Exponential models as generative models for image reconstruction

Bruno Galerne
bruno.galerne@univ-orleans.fr

Institut Denis Poisson
Université d'Orléans, Université de Tours, CNRS

Master MVA
Cours "Cours Modèles Stochastiques pour l'analyse d'images"
Jeudi 19 mars 2020 - **Confinement Covid-19 J4**

Outline

Last week:

- Mathematics for macrocanonical models: existence, entropy maximization, exponential models,. . .
- Sampling of macrocanonical models
- Lab session on sampling using Langevin dynamics
- Maximal entropy for texture synthesis

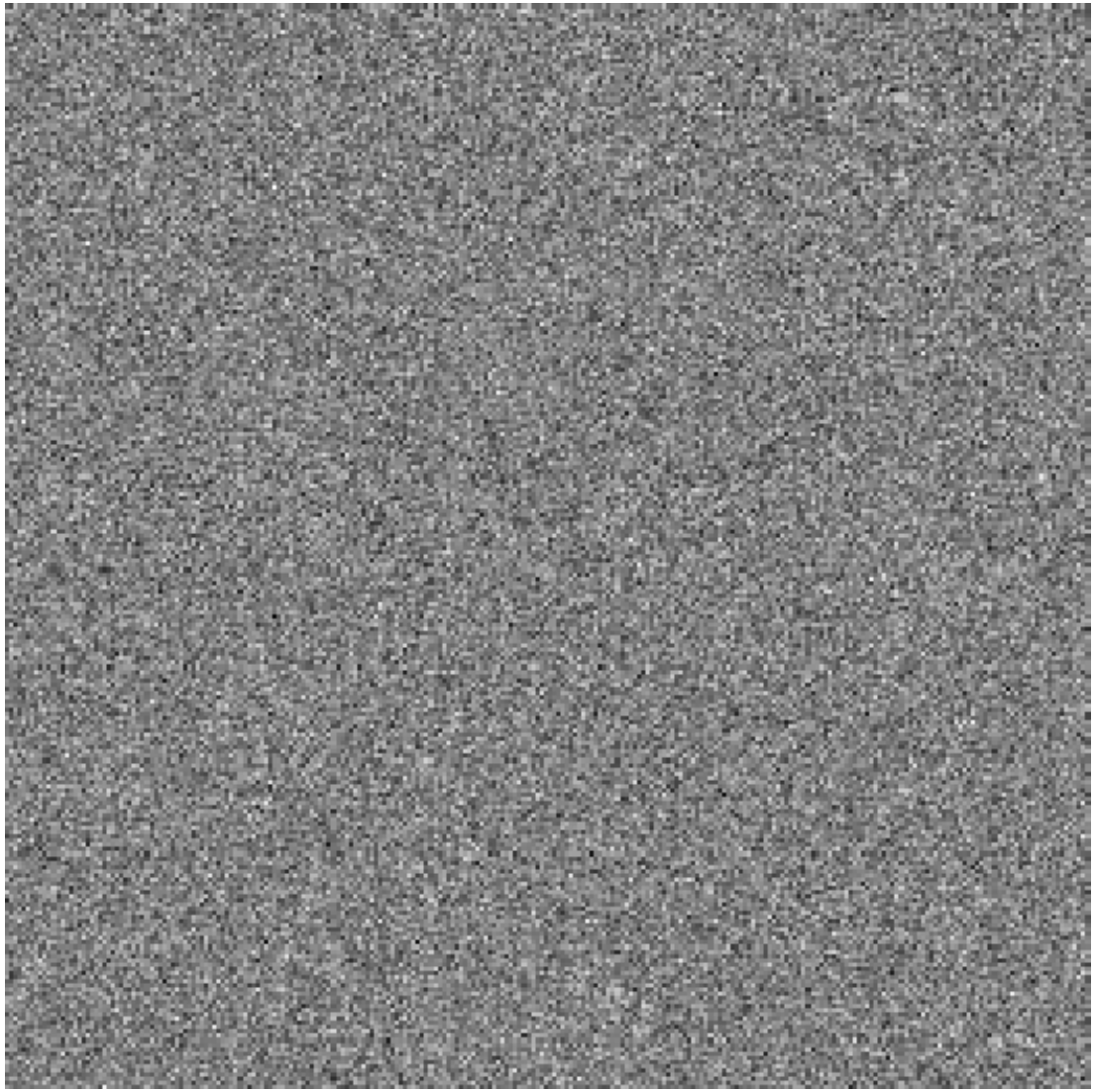
Today: Maximal entropy for image reconstruction
Two examples

1. Detection of geometric structures (line segments) in images, and reconstruction from them
2. Understanding image descriptors (SIFT)

Slides from Agnès Desolneux.

1. Reconstruction from Line Segment Detections (LSD)

Detecting geometric structures in images



What do you see ?

Helmholtz Principle (Non-accidentalness principle)

Two ways :

1. First way is common sense : *“we don’t see anything in a noise image”* (Attneave 1954)

F. Attneave, Some informational aspects of visual perception. *Psych. Rev.*, 1954.
A.P. Witkin and J. Tenenbaum. On the role of structure in vision, In *Human and Machine Vision*, 1983.

D. Lowe. *Perceptual Organization and Visual Recognition*, Kluwer Academic Publishers, 1985.

Helmholtz Principle (Non-accidentalness principle)

Two ways :

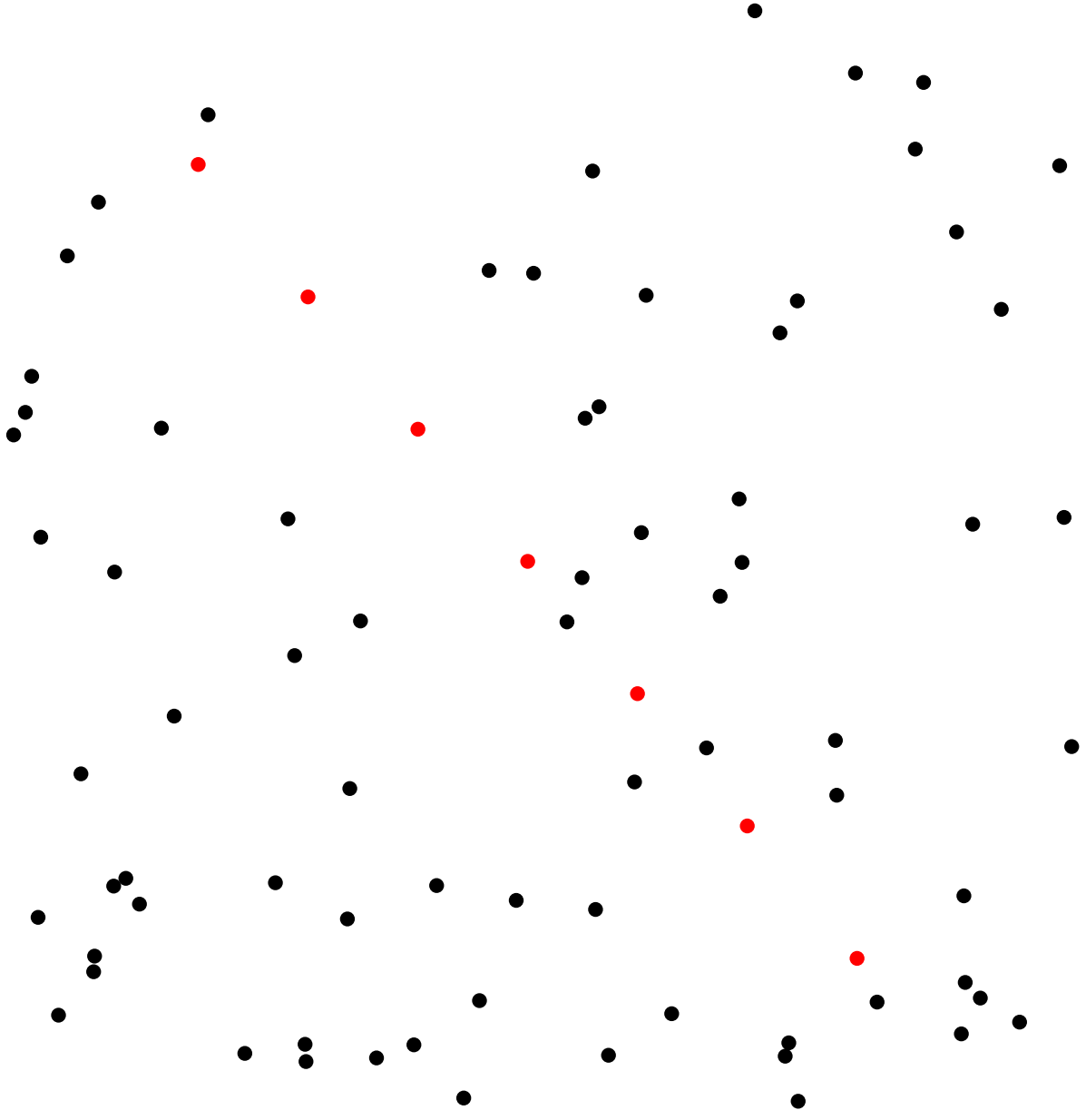
1. First way is common sense : *"we don't see anything in a noise image"* (Attneave 1954)
2. Stronger statement : *"we perceive what has a low probability of arriving by accident"*, in other words *"if a large deviation from randomness occurs, then a structure is perceived"* (Witkin and Tenenbaum 1983, Lowe 1985)

F. Attneave, Some informational aspects of visual perception. *Psych. Rev.*, 1954.
A.P. Witkin and J. Tenenbaum. On the role of structure in vision, In *Human and Machine Vision*, 1983.

D. Lowe. *Perceptual Organization and Visual Recognition*, Kluwer Academic Publishers, 1985.

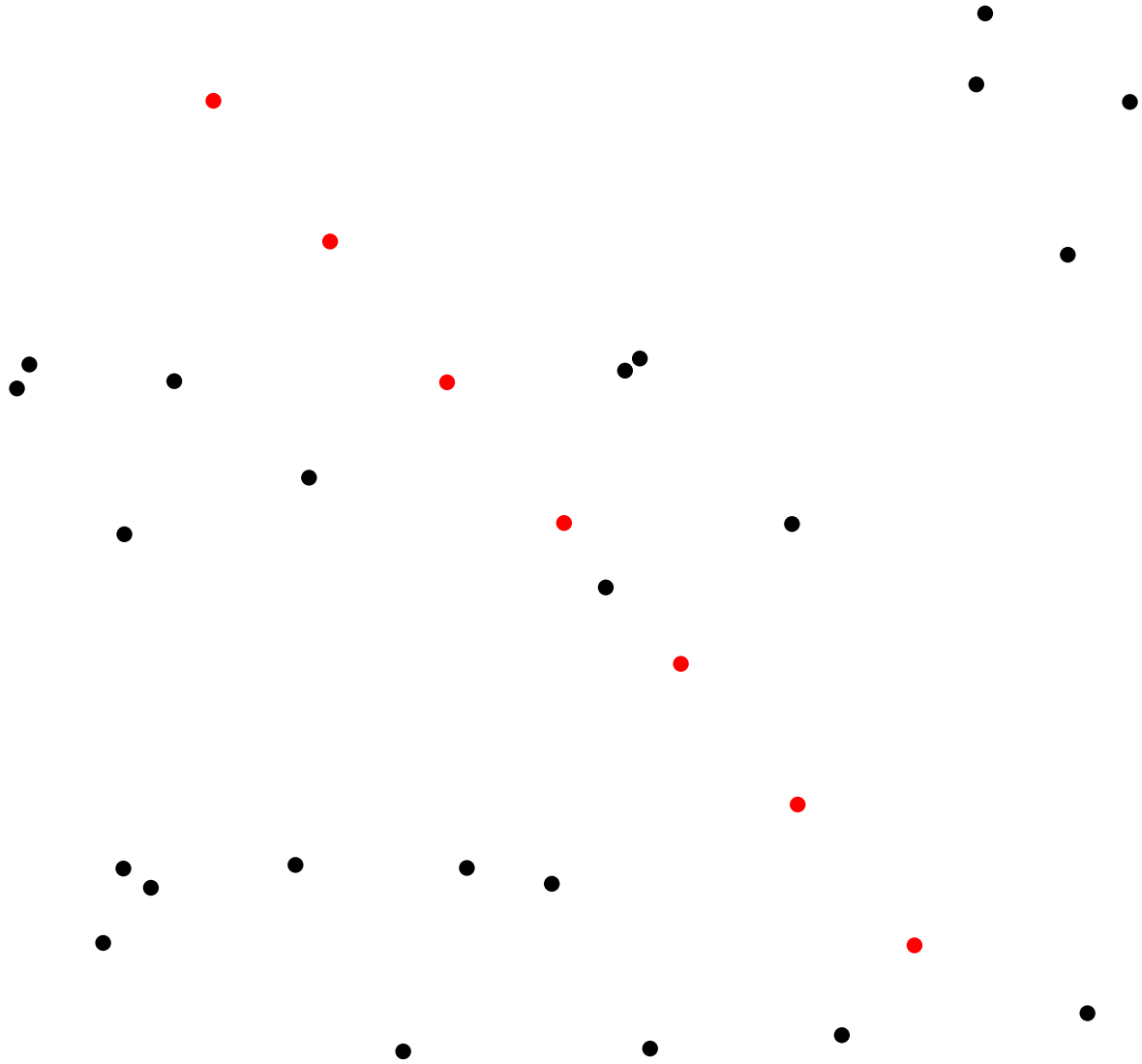
Illustrating example

Illustrating example



Illustrating example

Illustrating example



A contrario framework

General framework : just need an *a contrario* model of what the image is not (pure noise)

- Observe E a geometric event in an image

A. Desolneux, L. Moisan, and J.-M. Morel. *From Gestalt Theory to Image Analysis : A Probabilistic Approach*, Springer-Verlag, 2008.

A contrario framework

General framework : just need an *a contrario* model of what the image is not (pure noise)

- Observe E a geometric event in an image
- Compute $\mathrm{NFA}(E)$ that is the expected number of occurrences of E in an image following the a contrario model.

A. Desolneux, L. Moisan, and J.-M. Morel. *From Gestalt Theory to Image Analysis : A Probabilistic Approach*, Springer-Verlag, 2008.

A contrario framework

General framework : just need an *a contrario* model of what the image is not (pure noise)

- Observe E a geometric event in an image
- Compute $\mathrm{NFA}(E)$ that is the expected number of occurrences of E in an image following the a contrario model.
- **Definition :** Let $\varepsilon > 0$. If $\mathrm{NFA}(E) < \varepsilon$ then E is called *an ε-meaningful event*.

A. Desolneux, L. Moisan, and J.-M. Morel. *From Gestalt Theory to Image Analysis : A Probabilistic Approach*, Springer-Verlag, 2008.

Detection of straight segments in an image

LSD (Line Segment Detector) algorithm of Grompone et. al. (2010) :

Let $\Omega = \{1, \ldots, M\} \times \{1, \ldots, N\}$ be a discrete domain and let $u_0 : \Omega \to \mathbb{R}$ be an image. It orientation field is $\theta_0 : \Omega \to S^1 = [0, 2\pi)$ given by

$$\forall x \in \Omega, \quad \theta_0(x) = \frac{\pi}{2} + \operatorname{Arg} \frac{\nabla u_0(x)}{\|\nabla u_0(x)\|},$$

where

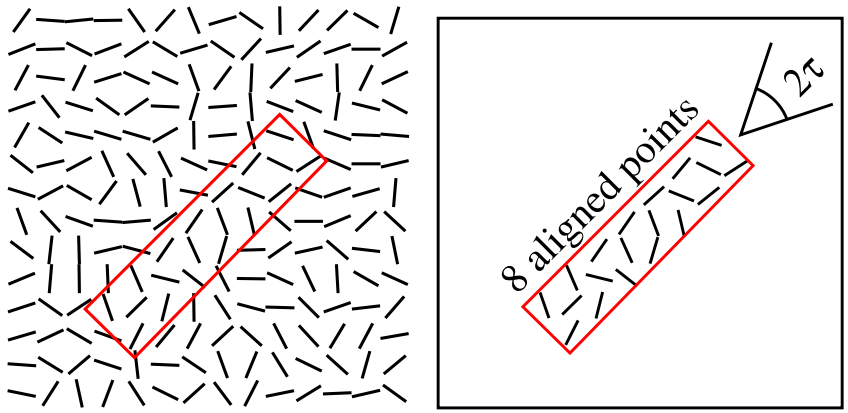
X1	X2
X3	X4

$$\nabla u_0 = \tfrac{1}{2} \begin{pmatrix} X_2 - X_1 + X_4 - X_3 \\ X_3 - X_1 + X_4 - X_2 \end{pmatrix}.$$

R. Grompone von Gioi, J. Jakubowicz, J.-M. Morel, G. Randall, A Fast Line Segment Detector with a False Detection Control, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010.

Let $r \subset \Omega$ be a rectangle with principal orientation $\varphi(r)$. Define the number of aligned pixels it contains by :

$$k(r;\theta_0) := \sum_{x \in r} \mathbb{1}_{|\theta_0(x) - \varphi(r)| \leqslant p\pi},$$



R. Grompone von Gioi, J. Jakubowicz, J.-M. Morel, and G. Randall, LSD : a Line Segment Detector, *Image Processing On Line (IPOL)*, 2012.

- Define a « pure noise model », that is a law P on orientation fields Θ, given here by : the $\Theta(x)$ are independent identically distributed uniformly on $[0, 2\pi)$.

- Define a « pure noise model », that is a law P on orientation fields Θ, given here by : the $\Theta(x)$ are independent identically distributed uniformly on $[0, 2\pi)$.

- Define the number of false alarms of the rectangle r in θ_0, under the a contrario model P by

$$\mathrm{NFA}_P(r;\theta_0) = N_{tests} \times \mathbb{P}_P[k(r;\Theta) \geqslant k(r;\theta_0)],$$

where N_{tests} is the number of tests, that is the number of rectangles in a $M \times N$ image ($\simeq (MN)^{5/2}$).

- In this definition, Θ is a random orientation field following the law P and $k(r;\Theta) = \sum_{x \in r} \mathbb{1}_{|\Theta(x) - \varphi(r)| \leqslant p\pi}$ is then a random variable following a binomial distribution of parameters $n(r) = \#r$ and p.

- Define a « pure noise model », that is a law P on orientation fields Θ, given here by : the $\Theta(x)$ are independent identically distributed uniformly on $[0, 2\pi)$.

- Define the number of false alarms of the rectangle r in θ_0, under the a contrario model P by

$$\mathrm{NFA}_P(r;\theta_0) = N_{tests} \times \mathbb{P}_P[k(r;\Theta) \geqslant k(r;\theta_0)],$$

where N_{tests} is the number of tests, that is the number of rectangles in a $M \times N$ image ($\simeq (MN)^{5/2}$).

- In this definition, Θ is a random orientation field following the law P and $k(r;\Theta) = \sum_{x \in r} \mathbb{1}_{|\Theta(x) - \varphi(r)| \leqslant p\pi}$ is then a random variable following a binomial distribution of parameters $n(r) = \#r$ and p.

- When $\mathrm{NFA}_P(r;\theta_0) < \varepsilon$, we say that the rectangle r is ε-meaningful.

Main property of the NFA

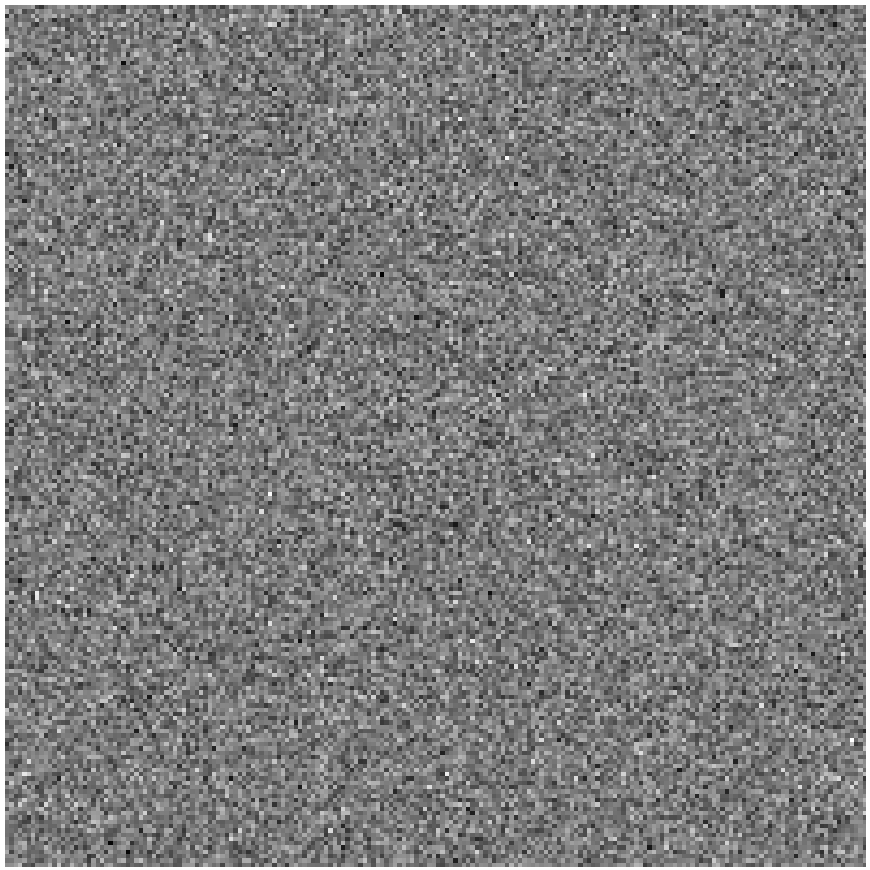
Proposition

When Θ is a random orientation field following the law P, then the $\mathrm{NFA}_P(r_i;\Theta)$, $1 \leqslant i \leqslant N_{tests}$, become random variables and

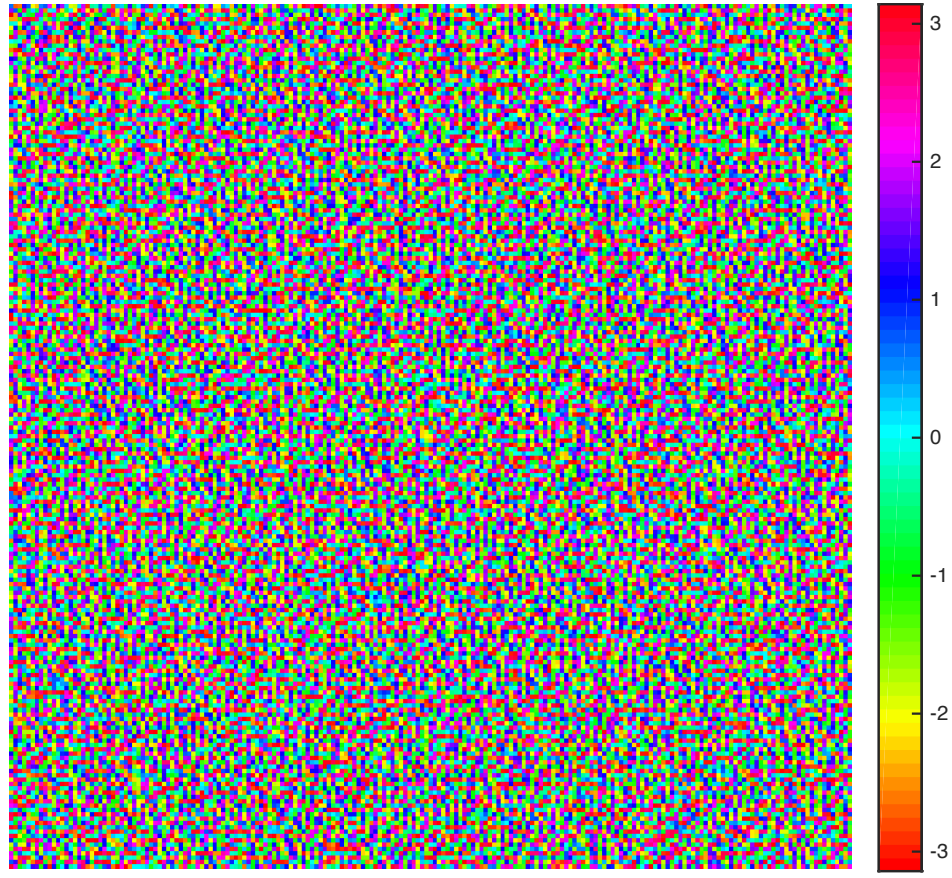
$$\mathbb{E}_P\left(\sum_{i=1}^{N_{test}} \mathbb{1}_{\mathrm{NFA}_P(r_i;\Theta)<\varepsilon}\right) < \varepsilon.$$

In other words, the expected number (under law P) of ε-meaningful segments is less than ε.

Example 0

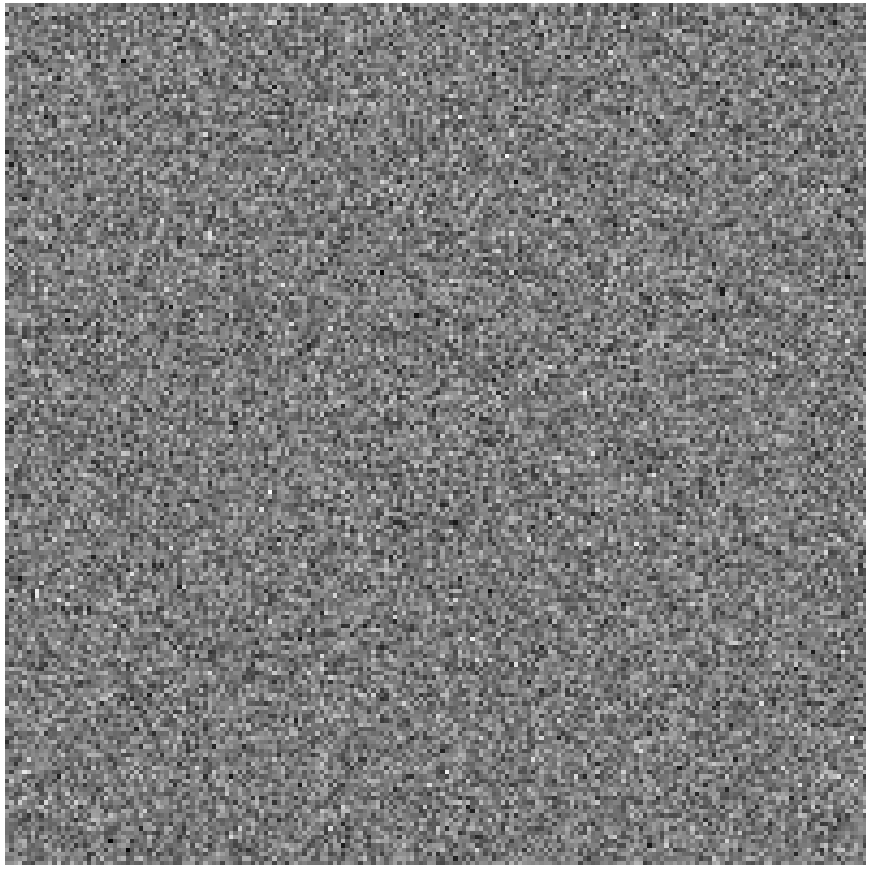


Original image u_0



Orientation field θ_0

Example 0

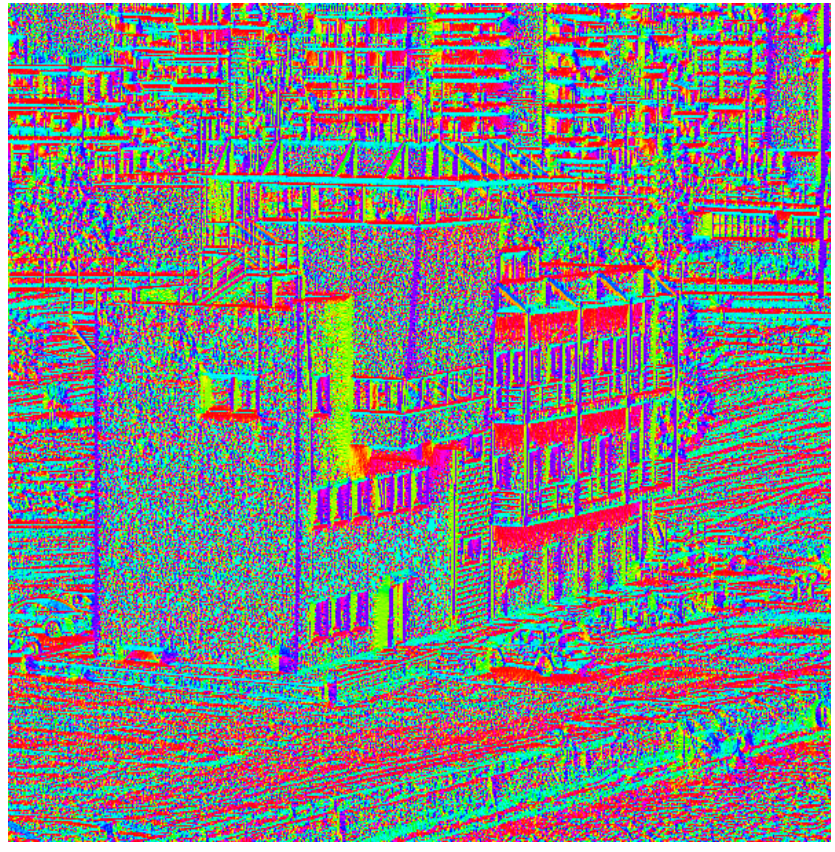


Original image u_0

LSD result

Example 1

Original image u_0



Orientation field θ_0

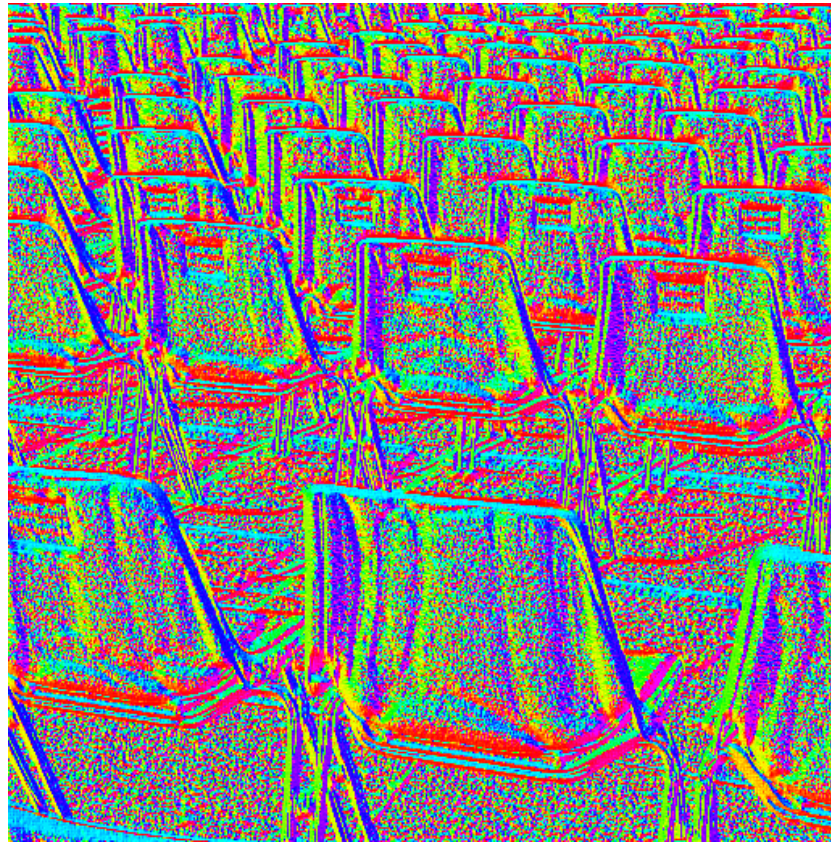
Example 1

Original image u_0

LSD result

Example 2

Original image u_0



Orientation field θ_0

Example 2

Original image u_0

LSD result

- A contrario approach : a computational approach to detect geometric structures in images.
- Many other applications : shape recognition, image matching, clustering, stereovision, denoising (grain filter), etc.
- Can we turn it into a generative approach ?

Some questions about the *a contrario* methodology

In the *a contrario methodology*, we use a noise model (null-hypothesis) H_0. But

- What happens if we change H_0 ?
- Given some observed geometric events, are there laws H_0 that can « explain » them ?
- If yes, which one is at the same time « as random as possible » ?
- What do the samples of it look like ?

A. Desolneux, When the a contrario approach becomes generative, *International Journal of Computer Vision*, 2016.

Changing the a contrario noise model

In the LSD algorithm, the law U (i.i.d. uniform) on orientation fields is used.
Let the precision p be fixed and let also $\varepsilon > 0$ be fixed.

Definition

Let $\theta^0 : \Omega \to S^1$ be an orientation field. Let $r_1, \ldots, r_m$ be the (disjoint) rectangles detected by the LSD in θ^0.
We define 3 sets of laws :

- Let $\mathcal{P}$ be the set of laws P on Θ such that none of $r_1, \ldots, r_m$ is meaningful under P :

$$\forall 1 \leqslant j \leqslant m, \quad \mathrm{NFA}_P(r_j; \theta^0) := N_{tests} \times \mathbb{P}_P[k(r; \Theta) \geqslant k(r; \theta^0)] \geqslant \varepsilon.$$

- Let $\mathcal{Q}$ be the set of laws Q on Θ such that θ is sampled from Q, then, « in most cases, the LSD will provide the same detections as in θ^0 ». That is :

$$Q \in \mathcal{Q} \iff \forall 1 \leqslant j \leqslant m, \quad \mathrm{Med}_Q(\mathrm{NFA}_U(r_j; \Theta)) \leqslant \mathrm{NFA}_U(r_j; \theta^0).$$

- Finally, let $\mathcal{I}$ be the set of laws on Θ such that the $\Theta(x)$ are independent (but not necessarily identically distributed).

Characterizing $\mathcal{P}$ and $\mathcal{Q}$

Characterizing $\mathcal{P}$:

$$P \in \mathcal{P} \iff \forall 1 \leqslant j \leqslant m,\ \mathbb{P}_P[k(r_j;\Theta) \geqslant k(r_j;\theta^0)] \geqslant \frac{\varepsilon}{N_{tests}}.$$

Characterizing $\mathcal{Q}$:

$$\begin{aligned} Q \in \mathcal{Q} &\iff \forall 1 \leqslant j \leqslant m,\ \mathrm{Med}_Q(\mathrm{NFA}_U(r_j;\Theta)) \leqslant \mathrm{NFA}_U(r_j;\theta^0) \\ &\iff \forall 1 \leqslant j \leqslant m,\ \mathbb{P}_Q[\mathrm{NFA}_U(r_j;\Theta) \leqslant \mathrm{NFA}_U(r_j;\theta^0)] \geqslant \frac{1}{2} \\ &\iff \forall 1 \leqslant j \leqslant m,\ \mathbb{P}_Q[k(r_j;\Theta) \geqslant k(r_j;\theta^0)] \geqslant \frac{1}{2}. \end{aligned}$$

This implies in particular that

$$\mathcal{Q} \subset \mathcal{P}.$$

Laws in $\mathcal{P} \cap \mathcal{I}$ and $\mathcal{Q} \cap \mathcal{I}$

When $\Theta \sim P \in \mathcal{P} \cap \mathcal{I}$, then

$$k(r;\Theta) = \sum_{x \in r} \mathbb{1}_{|\Theta(x) - \varphi(r)| \leqslant p\pi}$$

follows a Poisson binomial distribution of parameters $\{q_x\}_{x \in r_j}$ where

$$q_x := \mathbb{P}_P[|\Theta_x - \varphi(r_j)| \leqslant p\pi] = \int_{\varphi(r_j)-p\pi}^{\varphi(r_j)+p\pi} f_P^{(x)}(\theta_x)\, d\theta_x.$$

Laws in $\mathcal{P} \cap \mathcal{I}$ and $\mathcal{Q} \cap \mathcal{I}$

When $\Theta \sim P \in \mathcal{P} \cap \mathcal{I}$, then

$$k(r;\Theta) = \sum_{x \in r} \mathbb{1}_{|\Theta(x) - \varphi(r)| \leqslant p\pi}$$

follows a Poisson binomial distribution of parameters $\{q_x\}_{x \in r_j}$ where

$$q_x := \mathbb{P}_P[|\Theta_x - \varphi(r_j)| \leqslant p\pi] = \int_{\varphi(r_j)-p\pi}^{\varphi(r_j)+p\pi} f_P^{(x)}(\theta_x)\, d\theta_x.$$

Proposition

1. *The law $P_0 \in \mathcal{P} \cap \mathcal{I}$ with probability density $f_{P_0}(\theta) = \prod_x f_{P_0}^{(x)}(\theta_x)$ with*

$$f_{P_0}^{(x)}(\theta_x) = \begin{cases} \frac{1}{2\pi} & \text{si } x \notin \cup_{j=1}^m r_j \\ \frac{1}{2p\pi} B^{-1}_{n(r_j),k(r_j;\theta^0)}\left(\frac{\varepsilon}{N_{tests}}\right) & \text{if } x \in r_j \text{ and } |\theta_x - \varphi(r_j)| \leqslant p\pi \\ \frac{1}{2(1-p)\pi}\left(1 - B^{-1}_{n(r_j),k(r_j;\theta^0)}\left(\frac{\varepsilon}{N_{tests}}\right)\right) & \text{if } x \in r_j \text{ and } |\theta_x - \varphi(r_j)| > p\pi \end{cases}$$

is a local maximum of entropy in $\mathcal{P} \cap \mathcal{I}$.

2. *The law $Q_0 \in \mathcal{Q} \cap \mathcal{I}$ that has an analogous form as P_0, but where*

$$\text{we replace } B^{-1}_{n(r_j),k(r_j;\theta^0)}\left(\frac{\varepsilon}{N_{tests}}\right) \text{ by } B^{-1}_{n(r_j),k(r_j;\theta^0)}\left(\frac{1}{2}\right) \simeq \frac{k(r_j;\theta^0)}{n(r_j)},$$

is a local maximum of entropy in $\mathcal{Q} \cap \mathcal{I}$.

Example 1

Original image I^0

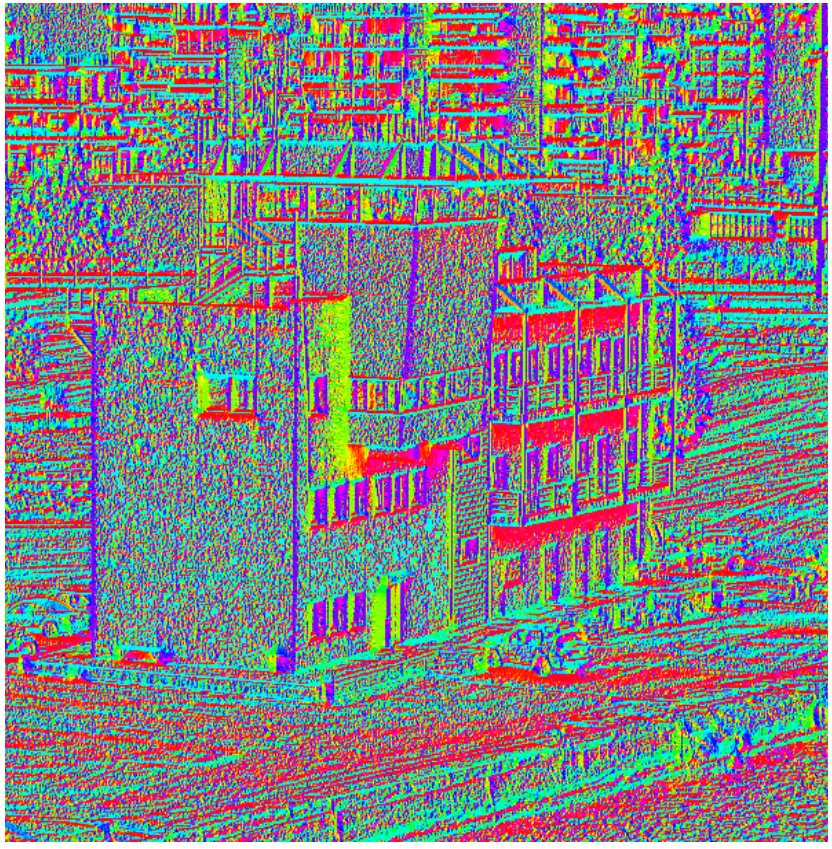
LSD result

Example 1

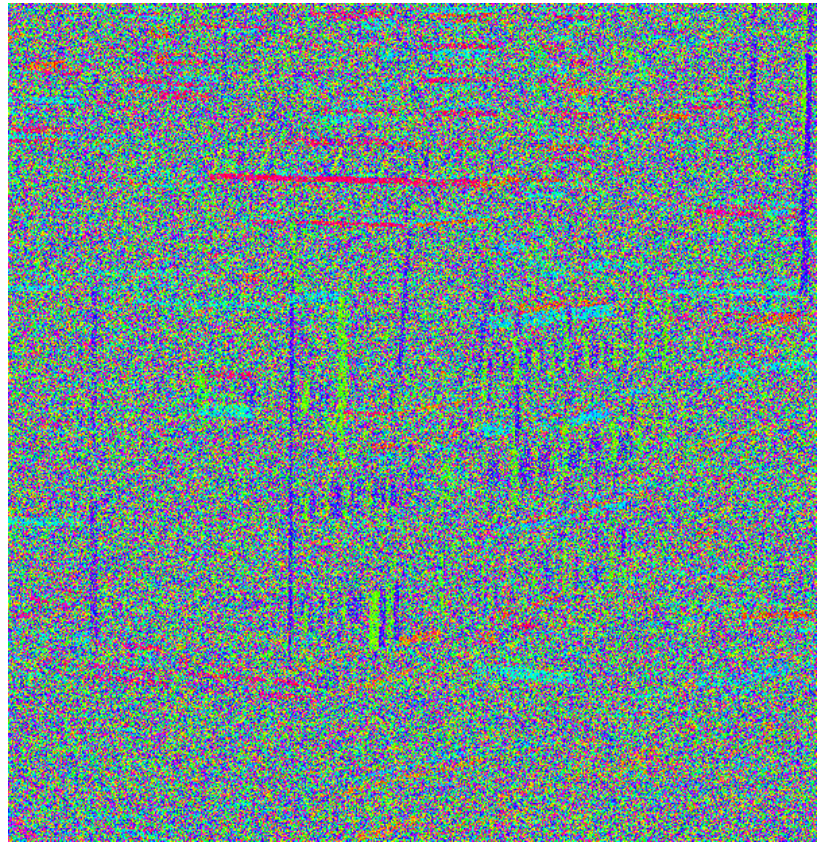
Original image I^0

Rectangles of the LSD

Example 1

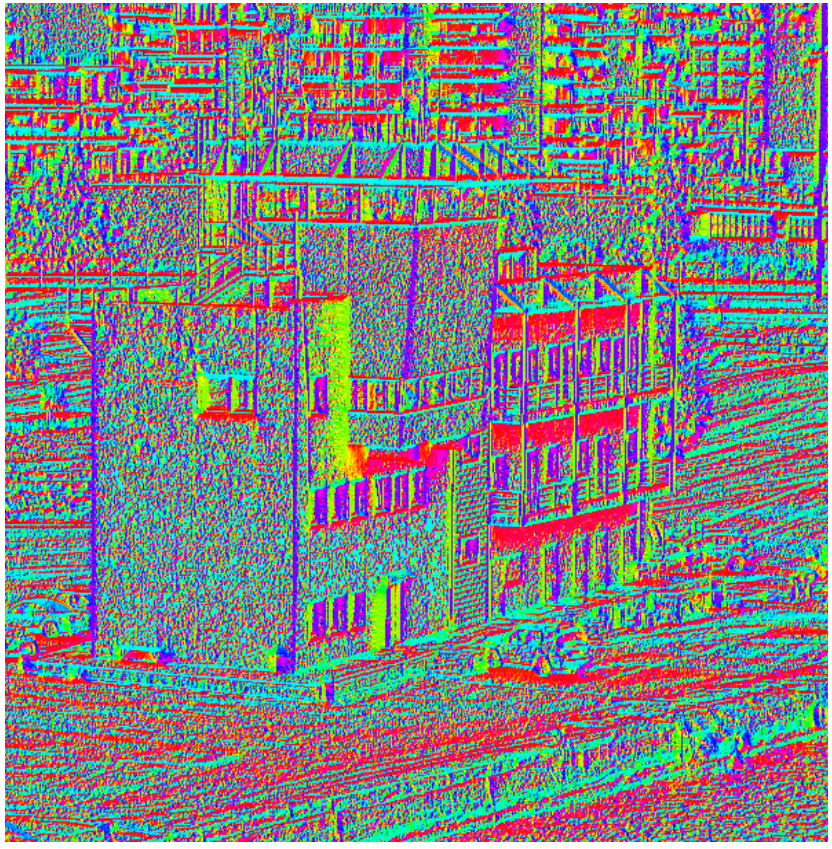


Orientation field θ^0

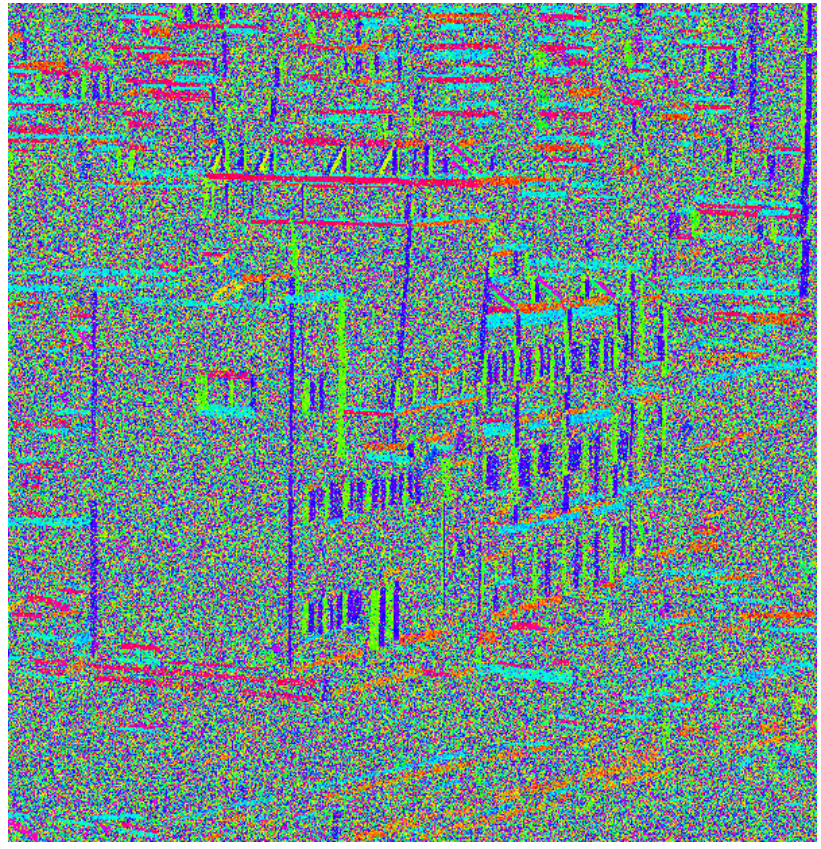


Sample from P_0

Example 1



Orientation field θ^0



Sample from Q_0

Example 2

Original image I^0

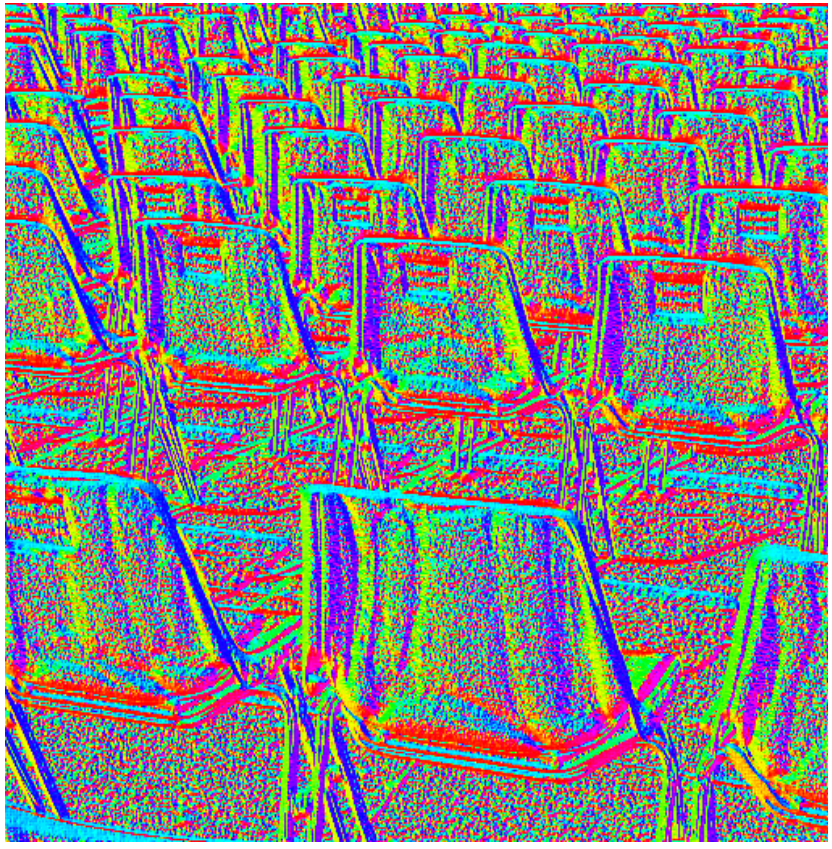
LSD result

Example 2

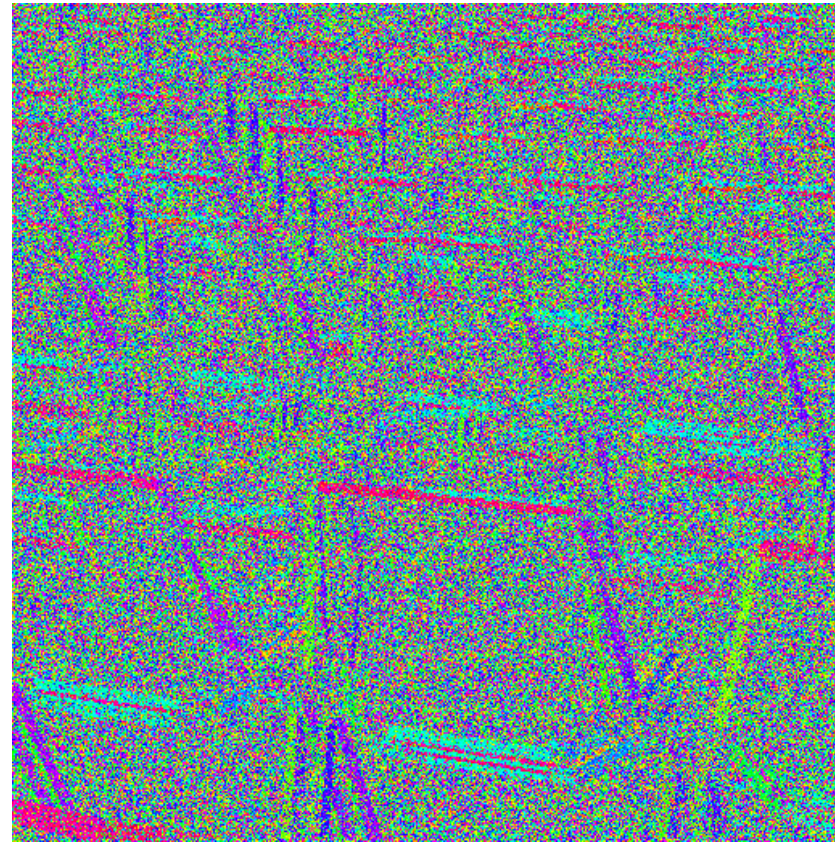
Original image I^0

Rectangles of the LSD

Example 2

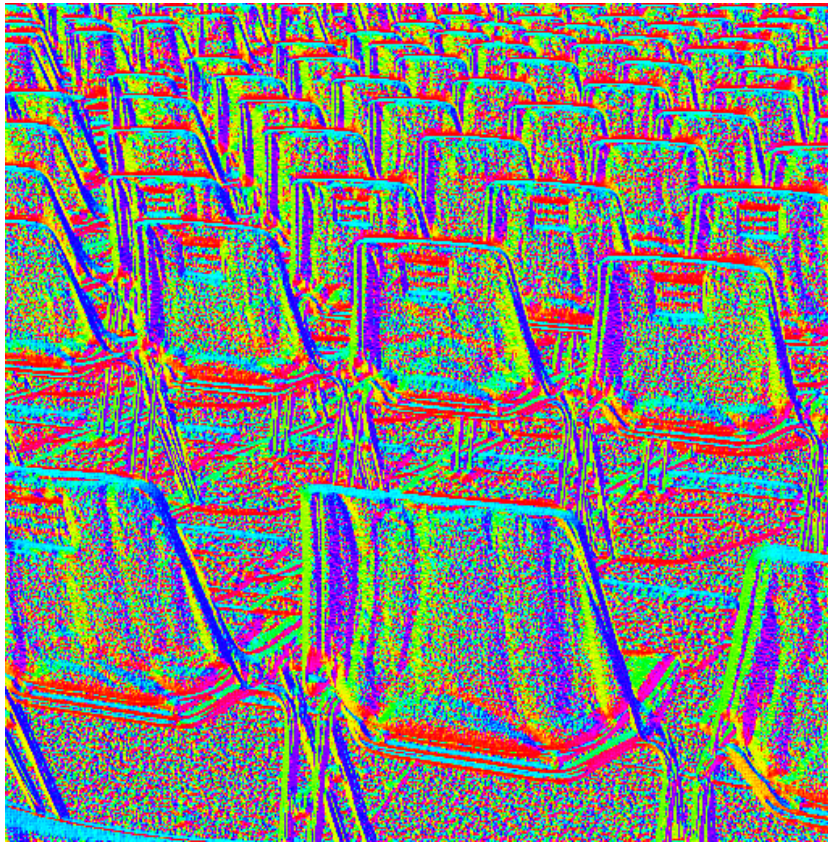


Orientation field θ^0

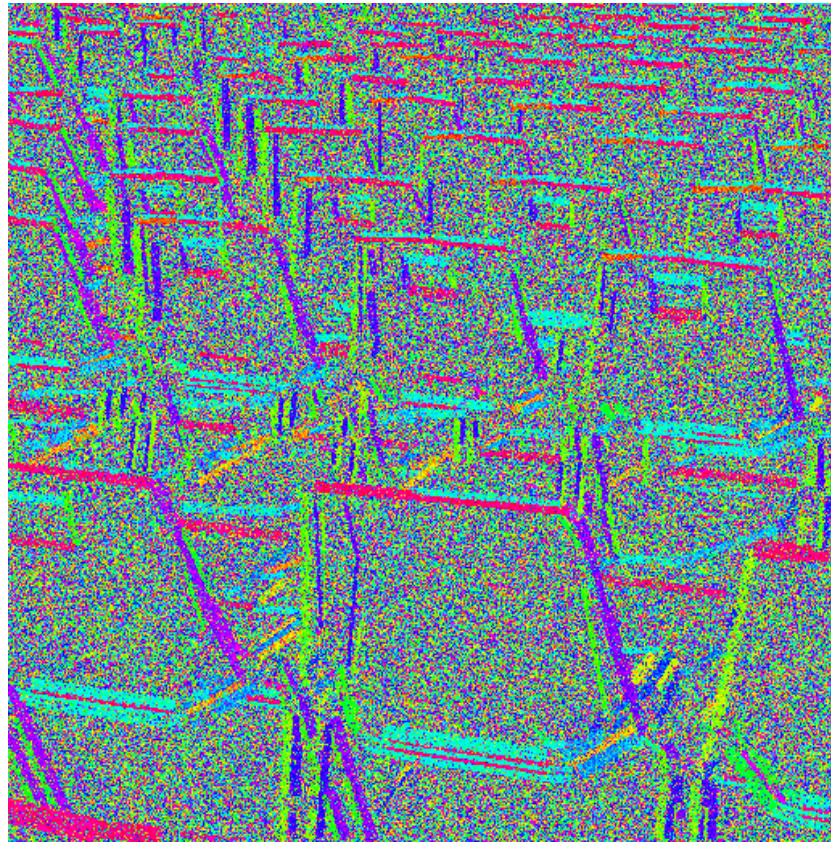


Sample from P_0

Example 2



Orientation field θ^0



Sample from Q_0

Example 3

Original image I^0

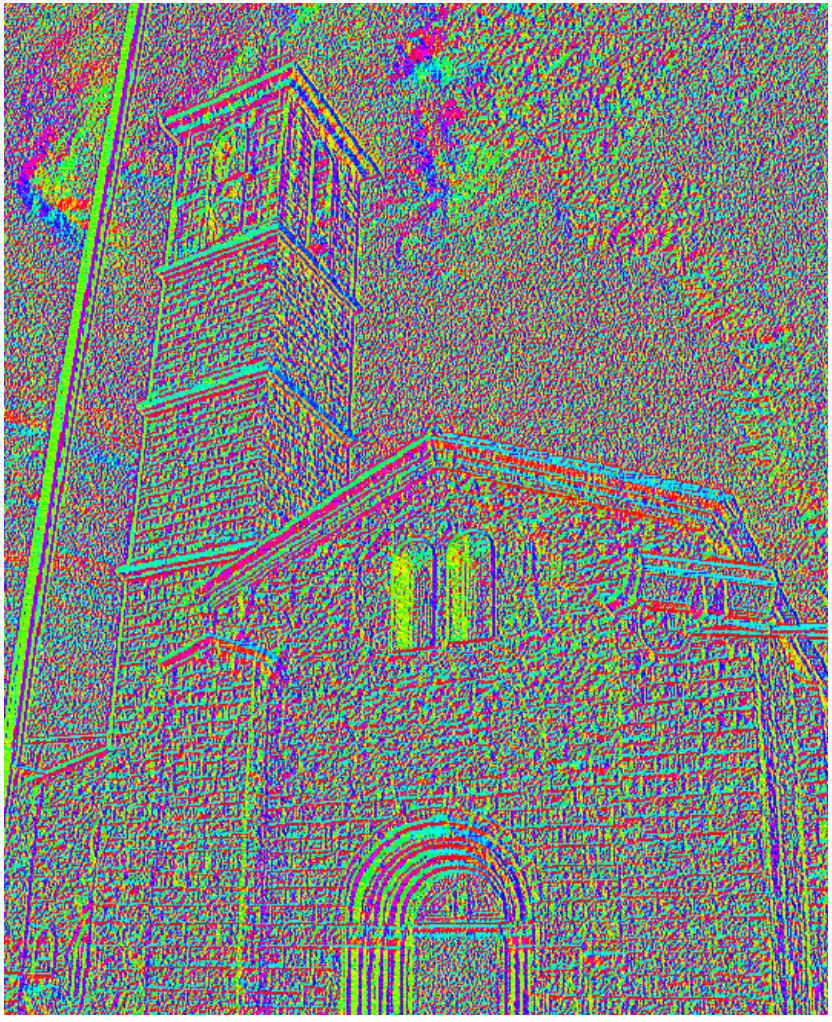
LSD result

Example 3

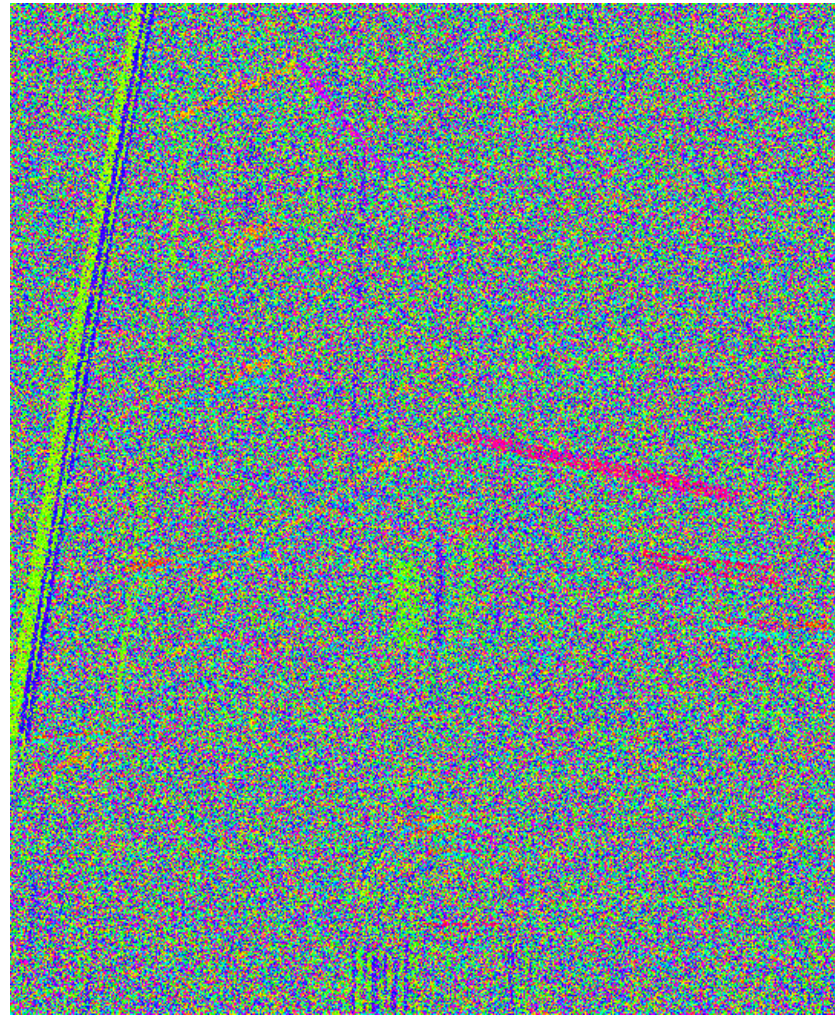
Original image I^0

Rectangles of the LSD

Example 3

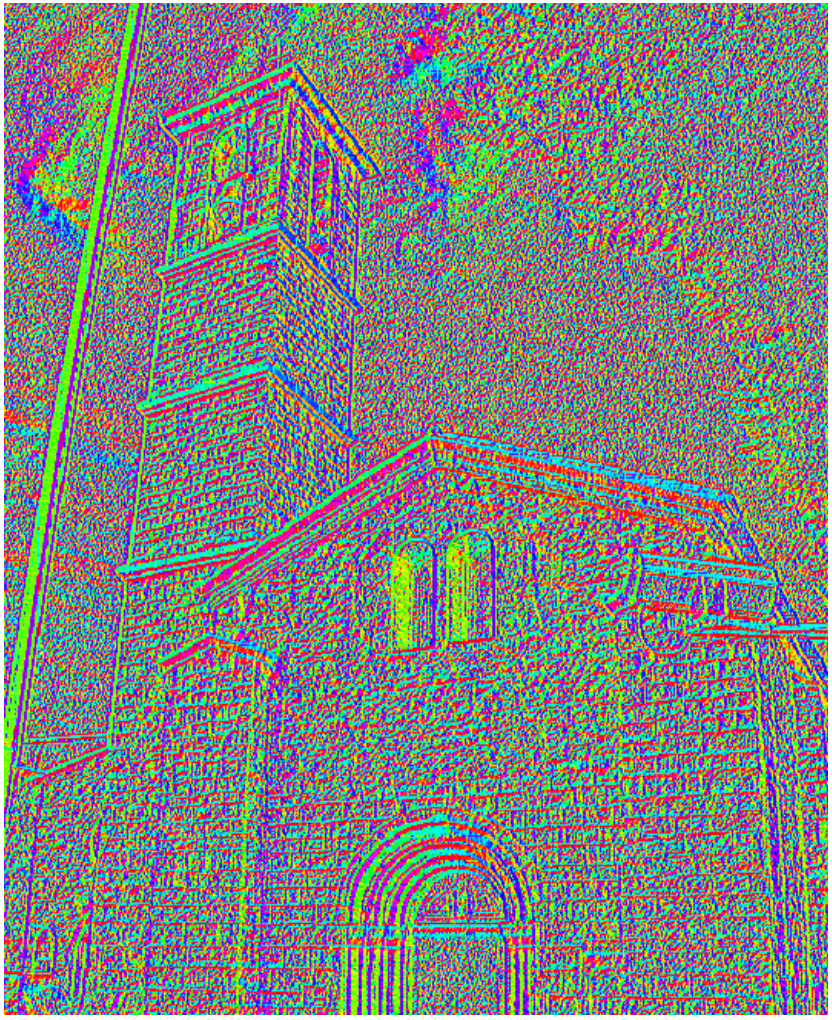


Orientation field θ^0

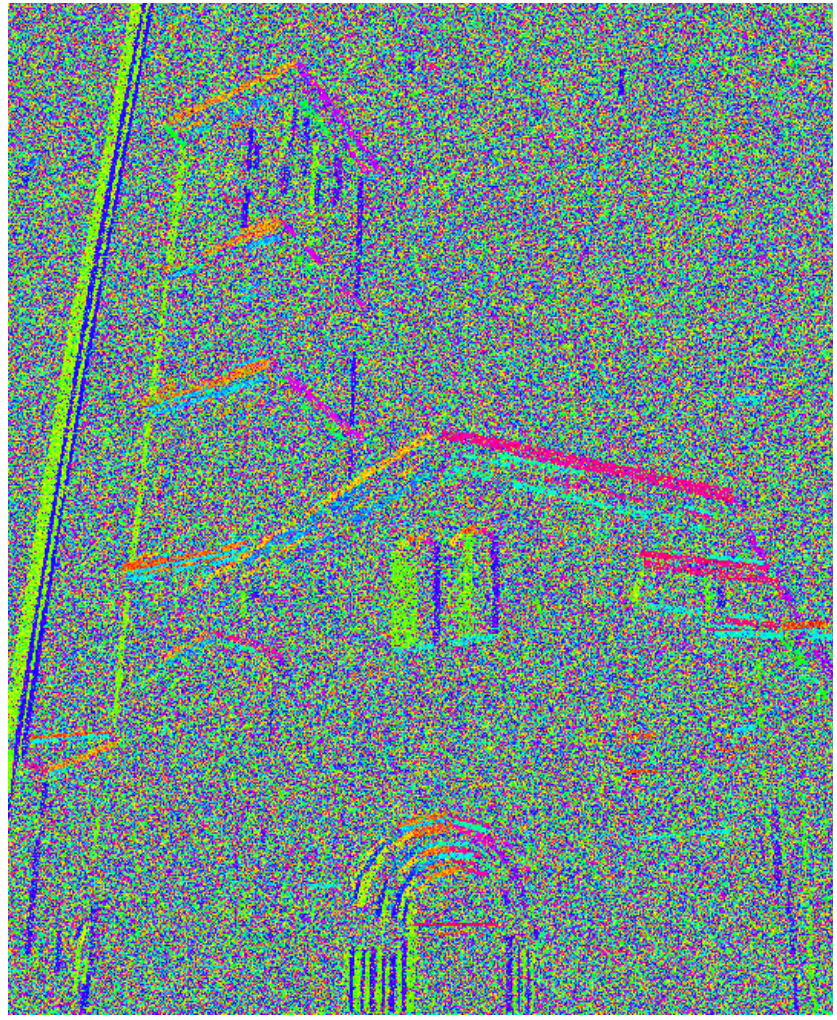


Sample from P_0

Example 3



Orientation field θ^0



Sample from Q_0

Reconstructing an image from an orientation field

Question : How to reconstruct an image from an orientation field θ ?

$\longrightarrow$ Use the Poisson editing method of Perez et. al. (2003) :
look for u image on Ω such that

$$\sum_{x\in\Omega} |\nabla u(x) - R(x)e^{i\theta(x)-i\frac{\pi}{2}}|^2 \quad \text{is minimal ,}$$

where the $R(x)$ are given gradient amplitudes.

Solution : solve $\Delta u = \operatorname{div}(Re^{i\theta-i\frac{\pi}{2}})$, easily with the discrete Fourier transform, by setting

$$\forall \xi \in \Omega \setminus \{0\}, \quad \widehat{u}(\xi) = \frac{\frac{2i\pi\xi_1}{M}\widehat{v_1}(\xi) + \frac{2i\pi\xi_2}{N}\widehat{v_2}(\xi)}{\left(\frac{2i\pi\xi_1}{M}\right)^2 + \left(\frac{2i\pi\xi_2}{N}\right)^2},$$

where $v_1 = R\sin\theta$ and $v_2 = -R\cos\theta$, and $\widehat{u}(0)$ is an arbitrary constant (mean value of the reconstructed image u).

Defining vector field norm R

Two strategies are used to define the vector field norm $R(x)$ to be associated with the orientation field θ:

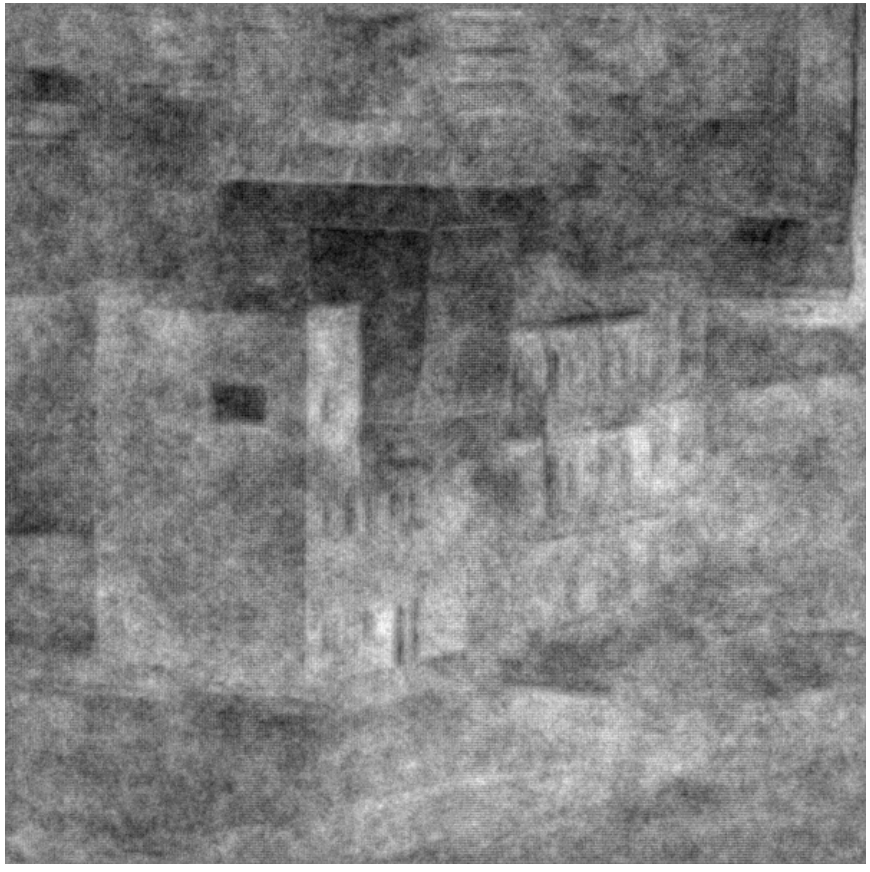
- R_{rand}: Take $R(x)$, $x \in \Omega$, i.i.d. with uniform distribution in $[0, 1]$.
- R_{100} : Defined by

$$R(x) = \begin{cases} 100 & \text{if } x \in \cup_{j=1}^{m} r_j \\ 1 & \text{otherwise.} \end{cases}$$

The second option ensures that detected segments would correspond to edges having a high contrast.

Results on example 1

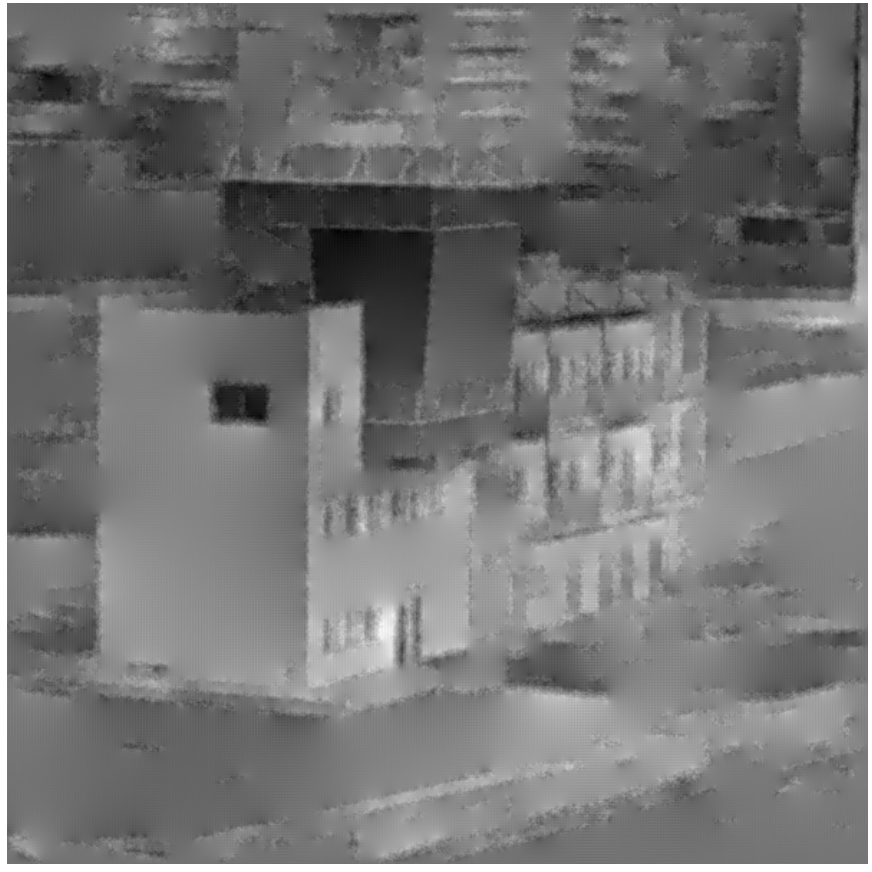
Original image I^0



Reconstruction from P_0 and R_{rand}

Results on example 1

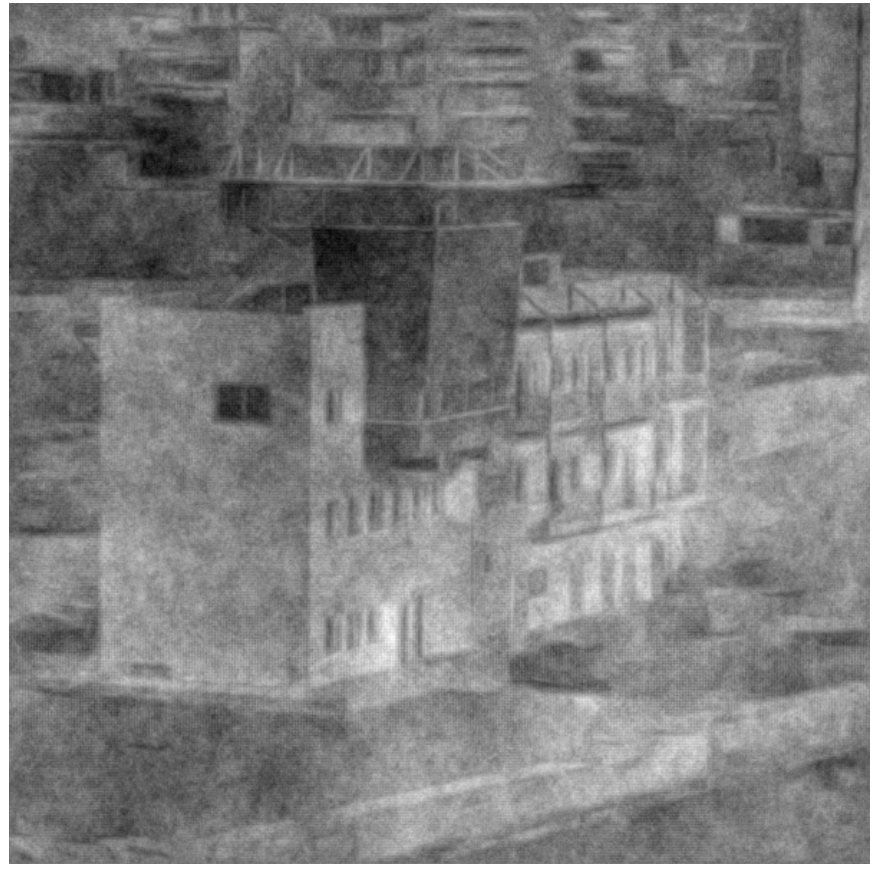
Original image I^0



Reconstruction from P_0 and R_{100}

Results on example 1

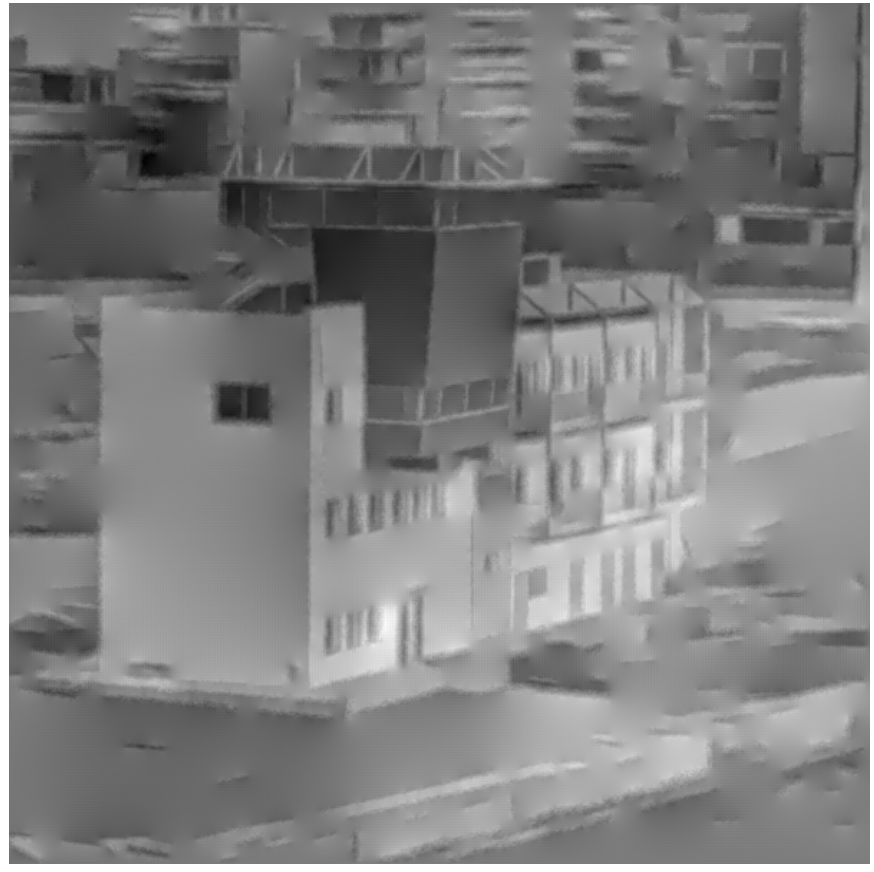
Original image I^0



Reconstruction from Q_0 and R_{rand}

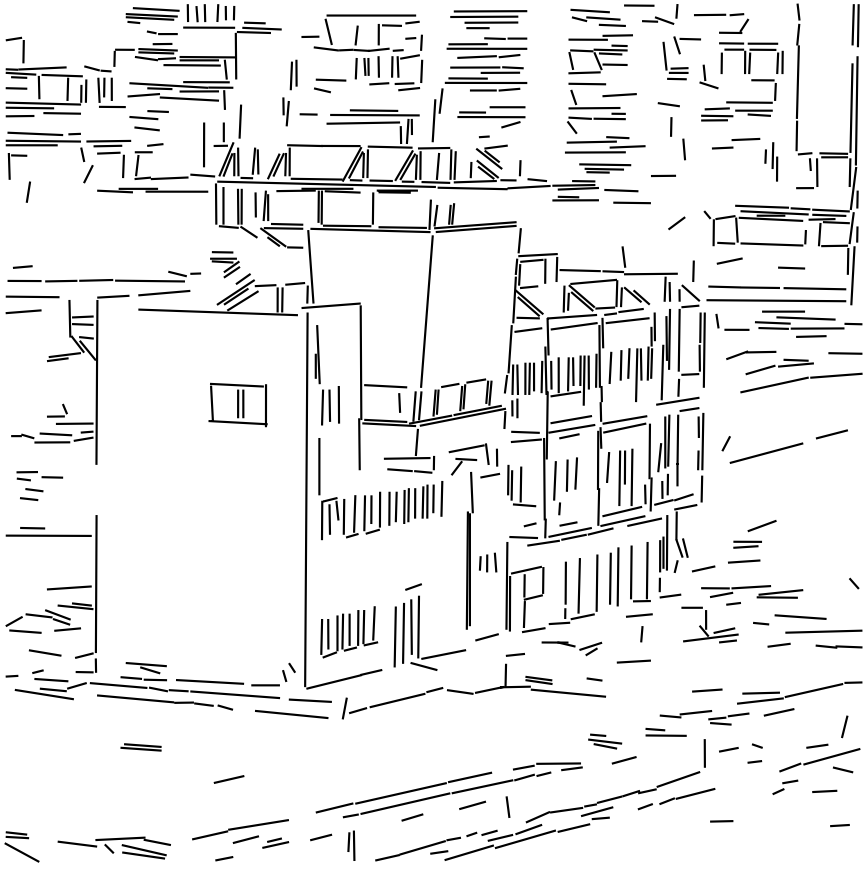
Results on example 1

Original image I^0

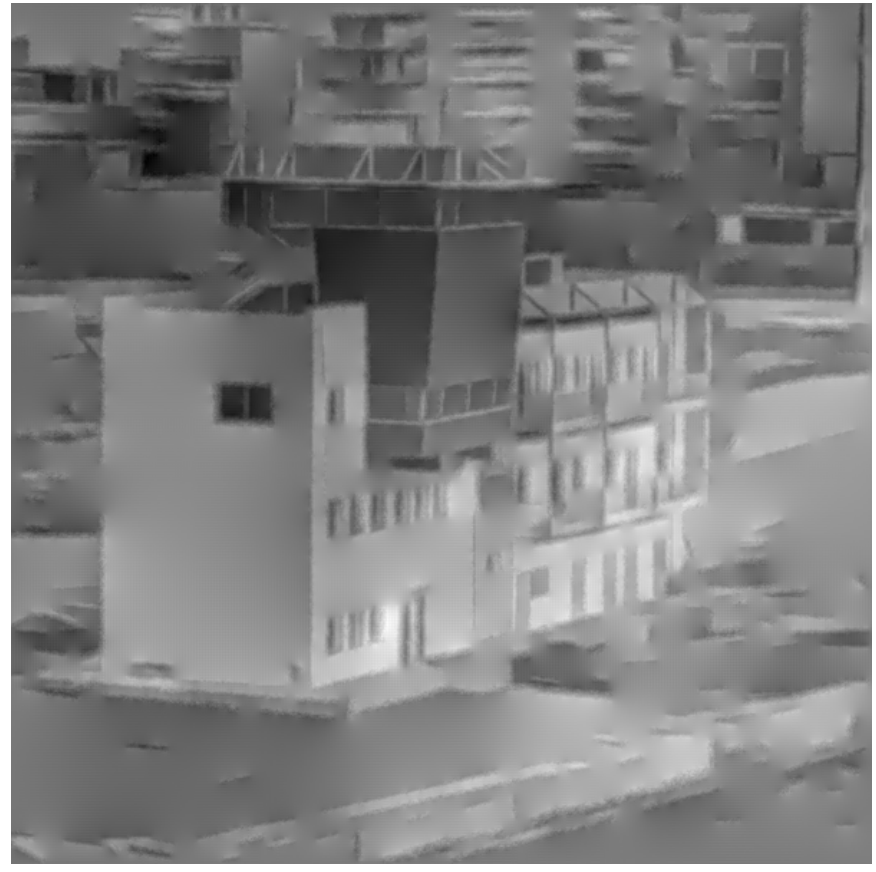


Reconstruction from Q_0 and R_{100}

Results on example 1



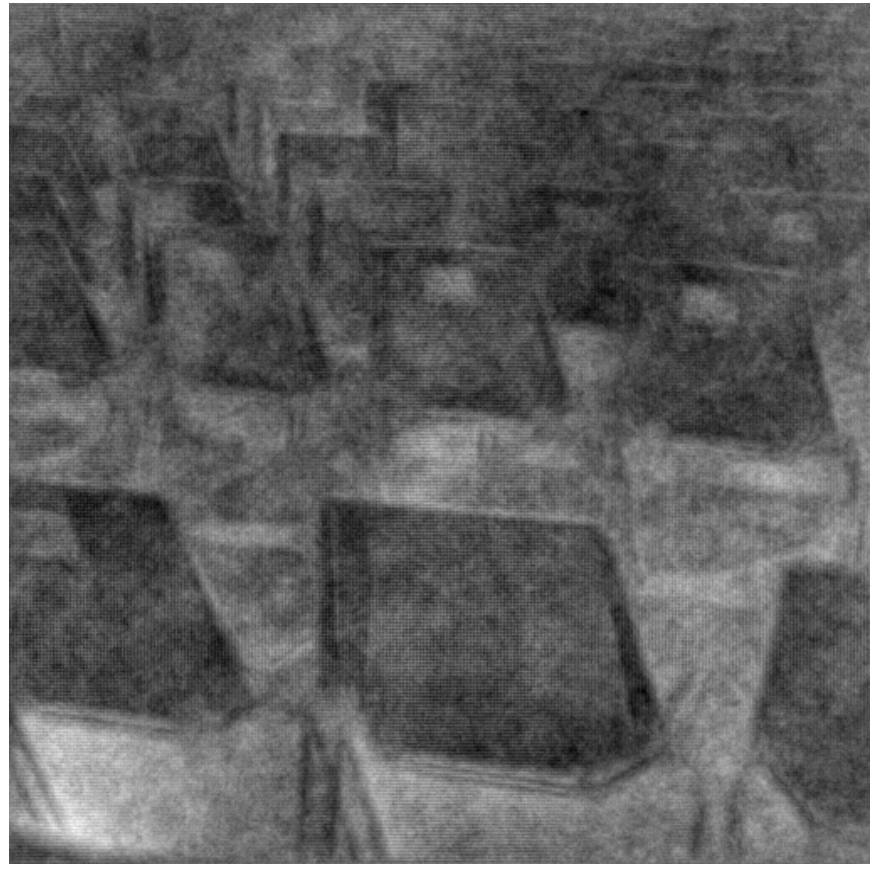
LSD result on I^0



Reconstruction from Q_0 and R_{100}

Results on example 2

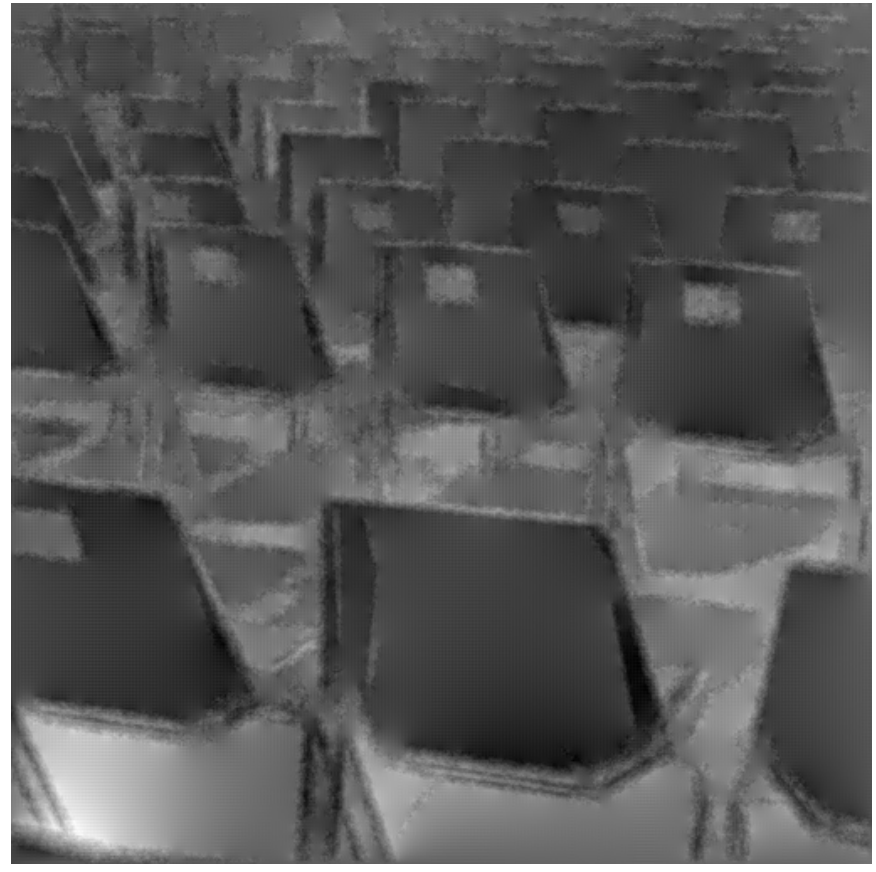
Original image I^0



Reconstruction from P_0 and R_{rand}

Results on example 2

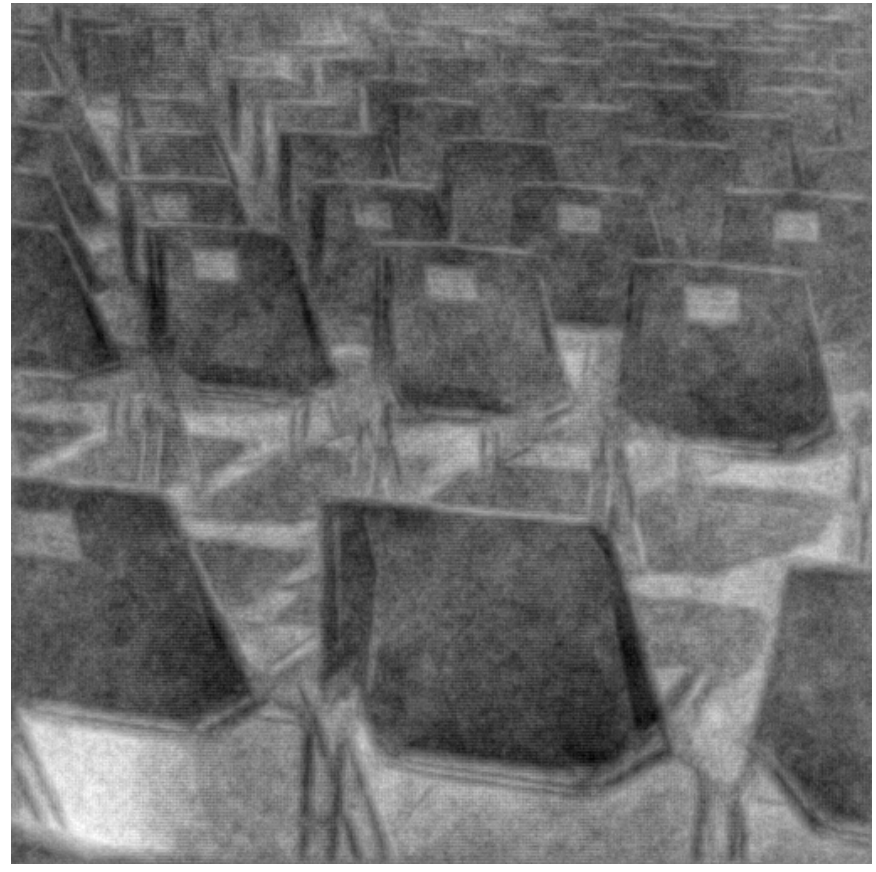
Original image I^0



Reconstruction from P_0 and R_{100}

Results on example 2

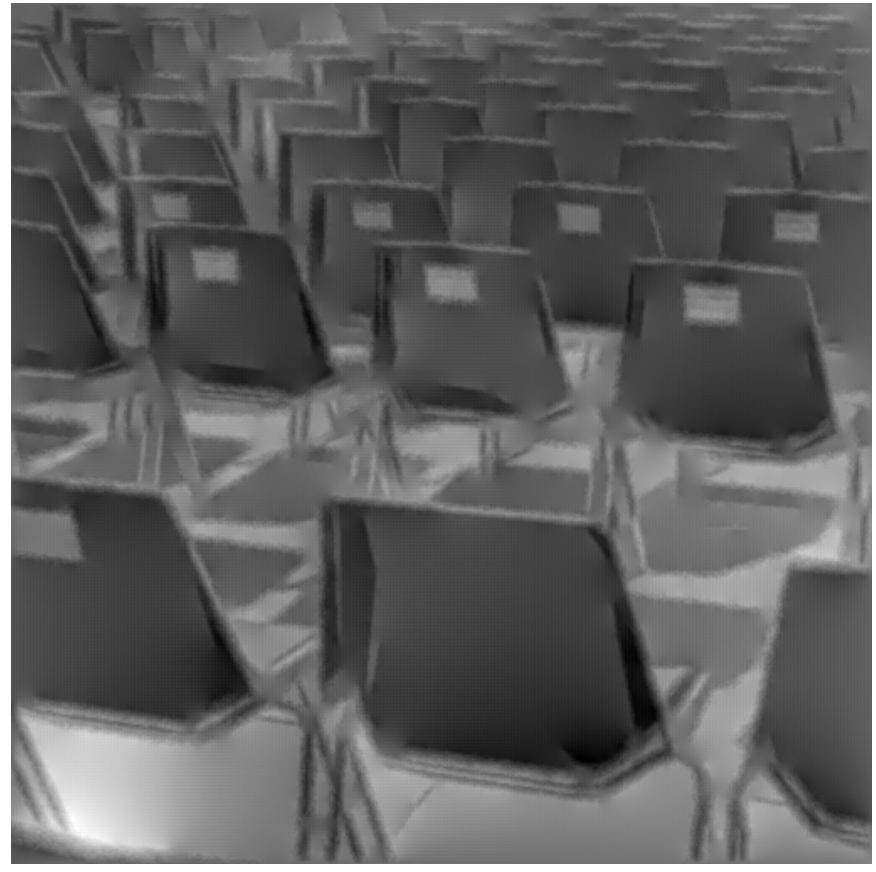
Original image I^0



Reconstruction from Q_0 and R_{rand}

Results on example 2

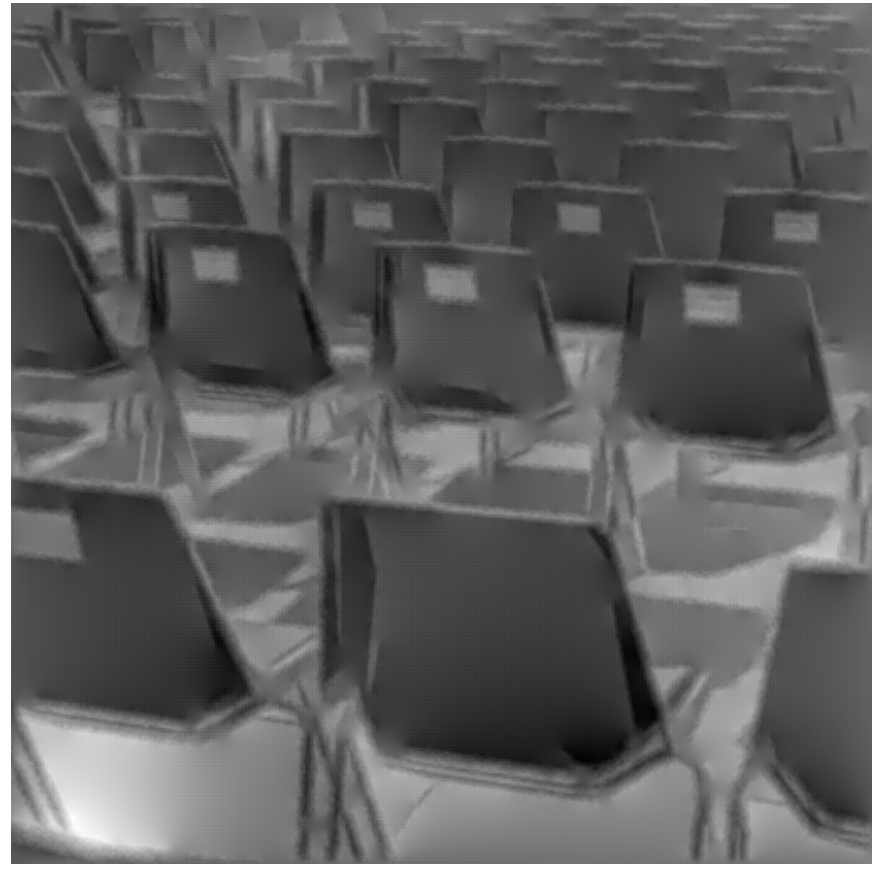
Original image I^0



Reconstruction from Q_0 and R_{100}

Results on example 2

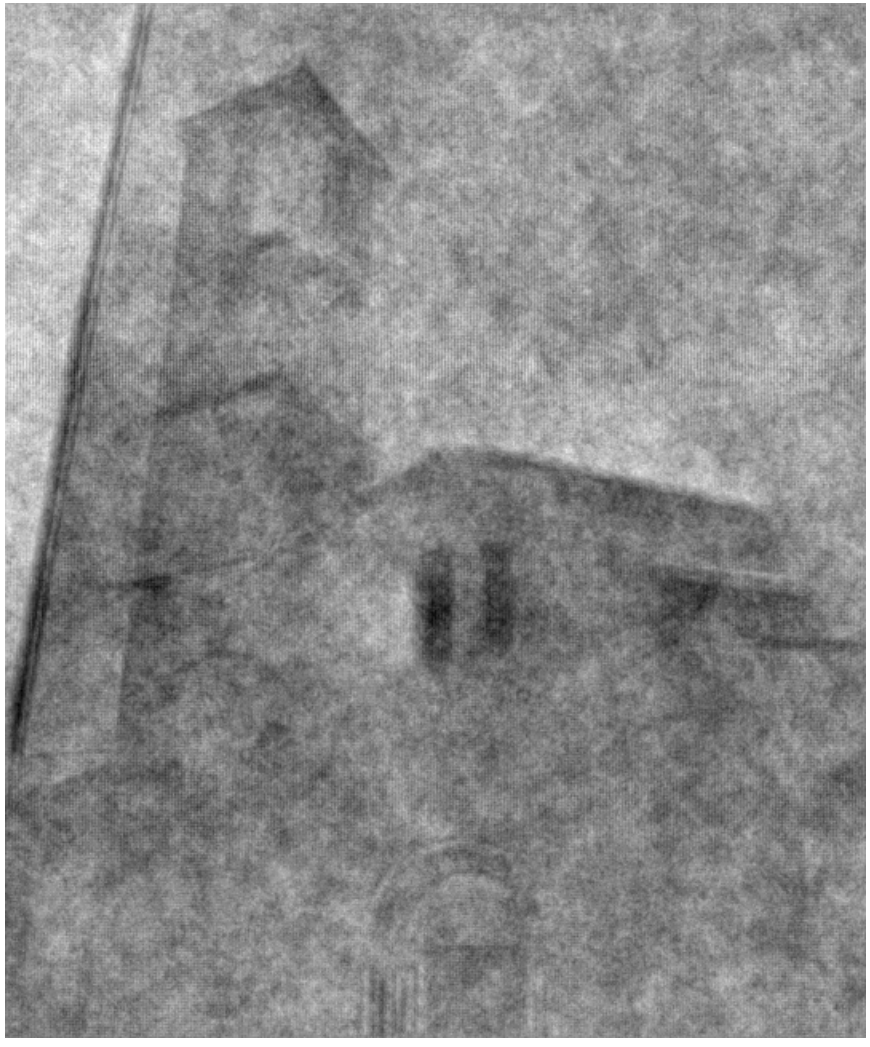
LSD result on I^0



Reconstruction from Q_0 and R_{100}

Results on example 3

Original image I^0



Reconstruction from P_0 and R_{rand}

Results on example 3

Original image I^0

Reconstruction from P_0 and R_{100}

Results on example 3

Original image I^0

Reconstruction from Q_0 and R_{rand}

Results on example 3

Original image I^0

Reconstruction from Q_0 and R_{100}

Results on example 3

LSD result on I^0

Reconstruction from Q_0 and R_{100}

LSD results on reconstructed images

FIGURE: LSD results on : P_0 and R_{100} (left), Q_0 and R_{100} (middle), and the original image (right).

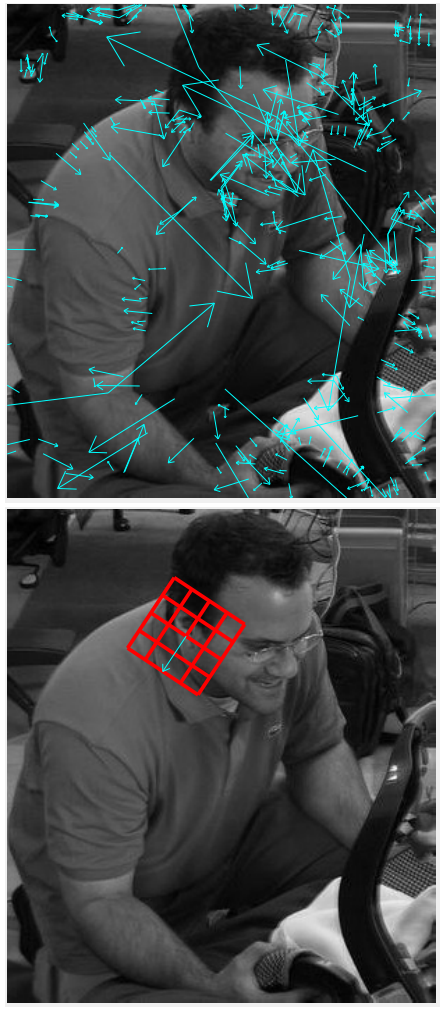
2. Reconstructing from SIFT descriptors

Understanding SIFT descriptors

SIFT descriptors (Lowe 1999) are widely used for image comparison. Now, what kind of information do they capture ?

Let $u_0 : \Omega \to \mathbb{R}$ be an image.

1. Computing SIFT keypoints :
   - 1.1 Extract local extrema of a discrete version of the Gaussian pyramid $(\mathbf{x}, \sigma) \mapsto \sigma^2 \Delta g_\sigma * u_0(\mathbf{x})$.
   - 1.2 Discard extrema with low contrast and extrema located on edges.
2. Computing SIFT local descriptors associated to the keypoint $(\mathbf{x}, \sigma)$:
   - 2.1 Compute one or several principal orientations θ.
   - 2.2 For each detected orientation θ, consider a grid of 4×4 square regions (of size $3\sigma \times 3\sigma$) around $(\mathbf{x}, \sigma)$, with one side parallel to θ. In each subcell compute the histogram of $\mathrm{Angle}(\nabla g_\sigma * u_0) - \theta$ quantized on 8 values $(k\frac{\pi}{4}, 1 \leqslant k \leqslant 8)$.
   - 2.3 Normalization : the 16 histograms are concatenated to obtain a feature vector $f \in \mathbb{R}^{128}$, which is then normalized and quantized to 8-bit integers. *(Step not considered here)*



Let $(s_j)_{j\in\mathcal{J}}$ be the collection of SIFT subcells, $s_j \subset \Omega$.
Let $(\mathbf{x}_j, \sigma_j)$ be the keypoint associated to a s_j, and θ_j the principal orientation.

In s_j we extract the quantized HOG at scale σ_j :

$$H_j^\ell = \frac{1}{|s_j|}\left|\left\{\mathbf{x} \in s_j; \mathrm{Angle}(\nabla g_{\sigma_j} * u_0)(\mathbf{x}) - \theta_j \in B_\ell\right\}\right|,$$

for $\ell = 1$ to 8, where $B_\ell := [(\ell-1)\frac{\pi}{4}, \ell\frac{\pi}{4}]$.

For $x \in \Omega$, we will denote by

$$\mathcal{J}(\mathbf{x}) = \{j \in \mathcal{J} \mid \mathbf{x} \in s_j\}.$$

I. Rey Otero and M. Delbracio, Anatomy of the SIFT Method, *Image Processing On Line*, 2014.

Reconstruction from HOG ?

The goal : reconstruct an image whose content « agrees » with the multiscale HOG H_j in the SIFT subcells s_j.

Difficulties :

- We just have histograms of the gradient orientations.
- The gradient magnitude seems to be completely lost.
- A point $\mathbf{x}$ can belong to several subcells s_j. How to combine the different informations given by the histograms H_j ?

A. Desolneux and A. Leclaire, Stochastic Image Models from SIFT-like descriptors, *SIAM Journal on Imaging Sciences*, 2018.

First solution : Multi-Scale Poisson reconstruction

Each subcell creates its own candidate gradient field :

$$\text{for} \quad j \in \mathcal{J}, \quad V_j(\mathbf{x}) = \frac{1}{\sigma_j} e^{i\gamma_j(\mathbf{x})} \mathbf{1}_{s_j}(\mathbf{x}).$$

The choice $1/\sigma_j$ is motivated by the homogeneity argument ($\nabla\big(u(\frac{\mathbf{x}}{\sigma})\big) = \frac{1}{\sigma}\nabla u(\frac{\mathbf{x}}{\sigma})$). And here γ_j is a random orientation field sampled from the probability density function H_j.

First solution : Multi-Scale Poisson reconstruction

Each subcell creates its own candidate gradient field :

$$\text{for} \quad j \in \mathcal{J}, \quad V_j(\mathbf{x}) = \frac{1}{\sigma_j} e^{i\gamma_j(\mathbf{x})} \mathbf{1}_{s_j}(\mathbf{x}).$$

The choice $1/\sigma_j$ is motivated by the homogeneity argument ($\nabla\big(u(\frac{\mathbf{x}}{\sigma})\big) = \frac{1}{\sigma}\nabla u(\frac{\mathbf{x}}{\sigma})$). And here γ_j is a random orientation field sampled from the probability density function H_j.

To recover an image, consider the following multiscale Poisson energy to be minimized :

$$G(u) = \sum_{j\in\mathcal{J}} \sum_{\mathbf{x}\in\Omega} \|\nabla(g_{\sigma_j} * u)(\mathbf{x}) - V_j(\mathbf{x})\|_2^2.$$

The solution is given by

$$\forall \boldsymbol{\xi} \neq 0, \quad \widehat{u}(\boldsymbol{\xi}) = \frac{\sum_{j\in\mathcal{J}} \widehat{g}_{\sigma_j}(\boldsymbol{\xi}) \Big(\overline{\widehat{\partial}_1(\boldsymbol{\xi})} \widehat{v}_{j,1}(\boldsymbol{\xi}) + \overline{\widehat{\partial}_2(\boldsymbol{\xi})} \widehat{v}_{j,2}(\boldsymbol{\xi}) \Big)}{\sum_{j\in\mathcal{J}} |\widehat{g}_{\sigma_j}(\boldsymbol{\xi})|^2 \Big(|\widehat{\partial}_1(\boldsymbol{\xi})|^2 + |\widehat{\partial}_2(\boldsymbol{\xi})|^2 \Big)}.$$

Now it may be useful to add a regularization term controlled by a parameter $\mu > 0$. Then, if we minimize

$$G(u) + \mu \|\nabla u\|_2^2,$$

we get the well-defined solution

$$\widehat{u}(\boldsymbol{\xi}) = \frac{\displaystyle\sum_{j \in \mathcal{J}} \widehat{g}_{\sigma_j}(\boldsymbol{\xi}) \left(\overline{\widehat{\partial}_1(\boldsymbol{\xi})} \widehat{v}_{j,1}(\boldsymbol{\xi}) + \overline{\widehat{\partial}_2(\boldsymbol{\xi})} \widehat{v}_{j,2}(\boldsymbol{\xi}) \right)}{\left(\mu + \displaystyle\sum_{j \in \mathcal{J}} |\widehat{g}_{\sigma_j}(\boldsymbol{\xi})|^2 \right) \left(|\widehat{\partial}_1(\boldsymbol{\xi})|^2 + |\widehat{\partial}_2(\boldsymbol{\xi})|^2 \right)}.$$

Second solution : combining the different subcells

The second reconstruction method (denoted by MaxEnt) will consist in solving a classical Poisson problem with one single objective gradient

$$V(\mathbf{x}) = \Big(\max_{j \in \mathcal{J}(\mathbf{x})} \frac{1}{\sigma_j} \Big) e^{i\gamma(\mathbf{x})} \mathbf{1}_{\mathcal{J}(\mathbf{x}) \neq \emptyset},$$

where γ is a sample of an orientation field which is inherently designed to combine the local HOG at the scale $\sigma = 0$.

Let θ be an orientation field on Ω. We then consider for all $j \in \mathcal{J}$ and $1 \leqslant \ell \leqslant 8$, the real-valued feature function given by

$$\forall \theta \in \mathbb{T}^{\Omega}, \;\; f_{j,\ell}(\theta) = \frac{1}{|s_j|} \sum_{\mathbf{x} \in s_j} \mathbf{1}_{B_\ell}(\theta(\mathbf{x})).$$

We are then interested in probability distributions P on $\mathbb{T}^{\Omega}$ such that

$$\forall j \in \mathcal{J}, \forall \ell \in \{1, \ldots, 8\}, \quad \mathbb{E}_P\big(f_{j,\ell}(\theta)\big) = f_{j,\ell}(\theta_0) (\simeq H_j^\ell) \;, \tag{1}$$

where θ_0 is the original orientation field.

Theorem (Exponential distribution)

There exists a family of numbers $\lambda = (\lambda_{j,\ell})_{j\in\mathcal{J},1\leqslant\ell\leqslant 8}$ *such that the probability distribution*

$$dP_\lambda = \frac{1}{Z_\lambda}\exp\left(-\sum_{j,\ell}\lambda_{j,\ell}f_{j,\ell}(\theta)\right)d\theta$$

satisfies the constraints (1) *and is of maximal entropy among all absolutely continuous probability distributions w.r.t.* $d\theta$ *satisfying the constraints* (1)*.*

Remark : One can show that the solutions P_λ are obtained by minimizing the smooth convex function

$$\Phi(\lambda) = \log Z_\lambda + \sum_{j,\ell}\lambda_{j,\ell}f_{j,\ell}(\theta_0).$$

Results

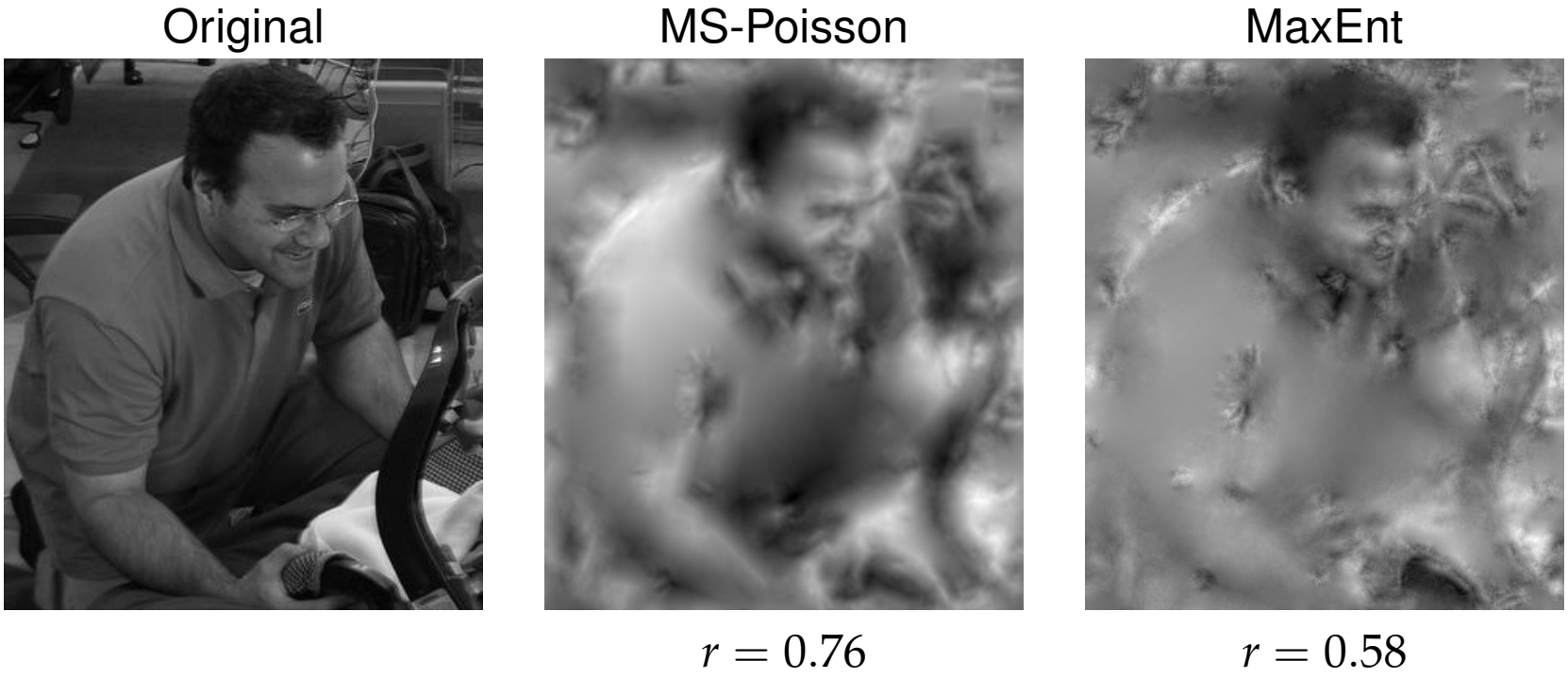


FIGURE: **Reconstruction results.** For each row, from left to right, we display the original image with over-imposed arrows representing the keypoints (the length of the arrow is 6σ which is the half-side of the SIFT cell, and twice the size of the corresponding SIFT subcells), the reconstruction with MS-Poisson, and the reconstruction with MaxEnt. (Image Credits Vondrick et al. 2013).

Results

Original

MS-Poisson

$r = 0.76$

MaxEnt

$r = 0.58$

FIGURE: **Reconstruction results.** For each row, from left to right, we display the original image with over-imposed arrows representing the keypoints (the length of the arrow is 6σ which is the half-side of the SIFT cell, and twice the size of the corresponding SIFT subcells), the reconstruction with MS-Poisson, and the reconstruction with MaxEnt. (Image Credits Vondrick et al. 2013).

Results

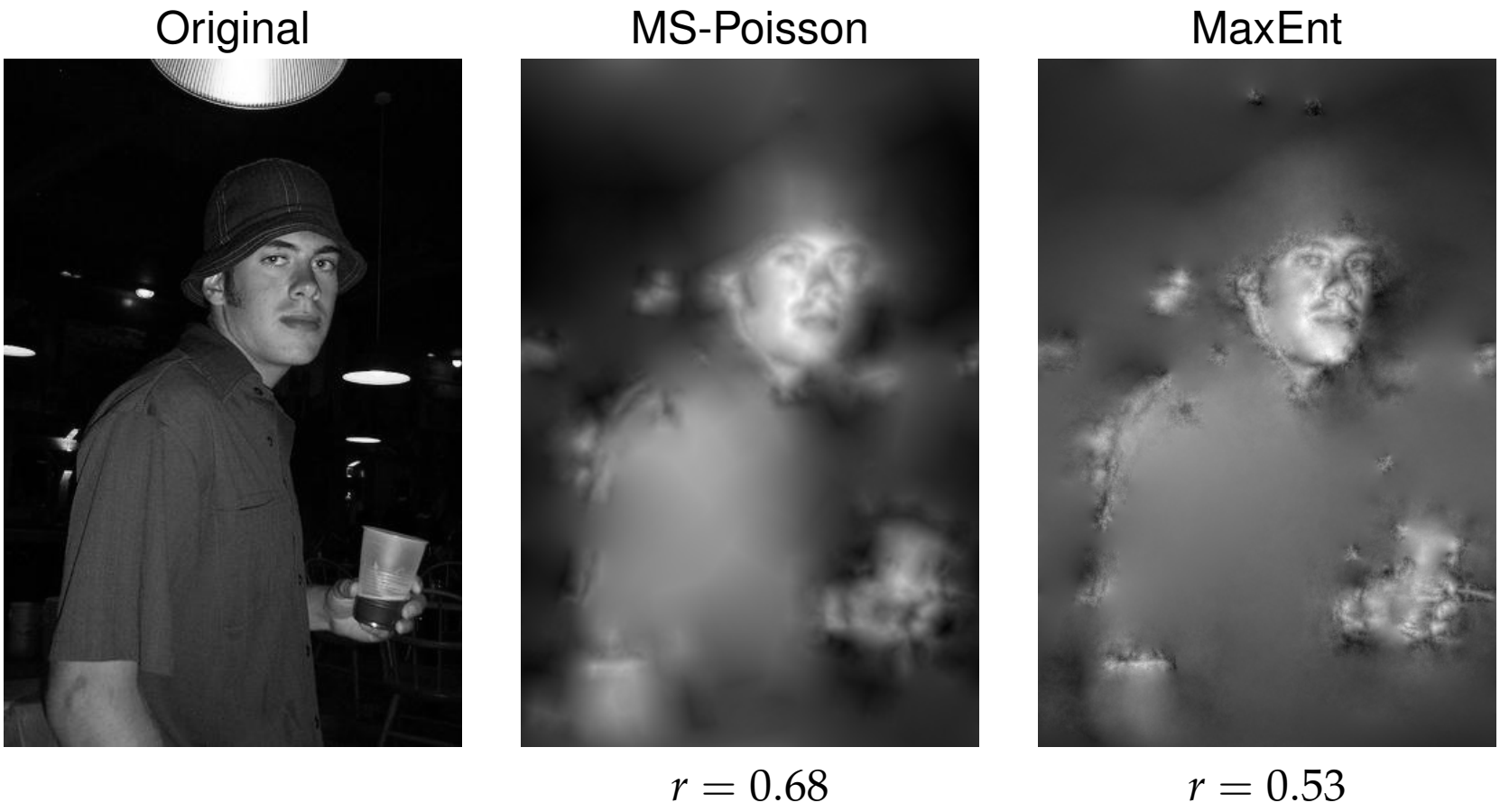


FIGURE: **Reconstruction results.** For each row, from left to right, we display the original image with over-imposed arrows representing the keypoints (the length of the arrow is 6σ which is the half-side of the SIFT cell, and twice the size of the corresponding SIFT subcells), the reconstruction with MS-Poisson, and the reconstruction with MaxEnt. (Image Credits Vondrick et al. 2013).

Results

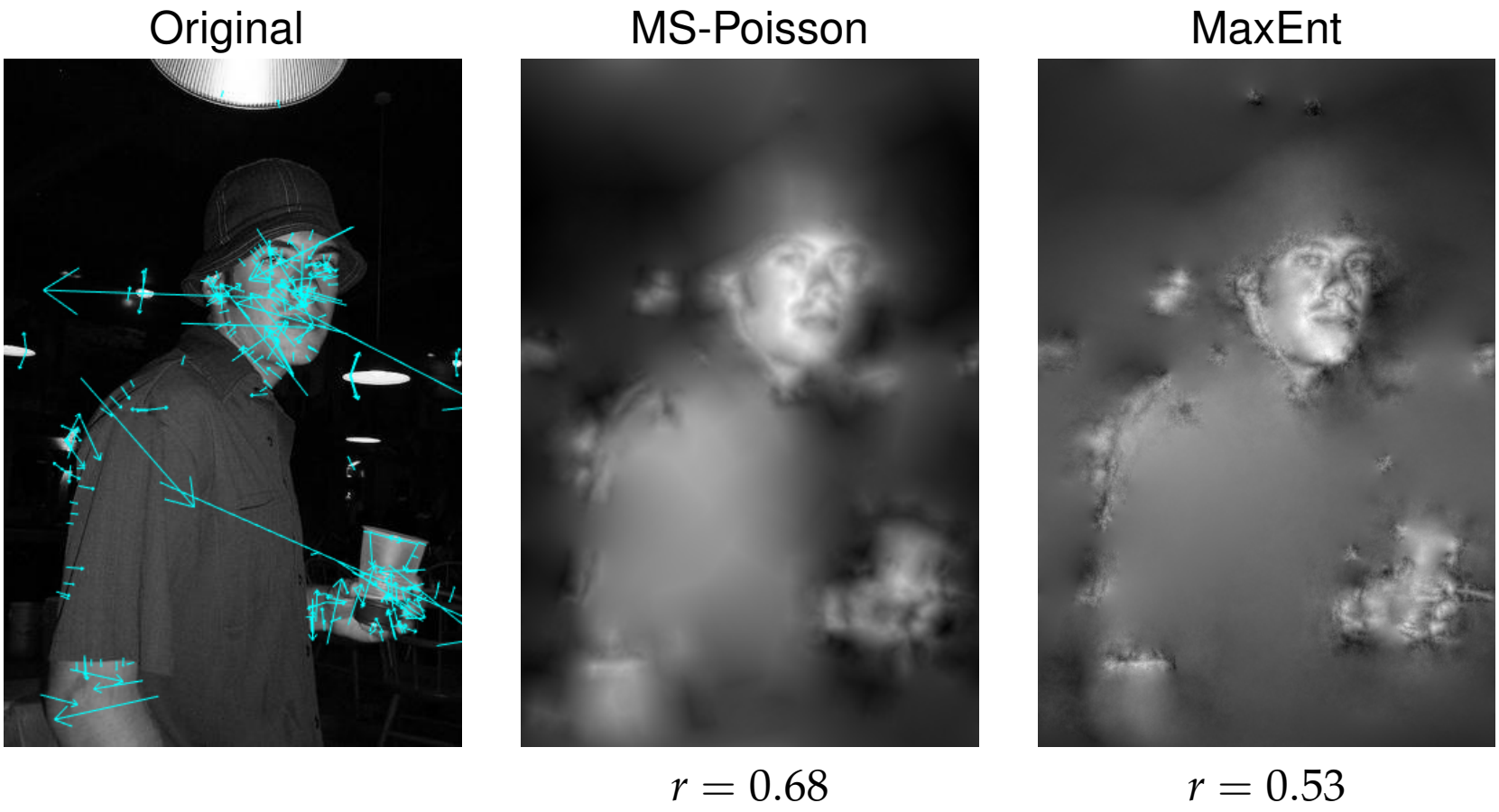


FIGURE: **Reconstruction results.** For each row, from left to right, we display the original image with over-imposed arrows representing the keypoints (the length of the arrow is 6σ which is the half-side of the SIFT cell, and twice the size of the corresponding SIFT subcells), the reconstruction with MS-Poisson, and the reconstruction with MaxEnt. (Image Credits Vondrick et al. 2013).

Results

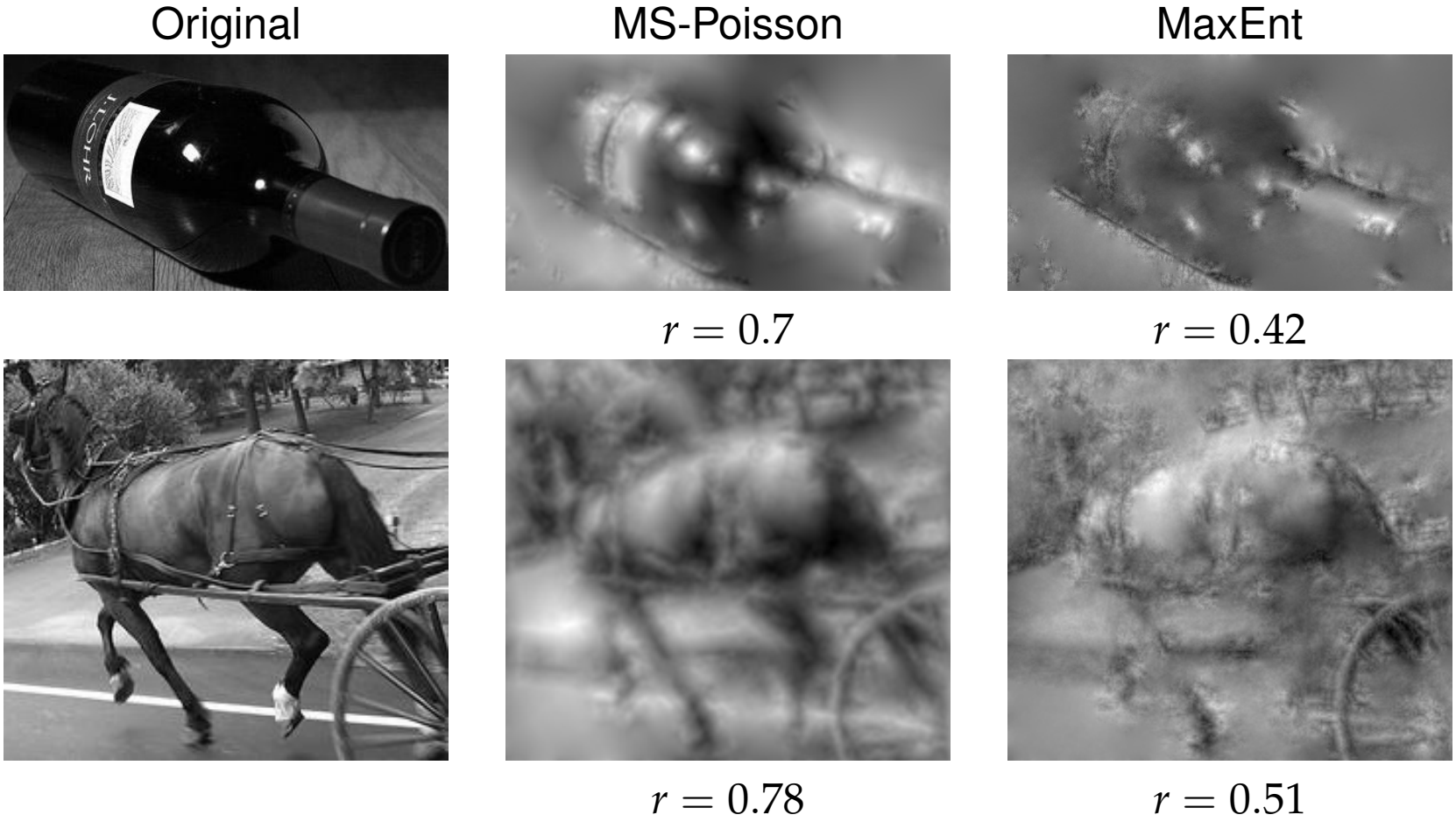


FIGURE: **Reconstruction results.** For each row, from left to right, we display the original image with over-imposed arrows representing the keypoints (the length of the arrow is 6σ which is the half-side of the SIFT cell, and twice the size of the corresponding SIFT subcells), the reconstruction with MS-Poisson, and the reconstruction with MaxEnt.

Results

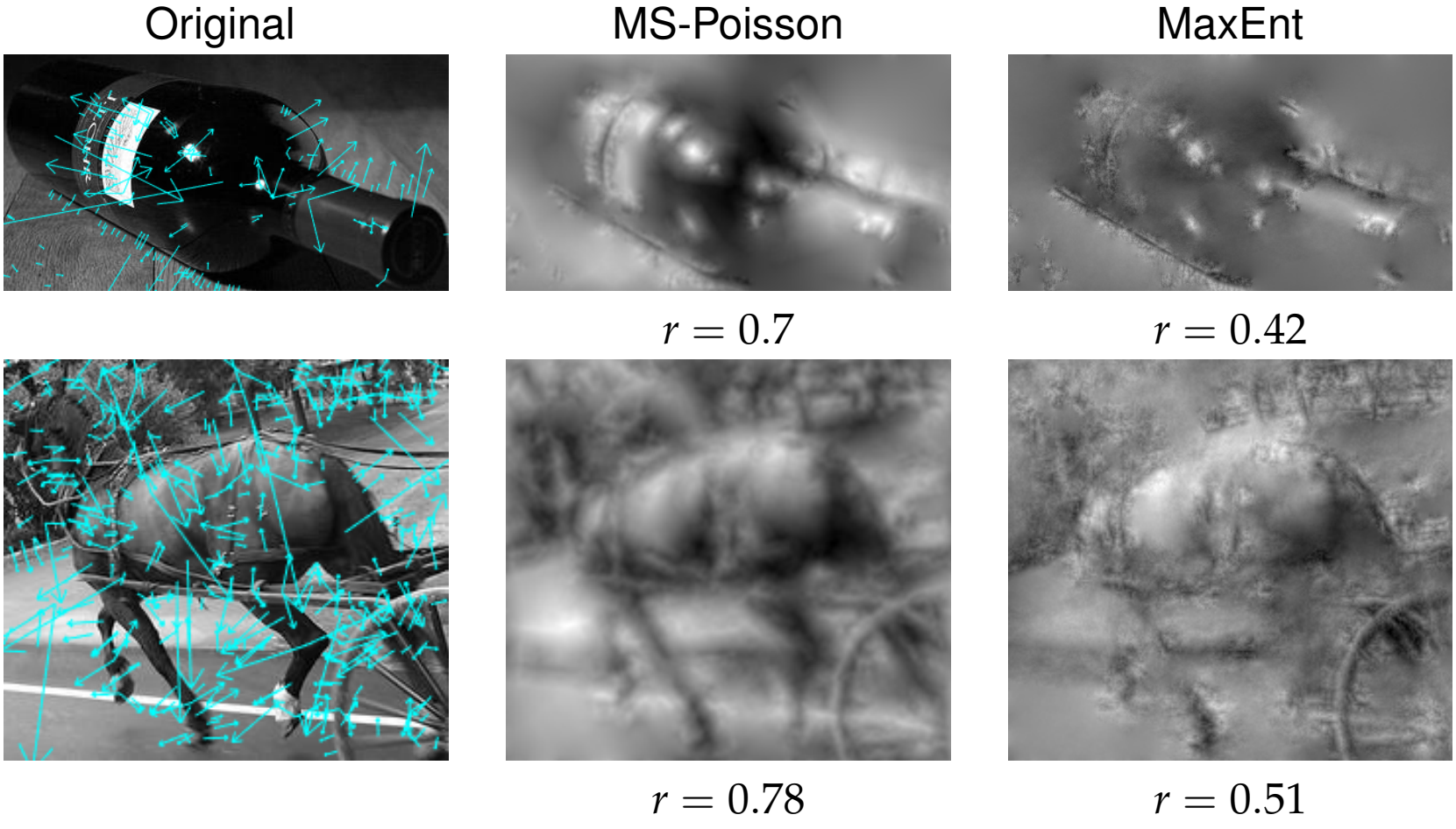


FIGURE: **Reconstruction results.** For each row, from left to right, we display the original image with over-imposed arrows representing the keypoints (the length of the arrow is 6σ which is the half-side of the SIFT cell, and twice the size of the corresponding SIFT subcells), the reconstruction with MS-Poisson, and the reconstruction with MaxEnt.