A. Prieto (Ed.)

Artificial Neural Networks

International Workshop IWANN '91
Granada, Spain, September 17–19, 1991
Proceedings

Springer-Verlag
Berlin Heidelberg New York
London Paris Tokyo
Hong Kong Barcelona
Budapest

Series Editors

Gerhard Goos
GMD Forschungsstelle
Universität Karlsruhe
Vincenz-Priessnitz-Straße 1
W-7500 Karlsruhe, FRG

Juris Hartmanis
Department of Computer Science
Cornell University
Upson Hall
Ithaca, NY 14853, USA

Volume Editor

Alberto Prieto
Departamento de Electrónica y Tecnología de Computadores
Facultad de Ciencias, Universidad de Granada, 18071 Granada, Spain

CR Subject Classification (1991): B.3.2, B.7.1, C.1.3, C.3, C.5, F.1.1-2, F.2.2, H.3.m, I.2.6-10, I.4-5, J.2, J.7

ISBN 3-540-54537-9 Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-54537-9 Springer-Verlag New York Berlin Heidelberg

Printed in Germany

Typesetting: Camera ready by author
Printing and binding: Druckhaus Beltz, Hemsbach/Bergstr.
2145/3140-543210 - Printed on acid-free paper

6. Applications

AN APPLICATION OF NEURAL NETWORKS TO NATURAL SCENE SEGMENTATION

Marcelino Vicens **
Jesús Albert *
Vicente Arnau *

**Institut de Robòtica
* Dept. Informàtica i Electrònica
Universidad de València
Av./Dr. Moliner, 50
46100 - Burjassot(València)

Abstract:

This paper introduces a method for low level image segmentation. Pixels of the image are classified corresponding to their chromatic features.

The classifier is implemented by means of a neural network trained using a manual clasification of the coloured images. The back-propagation algorithm is applied to one image which works as training set, and the results applied to the rest of them.

Because of the large amount of data contained in one image, a random selection of the points is made before training.

Finally, the paper shows the results of applying the method to fruits detection in natural scenes taken from the open field.

Introduction

Within the area of computer vision, one of the most important and complex problems is the one of segmentation. There is an extensive bibliography about different aspects of image segmentation [1][2] in which more or less general methods are exposed as well as numerous applications.

In this paper the task we are dealing with is the segmentation of colour natural scenes applied to the location of fruits for automatic harvesting. With this aim the variability introduced by natural lighting (uncontrolled) adds to the complexity of the underlying problem (chromatic characterization of fruits possibly in different ripe states) giving special connotations to the problem. Besides, due to the type of application, it will be necessary for the on-line segmentation procedure to be fast enough to work together with the corresponding robotic system.

According to the above mentioned the low level approximation has been chosen, so the bigger objects to be dealt with are the pixels of image and its associated features. This allows the application of simple and above all quick classification methods rather than the high level segmentation methods based on borders or regions. In following sections a method for low level segmentation of colour images based on an artificial neural network, previously trained with a image labelled by an observer in two classes will be introduced. The hardware implementation of the network will give us the required speed of clasification.

Considerations on Colour Representation

The experimental laws of colour-matching are summed up in the trichromatic generalization [3]. This states that over a wide range of conditions of observation many colours can be matched completely by additive mixtures in suitable amounts of three fixed primary colours. The choice of these primary colours, though very extensive, is not entirely arbitrary. Any set in which none of the primaries can be matched by a mixture of the other two may be used.

However, the triestimulus theory does not reproduce correctly the perceptive phenomena, since the eye has the ability to distinguish colours in terms of other different parameters, the chromaticness and the lightness.

In the experiments in colour discrimination, an observer assesses differences or equalities of his perception in certain perceptual attributes of the colour of the test samples. The principal psychological attributes for object colours are Hue, Saturation and Lightness. The scaling of the attributes has to be constructed from the judgement of the observers.

Consider a set of colour samples all having a particular fixed lightness; their representative points will generate a surface, and two pair of samples belonging to this surface that are judged by observers to present the same difference in chromaticness will not generally be represented by pairs of points separated by the same distance. The question is to find a transformation to another space such that these pairs of samples are represented by points that lie at equal distance. Most current arrangements (ISH) for grading chromaticness are organized in terms of loci of constant Hue and Saturation on a surface of constant Lightness; these loci form a polar coordinates system. The central point represents gray, points on a circle represent colours of constant Saturation, and points on a radial line starting at the center represent colours of the same Hue.

To concrete we will considerate now the system of coordinates employed in this paper which can be obtained from the colour signals generated by a TV camera [4]. TV cameras generate three signals R, G, B proportional to the red, green and blue measures in the formed image that are able to reproduce in a TV monitor an image apparently equal to the original one.

These tree signals, are not the best to identify the objects [5] in the scene, but are the ones supplied by a comercial TV camera, and our present objective is to try the identification of objects with only the above measurements. A possible improvement is to use a different set of colour coordinates, but the only experiences made are on the RGB set because we assume that the neural network is capable of doing a linear transform on the data, and most of the colour coordinates are linearly dependent of each other. The use of a diferent set of coordinates not linearly related to RGB, like the ISH, whill modify the metrics of the sample space to one directly related to human vision, and may improve the above results.

Purpose of the network.

The main purpose of the network is to discriminate between oranges and other objects from an image correspondig to an orange tree. The result is sent to a robot arm that performs harvesting of orange fruits. It is very important to classify the objects in a conservationist mode, in the sense that it is preferable leaving some fruits on the tree than sending the robotic arm towards a region with no fruits (sky or brillant objects). This action has a cost in number of fruits collected but bad classified oranges in one image can be detected in another take if a suitable strategy of image scaning is followed.

The neural network used.

In the present work, we use a three-layer perceptron with feedforward connections. The first layer of neurons accepts the three input RGB values digitized to 256 levels. There are two levels of hidden units with 16 neurons each one, and the output level with only one neuron that sends for each triplet of RGB pixels the indication of orange/no-orange to the external units.

The minimum nodes for a multilayer perceptron proven by Kolmogorov and described in [6] is N(2N+1) if N is the number of input variables. In our case 3(2*3+1) = 21 nodes. Experience has proved that the 21 neuron network is no trainable using the back-propagation algorithm (in a resonable time) with the data used, so we take a large network.

The training of the network is unpracticable with the data taken from an unique sample image because the number of points in the training set (128 x 128 = 16384) is too big in our case. The training set was made from a random sample of points from the training image. In this way, the magnitude of the problem is made tractable.

The training procedure.

The neural network was trained with the back-propagation algorithm, using an image previously classified by a human observer. The essays to train the network with a random sample of the points always failed because one of the phases classified well, and the other not classified at all. The conclusion we take is that the amount of pixels of one set is more important than the other 1/25 times or more, and the suposed alternancy between the samples fails at all.

To reduce the problem, we take a random sample of pixels over every one of the two sets, and the samples are presented to the training algorithm in an estrictly alternating sequence. In this way, the network trains adequately with the order of 10000 iterations.

The results of this training are statistically good, the order of only 10 % of the points are misclassified, but the plot of the segmentation was surprisingly bad, because 10 % of not orange points is greater than the number of orange points, due to the statistics of the two sets.

As final solution it was chosen to train partially the network with a reduced set of statisticaly taken points from each one of the sets and, when the training is done, introduce points randomly taken from the image, independently of the sets they belong to.

With the above training procedure, the number of points misclassified of each one of the two sets differ in percentage, the experimental results show that the greatest set is best classified in percentage, that is the result that we aim to obtain.

Experimental results

In the final images (Fig. 1, 2, 3, 4) we present four orange tree images classified with our algorithm. In all of them, the first quadrant is the original intensity image (obtained through mapping the RGB image to the intensity), the second cuadrant, is the manually classified image used for training, the third is the classified image whith our neural network, trained with a random sample of the same image, and the fourth one is a crossed clasification with images 1 and 2 in one case, and 2-3 in the other case.

In every figure, the classification is adecuated to the work, the percentages of misclassification in the set of no fruits are within the order of 0.1 to 0.75 for the closed test, and 0.24 to 1.25 for the open test.

The results for the set of oranges are apparently very bad in the percentage of points of orange class (5 to 28 % and more), but over the entire set of points in the image, the percentage must be divided by 25 over the entire image, due to the different number of points of the training set.

Current work.

Now we are doing experiments in non supervised training with a neural network like those reported by Kohonen [7] in order to avoid the manual clasification that includes numerous errors.

We are trying a training unsupervised algorithm based on the back-propagation algorithm that can work on a two layer network.

In other sense, there is the use of other colour coordinates not linearly related to RGB to represent the coloured image, and the use of local information of one pixel and his neighbours.

REFERENCES:

[1] Haralick, R. M. and Shapiro, L.G. "Image segmentation techniques". Computer Vision, Graphics and Image Processing, 29 (1985).
[2] Ballard, D.N. and Brown, C.M. "Computer Vision". Prentice Hall, 1982
[3] Wyszecky, G. and Stiles, V.S. "Color Science" John Wiley and Sons Inc, 1978
[4] Pratt, W.K. "Digital Image Processing" John Wiley and Sons, Inc. 1978
[5] Albert, J. Ferri, F., Domingo, J. and Vicens, M. "An approach to Natural Scene Segmentation by means of Genetic Algoritms with Fuzzy Data" .
[6] Lorentz, G.G. "The 13th. Problem of Hilbert" in F.E. Browder (Ed). Mathematical developments Arising from Hilbert Problems, American Mathematical Society, Providence, R.I. 1976.
[7] Kohonen, T. "Self-Organization and Associative Memory". Springer-Verlag, Berlin, 1984 and 88

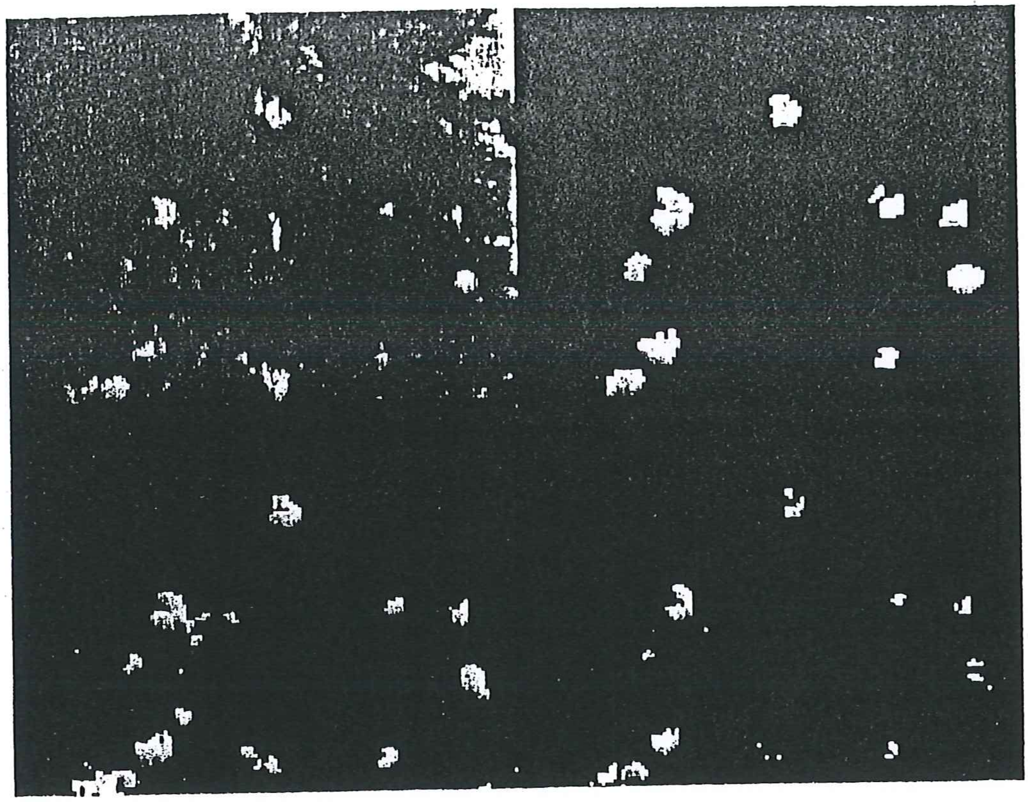

Fig. 1

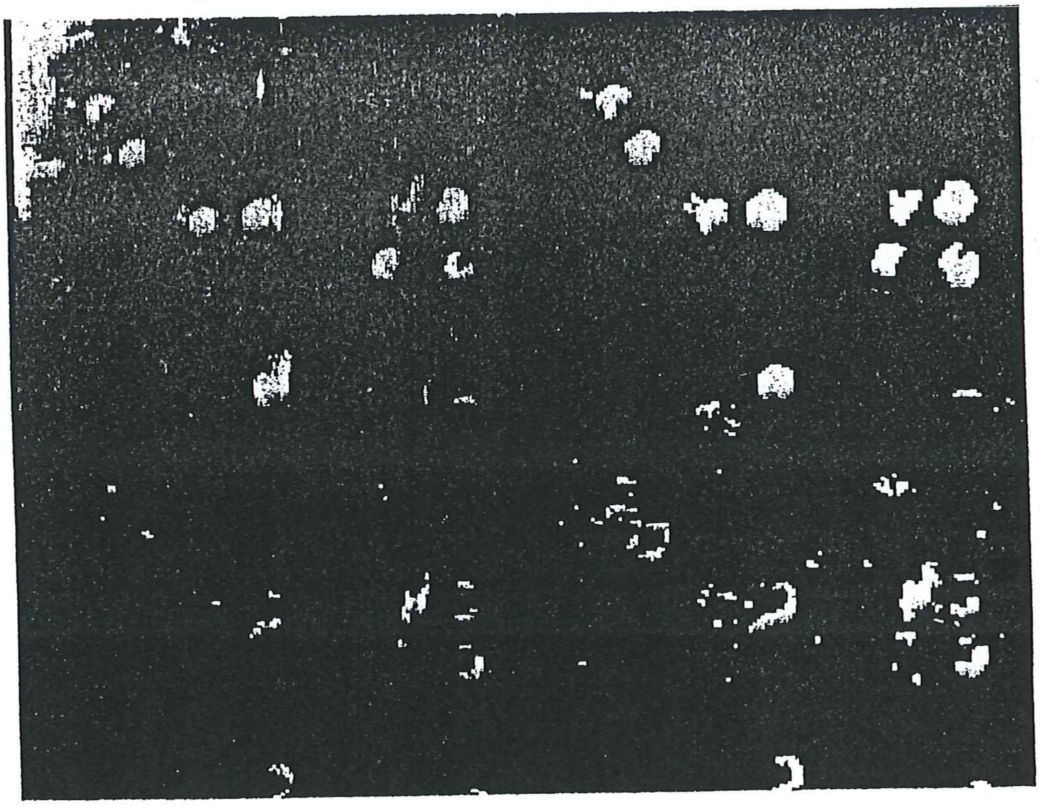

Fig. 2

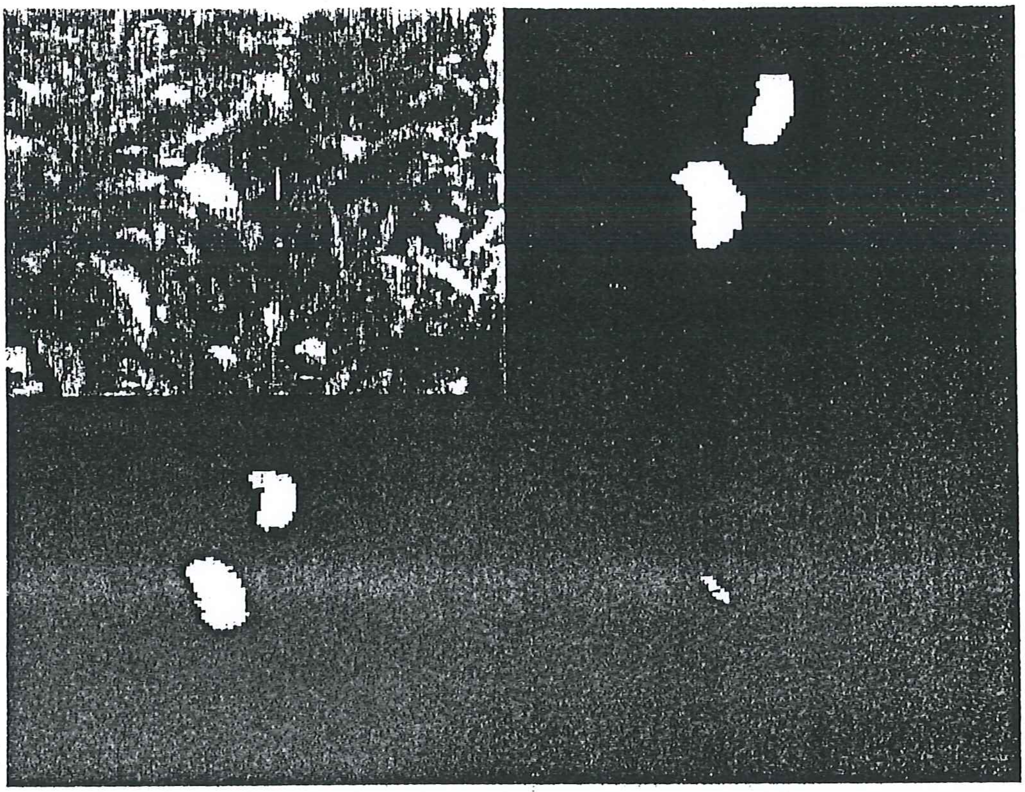

Fig. 3

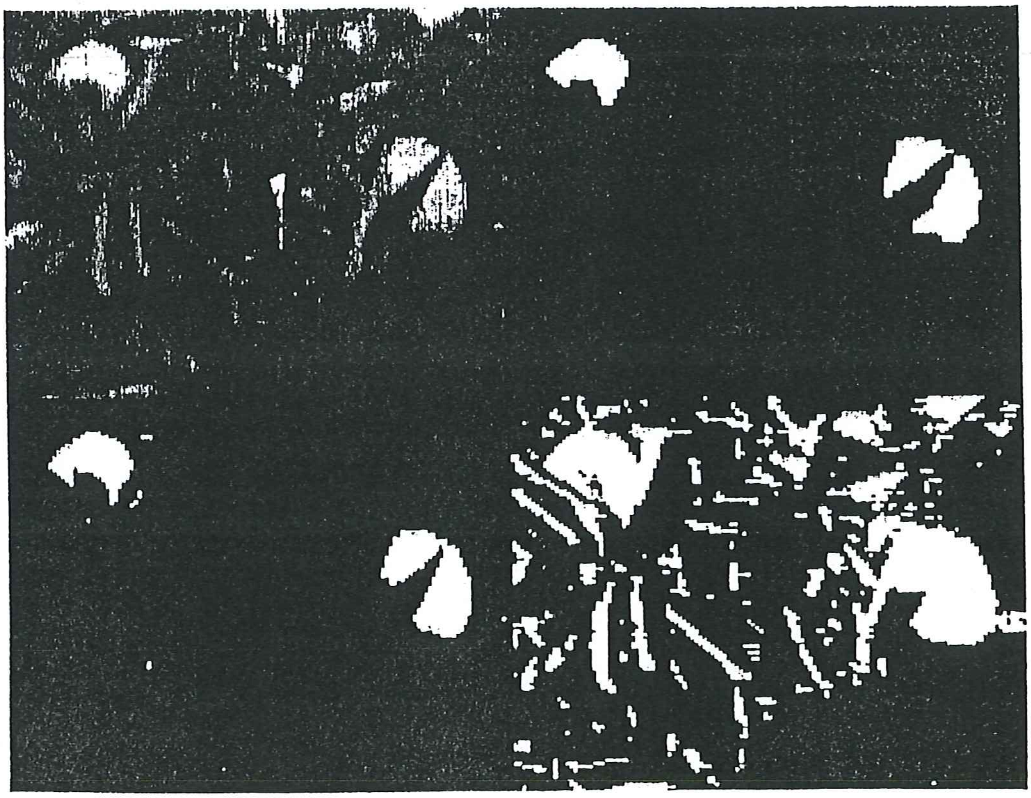

Fig. 4