CHAPTER I

INTRODUCTION

1.1 Introduction

Recovering of three dimension depth from images is a basic problem in computer vision and has important applications in robotics, scene understanding and three dimension reconstruction. The region depth ordering is a part of three dimension recovering work. The overview of the region's image depth ordering process can be shown in Fig.1.1. The process includes the 3D to 2D transformation, image plane, 2D to 3D transformation, and multi layer image. In the first process, several special depth acquisition instruments (such as monoscopic and stereoscopic camera, depth-range camera, etc) have been designed to recover the three-dimension scene structure. Our constraint is that a source of image must be a single monocular image and must be taken by the Single-Lens Reflex(SLR) camera. Usually the depth information of monocular image was lost during the transformation process, thus our work concentrates on recovering the depth order of all regions in an image and this process is included in the two dimension to three dimension transformation process as shown in Fig.1.1. The meaning of region depth ordering refers to the order of all regions in an image that is ordered by logical depth from camera to the surface points of each region.

Because there are several kinds of depth acquisition instruments, thus the

Figure 1.1: The overall processes of work.

two dimension to three dimension transformation algorithms depend on the type of instrument used. For example, a stereoscopic dual-camera makes use of a co-planar configuration of two separated monoscopic cameras, each camera captures one eyes view, and depth information is computed using binocular disparity. Another example is a depth-range camera which is a conventional video camera enhanced with an add-on

laser element, that captures a normal two-dimensional RGB image and a corresponding depth map. The laser element emits a light toward the real world scene, which hits the objects in the scene and reflected back. This information is subsequently registered and used for the construction of a depth map.

All the techniques described above are used to directly generate 3D content, on the other hand, our region depth ordering from a single monocular image is based on capturing depths and relationships between in-focus and blurred regions using the stationary wavelet transform. We applied a supervised neural network to learn the order of in-focus and out-of-focus regions by using wavelet coefficients of each region as input to the multi-layer neural network classifier. This technique allows us to order any regions image into a set of two-dimension region images that are ordered from near to far.

1.2 Statement of Problems

In this dissertation, efficient learning methodologies is proposed in order to find the index sequence of all regions in an image. The problem to be solved in this dissertation can be classified as follows:

1. The reconstruction of the order of all regions from near to far in a single monocular image is a challenging problem. Even though the depth information is lost during transformation process, there should be some relationships among information in an image that can be used to recover the depth information. How should a suitable feature be used in order to recover the depth information from a single monocular image, regardless of the actual depth but considering the relative depth instead?
2. Is there any neural network that can be applied to automate region's image depth ordering process in a single monocular image?

1.3 Objective of the Research

The objective of this dissertation prospectus is to develop a technique for training a computer to order all regions in an image which satisfies the following aspects:

1. A systematic process for ordering the regions in an image by logical depth.
2. An approach to arrange the training sets that includes the relationships between the in-focus, out-of-focus region images and their order in an image.

1.4 Scope of Work

In this dissertation, the scopes of work are constrained as follows:

1. Consider Single monocular image.
2. Consider only gray scale images and imaging acquisition is unknown.
3. Consider only images that have depth-of-field information.
4. Limitation of depth's order is to three orders.

1.5 Organization of Dissertation

The organization of this dissertation comprises of the followings: Chapter 2 summarizes background knowledge of the three dimension to two dimension transformation process, reviews some related works on two dimension to three dimension transformation methods, methodologies of multi-layer feed forward, and backpropagation neural network. Chapter 3 introduces materials and methods that are imposed on the training set images and the architecture of multi-neural networks that are used in this dissertation. Chapter 4 describes the experimental parameters and results. Finally, the conclusion of this dissertation is given in Chapter 5.