Classification of Elements in an 3D Urban Virtual Environment Using Artificial Neural Nets

Juliana Gouveia Denipote
Escola de Engenharia de São Carlos EESC/USP
judeni@sel.eesc.usp.br
Cientistas Associados Ltda.
juliana.gouveia@cientistasassociados.com.br

Rodrigo Assaf
Departamento de Computação – Universidade Federal de São Carlos UFSCar
assaf@polvo.ufscar.br
Cientistas Associados Ltda.
rodrigo.assaf@ cientistasassociados.com.br

Abstract

The long time spent during the 3D modeling of the environment objects is the main problem of creating Virtual Environment cities (VE). Automatic generation tools are possible ways to increase performance when a VE is been built. This work describes an automatic generation tool using Computational Vision and Artificial Intelligence techniques to create a VE from a real city photographs.

1. Introduction

Applications for VE that deal with accurate and photo-realistic 3D city models are used in many fields, such as entertainment, tourism, 3D geographic information systems, among other ones [1], [2].

Although VE has been used in many areas of applications, its creation is a complex task because of the many objects that can be included in the 3D environment such as houses, squares, buildings and avenues. A criterion of the detail level must be defined and used during the creation of the EV in order to make it more realistic as possible depending on the needs of the application.

Another important issue is to define how city objects will be modeled: automatically or manually. Automatic way is an excellent technique to create large virtual city areas in a shorter time than the manual way.

2. Methods

As a solution for the problem of automatic creation of a virtual city model, it was used photos of an urban area, acquired in an easy and cheap way to guarantee the viability of VE. The images were taken from Google Earth [3], which is a free for personal use tool, that contains photographs images taken by satellites and aircraft sometime in the last three years.

Figure 1 – São Paulo city aerial photograph

In those images are applied Computational Vision and Artificial Intelligence methods to obtain the waited result as showed in figure 2. In figure 2, firstly, the Image Processing using Computational Vision techniques must be made. The image processing is divided in three activities:

Pre-processing: the image's noises are decreased and the quality is enhanced by Sharpening method, that uses the gradient magnitude to enhance borders, and Thresholding method to enlarge the contrast of object colors.

Segmentation: objects of interest as buildings, pavements and green parks are found using Sobel method to find horizontal and vertical borders. In this task, it should be attempted to factors that difficult the objects identification, for example, clouds shades, and objects that do not have rectangular borders.

Characteristics Extraction: for the found objects, the color average of their pixels is extracted, and the number of pixels of the height and the width are calculated.

The characteristics found in the image processing will be the entrance patterns for the neural net. The neural net architecture used in this work was Multi Layer Perceptron (MLP) architecture. This architecture was created to solve the problem of patterns that are not linearly separable [4] and its training is supervised and has back propagation.

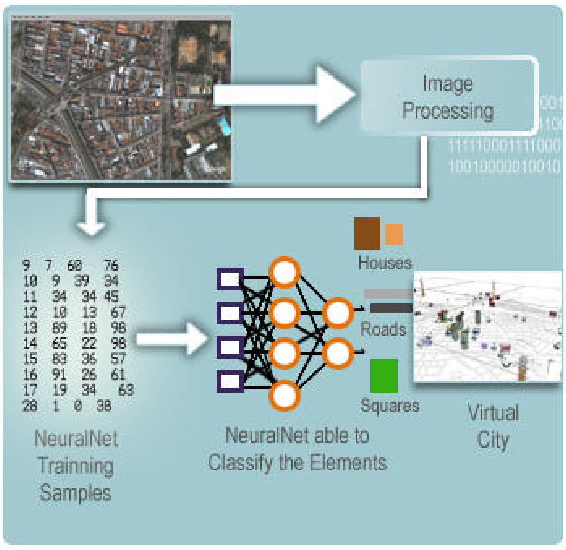

Figure 2 – Solution for creating a 3D Virtual City

There is a 3D models database to store models for the three classes of objects. There are ten different 3D models for houses, ten models for buildings and ten models for squares, because photos cover residential areas. As long as each object of the image was properly identified and classified, a 3D object is chosen in an aleatory way in the database. For example, if it is identified that the object is a house, one 3D model is chosen among the ten house models, and is inserted in the VE in the position corresponding in the image.

3. Results

Tests with 45 training patterns and 15 tests patterns, which should be classified as houses, buildings and squares, were made and normalized in the interval [0,1]. It was observed that the network reached an excellent classification with a 9% of estimative error for houses class, 12.28% for buildings class and 0,29% for square class. During the training stage, a momentum term was used to accelerate the convergence of the Mean Square Error (MSE) process. The training resulted in 722 epochs with the MSE graphic for every epoch represented on figure 3. Note that for each epoch, the error decreases, converging for a small and acceptable MSE value.

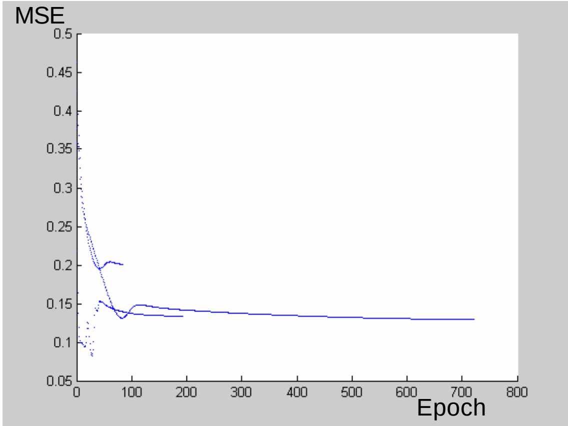

Figure 3 – MSE x Epoch

Analyzing the results, it is possible to observe that the samples quality influence the pattern classification. We found a problem with Sobel segmentation method for those objects that do not have rectangular borders, because Sobel uses the gradient direction to find similarity on the pixels that belong to the border. Other segmentation methods are being tested to improve the results of the aerial photographs´ segmentation. As an addition, the patterns normalization is a way to enhance the samples, and it was confirmed that normalizing the dates, the accuracy of classification reached the rate of 100% of the tested samples, while without the normalization, the rate was 80%.

4. Conclusion

For a best performance of the MLP training, the image processing must reach a satisfactory accuracy during the segmentation stage in order to result on high quality samples for the neural network. With a good set of samples, the MLP network is a good solution for objects classification for creating a 3D virtual city. The chosen neural net architecture and segmentation method were good solutions for the problem of the automatic VE creation with low error rates.

5. References

[1] A. Huguet, R. Carceroni and A. Albuquerque. Towards Automatic 3D Reconstruction of Urban Scenes from Low-Altitude Aerial Images. ICIAP (2003).

[2] Sistema Computacional para Redução de Perdas em Redes de Distribuição de Energia com Interface em Realidade Virtual, PIPE FAPESP Project, Executors ICMC/USP and Scientists Associates Ltda. http://www.icmc.usp.br/~energ_rv/, 2004.

[3] Google Earth Beta (2005). Available at: http://earth.google.com/earth.html. [Last visited on July, 2005]

[4] Haykin, Simon, “Redes Neurais – Princípios e Prática”, Ed. Bookman, 2ª Edição, Porto Alegre, 2001.