The SOM Algorithm Applied to Stereo Vision

RAPHAELA GALHARDO FERNANDES[1], RICARDO WENDELL R. SILVEIRA[1], ADRIÃO DUARTE D. NETO[1]

[1]DCA – Departamento de Engenharia de Computação e Automação - UFRN, Natal, RN, Brasil
{raphaela,wendell,adriao}@dca.ufrn.br

Abstract. This paper describes a technique based on self-organizing neural networks, in order to perform the matching procedure on stereo vision problems. The technique is based on the SOM (Self Organized Maps) algorithm, also known as Kohonen network. It maps the captured image to the neural network and proceeds with the matching in that space.

1 Introduction

Systems based on stereo vision are able to estimate objects depths, using matching procedures. The matching consists on finding the corresponding pixels between a pair of stereo images.

There are many traditional techniques to solve the problem of matching [1]. Many of these techniques require a heavy computational effort. In this paper it will be presented a technique that aims making an intelligent matching procedure, using a neural network, particularly the Kohonen network, which reduces the problems that are found on traditional methods. The proposed method is described briefly as follows.

2 Proposed Method

Considering initially the images captured by two cameras, a ”hi-pass” filtering is made on both of them, so that image pixels are solely those that belong to object edges.

The following step consists of training the Kohonen network, in which neurons are arranged in a 2-dimensional grid. The training data are five descriptors of each pixel in one of the two images, in this case the right image was chosen for no particular reason. The used descriptors were: orientation and variation (given by Sobel’s operator), intensity, laplacian of gaussian and the y coordinate of the image, aiming to respect the epipolar line constraint.

Once the network is trained, input data are distributed throughout the grid of neurons. The next step is finding the winning neuron for each pixel of the right image. When the winning neuron is found, the pixel is associated to it.

The same procedure is then performed on pixels of the left image. In this step, after finding the winning neuron, it is also computed which pixel of the right image (among all that are associated to the winning neuron) is the most similar to the pixel of the left image. Equivalently, a ”winning data” is chosen. By following this procedure, the matching between pixels of the pair of images is accomplished.

This technique was developed to be used on stereo systems that use a parallel configuration of cameras. Therefore, when the ”winning data” is being computed, epipolar line and geometric constraints have to be respected, so that the search for the ”winning data” is made only among those data on the right image that respect those constraints. This procedure minimizes the required effort of the matching process when compared to the work of Wang et al. [2], which used a supervisioned neural network.

3 Results and Conclusions

In order to evaluate the proposed method, many tests with many images have been made. Figure 1 shows one of the tests. The arrows point to the position of the matched pixels. The matching was successful, but some adjustments are necessary to speed up the process of generating the SOM and further improve the procedure.

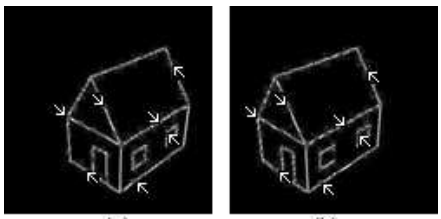

Figure 1: Illustration with some matched pixels. (a) left image. (b) right image

References

[1] D. Marr, *Vision. A Computational Investigation into the Human Representation and Processing of Visual Information*, MIT Press (1982).

[2] J.H. Wang and C.P. Hsiao, *On Disparity Matching using Vision via a Neural Network Framework*, Proc. Natl. Sci. Counc., Vol23, N 5, pp.665-678 (Taipe, Taiwan, 1999).