VISUAL PERCEPTION THEORY GUIDED IMAGE PERCEPTUAL DEPTH ESTIMATION

BING LI, DE XU, SONGHE FENG AND AIMIN WU

Institute of Computer Science and Engineering
Beijing Jiaotong University
Beijing 100044, P. R. China
bjtulb@gmail.com; dxu@bjtu.edu.cn

ABSTRACT. *Depth estimation is a classical problem in computer vision. The depth of image is conventionally defined as the distance between the corresponding scene point of the image and the pinhole of the camera, which is not in harmony with the depth perception of human vision. In this paper we define a new perceptual depth of image which is perceived by human vision. The traditional computation models of image depth are all based on the physical imaging model, which ignores the human depth perception. This paper presents a novel computation model based on the visual perception theory. In this approach, using Pictorial Cues like sky and ground, we can get the relative perceptual depth from a single 2-D image. Experimental results show that our model is effective and corresponds to the human perception.*
Keywords: Visual perception, Depth estimation

1. **Introduction.** The depth information of a scene is very important in computer vision. It is a useful clue, for instance, for the purpose of object recognition and scene interpretation. The traditional definition of the depth of a visible surface is the distance from the surface to the camera. i.e., it is the distance between the surface to the thin convex lens [1]. This definition is not in harmony with the human depth perception, thus we redefine a perceptual depth conception based on the visual perception in this paper.

So far, various methods for depth estimation have been proposed. The method of stereopsis [2] is the most popular of them. In this method, the same scene is imaged by two different cameras at two different positions. And then we can get the depth from the binocular disparity. In addition to this method, researchers have used the camera focus for depth recovery, which is called depth-from-focus (DFF) method [1], [3], [4], [5], [14], citelu. This method is based on the defocus and blurring in the image to determine its depth. Besides these, A. Torralba and A. Oliva proposed another depth estimation algorithm called depth estimation from image structure [6], which is based on the whole scene structure that does not rely on specific objects. In these methods, some just use a sin-gle 2-D image to estimate the depth [1], [6], [7]. Depth estimation from a single image is simple and more usually used in many situations. In our method, we also only use a single image.

All these depth estimation methods are only derivate from physical image model, which neglect the depth perception of human beings. The depth of image is just the depth which