Lectures on
Stochastic Methods for Image Analysis

Agnès Desolneux

CNRS, CMLA, ENS Cachan

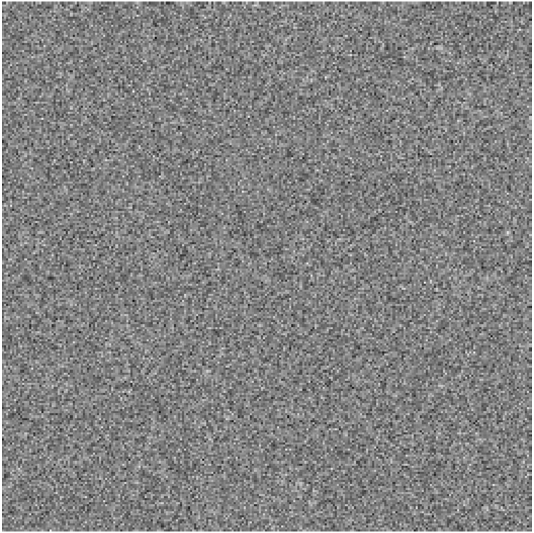
4th Stochastic Geometry Days,
Poitiers, 24-26 August 2015

LECTURE 1

VISUAL PERCEPTION :
GESTALT THEORY AND THE NON ACCIDENTALNESS PRINCIPLE

PART 0 : INTRODUCTION

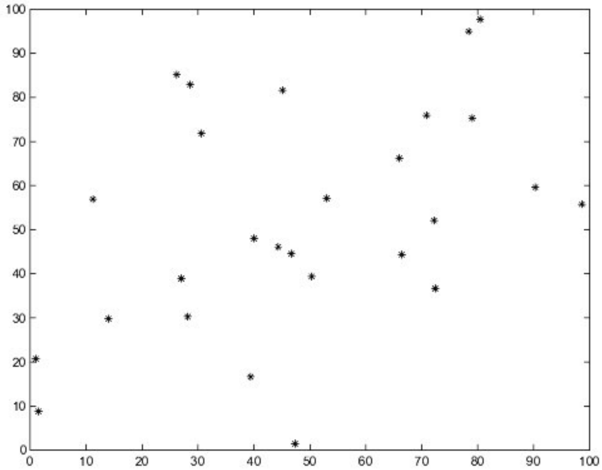
Introduction

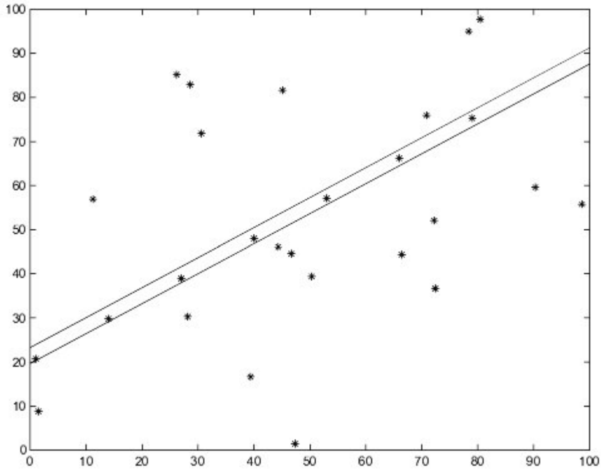


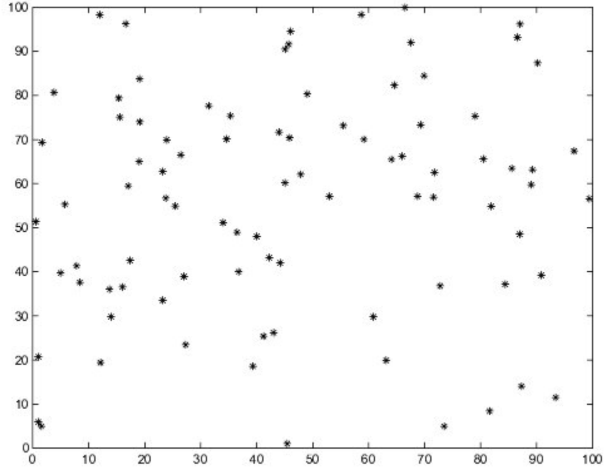
Helmholtz Principle (non-accidentalness principle)

Helmholtz Principle, also called the **non-accidentalness principle**, can be stated in two different ways :

1. The first way is common sensical. It simply states that *"we do not perceive any structure in a uniform random image"*.
   (In this form, the principle was first stated by Attneave in 1954).
2. In its stronger form, the Helmholtz principle states that whenever some large deviation from randomness occurs, a structure is perceived. In other words : *"we immediately perceive whatever has a low likelihood of resulting from accidental arrangement"*.
   (Stated in Computer Vision by S.-C. Zhu or D. Lowe)

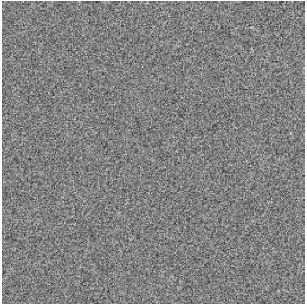

100
90
80
70
60
50
40
30
20
10
0
0 10 20 30 40 50 60 70 80 90 100


100
90
80
70
60
50
40
30
20
10
0
0
10
20
30
40
50
60
70
80
90
100

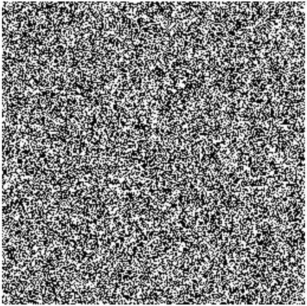
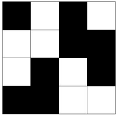
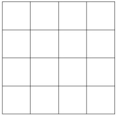


100
90
80
70
60
50
40
30
20
10
0
0 10 20 30 40 50 60 70 80 90 100

What's make an image an image and not noise ?

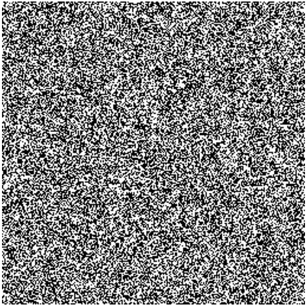
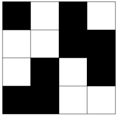
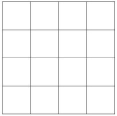


What structures are we looking for ?

 contains   ?

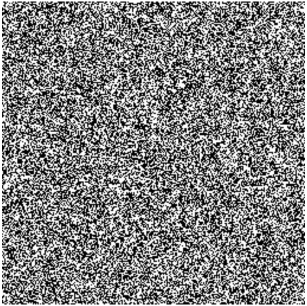
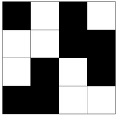
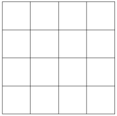
- Not all possible structures are relevant for visual perception.

What structures are we looking for ?

 contains   ?

- Not all possible structures are relevant for visual perception.
- They have to be fixed before - and not after the observation.

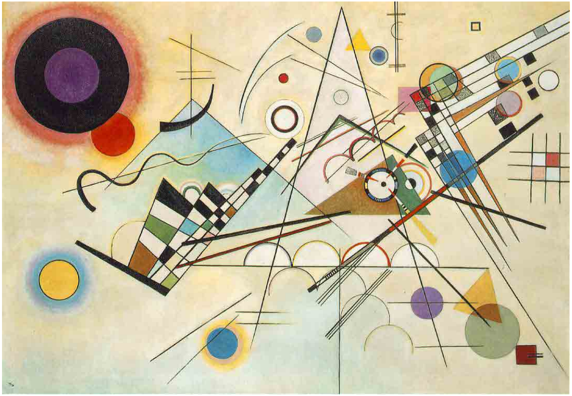
What structures are we looking for ?

 contains   ?

- Not all possible structures are relevant for visual perception.
- They have to be fixed before - and not after the observation.
- The relevant structures for visual perception have been studied by the Gestalt School of Psychophysiology.

Visual Perception

How do we perceive geometric objects in images ?



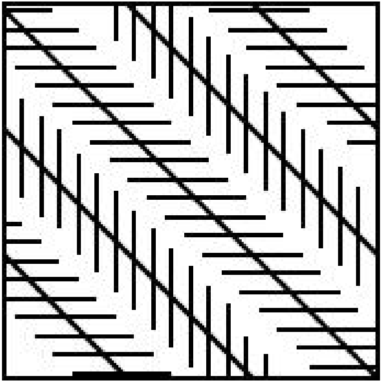
What are the laws and principles of visual construction ?
In other words, how do you go from pixels (or retina cells) to visual objects (lines, triangles, etc.) ?

PART I : GESTALT THEORY OF VISUAL PERCEPTION

Before Gestalt Theory : optic-geometric illusions

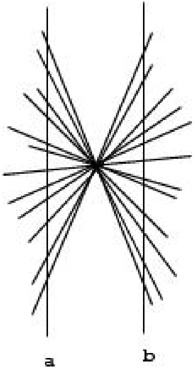
The aim of these illusions is to ask : “what is the reliability of our visual perception ?”

A first example : Zoellner’s Illusion (1860)



Other examples of optic-geometric illusions

Hering's Illusion (1861)



Are lines a and b straight ?

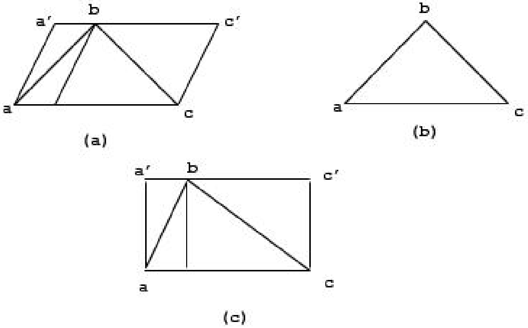
Other examples of optic-geometric illusions

Müller-Lyer's Illusion (1889)

(a) (b)

Other examples of optic-geometric illusions

Sander 's Illusion



Gestalt Theory

- Gestalt theory does not continue on the same line. The question is not why we sometimes see a distorted line when it is straight ; the question is why we do see a line at all. This perceived line is the result of a construction process whose laws it is the aim of Gestalt theory to establish.

- M. Wertheimer *Unterzuchungen zur lehre der gestalt*, Psychologishe Forshung (1923).
- G. Kanizsa, *Grammatica del Vedere / La Grammaire du Voir*, Éditions Diderot, arts et sciences, 1980 / 1997.
- W. Metzger, *Gesetze des Sehens*, Kramer, 1953.

Gestalt Theory

- Gestalt theory does not continue on the same line. The question is not why we sometimes see a distorted line when it is straight ; the question is why we do see a line at all. This perceived line is the result of a construction process whose laws it is the aim of Gestalt theory to establish.
- Gestalt theory (Wertheimer, Metzger, Kanizsa) starts with the assumption of (a small list of) active grouping laws in visual perception : vicinity, same attribute (like colour, shape, size or orientation), alignment, good continuation, symmetry, parallelism, convexity, closure, constant width, amodal completion, T-junctions, X-junctions, Y-junctions.

- M. Wertheimer *Unterzuchungen zur lehre der gestalt*, Psychologishe Forshung (1923).
- G. Kanizsa, *Grammatica del Vedere / La Grammaire du Voir*, Éditions Diderot, arts et sciences, 1980 / 1997.
- W. Metzger, *Gesetze des Sehens*, Kramer, 1953.

Gestalt Theory

- Gestalt theory does not continue on the same line. The question is not why we sometimes see a distorted line when it is straight ; the question is why we do see a line at all. This perceived line is the result of a construction process whose laws it is the aim of Gestalt theory to establish.
- Gestalt theory (Wertheimer, Metzger, Kanizsa) starts with the assumption of (a small list of) active grouping laws in visual perception : vicinity, same attribute (like colour, shape, size or orientation), alignment, good continuation, symmetry, parallelism, convexity, closure, constant width, amodal completion, T-junctions, X-junctions, Y-junctions.
- The above listed grouping laws belong, according to Kanizsa, to the so called primary process, opposed to a more cognitive secondary process.

- M. Wertheimer *Unterzuchungen zur lehre der gestalt*, Psychologishe Forshung (1923).

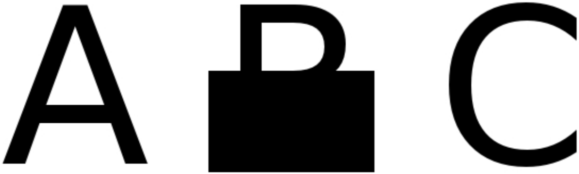
- G. Kanizsa, *Grammatica del Vedere / La Grammaire du Voir*, Éditions Diderot, arts et sciences, 1980 / 1997.

- W. Metzger, *Gesetze des Sehens*, Kramer, 1953.

Seeing and thinking

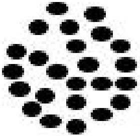
Figure : Grammatica del Vedere, Gaetano Kanizsa.

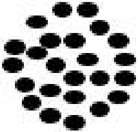
Seeing and thinking



Elementary grouping laws

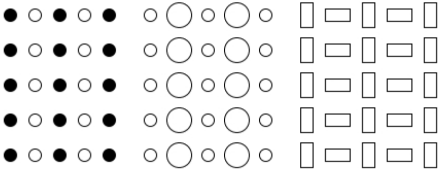
Vicinity





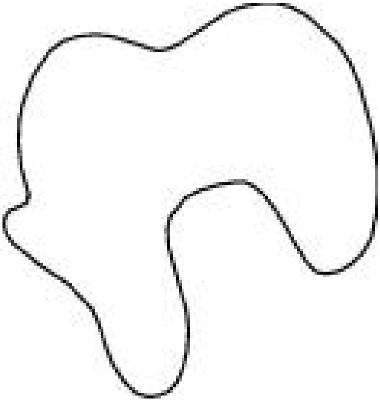
Elementary grouping laws

Same color/size/orientation



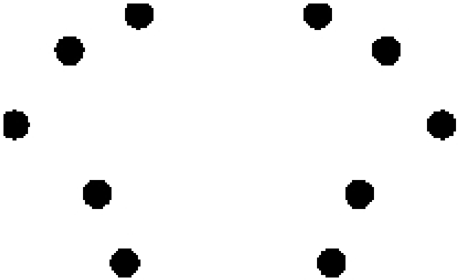
Elementary grouping laws

Closure



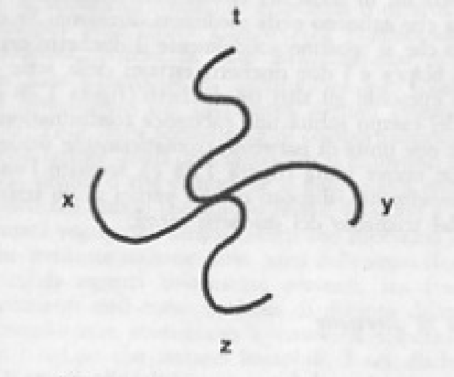
Elementary grouping laws

Symmetry



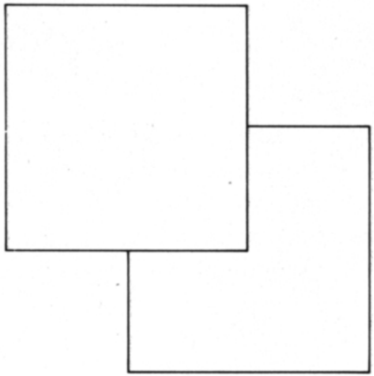
Elementary grouping laws

Good continuation



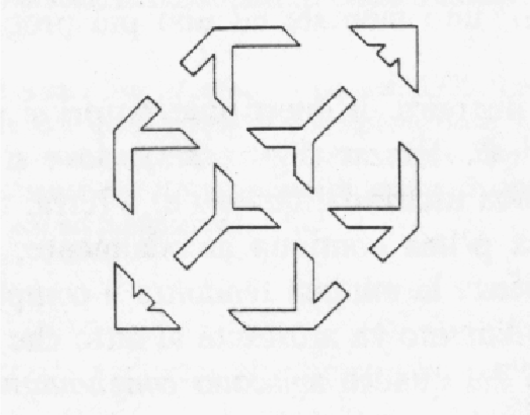
T- and X- junctions

T-junctions



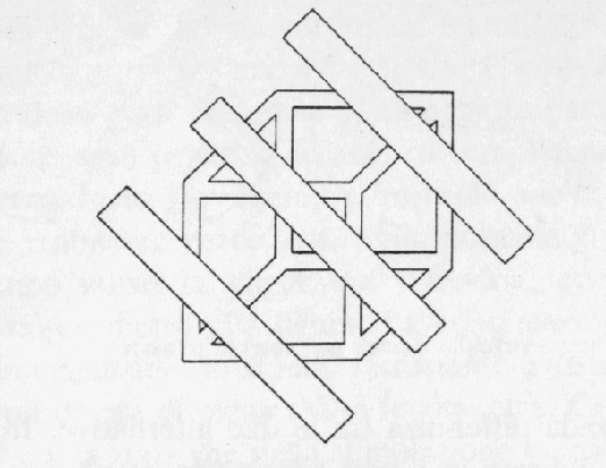
T- and X- junctions

T-junctions



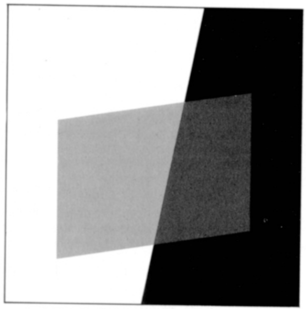
T- and X- junctions

T-junctions

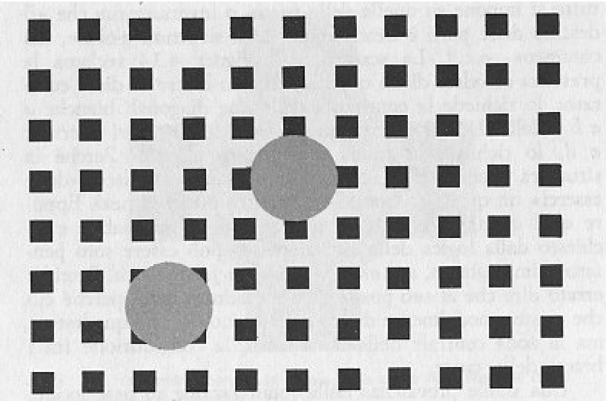


T- and X- junctions

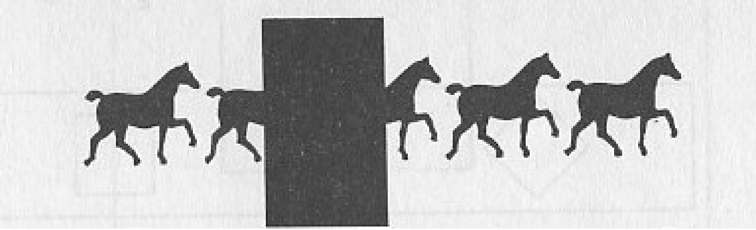
X-junctions



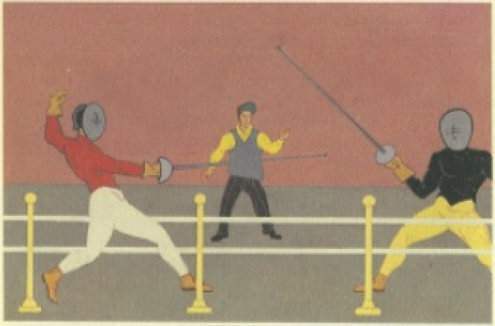
Amodal Completion



Amodal Completion



Amodal Completion



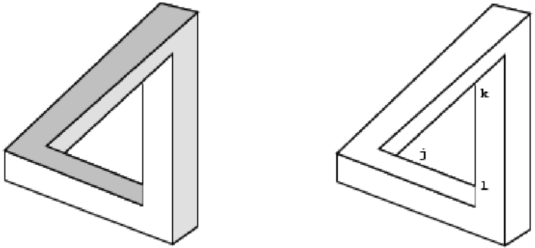
3.17 L'arbitro infilzato.

Impossible Figures

Visual grouping laws are stronger than physical sense.

Perspective effect is created by **Y-junctions**.

Penrose Triangle

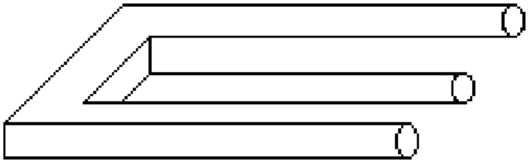


Impossible Figures

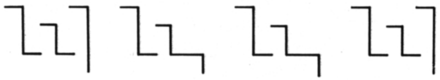
Visual grouping laws are stronger than physical sense.

Perspective effect is created by **Y-junctions**.

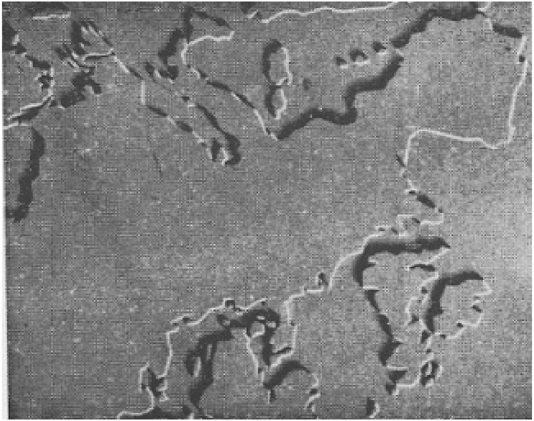
Penrose Fork



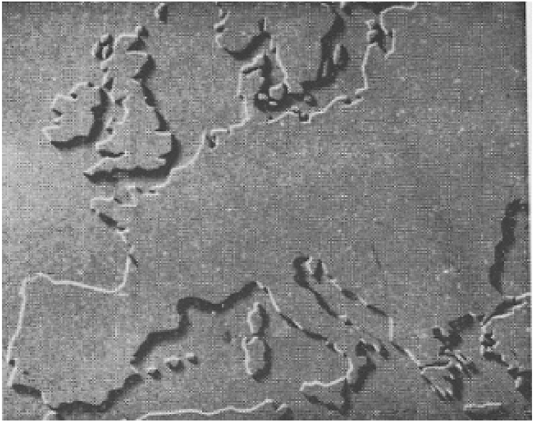
... and past experience ?



... and past experience ?

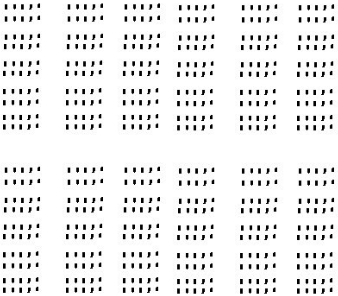


... and past experience ?



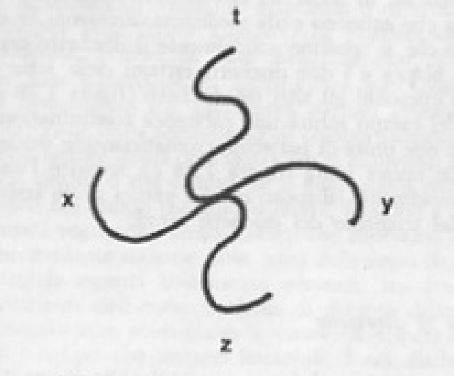
Recursivity

All grouping Gestalt laws are *recursive* : they can be applied first to atomic inputs and then in the same way to partial Gestalts already constituted.

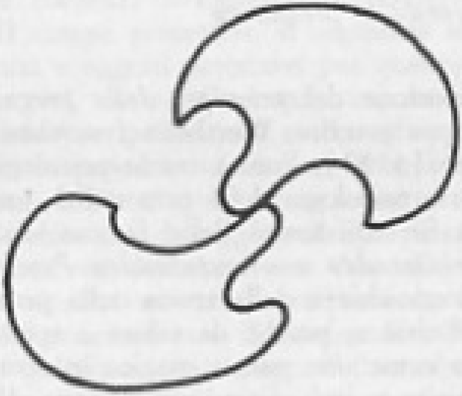


The same partial Gestalt laws namely alignment, parallelism, constant width and proximity, are recursively applied not less than six times.

Conflicts between grouping laws

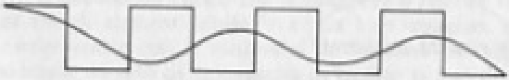


Conflicts between grouping laws



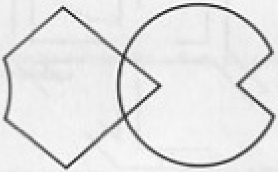
Chiusura contro buona continuazione.

Conflicts between grouping laws



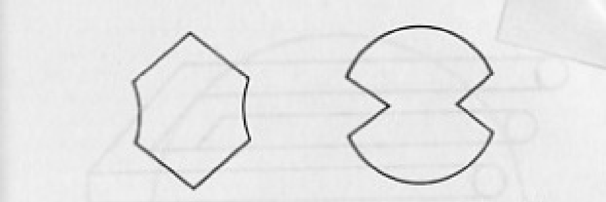
1.55 La continuità di direzione è più forte della chiusura.

Conflicts between grouping laws



2.29 Se la tendenza alla massima simmetria avesse la stessa importanza nel vedere che ha nel pensare, questa configurazione sarebbe vista come la giustapposizione delle due forme simmetriche di figura 2.30.

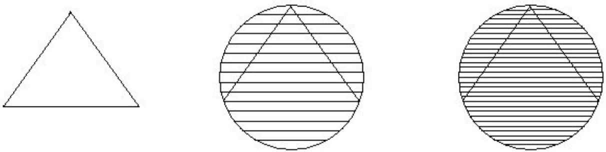
Conflicts between grouping laws



2.30 La figura 2.29 è costituita da queste due forme simmetriche. Perché è tanto difficile vederle?

Masking phenomenon

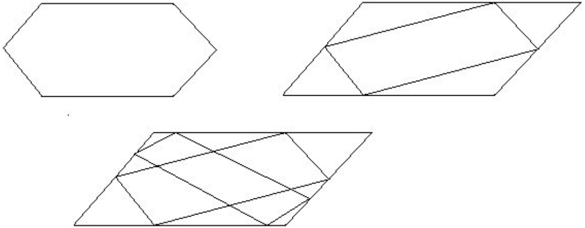
Masking **by texture**



“The whole masks its parts”.

Masking phenomenon

Masking **by addition**



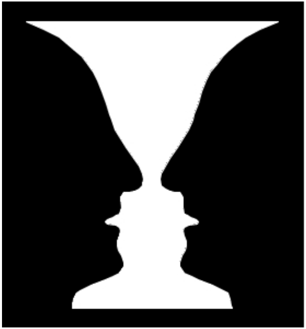
Masking phenomenon

Masking **by subtraction**

A masking is efficient if a new grouping law appears.

Masking phenomenon

Masking **by articulation shape/background**



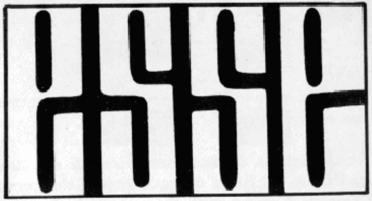


FIG. 5.22. Figure nere su sfondo bianco o lettere bianche su sfondo nero.

Kanisza : "*il fondo non é forma*" (« The background has no shape »).

PART II : THE NON-ACCIDENTALNESS PRINCIPLE

Helmholtz Principle (or Non-accidentalness Principle)

Helmholtz Principle :

- « We don't perceive any structure in a noise image » (Attneave).
- « We immediatly perceive any structure that is too regular to be there just by chance » (Zhu, Lowe)

Examples :

- parallel straight lines in 2D are perceived as the projection of parallel 3D lines.
- if you play dice and you obtain a sequence 6,6,6,6,6 - you will certainly notice it !

$\longrightarrow$ The principle is used to compute *detectability thresholds*.

Example : Birthdays in a class

Question : In a class of 30 students, is it surprising if n of them have the same birthday ?

Hypothesis : birthdays = random variables, independent and identically uniformly distributed on the 365 days of the year.

Let C_n be the number of groups of n students having the same birthday.
Let $\mathbb{P}_n = \mathbb{P}[C_n \geqslant 1]$ and $p_n = \mathbb{P}[C_n \geqslant 1$ and $C_{n+1} = 0]$.

We have $\mathbb{E}[C_n]$ = expected value of the number of groups of n students having the same birthday.

$\longrightarrow$ What are the values of $\mathbb{P}_2$, $\mathbb{E}[C_2]$? And more generally of $\mathbb{P}_n$, $\mathbb{E}[C_n]$?

Example : Birthdays in a class

Question : In a class of 30 students, is it surprising if n of them have the same birthday ?

Hypothesis : birthdays = random variables, independent and identically uniformly distributed on the 365 days of the year.

Let C_n be the number of groups of n students having the same birthday.
Let $\mathbb{P}_n = \mathbb{P}[C_n \geqslant 1]$ and $p_n = \mathbb{P}[C_n \geqslant 1$ and $C_{n+1} = 0]$.

We have $\mathbb{E}[C_n]$ = expected value of the number of groups of n students having the same birthday.

$\longrightarrow$ What are the values of $\mathbb{P}_2$, $\mathbb{E}[C_2]$? And more generally of $\mathbb{P}_n$, $\mathbb{E}[C_n]$?

One can compute : $\mathbb{P}_2 = 1 - \frac{365 \times 364 \ldots 336}{365^{30}} \simeq 0.706$.
$\longrightarrow$ it's not a big surprise to have two students with the same birthday !

Birthdays in a class

We can also compute :

$$\mathbb{P}_3 = \mathbb{P}_2 - p_2 = \mathbb{P}_2 - \frac{1}{365^{30}} \sum_{i=1}^{15} \left[\frac{\prod_{j=0}^{i-1} \binom{30-2j}{2}}{i!} \prod_{k=0}^{29-i} (365-k) \right] \approx 0.028.$$

$\rightarrow$ As n grows, $\mathbb{P}_n$ is more and more uneasy to compute.

Birthdays in a class

We can also compute :

$$\mathbb{P}_3 = \mathbb{P}_2 - p_2 = \mathbb{P}_2 - \frac{1}{365^{30}} \sum_{i=1}^{15} \left[\frac{\prod_{j=0}^{i-1} \binom{30-2j}{2}}{i!} \prod_{k=0}^{29-i} (365-k) \right] \approx 0.028.$$

$\rightarrow$ As n grows, $\mathbb{P}_n$ is more and more uneasy to compute.

But the **expectation** is much simpler :

$$\begin{aligned} \mathbb{E}[C_n] &= \mathbb{E}\left[\sum_{1 \leqslant i_1 < \cdots < i_n \leqslant 30} \mathbb{1}_{\{i_1, \ldots\, i_n \text{ have the same birthday}\}} \right] \\ &= \sum_{1 \leqslant i_1 < \cdots < i_n \leqslant 30} \mathbb{P}[i_1, \ldots, i_n \text{ have the same birthday}] = \frac{1}{365^{n-1}} \binom{30}{n}. \end{aligned}$$

Birthdays in a class

We can also compute :

$\mathbb{P}_3 = \mathbb{P}_2 - p_2 = \mathbb{P}_2 - \frac{1}{365^{30}} \sum_{i=1}^{15} \left[\frac{\prod_{j=0}^{i-1} \binom{30-2j}{2}}{i!} \prod_{k=0}^{29-i} (365-k) \right] \approx 0.028.$

$\rightarrow$ As n grows, $\mathbb{P}_n$ is more and more uneasy to compute.

But the **expectation** is much simpler :

$$\begin{aligned} \mathbb{E}[C_n] &= \mathbb{E}\left[\sum_{1 \leqslant i_1 < \cdots < i_n \leqslant 30} \mathbb{1}_{\{i_1, \ldots\, i_n \text{ have the same birthday}\}} \right] \\ &= \sum_{1 \leqslant i_1 < \cdots < i_n \leqslant 30} \mathbb{P}[i_1, \ldots, i_n \text{ have the same birthday}] = \frac{1}{365^{n-1}} \binom{30}{n}. \end{aligned}$$

Thanks to **Markov Inequality**, we always have $\mathbb{P}_n \leqslant \mathbb{E}[C_n]$.

Numerical computations give :

$$\begin{gathered} \mathbb{P}_2 = 0.706, \text{ and } \mathbb{E}[C_2] = 1.192\,; \\ \mathbb{P}_3 = 0.0285, \text{ and } \mathbb{E}[C_3] = 0.0347\,; \\ \mathbb{P}_4 = 5.3 \times 10^{-4}, \text{ and } \mathbb{E}[C_4] = 5.6 \times 10^{-4}\,. \end{gathered}$$

PART III : COMBINING GESTALT THEORY AND THE NON-ACCIDENTALNESS PRINCIPLE

General formulation of *a contrario* methods

Given : n geometric objects $O_1, \ldots O_n$. Let X_i be a random variable describing an attribute of the O_i (for instance : position, color, orientation, size, etc...).

Hypothesis $\mathcal{H}_0$ (also called *background distribution* or *noise model* or *a contrario model*) : $X_1, \ldots X_n$ are independent identically distributed.

Observed event E on $X_1, \ldots X_k$ (ex : $X_1, \ldots X_k$ are similar).

Can this observed event happen by chance ? (= how likely is it under the null hypothesis $\mathcal{H}_0$?)

Test :

$$\mathrm{NFA}(E) := \mathbb{E}_{\mathcal{H}_0}[\text{nb of occurrences of } E] \leqslant \varepsilon$$

If the test is positive, then the oberved event E is said to be an ε-meaningful event.

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

Example : A black square in a Bernoulli noise

Hypothesis $\mathcal{H}_0$: binary image of size $L \times L$, pixels i.i.d. $\sim$ Bernoulli of parameter p.

Observation : we observe a black square of size $l_0 \times l_0$ pixels somewhere in the image.

$\mathbb{P}[\text{a given square of side length } l \text{ is all black}] = p^{l^2}$.

$\mathbb{E}[\text{Nb of black squares with side length } l \text{ in the image}] = p^{l^2}(L-l)^2$.

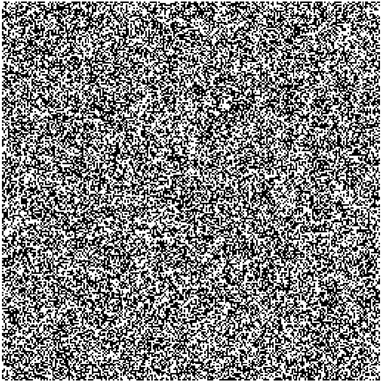
$\mathbb{E}[\text{Nb of black squares with side length } l \geqslant l_0] = \sum_{l \geqslant l_0} p^{l^2}(L-l)^2$.

Questions :

- How to fix the value of p ?
- Why considering only squares ?

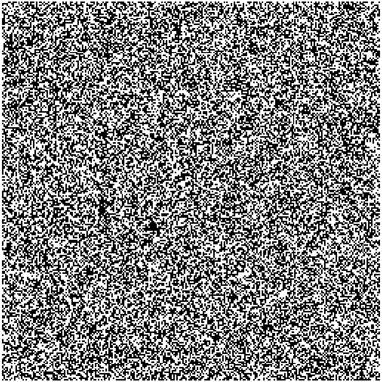
Example : A black square in a Bernoulli noise

$L = 256$, $p = 0.5$
$l_0 = 3$. NFA $= 125.9$



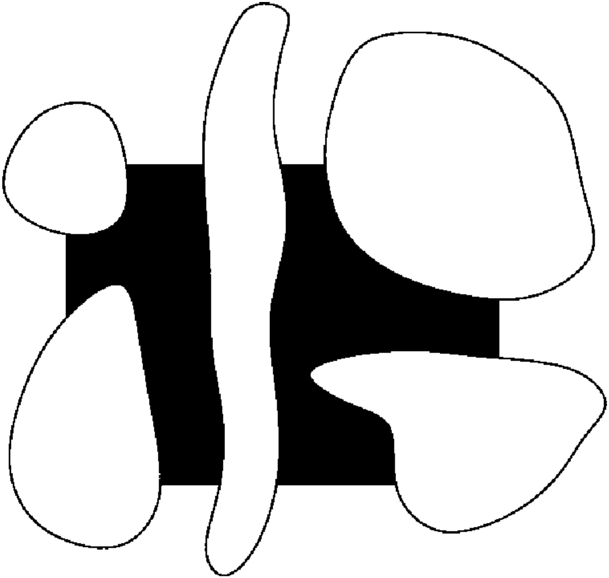
Example : A black square in a Bernoulli noise

$L = 256$, $p = 0.5$
$l_0 = 5$. NFA $= 0.001$



EXERCISES

Exercise 1 : Gestalt comment of this figure



Exercise 2 : Gradient orientation in noise image

The gradient in a (discrete) image u is computed at a pixel of coordinates (x, y) by

$$\begin{aligned}\nabla u(x,y) &= \frac{1}{2}\begin{pmatrix} u(x+1,y+1)+u(x+1,y)-u(x,y+1)-u(x,y) \\ u(x+1,y+1)+u(x,y+1)-u(x+1,y)-u(x,y) \end{pmatrix} \\ &= \|\nabla u(x,y)\| e^{i\theta_u(x,y)}\end{aligned}$$

1) What is the law of $\theta_u(x, y)$ when u is a Bernoulli noise image ?

2) What is the law of $\theta_u(x, y)$ when u is a white noise image (meaning that the grey levels are i.i.d. $\mathcal{N}(0, \sigma^2)$) ?