MVA'90 IAPR Workshop on Machine Vision Applications Nov. 28-30,1990, Tokyo

REGION TRACKING THROUGH NEURAL CLASSIFIER

Patrick Thévenoux[1 2], Pascal Vincent [1]

(1) aerospatiale DE/E/ETRI
Centre de Recherche de Magny-les-Hameaux
F-78 114 Magny-les-Hameaux, France

(2) ETCA/CREA/SP
16 bis avenue Prieur de la Côte d'Or
F-94 114 Arcueil, Fra,ce
Email: pth@etca.fr

ABSTRACT

We present herein a classification of regions from an image which is based on textural measurements. It aims to distinguish groups of regions having the same class among a pre-established set of categories, that may be potential focus areas for an aerial mobile. The presented classifier relies on a neural architecture.

INTRODUCTION

While flying, an enormous amount of supporting information is now available to control and correct navigation. However, little part is devoted currently to visual perception. Using a camera aboard allows to confront reality to flight planning, satellite information, digital terrain models...

One of the source of the lack of vision influence in this domain is the difficulty to limit the focus on one single object in the scene. Within this frame, three solutions are commonly available: perceptual organization[1] which searchs for structural entities which are, in outdoor scene, fragile elements and of little use for relative navigation, multi-sensor fusion which eases the extraction of volumes for exemple[2,3] and finally, textural analysis which provides an a priori classification of zones. We present herein an experiment of this latter technique dedicated to flight navigation thanks to a neural classifier.

PRINCIPLE

Basic information derives from the infrared image flow. Ground is observed obliquely from the mobile, which has important impacts on image resolution that becomes sight-related. Then images are then segmented into regions to compress the information according to the scene content. Then each region is associated to textural measurements which are the neural network inputs.

The use of the network is then two-phased. First and off-line, parameters are determined according to flight conditions: altitude, sensor particularities, weather and time constraints...

At this learning stage, an operator is required to give some examples of region classifications. Then the combination of this manual classification and of the textural measurement provides the initialization of the network weights. After learning, the classification is automatic.

SEGMENTATION

Within this frame, the image segmentation we use is oriented towards object recognition and identification[2]. Therefore, a dedicated textural segmentation has not been developed in order to maintain linear information which is often reduced by region-oriented methods. Briefly, segmentation relies on contours detected by a first order derivative operator after filtering against impulsion noise. Regions are the dual primitives of closed contours.

As ever, results are not ideal but quite correct. In some cases, regions are artificial while others are missing. This leads to group mistakenly some regions of the images which have different classifications in reality.

REGIONAL MEASUREMENTS

Once the region image is available, textural measurement are computed. These attributes to the regions are the classifier (i.e. the neural network) inputs.

Generally, texture is viewed as a combination of tightly interlaced elements, a description of both the size and the organization of small entities constituting a substance, a visual or tactile characteristic of a surface.

Homogeneity differences and internal structures are to be detected in order to define a texture. In theory, this leads to complex models[4] we cannot afford.

Therefore, we limit ourselves to low-order descriptors. The 14 parameters we chose with predictably more or less success according to[5] are the regional average, the standard deviation, the local extremas, the horizontal and vertical run lengths, the horizontal and vertical differences between pixels, the gradient along the four main directions and the unitary translations.

NEURAL NETWORK

The study on neural networks relies on the will to understand and imitate the human brain.

The descriptions of the basic elements of the brain, made of 100 billions of neurons, and of the architecture, the way the neurons are bound together, allows to create architectural models related to some interesting properties of the different parts in the brain concerned by perception.

Since the first important works (1943: McCulloch and Pitts[6], 1958 Rosenblatt's PERCEPTRON[7]) and despite of the limitations on the PERCEPTRON that Minski and Papert[8] pointed out in 1969, numerous applications rely on this principle, in particular in Image Processing.

The interest devoted to a neural classifier is due to the following properties:

- programmable by learning: the parameters of the neural model are identified through given examples to process. This is a fundamental point because it eases the setting up of the classifier when a correct learning base is prepared;
- delocation of the information which induces tolerance towards breakdowns. Information is distributed among the connections between neurons. The loss of some cells does not damage definitely the memory;
- possibilities of generalization and restoration. It is not necessary to learn every configuration to be able to classify a set of data;
- massively parallel processing of the information related to the intrisic architecture of the network.

Network description: our network is very classic and is defined as followed by:

⇨ the neuron (figure 1):

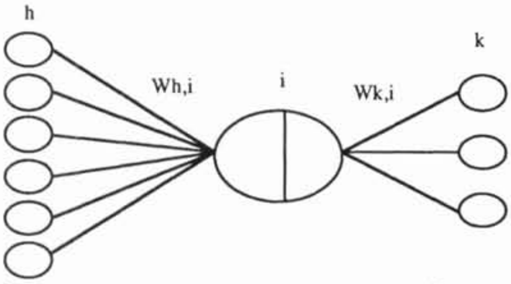

Figure 1: Basic cell

On the m input neurons:

$$I_i = \sum_{i=1}^{n} W_{i,h} * Y_h \qquad h \text{ in } [1..m]$$

$Y_i = f(I_i)$ f: neural fonction

Y_i : the output of cell i.

$W_{i,j}$: weigth between cells i and j.

The neural model is equivalent to a weighted sum of the inputs through the synapses. Then, this sum leads to determine the neural state or activation. Fonction f, which must be continuous because of the chosen learning phase, is evaluated empirically as:

$$f(x) = \frac{1}{1 + e^{-x}}$$

⇨ type: a layered network (figure 2) is chosen according to their generality and their similarity with the visual cortex.

The number of layers is three. Various attempts have been tried to get the best results at the learning and recogntion phase with respect to this number. It seems to be a good compromise[9] between a low number with little discriminancy power and a high number which focuses mainly on details;

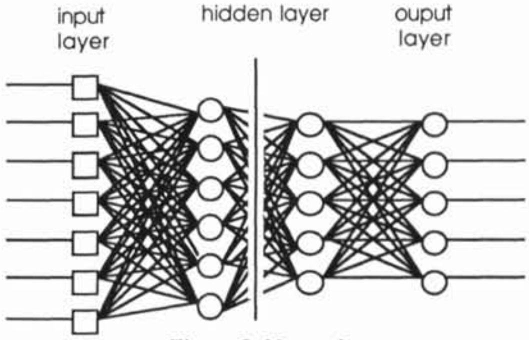

Figure 2: Network

⇨ the learning phase is supervised: at every step a couple of input and output data is given to the network;

⇨ the I/O of the network are imposed by the application. The entries are the textural parameters of the regions. The number of parameter defines the number of neurons of the input layer. In the study, this number has varied to estimate their respective discriminancy powers. The output number of neurons is the number of classes which are in this application buildings, roads, fields and forests;

⇨ The chosen learning law is the back-propagation of the error gradient[10]. At every given example, it consists in:

- computing the network output in propagating the input data from the input layer towards the output layer;
- ⇨ computing the error (quadratic error on each output cell);
- ⇨ propagating the error back towards the input layer in modifying the weights $W_{i,j}$.

Learning: the most delicate part is to prepare a learning base. The base contributes to the ulterior efficiency of the classifyer.

Learning is done on one image of the sequence for which a manual classification is relatively easy to perform.

Every parameter is normalized between 0 and 1 to get rid of the sensor inner characteristics in the learning image. Maximal and minimal values are used to shrink the corresponding parameter dynamic in the other images.

RESULTS

Discriminancy power of the textural parameters: The average parameter is mainly linked to the region intensity. Natural elements (fields and forests) are characterized by their low emission of heat. It is confirmed by the extrema parameters.

Standard deviation results are relatively paradoxal. Buildings and roads are the less uniform zones! In this version, building are related to a unique class but this result leads to divide it into two classes: walls and roofs. In fact, the nature of these elements is very different. Therefore, learning and recognition do not give the expected results. As far as the roads are concerned, the side effect due to the road limits (ditches, security rails...) on these linear elements is primordial and causes the perturbation. The unitary translations, the gradients and the pixel differences show important deviations. It is likely that these parameters have little influence in the classification.

Eventually, the run lengths show a good uniformity over each class. But the measures on the classes are relatively near. Among the set, roads seem to be the most separable class.

Groupings: The presented images correspond to an airport base overflight (figure 3). Every class is present in them.

25 regions of the middle image were used during the learning phase. Different grey levels are linked to the classes.

The initial classifications are maintained, which validates the learning phase.

One attempt was made to keep only 8 parameters (average, standard deviation, extrema, differences and run lengths) but the network learned with a lowest precision (1 erroneous classification).

Globally, results are satisfying. However some errors are reported on the last image of the sequence. Regions in this image have a larger surface than in the other images. Thereafter, the parameters are highly averaged and less representative of each class. It seems that oversegmentation is a criterion to reinforce the quality of the results. The learning phase can take benifit of it too, by assuring breaks between regions of different classes.

CONCLUSION

Simple textural measurements may bring positive effects within the frame of aerial autonomous navigation.

It allows to group in realistic conditions the zones of an image related to the navigation problem. By setting up two classical methods in their domains, a constraint of Image Processing concerning the duality between edges and texture is overpassed to allow a valuable cooperation.

Next developments concern the use of an explicit dependancy of the parameters relatively to the sight in the image, which is typical of aerial navigation.

ACKNOWLEDGEMENTS

The authors wish to thank Pr. G. Stamon from the *Ecole des Hautes Etudes en Informatiques* in Paris for his useful and valuable comments.

REFERENCES

1. Lowe D.G. (9/84). *Perceptual organisation and visual recognition*. Ph.D. Thesis, Stanford University, Stanford (Ca), USA, 142p.

2. Thévenoux P., Zavidovique B., Stamon G. (9/89). LOAS: Automatic site localisation and orientation using a 3D model. *5th ICIAP*, Positano, Italy, 501-508.

3. Blanquart J., Orgiazzi P., Grenier G. (4/90). Feature association from multi-sensor data fusion. *SPIE Aerospace Sensing*, **1306**, Orlando (Fl), USA.

4. Gagalowicz A. (5/1983). Vers un modèle de textures (in French). *Thèse de Doctorat d'Etat*, Université Pierre et Marie Curie (Paris VI), Paris, France, 351p.`

5. Haralick R. (5/79). Statistical and structural approaches to texture. *Proc. IEEE*, **67**, n°5, 786-804.

6. Mc Culloch W.S., Pitts W. (1943). A logical calculus of the ideas immanent in nervous activity. *Mathematic Journal of Biophysics*, **5**, 115-133.

7. Rosenblatt F. (1958). The perceptron: a probalistic model for information storage and organization in the brain. *Psychological review*, **65**, 386-408.

8. Minsky M., Papert S. (1969). *Perceptrons*. Editions MIT Press, Cambridge, USA.

9. Bunpei I., Miyake S. (7/87). Capabilities of three-layered perceptrons. *IEEE Int. Conf. on Neural Networks*, **1**, 641-648.

10. Rummelhart, Hinton, Willems (1986). Learning internal representation by error propagation. *Parallel Distributed Processing*, **1**, MIT press, Cambridge (Ma), USA.

Figure 3: Image sequence, manual and automatic classification