# IMAGE - SHAPE RECOGNITION

**Eugen Ganea**, M.Sc., **Mihai L.Mocanu,** Ph.D.

*Department of Software Engineering, School of Automation, Computers and Electronics, University of Craiova, 13 A.I. Cuza Str., 1100 Craiova, Romania*
Phone: (0251) 435724/ext.163, Fax: (0251) 438198, *Elm: ganeae@nt.comp-craiova.ro*

**Abstract:** The paper focuses on the specific software architecture of some image processing applications. Image processing related fields are vast and based on a very precise mathematical foundation. An application written in Java, called JImageProcessing, is presented in the final part of the paper.

***Key words:*** image processing, shape recognition

## Introduction

Managing images is a large and independent area. This domain's base is a very precise mathematic theory, but, generally, the implementations on different digital systems are huge consumers of resources (such as memory and computing power), especially if we consider real time applications obtaining and using image information.

In a very large sense, an image is a description of the variation of a parameter on a surface. For example, the images are the result of the variation of the intensity of light in a bi-dimensional plane (the classic definition). This parameter (intensity of light) is not the only one used. An image can also be generated by the temperature of a digital circuit, or, at the other end, by the emissions of radiations (of different wavelengths) of different galaxies, etc. What is still a common pattern is that these kinds of images are usually converted to the classic images, i.e. color is changed in a way or another with special algorithms. So the human operator can do a visual evaluation of the variation of some parameters.

This is why in this study we are going to refer only to the "classic" images. Within the classic conception, an image is a bi-dimensional signal; naturally, the processing of the images would be consider as a branch of digital signals processing (which also includes telecommunications and audio signals processing).

## Image Processing in the Large

Processing the images and the digital processing of the signals require many hardware resources. The algorithms can be implemented on common digital systems (PC – the evolution of microprocessors offers the necessary calculation power), but for dedicated real time systems are used dedicated microprocessors – Digital Signal Processors. This kind of processors implements (hardware) different optimizations and parallelisms to offers the necessary calculation power (usually with lower frequencies than the classic microprocessors). Obviously, they need a lower power level (the power consumed is reduced).

Processing the images includes or is related to others domains:

- Acquisition, compression and storing the images
- Restoration images with geometrical corrections, noise filtrations and contrast adjustments
- Photogram (phenomena and objects made by photos) measurement
- Pattern matching, shape recognition, face recognition
- Artificial vision (computer vision, robot vision)
- Artificial intelligence
- Image synthesis
- Computer image generation, a.o..

Artificial intelligence and image processing are related domains. An important number of powerful algorithms of image processing domain use methods and artificial intelligence techniques and concepts such as neuronal networks, fuzzy logic, etc. On the other hand, artificial intelligence supposes designing and construction of some systems capable of performing human activities: learning from experience, understanding natural language, taking decisions or using analogisms to succeed in resolving different problems.

All of these require a minimum quantity of information (such as environmental information). This information is obtained by intelligent systems through the sensors and creates an image of the environment at a given time (snapshot). Then, useful information still needs to be extracted. This technique is related to the artificial vision domain (Computer Vision, Robot Vision) – this is a discipline commune to both artificial intelligence and image processing and tries to find answers to some questions such as:

- What kind of information must be extracted from the acquired images?
- How can be extracted this information?
- How can this be represented?
- How can this be used to accomplish some purpose?

Shape recognition represents a procedure employed to extract information from acquired images. It is a very large field that include human face recognition, handwriting recognition, finger – prints recognition, etc. Shape recognition means a classification and/or a description of the image contain. The classification consist in attribution of an unknown shape from the acquired image to a class from a set of predefined classes; the classifying operation produce an output image which represents a map of the objects of the scene in the image. In the new image, the values of the pixels represent actually the codes associated to corresponding classes. This classification uses mathematic methods of theoretic decisions or statistics, methods that are based on a few elements of the theory of statistic decisions.

The classifying algorithms divide in two categories:

- Supervised algorithms (semi-automatic algorithms) – the presence of a human operator is necessary at the beginning of the classifying process; the human operator has to specify the number of classes to result, a few characteristics of the classes, etc.;
- Automatic algorithms (self-loading) – classify the image but without knowing the meaning of each class .

Classifying algorithms are based on the extraction of the features on a measure of similarity (a distance for example).

An important step in designing the automatic classification systems is the selection of characteristics (features) because the components of the features vector suppose the presence of a big quantity of information. This selection is independent from the number of classes and the analyzable shapes. A classifier generally contains three modules: the proper classifier module, a learning module (that suppose the presence of a set of samples and images to practice on) and a module of selection and features extraction. The second module may contain a database with different shapes features that can appear in the scene (spectral or geometrical signatures for example). Learning suppose the existence of this set of "training" that contains samples for that the appurtenance of classes is well known. The presence of the learning module is required for the supervised classification.

When the shape to recognize are very complex or the number of classes is high and cannot be determined (the elements - part of the object can combine in an infinity of ways), then the theoretic-decisional methods cannot be used. This is the case of the scenes, for example in robotics, when the objects arranged one by another must be recognized and extracted one by one from the stack. The complex form is seen as a combination of simpler forms which are recognized one by one and than are established the relationships between them, on regard of describing the object. If every of this form is, at its turn, a complex form, than simpler forms, represent it until the most simpler forms, called *primitives*.

The next figure represents a structural-syntactic block of recognition of shapes:

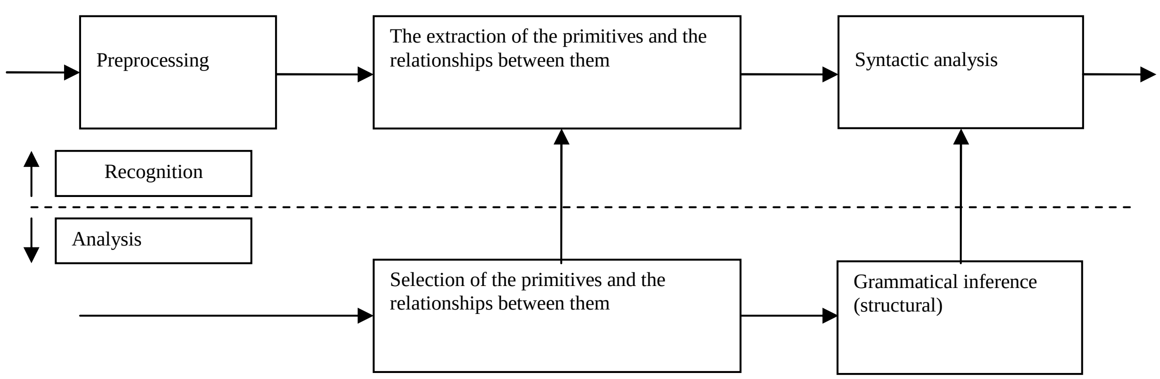


**Figure 1.** Shape recognition block

The recognize section is represented by processing, the extraction of primitives and the structural – syntactic analysis, including the selection of the primitives and the structural inference. In the syntactic recognition, a form is represented by a set of primitives. The rules followed by the components of the shapes are specified by the grammars of the forms description language. The structural recognition may use a relational graph or another type of representation to specify the relationships between shapes on regard of describing the object or the shape.
There is no general rule to select the primitives. Usually, the primitives are easily recognized. Many times, a compromised must be made between the complexity of the elementary shape and the complexity of the description grammar. Generally, as the shapes are more complex, the relationships between them are simpler.

After the selection of the primitives, the next step is the construction of some grammars capable of generating one or more languages to describe the studied shapes. The selection of one grammar depends of the selected primitives and of the compromise between the descriptive power of the grammar and the efficiency of the syntactic analysis.
The decision of recognition or acceptation of a shape is taken by syntactic analyzer. We usually use one or more grammars for each of the class or shape, the classification meaning the determination of the grammar that accepts a representation of a shape as correct.
On regard of the information presented before, the main methods of shape recognition can be organized as in the following scheme:

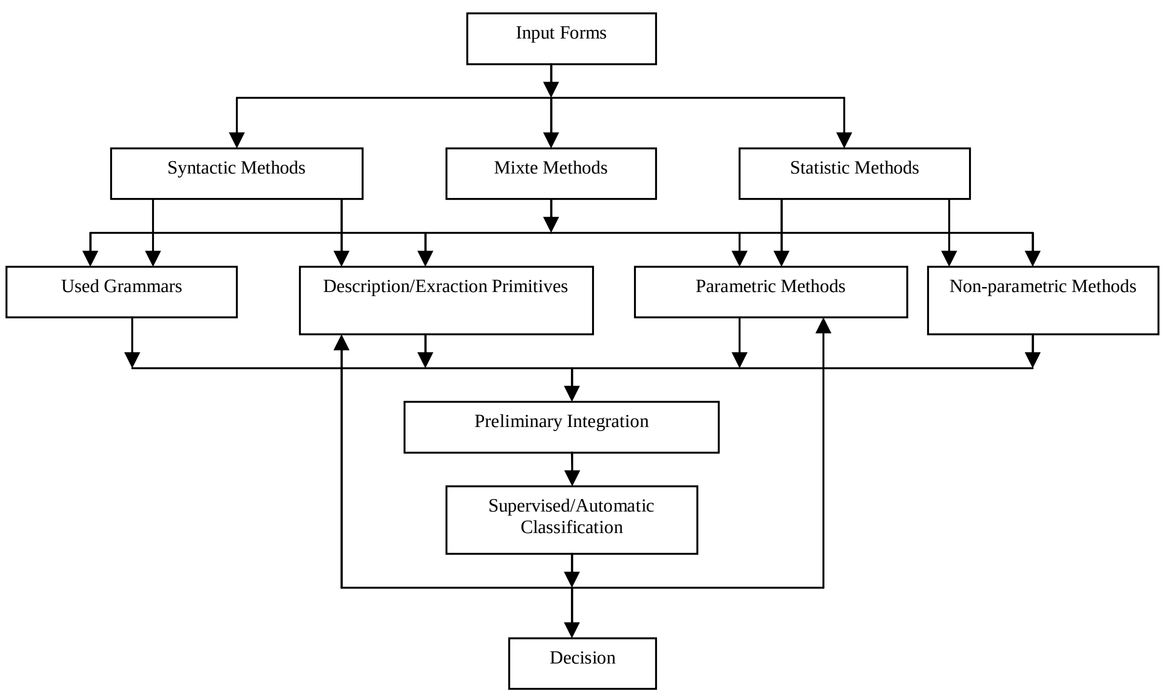


**Figure 2.** Shape recognition methods taxonomy

**Java Image Processing**

JImageProcessing is a Java application, which has as main purpose shape recognition in an image. The use of Java is intended to ease the future embedding of this application in a large category of applications dealing with images over the Internet (i.e. browsers).
The techniques used are very basic at this moment. To extract the typical information, a series of filters is first applied to the image. These filters transform the rough image in temporary images that serve to the final purpose.
The following filter types have been used:
- Low filter – used to eliminate the noise; the specter of the image is homogenized
- High filter – usually used to mark the outlines owing to its behavior of derivation.

The first filtering operation applied is used to eliminate the noise and the useless information.

The next step is the binary encoding of the image, that is, the colored image is transformed into an image with different levels of gray. Five different methods of edge separation (determination) were employed, thus making a comparative study. These five methods can be structured into two categories:

1. Dashed (linear) masking, based on:

- gradient methods for vertically edges detection; the gradient for horizontally edges detection;
- directional operators methods for horizontal lines detection;
- directional operators methods for 45 degrees inclined lines.

2.Non-liniary masks based on methods:

- the Sobel masks for horizontal detection, for vertical detection;
- the Prewitt masks for every direction : horizontal and vertical.

Those operators are gliding window type. The extraction of the edge with the help of those operators is a succession of convolutions from the initial image and the mask (kernel) of the operator. Optionally, for each operator, in the end of the convolutions, a limit segmentation can be done to obtain a binary image of the maps of the outlines. These image occupies less memory, the shapes recognition algorithms are faster.

In the gradient case, for example, the edges from an image have been extracted by comparing the norm of the gradient with a limit T:

norm $\geq$T – the point is on the outline

norm $<$ T – the point is not on the outline

On these outlines obtained by applying one of the operators enumerated, I applyed algorithms to make them thin or to fence them. For that, we use morphological operators of erosion and dilation. To define them, we consider a binary image. This image contains the interes object and the background (we are not interested of the background), for example the background is white and the object is black. By erosion every pixel of the object that touches the background is made background–pixel. By dilation, every background pixel that touches the object is made object–pixel. The opening is a erosion followed by a dilation and the closing is a dilation followed by a erosion. The first operation is used to eliminate from the image the little islands of object's pixels (these islands usually appear at the binarisation operation because of a bad choice of the limit). In the same way, the closing eliminates such islands from the background. These technics are used to process the noisy images where a part of the pixels may have a bad binary value.

The difference between edges detection and outline extraction is, more often, essential. Generally, an edge detector supplies a binary image, containing one pixel breadth lines in the positions in which the presence of a an edge or the border between two different regions is estimated. Spacial edges detectors realise the extraction of the outline if they are followed by other operations mean to assure binary image containing one pixel breadth lines.

After the identification of the outlines in the image, an outlining algorithm was implemented (called suggestively "the bug walk"). The steps of the algorithms are the following:

1. choose a way to track the outlines
2. localize a first pixel of the object, usually by line by line or column by column scavenging . This pixel is marked as the start pixel. It becomes the current pixel.
3. scan the neighbours of the current pixel, in the chosen way of following the outlines, until a new pixel-object is found. This one is marked and become the current pixel.
4. the third step is repeated until the outline is full covered – until the start-pixel become again the current pixel.
5. if the localisation of another objects in the images is wanted, the algorithm is repeated, beginning with the second step.

The following figure presents the result obtained after the application of the edge detection method and this algorithm:

**Figure 3.** Object separation

The next important problem in the image processing for shape recognition is the separation of the different objects in an image. This operation is called *segmentation*.

There are many methods for image segmentation. One of them, the one I have implemented, is that that „grows" the areas around some „centers".

- by a certain method, we mark the centers of the main objects of the image. We can do that with information from the user , by choosing the maximum contrast points, or with other information, depending on the application
- next, we calculate the "distance" between each two points in the image (we can use four or eight neighbours) : the more different are the colors, the greater is the distance
- we follow next an iterative calculation process that takes each point and attribute it to the same area as the closest neighbour point. This process can be done efficiently on a multi-processor machine
- the algorithm ends when every point is attributed to one of the initial areas.

For objects recognition, we use a set of models of the objects we want to recognize. We have build a "matching" filter – it "matches" the objects obtained after the application of the exposed algorithms and the models of the objects to recognize. The "perfect match" is very rarely obtained, because of the presence of noises, digitized image effects , etc.

## Conclusions

Images computer processing finds applicability in many various domains such as: radiographic images interpretation, the study of pollution using images acquired by the satellites, archeological objects image restoration. Another category is the processing and the extraction of some features used to recognize objects and automatic classification: automatic recognition of handwriting, identification of targets in military applications, automatic control of products quality, the automatic processing of satellite images to establish the weather forecast.

## References


1. K. R.: Digital Image Processing, Prentice Hall, Englewood Cli.s, NJ,1996
2. Image Processing, Analysis and Machine Vision, M Souka, V Hlavac and R.Boyle, Chapmann & Hall.
3. I. T. Young, J. J. Gerbrands, L. J. van Vliet, Fundamentals of Image processing
4. T. Pavlidis, Algorithms for Graphics and Image Processing, Computer Science Press, 1982
5. L. Hong, Y. Wan, A. Jain, Algorithm and Performance Evaluation, Pattern Recognition and Image Processing Laboratory, Department of Computer Science, Michigan Stat University, USA
6. Cocquerez, J. P., Philipp, S. (coord.): Analyse d'images: filtrage et segmentation,Masson, Paris, 1995
7. Dougherty E. R., Giardina, C. R.: Image Processing - Continous to Discrete, vol. 1,Geometric, Transform and Statistical Methods, Prentice Hall Inc., Englewood Cli.s,1987
8. Gonzales, R. C., Woods, R. E.: Digital Image Processing, Addison Wesley, Reading MA, 1992
9. Haralick, R. M., Shapiro, L. G.: "Glossary of Computer Vision Terms", în Pattern Recognition, vol. 24, no. 1, pag. 69-93, 1991
10. Haralick, R. M., Sternberg, S. R., Zhuang, X.: "Image Analysis using Mathematical Morphology", în IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 9, no. 4, Iulie 1987, pag. 532-549
11. Jain, A. K.: Fundamentals of Digital Image Processing, Prentice Hall, Englewood Cli.s NJ, 1989
12. Pitas, I., Venetsanopoulos, A. N.: Nonlinear Digital Filters — Principles and Applications, Kluwer Academic Publ., Norwell MA, 1990
13. Press, W. H., Flannery, B. P., Teukolsky, W. T., Vetterling, W. T.: Numerical Recipes in C. The art of scientific computing, Cambridge University Press, 1988
14. Serra, J.: Image Analysis and Mathematical Morphology, Academic Press, London,1982