FEED FORWARD NETWORK FOR VEHICLE LICENSE CHARACTER RECOGNITION

F. Lisa
J. Carrabina
C. Pérez-Vicente
N. Avellana
E. Valderrama

Centre Nacional de Microelectrònica
Universitat Autònoma de Barcelona.
08193 Bellaterra. Barcelona.
SPAIN
Tel. +343 5802625
Fax +343 5801496
E.mail. neures@cnmvax.uab.es

Abstract

This paper describes the construction of a system that recognizes vehicle license numbers using feed forward neural networks, once they have been extracted using classical methods. The system has been trained and tested on real-world data. In order to reduce the total amount of required memory and increase the process speed, an additional step has been added to the learning algorithm, that produces low precision weights {+1,0,-1}. The network obtained after this training process has a similar behaviour to those networks using a floating point representation for weights. A special hardware accelerator has been developed to achieve high speed recognition.

1. Introduction

Image recognition is one of the fields of research where neural networks show a higher degree of competitiveness against the classical methods, due to the requirements of high speed processing of large amount of data. Parallelism, inherent in neural network processing, could be more efficient for these purposes.

Vehicle license recognition is usually viewed as a two-step process: segmentation of the image and recognition of each symbol. This scheme is still the same in both the classical and the neural approaches.

In the segmentation process, the starting image, usually represented as a gray scale, is processed until the isolation of several binary images representing individual symbols. The main problem is the difficulty in extracting the vehicle license number under conditions of changeable and uneven illumination, so that one part of the license plate is darker than the others, besides the problem of getting the image from the vehicle.

Several methods are used in order to obtain binary images. These methods mainly deal with the choice of the threshold between the two states. Otsu's method [1] obtains the threshold after inspecting a grey level histogram of the image, whereas later improvements uses a variable threshold to overcome uneven illumination problems. These methods require several iterations through the whole image, and this makes them slower without parallelism.

Neural network approaches to solve image segmentation use several kernels, which represent features of the image such as edges, corners or others obtained through the training process. The presence of these features can be computed in parallel for all of them in a portion of the image. Using specialized neural network chips, the AT&T system can process 32*32 pixels in parallel [2]. For an image of 512*512 pixels, they need 256 executions of the recall phase of the network. Very simple one-layer networks, with 20 input units, suffice to determine the presence of a character from the feature representation.

The symbols obtained from the segmented image are in a fixed window ranging from 5*8 to 20*20, so that compression or expansion should be necessary before the recognition phase. This is true for both the classical approach, mainly consisting in the evaluation of the Hamming distance to the different patterns, and the neural approach.

There are different neural approaches ranging from those that operate over the whole image, to the ones that use feature extraction. The methods that deal with the whole image, which are of small size, use classical feedback networks -i.e Hopfield networks or BAM networks- or feed forward networks -i.e. multilayer perceptron networks, trained with the backpropagation algorithm-. In these cases, images usually do not come from real images, since compression process highly degrades theyr quality.

For large size windows, these methods are not useful since the size of the network becomes too large for the most powerful learning algorithms. The most common way to proceed is to use a two-step approach. There is a first step for the extraction of basic features, and a second step for the recognition of the set of features.

The highest degree of parallelism can be achieved with the smallest resolution for weight and state values, which is specially useful for analog chips, but also for the digital ones since the throughput is increased. This requires a slightly different learning scheme in order to obtain this set of weights. In the case of the AT&T chip [3], weights are represented by 6 bits and states by 3, and it can process 1000 characters per second.

2. Network topology.

The system developed at our laboratories focusses on the recognition of characters of vehicle licenses. These characters came from a segmentation process that uses a non-neural method developed at the Centre de Tractament d'Imatges at the UAB [4]. This method is a knowledge-based algorithm that warranties the correct segmentation of the images independently of the position and orientation of the vehicle.

Our network recieves an image of 20x30 binary pixels for each character. We use a two-level strategy to reduce the complexity of the learning process. Our basic goal is to obtain low resolution weights for the whole network and, therefore, reduce the processing time and complexity.