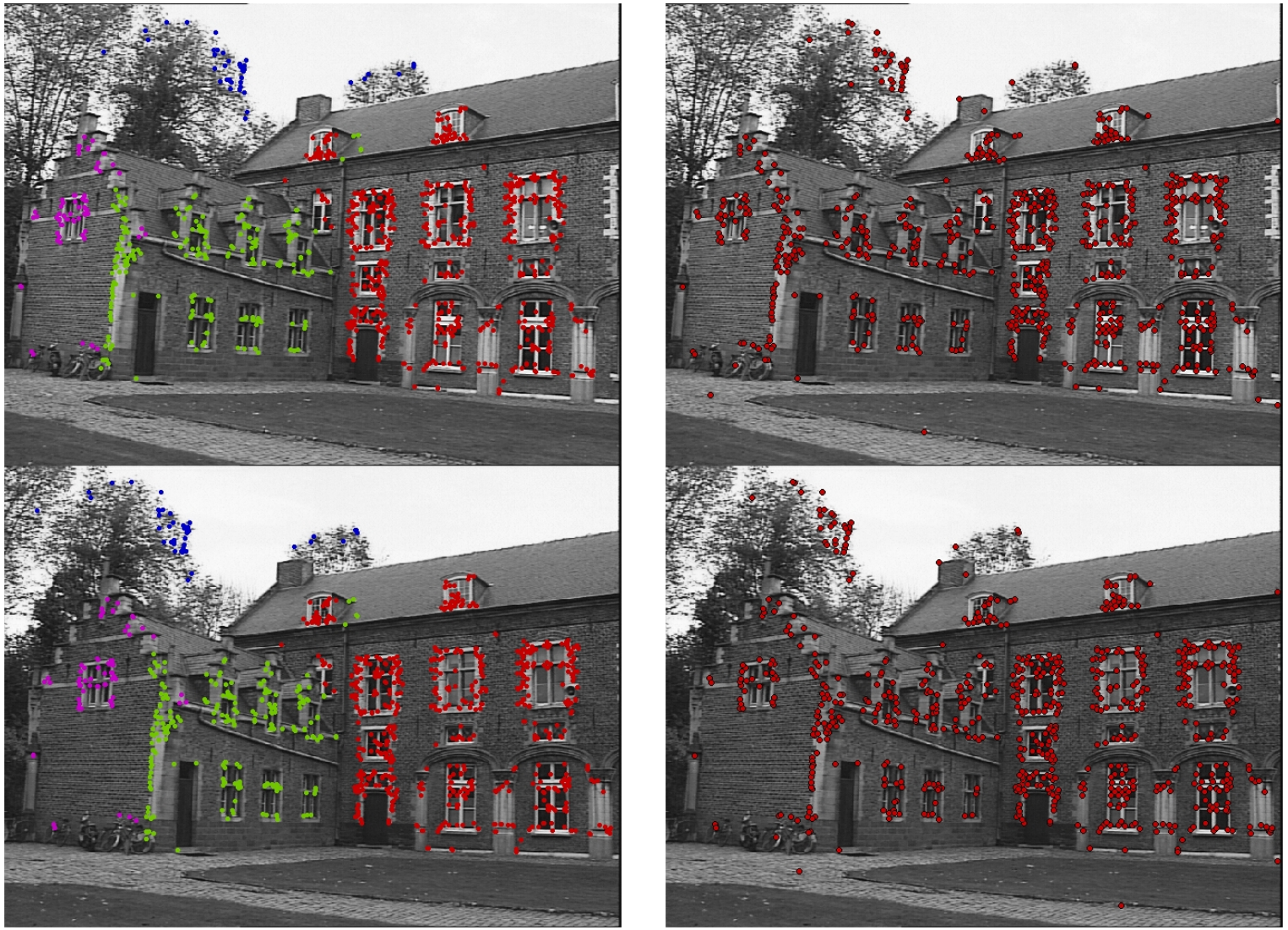
Appendix

Some more experiments (repetitive structures)

Grouping with projective or epipolar geometry constraint

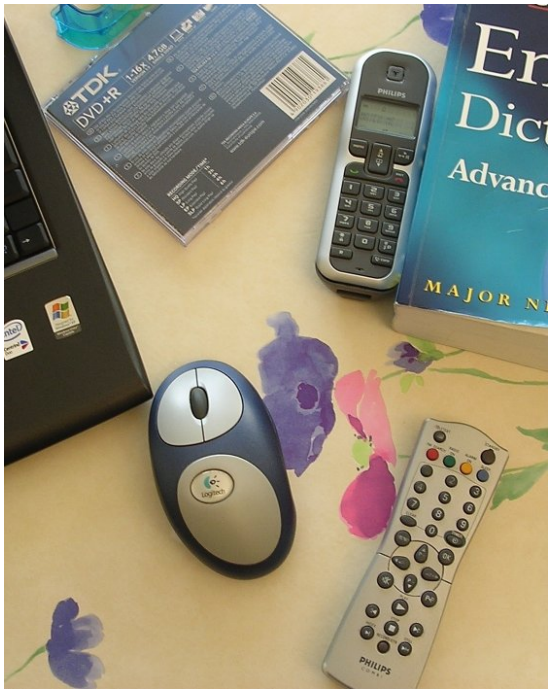
Some more experiments (repetitive structures)

Grouping with projective or epipolar geometry constraint



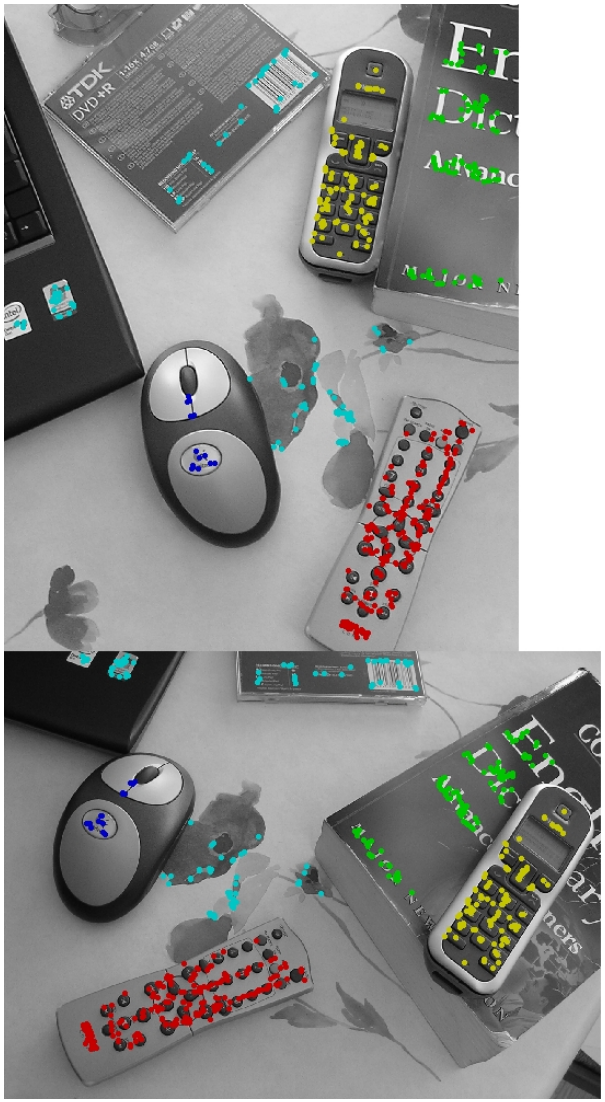
Some more experiments

Grouping with projective or epipolar geometry constraint



Some more experiments

Grouping with projective or epipolar geometry constraint



Some more experiments (Multiple occurrences)

Grouping with projective geometry constraint

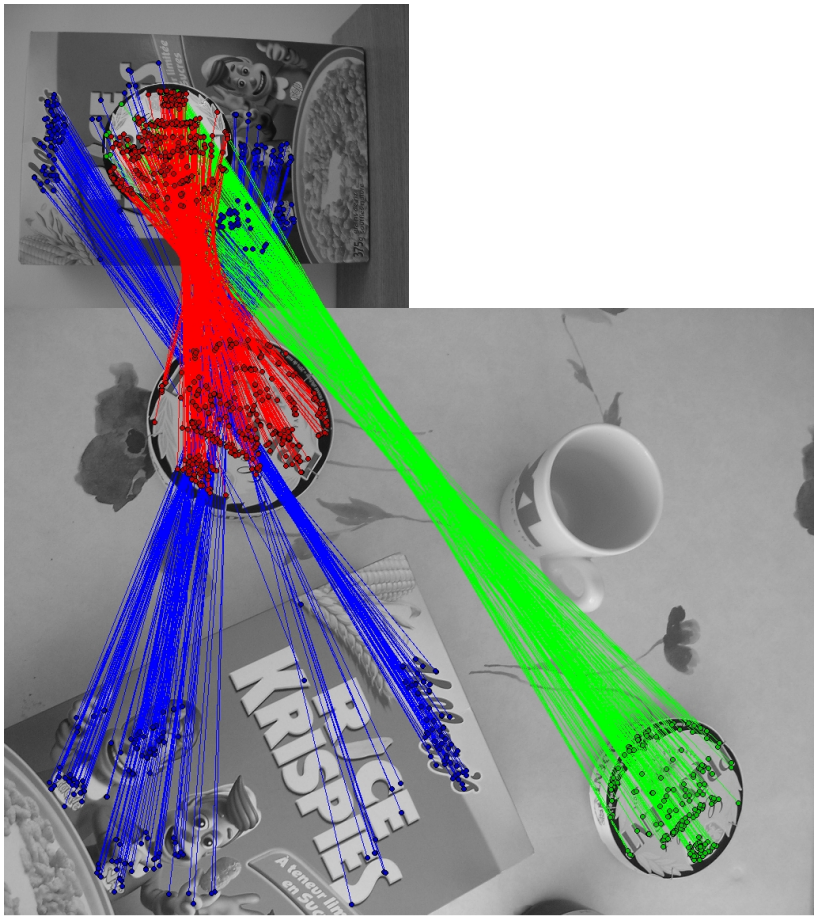
Some more experiments (Multiple occurrences)

Grouping with projective geometry constraint

Folded card

Salle Richelieu
COMÉDIE-FRANÇAISE
1680
Pedro et le commandeur

Salle Richelieu
COMÉDIE-FRANÇAISE
1680
Pedro et le commandeur

Folded card

Cheese boxes