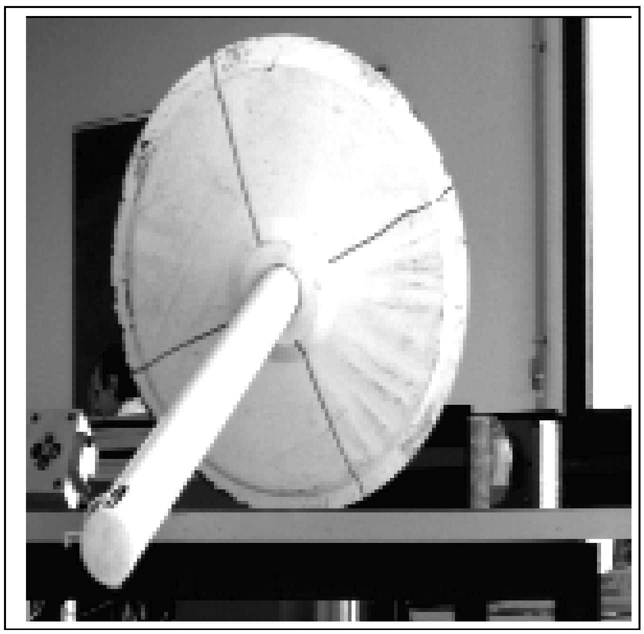
Recursive Perceptual Grouping for 3D objects reconstruction from 2D scenes

Laurent Alquier, Philippe Montesinos

LGI2P - Parc Scientifique G.BESSE
NIMES, F-30000
E-Mail : {montesin, alquier}@eerie.eerie.fr

URL : http://www.eerie.fr/~alquier/alquier.html



An example of scene to describe

Background

The context of this work is the following:

- **grey level pictures**
- **no prior knowledge of the scene**
- **objects with both polyhedric and curved edges**

The expected objective is the extraction of caracteristic shapes from a scene.

These shapes define the description of the scene for further applications:

- **3D modeling**
  to extract a 3D description of the scene from multiple views.
- **Virtual camera**
  to compute a missing view of a scene for different points of view.
- **Shape recognition**
  to detect similar scenes within a database of pictures.

Overview of the system

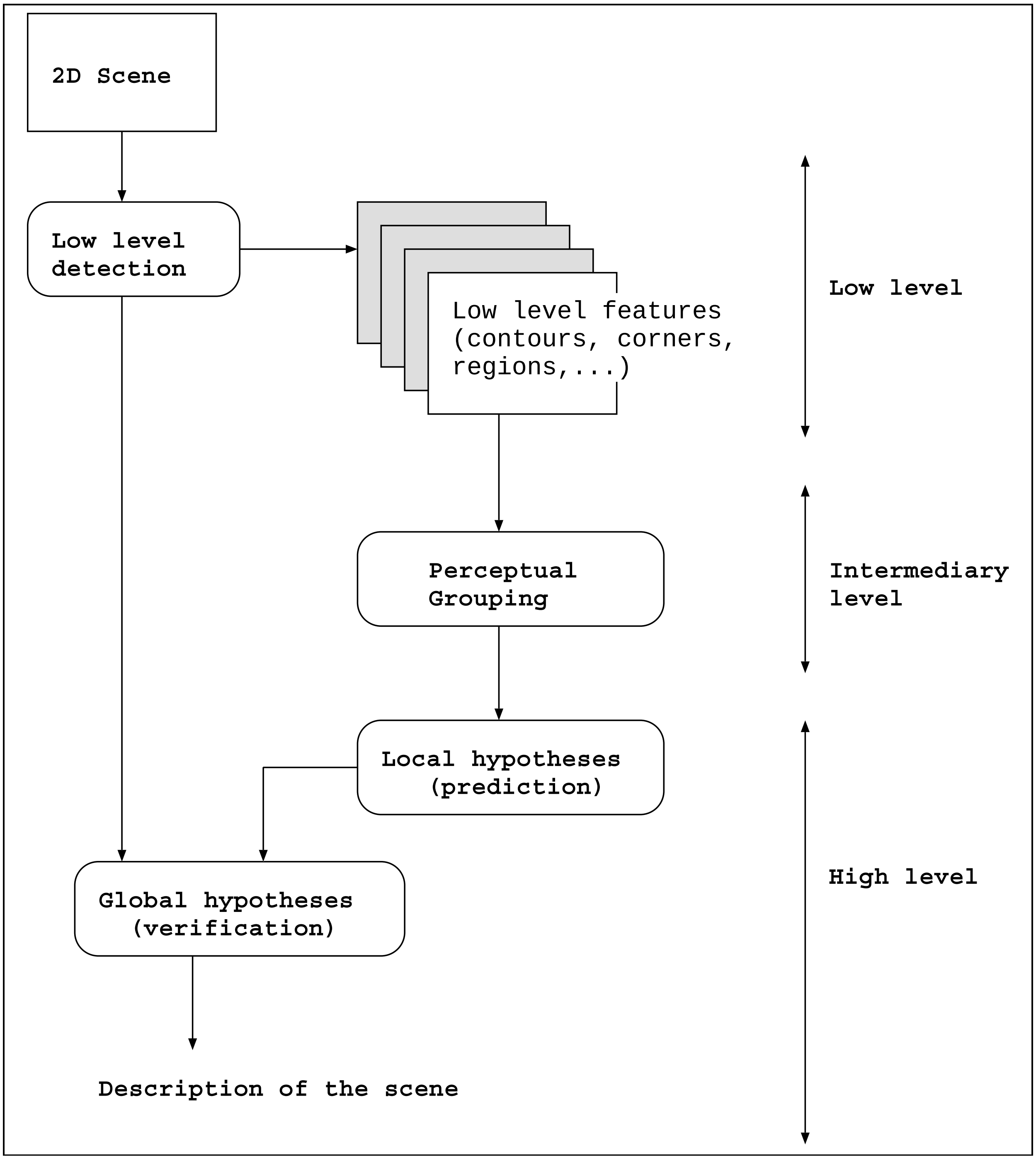


FIG. 1 - *Perceptual grouping for the description of a scene*

Low level processing

Low level features used to extract the elements of scene description :

- **contours / crest lines**
- **regions**
- **corners**
- **etc...**

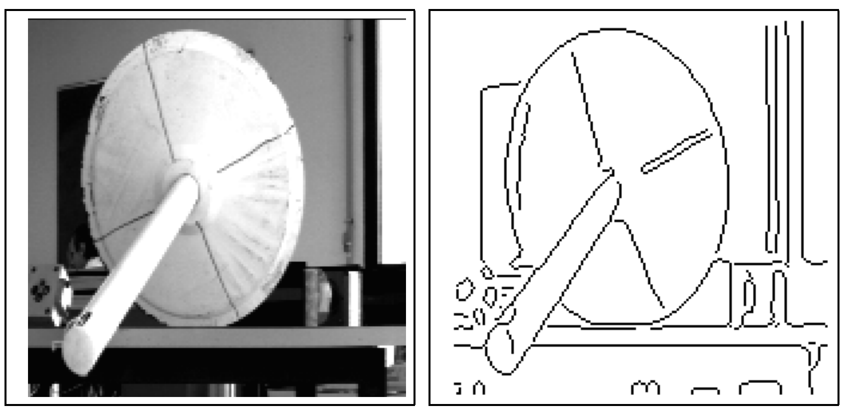


Fig. 2 - *Example of contours detection*

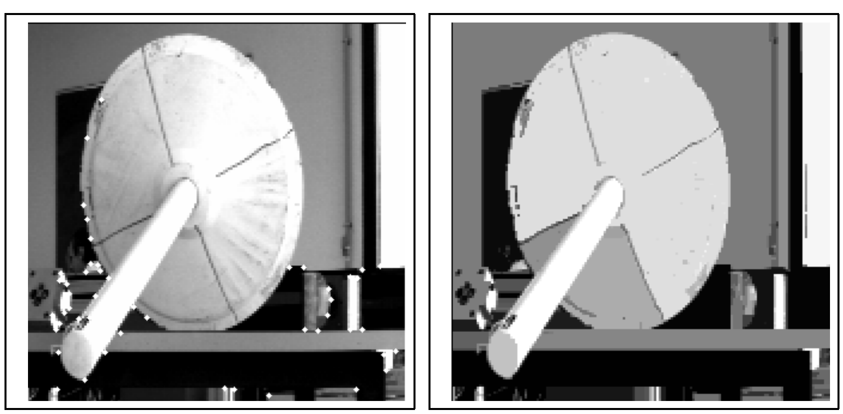


Fig. 3 - *Corners detection and Regions*

Intermediary level : Perceptual Grouping

Among the low level features, the most suited for shape description are lines provided by contours or crest-lines detection. Other features (such as corners) are used at an higher level for verification of hypotheses.

The objectives of this level are the following :

- **extract the most salient shapes**
- **fill discontinuities**
- **ignore artefacts produced by noise in low level processing**

Our method derives from principles of **Perceptual Grouping**. Shapes of a growing complexity are detected and organised incrementally from a local to global level according to visual relationships.

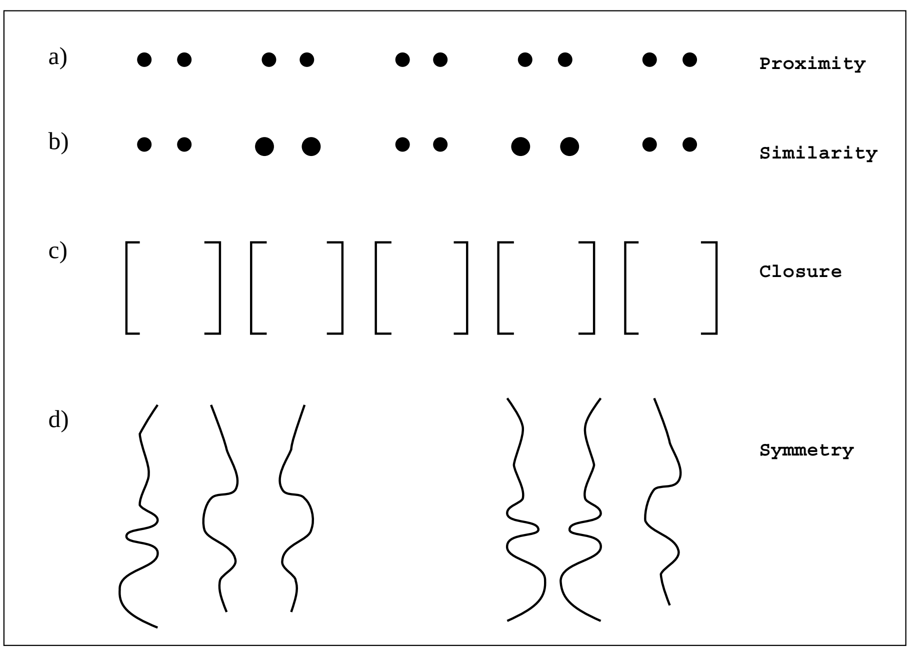


FIG. 4 - *Examples of visual properties*

Visual properties and Elements of description

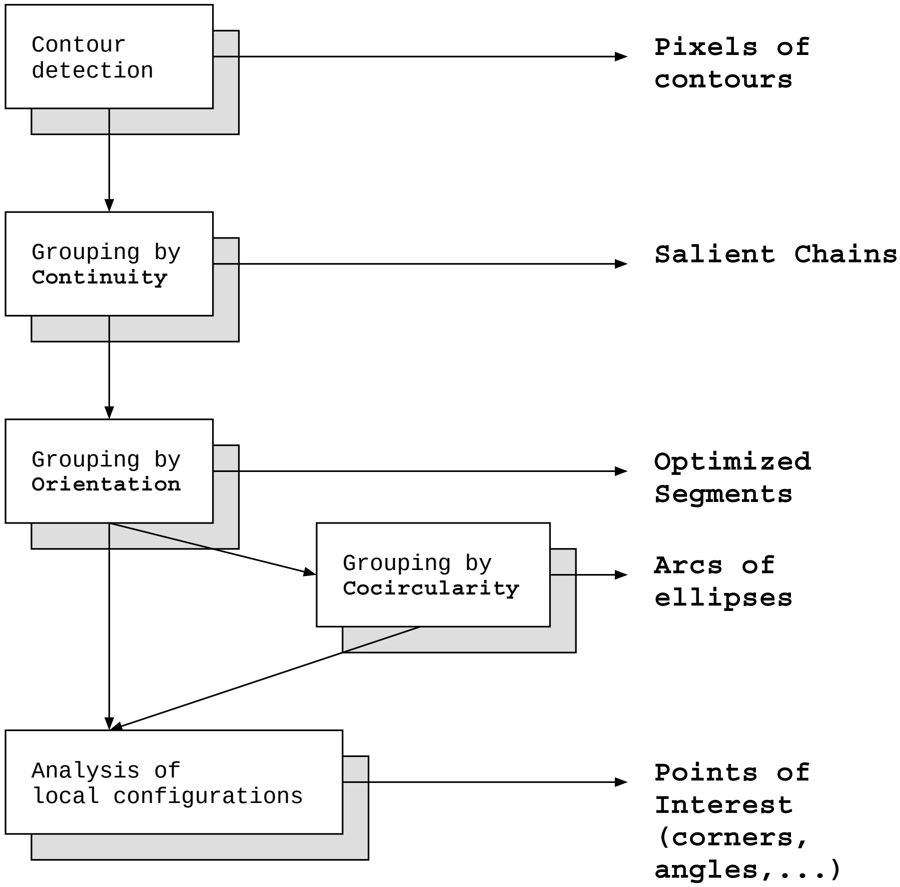


FIG. 5 - *Levels of Perceptual Grouping*

The first levels of Perceptual Grouping decomposes the scene into geometric primitives:

- Salient **CHAINS** of contours
- **POINTS** of interest
- Optimized **SEGMENTS**
- **ARCS** of ellipses

Grouping by Continuity

This first level is important in order to reduce the complexity of the task of scene description.

Its purpose is to extract the most salient structures from the image after detection of contours or crest-lines.

The reduced number of salient grouping (opposed to the number of edge elements) plays a role of **focus of attention** for the further levels of description.

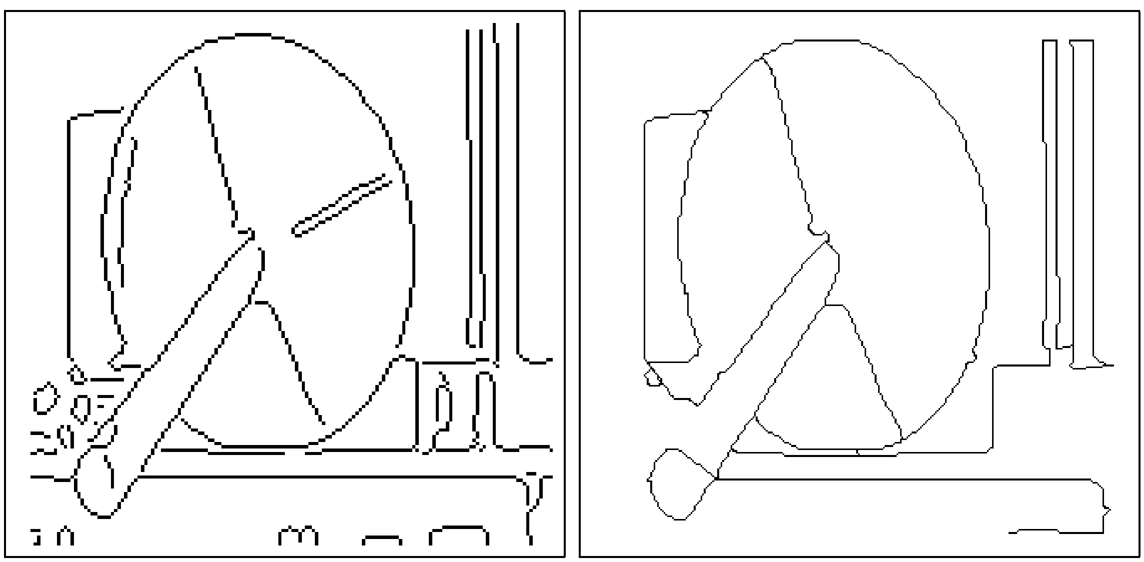


FIG. 6 - *Picture of edges 316x316 pixels - selection of 6 salient groupings after optimization*

Principle : Recursive optimization

At this level, the primitives grouped together are **pixels of contours** or **chains of pixels** .

The primitives are grouped together locally, according to visual properties described by a **quality function**. The optimization of this function gives the most salient groupings.

The quality function for a grouping takes into account a local term and a global contribution of each primitive. Along the iterations, salient structures receive high global contributions.

Finally, the most salient structures are detected by following the optimized local groupings from a primitive to the other.

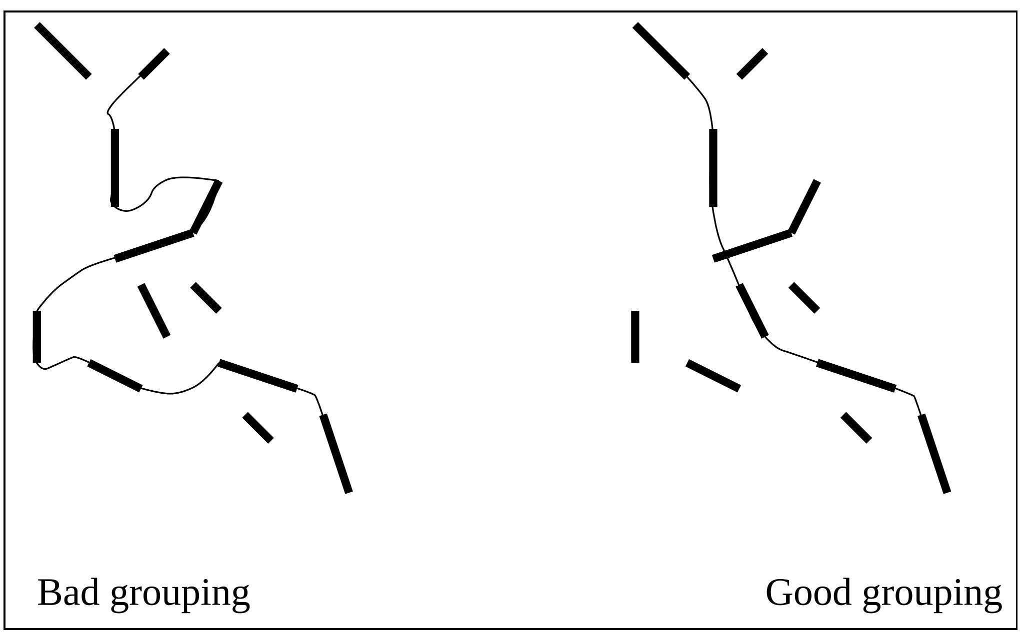


FIG. 7 - *Possible groupings of Continuity*

Formalisme : Recursive Optimization

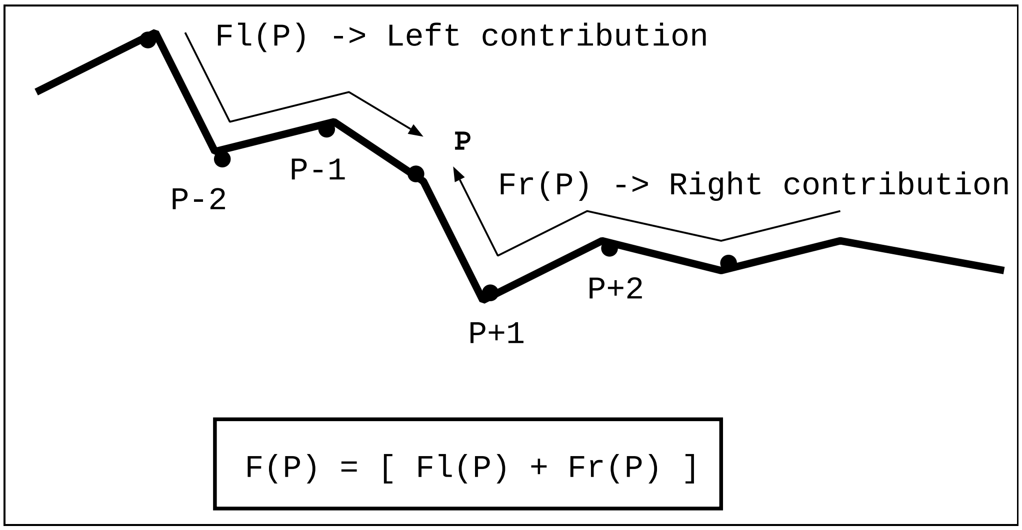


FIG. 8 - *Notations used for a grouping during the optimization*

For a given primitive, each term of the Quality Function $\mathcal{F}$ is written as a **bi-lateral function**:

$$\mathcal{F}(P) = (\mathcal{F}_r(P) + \mathcal{F}_l(P)) \quad (1)$$

with a path $\mathcal{F}_l(P)$ arriving in P and a path $\mathcal{F}_r(P)$ going from P.

Formalisme : Recursive optimization

Each lateral contribution is the sum of local contributions :

$$\mathcal{F}_l(\,P\,) = \begin{array}{l} \frac{1}{2}\cdot Q(\,P\,) \\ +\rho\cdot Q_P(\,P-1\,) \\ +\rho^2\cdot Q_{P-1}(\,P-2\,)+... \end{array} \tag{2}$$

Or, written in a recursive way. for a distance n from P :

$$\mathcal{F}_l{}^{(\,n\,)}(\,P\,) = Q_P(\,P\,) + \rho\cdot\mathcal{F}_l{}^{(\,n-1)}(\,P-1\,) \tag{3}$$

where $Q(\,P\,)$ is the local term in P and $Q_P(\,P-1\,)$ the contribution P provided by $(\,P-1\,)$.

Grouping between Pixels of Contours

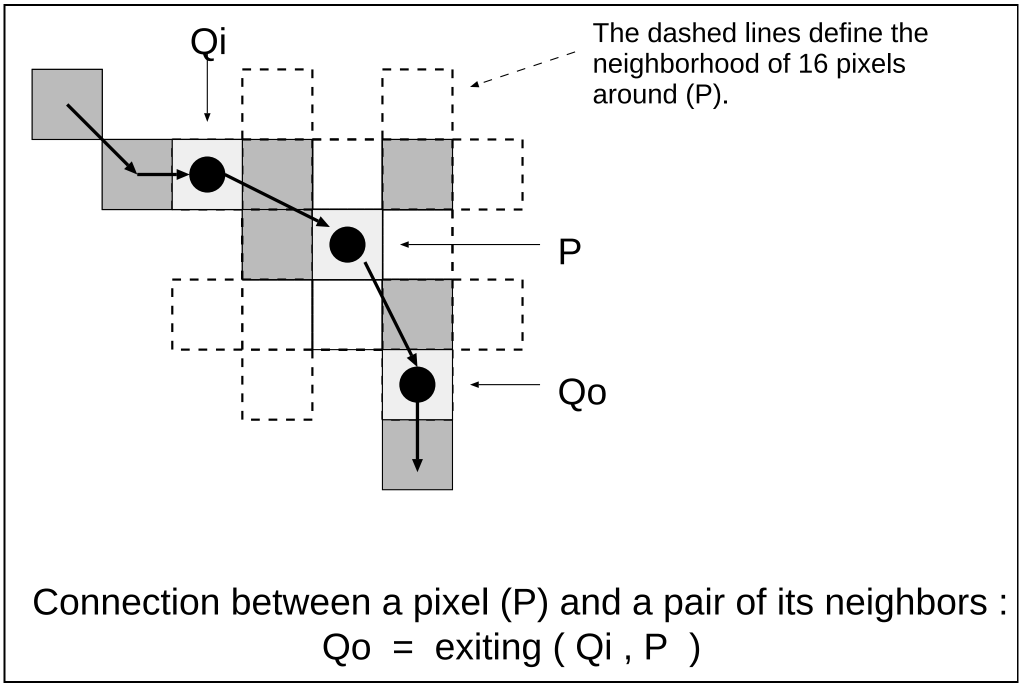


FIG. 9 - *Example of a grouping of pixels*

For each pixel :

- A grouping is defined by the association of two neighbors according to the best values of the quality function.
- Along the iterations, the importance of isolated primitives is decreased with respect to primitives which are part of larger structures.

Grouping between Chains of Contours

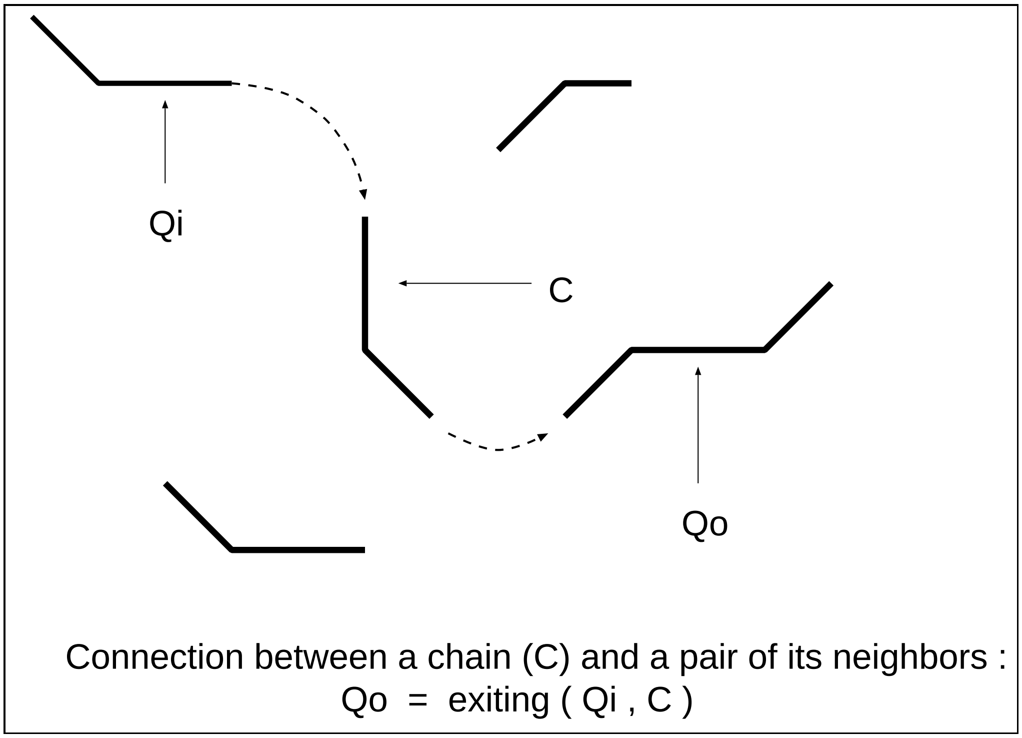


FIG. 10 - *Example of a Grouping of Chains*

Cas du groupement de chaines :

- The principle is the same than with pixels. The neighborhood is variable in this case.
- The extreme points of chains are grouped together.

  It is necessary to segment chains in such a way that "T" Junctions become possible.

Results with Noisy Images

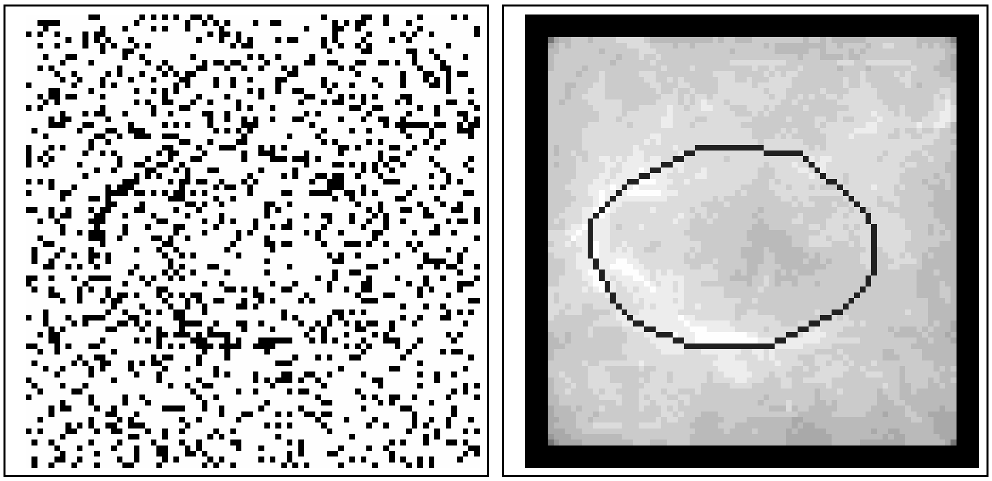


FIG. 11 - *Ellipse with white noise*

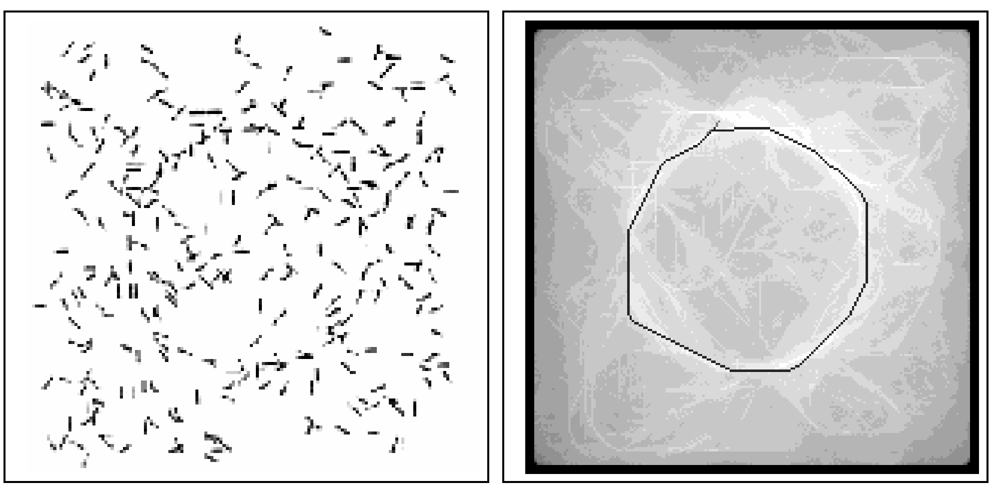


FIG. 12 - *Cercle with "oriented noise"*

Results with real situations : Satellite pictures

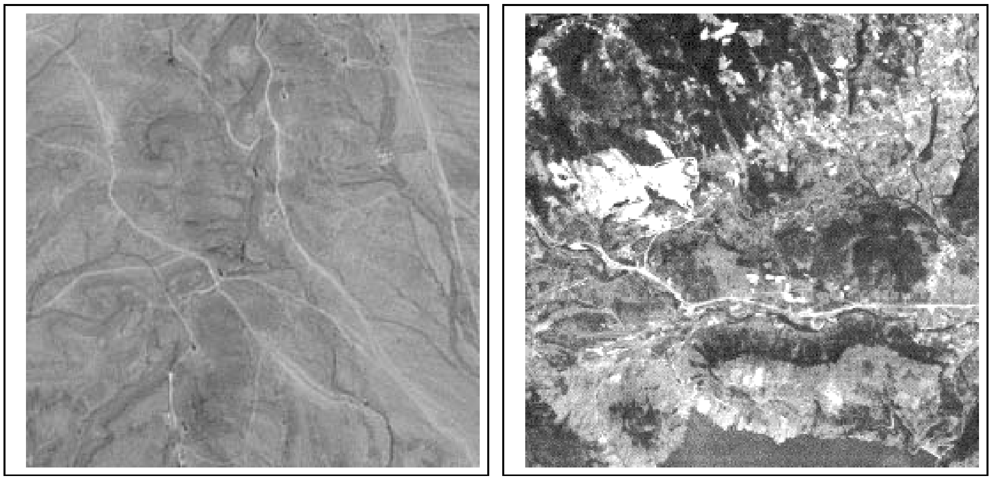


FIG. 13 - *Satellite Pictures - Original Images*

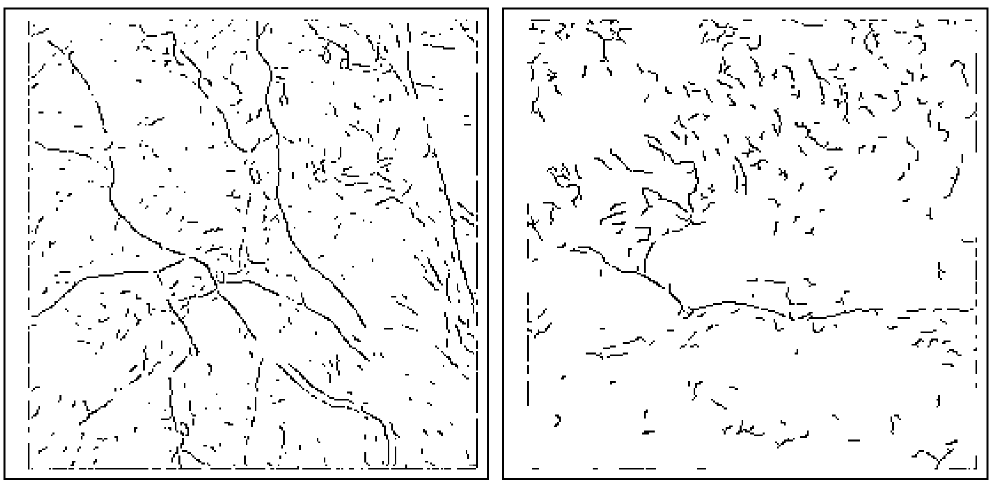


FIG. 14 - *After crest-lines detection*

FIG. 15 - *Selection of the main groupings according to their global quality*

Results with real situations : Indoor scenes

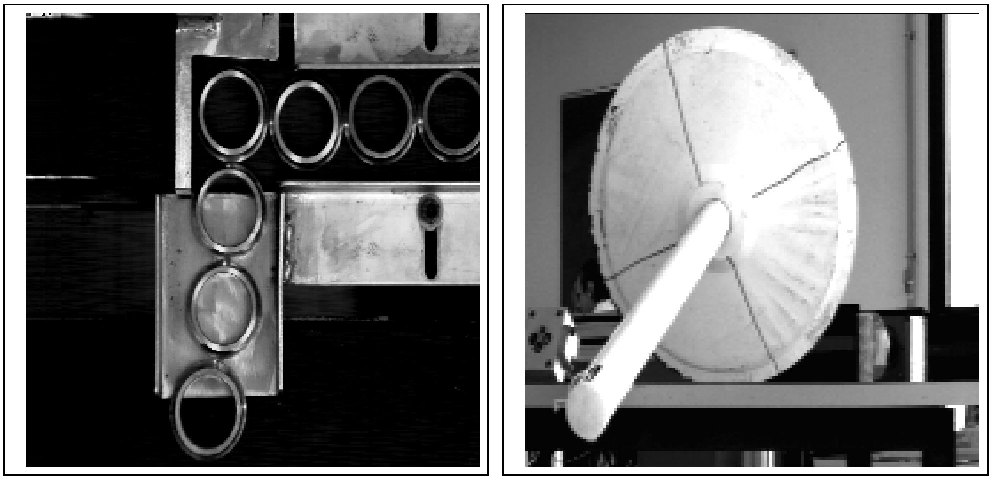


FIG. 16 - *Indoor Scenes - Original Images*

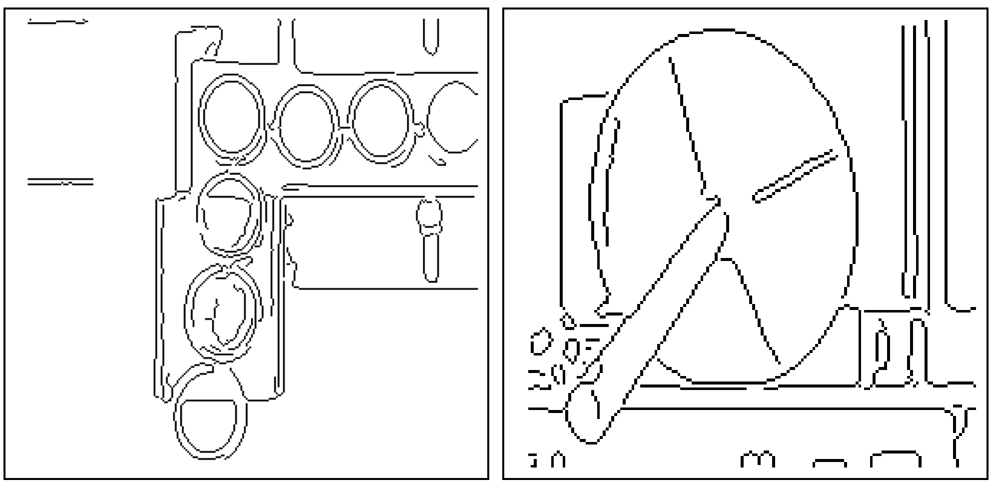


FIG. 17 - *After contour detection*

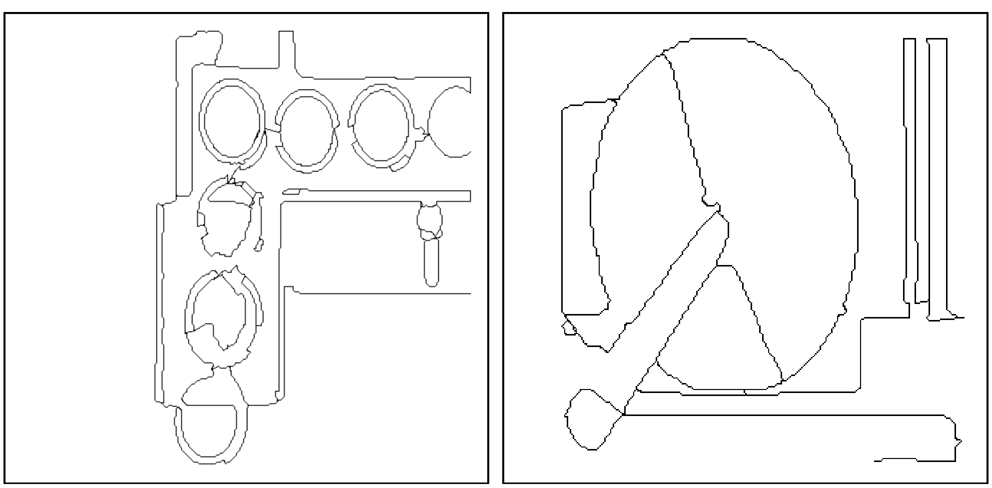


FIG. 18 - *Selection of the main groupings according to their global quality*

Conclusions for Grouping by Continuity

– The most salient structures are extracted without prior knowledge of the scene.

– The method is robust to noise and stable according to the choice of parameters.

– The salient groupings provide a reduced and optimized number of chains for the further levels of description of the scene.

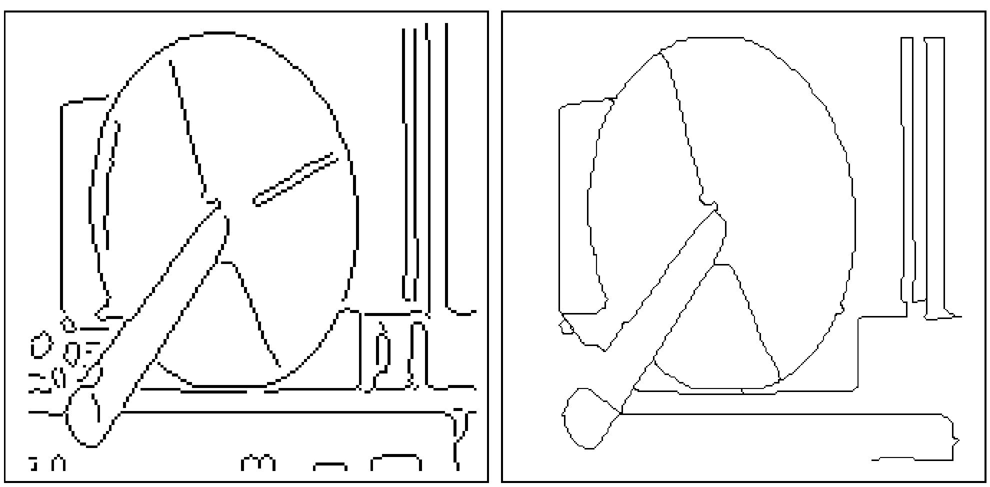


Fig. 19 -

Grouping by Colinearity / Cocircularity

Each salient chain is segmented into parametric primitives : Points of Interest, Segments and Arcs.

This level is achieved incrementally :

- **Detection of Segments - Colinearity**

  The segments are detected by polygonal approximation of the chain.

- **Detection of Arcs - Cocircularity**

  The junctions between each segment provided by the polygonal approximation are analysed in order to detect the Arcs.

- **Interest Points**

  The interest points are detected from the junctions between Segment-Segment, Arc-Arc or Arc-Segment (Corners, Tangent Points, Inflexion Points).

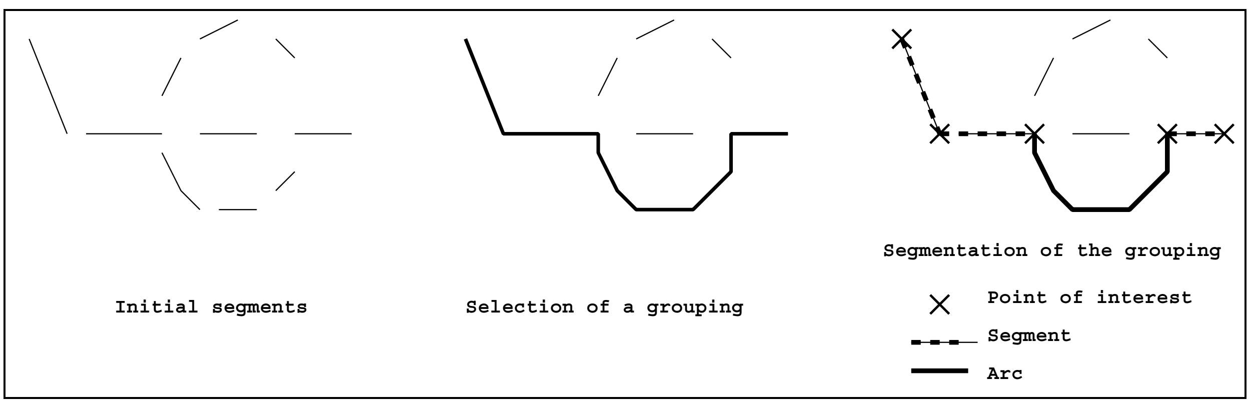


FIG. 20 - *Selection and segmentation of a grouping*

Colinearity : Detection of Segments

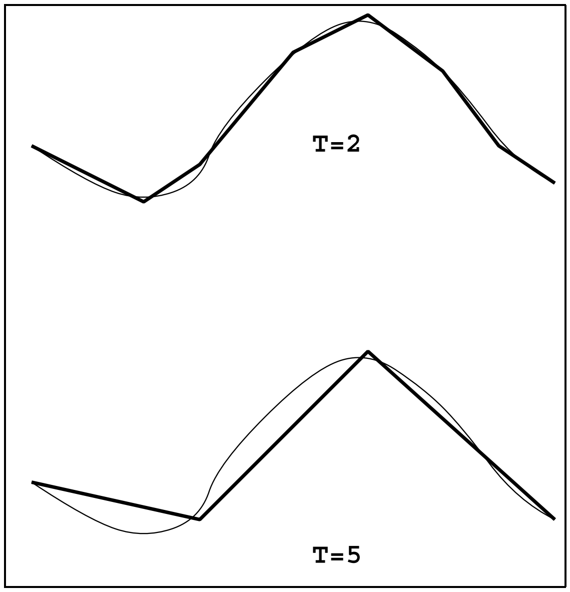


FIG. 21 - *Polygonal Approximation*

- **Detection of "cut" points**

  Quick and straight forward polygonal approximation (estimation of the surface between the chain and the current segment).

- **Grouping of segments**

  A coarse approximation misses circular arcs. It is necessary to detect circular parts along the chains with precise polygonal approximation.

  Colinear segments are then grouped together into optimized segments.

Cocircularity : Detection of Arcs

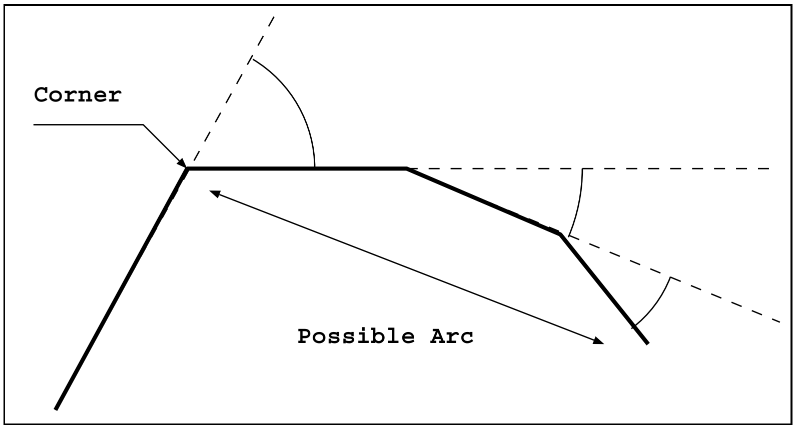


FIG. 22 - *Detection of Arcs*

- **Junctions between segments**

  Corners or possible arc are given by the amplitude of the angle between two following segments. Above a fixed threshold, the junction is considered as a **corner**. Otherwise, an hypothesis for an arc is started.

  The sign of this angle gives the changes in curvatures and define the presence of an **inflexion point**.

- **Detection of Arcs**

  The hypothesis for an arc is closed when a new corner or inflexion point is found. If a parametric equation can be found for this hypothesis, an **arc** is detected. If not, the hypothesis is considered as an undetermined curve.

Finally, **tangent points** are detected by comparing the extreme points of arcs and segments.

Results on the test scene

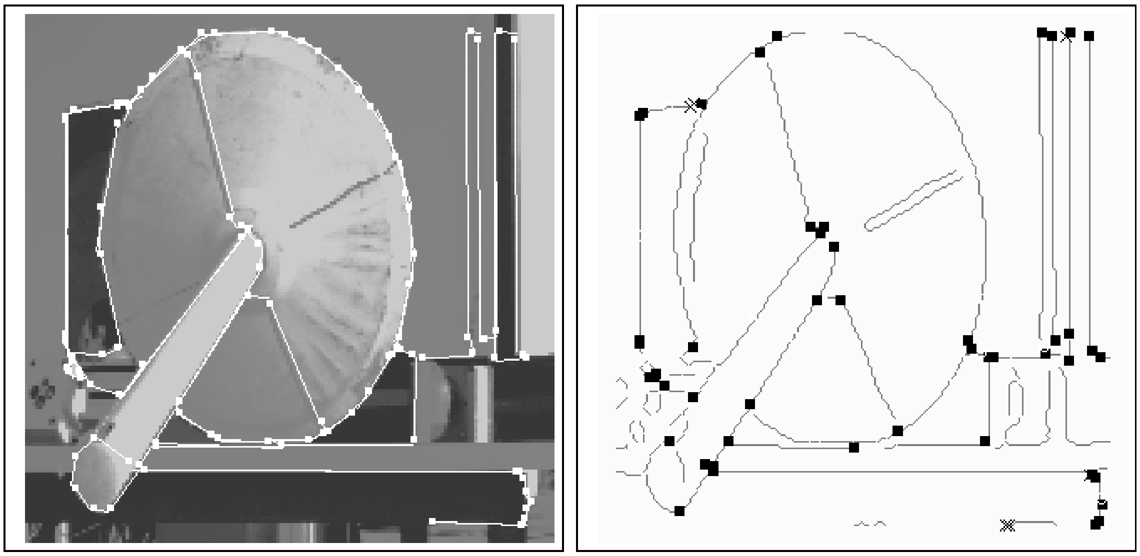


FIG. 23 - *Segments and Interest Points*

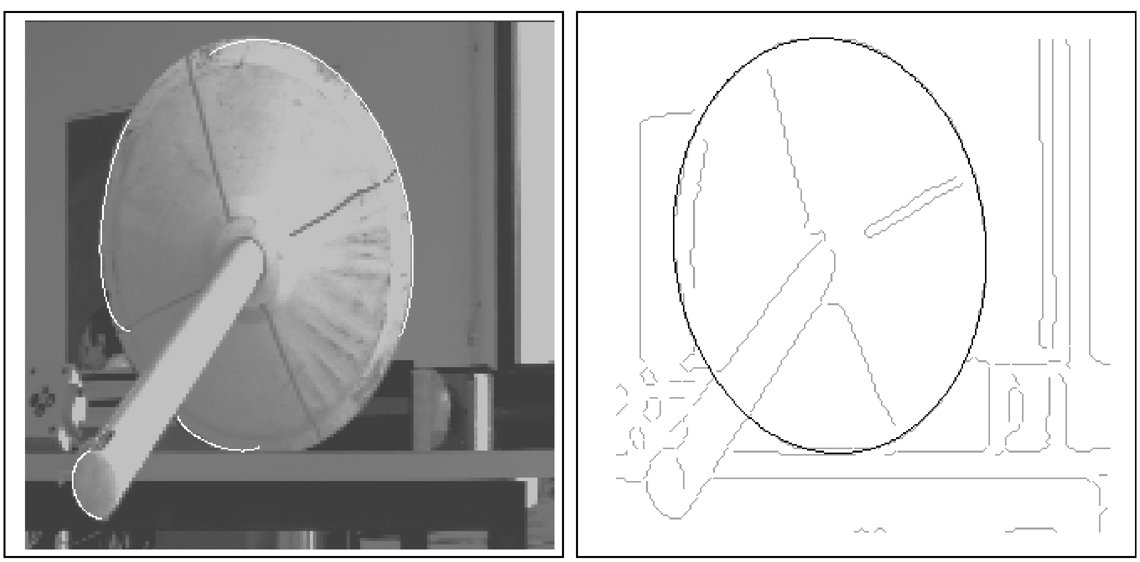


FIG. 24 - *Detection of Arcs*

Detection of Arcs and Ellipses

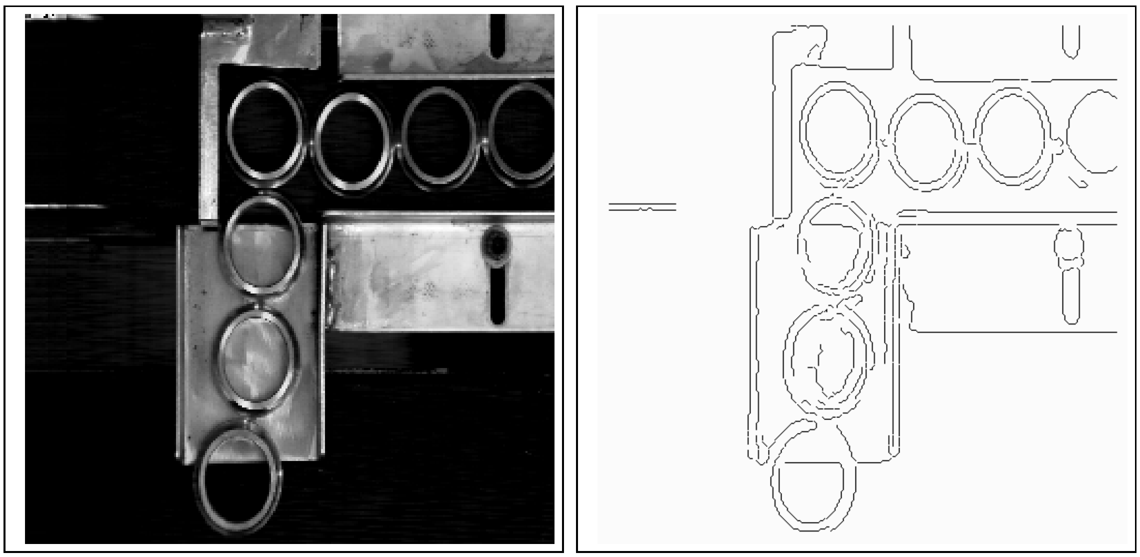


FIG. 25 - *Other example for the detection of ellipses*

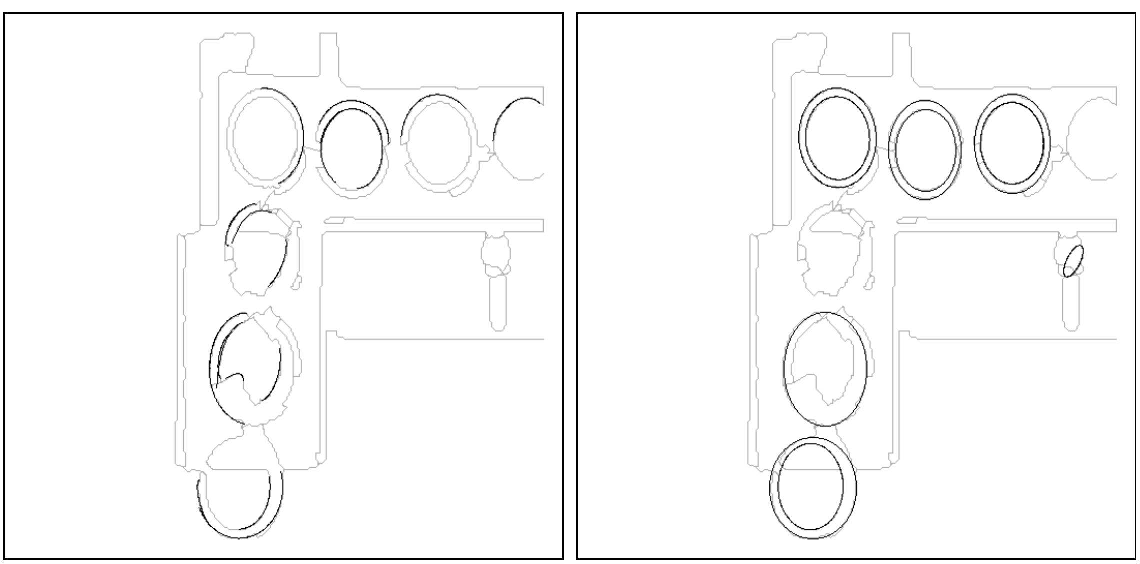


FIG. 26 - *The ellipses have been completed from the best arcs*

Perspective : Feature Matching

The main application of this description is to help the feature matching process in a 3D perception of the scene.

Current work is developping the elements of matching :

- **Detection and classification of Junctions**

  Each point of interest is analysed and associated to a Junction according to the configuration of segments and arcs in its neighborhood.

- **Matching of Junctions**

  By a relaxation process on the length and angles of segments around each junction.

- **Detection of complex features**

  The reduced number of segments (with respect to the usual segments of contours) helps with the construction of convex groups of segments (polygones) and with the detection of higher level visual properties (such as symmetry).

Conclusion

We presented a complete scheme for the description of a scene into geometric features.

- **Incremental Perceptual Grouping**
  Geometric primitives detected from a local to global point of view.

- **Efficient extraction of salient structures**
  Due to a robust preliminary level of Grouping by Continuity.

- **Focus of Attention**
  This reduced number of groupings has the function of *focus of attention* as it restricts the shape extraction process to the salient structures only.

- **Description of the scene**
  Provided as a graph of relationships between geometric primitives.

Future developments expected

- **Possible contributions from other cues**

  Such as other low level processing (a segmentation in regions could be useful to validate some hypotheses) or possibly other views of the same pictures with different scales.

- **Grouping according to more complex relationships**

  Such as Convexity or Symmetry.

- **Feature Matching**

  Between multiple images for a 3D reconstruction of the scene or a virtual camera.

- **Image Indexing**

  To help the retrieval process in a database.