Comparing Neural Networks, Invariant Moments and Mathematical Morphology Performances for the Automatic Object Recognition

G. Santilli [a,*], G. Laneve [b]

[a] School of Aerospace Engineering, University of Rome "La Sapienza", Rome, Italy – santilli@psm.uniroma1.it
[b] Centro di Ricerca Progetto San Marco, University of Rome "La Sapienza", Rome, Italy – laneve@psm.uniroma1.it

Abstract – Pattern recognition is an essential part of any high-level image analysis system. The CRPSM, in the framework of the European Projects GMOSS and G-MOSAIC, has developed some techniques able to automatically recognize and extract potential made-man structures which could be present in complex aerial and satellites images. In particular, this paper aims at describing several of the developed techniques which allow the automatic detection of given objects of interest. These techniques are based on different approaches, therefore the results provided by them are compared and the their advantages and disadvantages are highlighted.
The purpose described above is obtained by using several algorithms developed by CRPSM in the last few years based on the Mathematical Morphology, Geometrical Moment Invariant and Neural Networks approaches.

Keywords: Image processing, Mathematical Morphology, Moment Invariants, Neural Networks.

1. INTRODUCTION

Pattern recognition is an essential part of any high-level image analysis system. Most of these systems share a general structure made of four building blocks: image acquisition, pre-processing of the images, feature extraction, and classification (Khotanzad & Lu, 1990). To improve the classification results in various conditions, many approaches have been studied including decision theory, feature selection, optimization, learning and so on (Kim & Han, 1995).

Several European projects as GMOSS (Global Monitoring for Stability and Security, VI Framework Programme) and G-MOSAIC (GMES services for Management of Operations, Situation Awareness and Intelligence for regional Crises, VII Framework Programme) are devoted to show the applicability of the remote sensing images in security related applications. In fact, during the accomplishment of these projects several scenarios have been considered (hypothetical and real situations regarding disastrous events requiring a prompt delivery of accurate information were and are carried out). Among the many objectives of the Network the work packages (WPs) in which CRPSM (Centro di Ricerca Progetto San Marco - Sapienza Università di Roma) has been involved regard the detection of informal settlements, the monitoring of refugee camps (Laneve et al., 2006), the monitoring of national borders (WP Monitoring Population and Borders) and the monitoring of nuclear plants, ports and airports (WP Monitoring of Critical Assets). In the mainframe of these WPs the CRPSM has developed some techniques able to automatically recognize and extract potential made-man structures (Laneve & Santilli, 2006) which could be present in complex aerial and satellites images.

This paper aims, mainly, at describing the developed techniques which allow the automatic detection of given objects of interest. These techniques, based on different approaches, are compared and the advantages and disadvantages of each of them are highlighted. The algorithms, based on the theory of Mathematical Morphology (Dougherty & Lotufo, 2003) Geometrical Moment Invariants (Hu, 1962) and Neural Networks (Paschalakis, & Lee, 1999) were applied to HSR (High Spatial Resolution) or VHSR (Very High Spatial Resolution) satellite images like Quickbird, Ikonos, SPOT, ASTER and aerial images.

These algorithms have been developed exploiting the functions available in several Matlab Toolboxes (Gonzalez and Woods, 2002, Gonzalez et al. 2004). The final objective of this process is the development of an efficient and flexible recognition system that can use, in automatic way, the best technique to recognize and to extract, as accurately as possible, the required objects present in the images, regardless of orientation, size and position of them.

2. DATA AND METHODS

This paper aims at showing the results obtained by applying Mathematical Morphology, Moment Invariants and Neural Networks techniques to optical satellite and aerial images in order to compare all these techniques and to highlight the advantages and disadvantages of each of them.

2.1. Mathematical Morphology

The first method used by CRPSM, to extract the object of interest, relies on an object-oriented approach based on a theory for the analysis of spatial structures called *Mathematical Morphology* (Giada et al. 2002). It is called morphology because it aims at analysing objects shape and size. It is mathematical in the sense that the analysis is based on the set theory, integral geometry, and lattice algebra.

Mathematical morphology has proven to be a powerful image analysis technique. In mathematical morphology, two-dimensional grey tone images are seen as three-dimensional sets by associating each image pixel with an elevation proportional to its intensity level. An object of known shape and size, called the *structuring element*, is then used to investigate the morphology of the input set. This is achieved by positioning the origin of the structuring element to every possible position of the space and testing, for each position, whether the structuring element either is included or has a non-empty intersection with the studied set. The shape and size of the structuring element must be selected according to the morphology of the searched image structures.

The basic idea of this first method is to extract, in automatic way, information about objects of interest from satellite images exploiting not only the spectral characteristics, but also the morphological characteristics in order to simplify the images analysis by the users. The procedure results in nonlinear image operators which are useful for exploring geometrical and topological structures. A succession of such operators is applied to an image in order to make certain features apparent (Soille & Pesaresi, 2006). The methodology of the processing chain applied in the case of the optical images includes two phases (the *pre-processing* phase and the *processing* phase) which can be summarized as follows:

the *pre-processing phase,* necessary for removing the noise presents on the original image and for improving the contrast by allowing, to the morphological operators sequence, enhanced performances. This phase comprises following steps:

- *Contrast improvement*: it is obtained by a transformation made by a convolution between the original image and the suitable filters:
- *Histogram equalization*: this transformation aims at further improving the contrast image by an equalization made on areas smaller than the whole image area;
- *Gamma correction*: this transform is able to generate an improved gray-tone image, as in the case of roads shown in Fig. 1.

The *processing phase*, if the objective is the automatic roads extraction (Fig. 1), foresees the following steps:

- Definition of a linear and narrow structuring element (SE) since these are the main characteristics of the searched object (roads);
- Image analysis by a structuring element having proper length (in agreement to the roads length) and variable direction between 0° and 180°, with suited step (in agreement to the roads thickness). The complete image is studied and the roads with a longer length than the SE, in any direction, are found.

2.2. Geometric Moment Invariants

The second method used by CRPSM, to extract the objects of interest, is based on the *Geometric Moment Invariant* (GMI) theory**.** This technique has been chosen to extract image features since the generated features are Rotation, Scale and Translation (RST)-invariant (Saad, 2004). This technique is widely used to extract global features for pattern recognition due to its discrimination power and robustness. In this study, GMI is used to characterize and identify several aircrafts from satellite/aerial images. Geometric moments and their invariance properties have been successfully used in different pattern recognition tasks and are one of the most popular and widely used shape descriptors introduced by Hu (Hu, 1962).

Two-dimensional moments of a digitally sampled $M \times N$ image described by means of the gray function $f(x, y), (x, y = 0, .. .M - 1, N-1)$ is given as,

$$m_{pq} = \sum_{x=0}^{x=M-1} \sum_{y=0}^{y=N-1} (x)^p.(y)^q.f(x,y) \qquad (1)$$

$p,q=0,1,2,3\ldots$

The moments of $f(x, y)$ translated by an amount (a, b), are defined as,

$$\mu_{pq} = \sum_x \sum_y (x + a)^p.(y + b)^q.f(x,y) \qquad (2)$$

Thus the central moments m_{pq} or μ_{pq} can be computed from (2) by substituting $a = -\bar{x}$ and $b = -\bar{y}$ where,

$$\bar{x} = \frac{m_{10}}{m_{00}} \quad \text{and} \quad \bar{y} = \frac{m_{01}}{m_{00}}$$

$$\eta_{pq} = \sum_x \sum_y (x - \bar{x})^p.(y - \bar{y})^q.f(x,y) \qquad (3)$$

When a scaling normalization is applied the central moments change as,

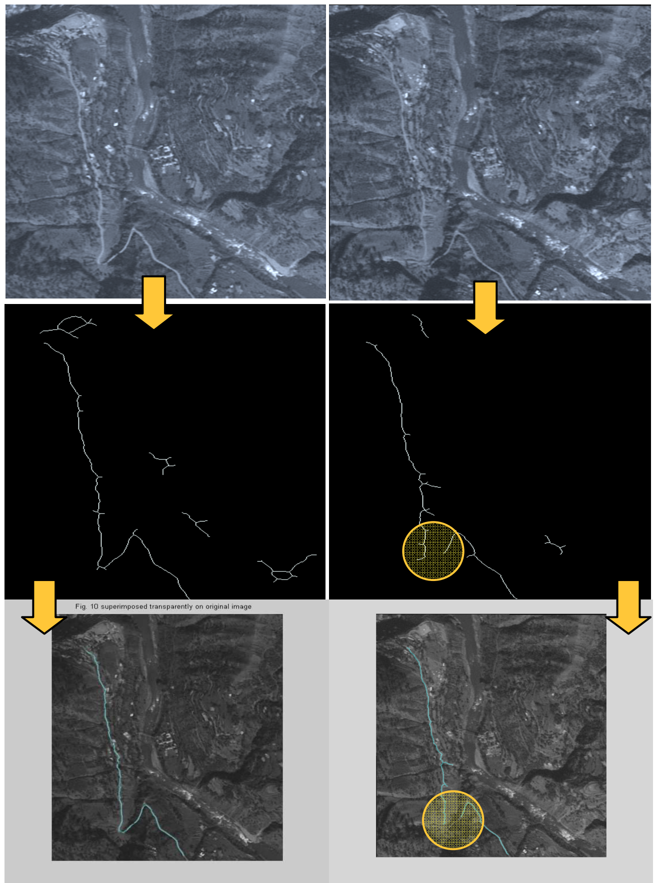

Figure 1. Automatic procedure to search the roads present on the SPOT 5 images (5 m resolution). The technique was applied before (above) and after (below) the earthquake that struck the Kashmir region on October 2006. This technique allowed to find the damages on the infrastructures.

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}}, \qquad \gamma = [(p + q)/2] + 1 \qquad (4)$$

Individual moments values do not have the descriptive power to uniquely represent arbitrary shapes, nor do they possess the required invariance characteristics, but, sets of functions based on these moments can be determined which do. In particular, Hu defines a set of seven values, computed by normalizing central moments through order three, that are invariant to object scale, position and orientation. In terms of the central moments, the seven moments are given as:

$$M_1 = \eta_{20} + \eta_{02} \qquad M_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$
$$M_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$$
$$M_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$$
$$M_5 = (\eta_{30} + 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} - \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} - \eta_{03})^2]$$
$$M_6 = (\eta_{20} - \eta_{02})[(\eta_{30} - \eta_{12})^2 - (\eta_{30} - \eta_{12})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} - \eta_{03})$$
$$M_7 = (3\eta_{21} - \eta_{30})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{30})^2] + (3\eta_{12} - \eta_{30})(\eta_{12} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} - \eta_{03})^2] \qquad (5)$$

Classification is achieved by matching a shape vector extracted from an image with previously encountered shape vectors from the training set. The moments described above can be calculated either from an image or from a shape's boundary. Jiang &

Bunke, (1994) have show that the two different calculations are mathematically equivalent.

2.3. Neural Networks

The third method used by CRPSM, to extract the objects of interest, is based on the *Artificial Neural Netwoks* (ANN) theory**.** The manner in which the neurons of a neural network are structured is intimately linked with a learning phase used to train the network (Dutta, 2007). Fig.2 presents the architectural layout of a multilayer perceptron.

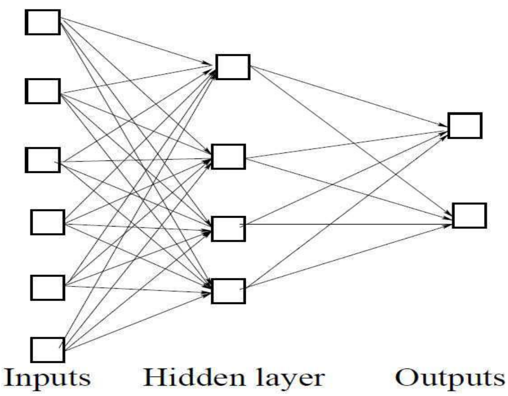

Figure 2. Schematic diagram of typical a multiplayer feed-forward network.

This class of feed-forward neural network distinguishes its-self by the presence of one or more hidden layers, whose computation nodes are correspondingly called hidden neurons or hidden units. The function of hidden neurons is to intervene between the external input and the network output in some useful manner. By adding one or more hidden layers, the network is enabled to extract higher-order statistics. The ability of hidden neurons to extract higher-order statistics is particularly valuable when the size of the input layer is large.

The back-propagation algorithm (Haykin, 2004) is used to train a given feed-forward multilayer neural network for a given set of input patterns with known classifications. When each entry of the sample set is presented to the network, the network examines its output response to the sample input pattern. The output response is then compared to the known and desired output and the error value is calculated. Based on the error, the connection weights are adjusted. The backpropagation algorithm is based on Widrow-Hoff delta learning rule in which the weight adjustment is done through mean square error of the output response to the sample input. The set of these sample patterns are repeatedly presented to the network until the error value is minimized. The sequential updating of weights is the preferred method for offline implementation of the back propagation algorithm. For this mode of operation, the salient steps of algorithm can be summarized as follows:

1.Initialization of weights.
2.Presentations of training examples.
3.Forward Computation.
4.Backward Computation.
5.Repeat until the output error is within or preselected threshold.

3. RESULTS AND DISCUSSION

Figure 1 shows clearly as roads of any shape can be searched in an automatic way on satellite image (obviously with a suitable spatial resolution). In particular, the two SPOT 5 images refer to an area of Kashmir hit by an earthquake in October 2006. The roads detection algorithm applied to the two images, acquired before and after the event, allows to monitor the status of the road network after the event.

The figure 3 shows the gray scale image used to drive the searching and extraction of the desired objects. Such image (aerial picture downloaded from Internet) presents different kind of airplanes and we chose the Concorde airplane as sample to extract by applying Geometrical Moments Invariant technique. In fact, in figure 4 we have the original sample (Concorde) and the corresponding modified images (scaled, mirrored and rotated) used in this case. Starting from the original image we can calculate the corresponding Moment Invariant for the considered object, which fully characterize the desired object. The power of this technique is the ability to provide the same values for the related moments regardless of scaling, translation and rotations. The figure 5 (right) shows the perfect coincidence between the Moment Invariant of the original and modified images, and this property can be exploited to identify and extract any object on the image.

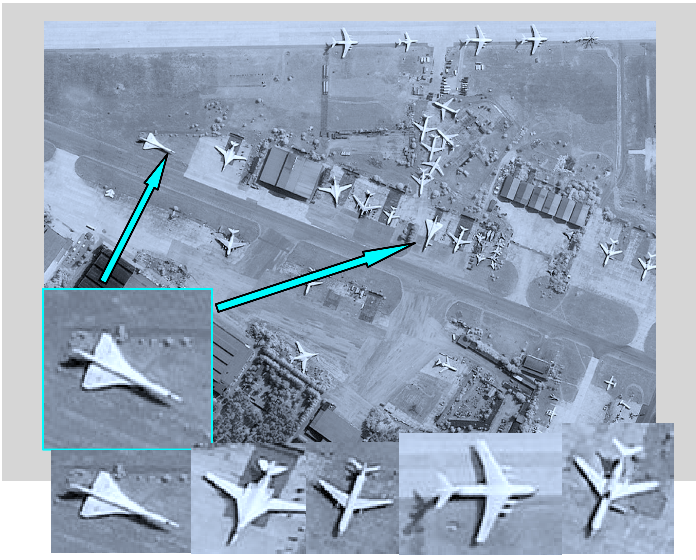

Figure 3. EROS grayscale image (1,8 m resolution) representing a graveyard of planes. At the bottom are shown some of these obsolete aircraft models.

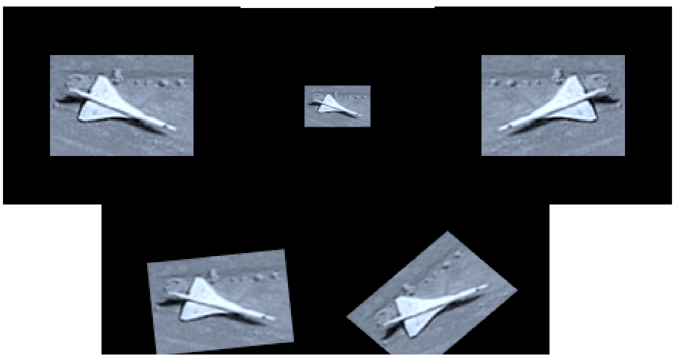

Figure 4. Concorde used as sample in order to extract it by Geometrical Moment Invariant. The shape signature of this object is always the same regardless of the applied rotations (below), scale factor (above) or translations.

The figure 6 shows an aerial panchromatic image (provided by DLR, 0,50 m of resolution) on the Munich-Salzburg highway. From that image we got a gray-scale image by using PCA technique. Afterwards we have trained the described ANN with our samples (cars with different sizes) to get the final configuration for neural network. After this operation we passed to ANN the binary image with the cars and others unwanted objects. At the end of the process the ANN was able to recognize different kind of cars, separating them by size as the figure 6 shows.

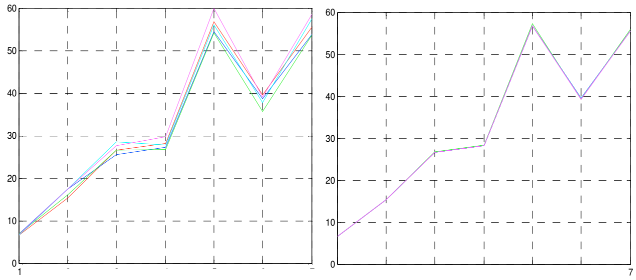

Figure 5. Moment Invariant calculated for every sample plane of figure 2 (left). Moment Invariant calculated for each Concorde considered in figure 4 (right). We can observe how all seven moments for the modified images coincide with those calculated for the original.

3. CONCLUSIONS

Through this work we presented a set of techniques and related algorithms, developed at CRPSM during its participation in several European projects, able to identify and extract objects in high and medium resolution satellite/aerial images. The examples described in the paragraphs above allow us to conclude that all the tried techniques (it is worthwhile to remember that many other techniques exist which have been neglected in this work) are able to achieve our goal, although there is no universal technique able to solve all the problems of this nature.

In particular, through the cases here-presented we saw that:

the Mathematical Morphology is very powerful and versatile but it is very sensitive to the image pre-processing and orientation of the objects sought,

the Invariant Moments technique is very powerful because they work well when the same object can be present with different scale, and orientation characteristics. However, the technique is very sensitive to the background around the wanted object (since this affects the calculation of the moments), and

the Neural Networks are very robust but the achievement of the "optimum net" in various situations can be very complex due to the training of itself by the samples.

All these considerations lead us to conclude that the ideal system for the automatic object extraction should include a set of techniques (not just those we described in this paper) which, depending on conditions and the complexity of the image, can be used to reach the goal.

4. REFERENCES

A. Khotanzad, J-H Lu, "Classification of invariant image representations using a neural network". IEEE Trans. Acoust., Speech Signal Process. 38: 1028–1038, 1990.

Y. K. Kim, J. H. Ham, "Fuzzy K-NN algorithm using modified K-selection". *Proc. Int. Joint Conf.* 4th IEEE International Conf. on Fuzzy Systems (FUZZ-IEEE/IFES'95) 3: 1673–1680, 1995.

G. Laneve, G. Santilli, "Automatic Detection of Man-Made Structures in Big Size Mosaics of Spaceborne Images," Image information Mining for Security and Intelligence ESA-EUSC Conference, Madrid – Spain, 2006.

G. Laneve, G. Santilli, I. Lingelfelder, "Development of Automatic Techniques for Refugee Camps Monitoring using Very High Spatial Resolution (VHSR)," IGARSS, Denver - U.S.A, 2006.

E. R. Dougherty, L. A. Lotufo, "Hands-on Morphological Image Processing," SPIE - International Society for Optical Engin., 2003

M.K. Hu, "Visual pattern recognition by moments invariants". IRE Trans. Information Theory, 8: 179- 87, 1962.

S. Paschalakis, P. Lee, "Pattern recognition in grey level images using moment based invariant features", Image Processing and Its Application. Seventh Intl. Conf. (Conf. Publ. No. 46: 245-249, 1999.

R. Gonzalez, R. Woods, "Digital Image Processing", 2nd ed., Prentice Hall, Upper Saddle River, NJ, 2002.

R. Gonzalez, R. Woods, and S. Eddins, "Digital Image Processing using Matlab", Pearson-Prentice Hall, Upper Saddle River, NJ, 2004.

S. Giada, T. De Groeve, D. Ehrlich, and P. Soille, "Information extraction from very high resolution satellite imagery over Lukole refugee camp, Tanzania," Int. J. Remote Sensing, vol. 24, N° 22, pp. 4251-4266, 2002.

P. Soille, M. Pesaresi, "Advances in mathematical morphology applied to geoscience and remote sensing," IEEE Transactions on Geosciences and Remote Sensing, vol. 40, pp. 2042-2055, 2006.

X.Y. Jiang, H. Bunke, "Fast segmentation of range images into planar regions by scan line grouping", Machine Vision and Applications, 7, 115-122, 1994.

S. Dutta, P. Nanda, " Moment based object recognition using artificial neural network". National Conference on Intelligent Systems, Hyderabad, 2007.

S. Haykin, "Neural Networks, Pearson Education Inc.,2004

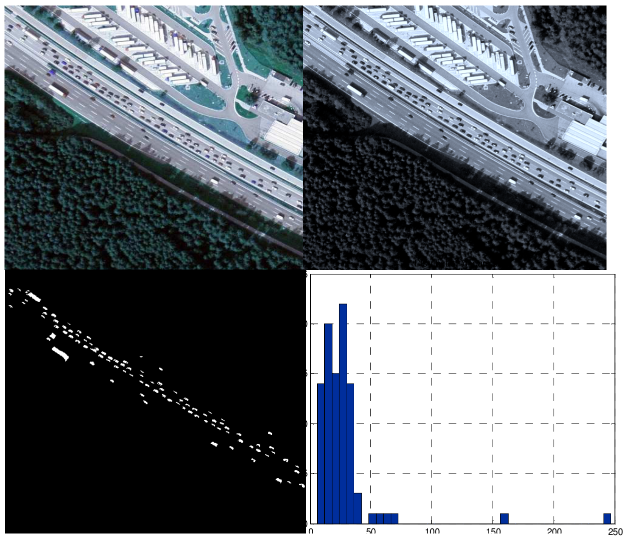

Figure 6. Aerial multispectral photo (top left) on Munich-Salzburg highway with 3K Camera (courtesy DLR). With a multiplayer feed-forward network cars on the highway can be extracted and classified by size (bottom right)