Effect of Various Edge Detection Techniques for Depth Estimation Based on Defocus Monocular Cue

R.M. Mulajkar, V.V. Gohokar

ABSTRACT
Depth estimation is an important step to generate 3D structure. The number of algorithm has explored to estimate depth based on Binocular cues (using two or more images). But very less work has done to estimate depth from monocular cue (using single image). This paper uses monocular cue as defocus to estimate depth. Also we compare different edge detection algorithms used for depth estimation using monocular defocus cue from single image. The edge detection is an important step in an image processing. The various edge detectors used are Roberts Kernel, Prewitt Kernel, Sobel Kernel, Canny edge detector. The experimental results show that canny edge detector is superior as compared to other and shows better depth for given image.

Keywords:
Image processing, Depth, Defocus, Monocular.

1. INTRODUCTION
In the field of image and video processing 2D-TO-3D conversion has recently shows tremendous evolution. There is a large demand to have efficient algorithm for 2D-TO-3D conversion. A method of 2D-TO-3D video conversion can be classified into two groups, the Automatic 2D-TO-3D video conversion and semi-automatic 2D-TO-3D video conversion. A typical 2D-TO-3D conversion process consist of two steps: depth estimation for a given 2D image and depth image based rendering [DIBR]. The DIBR is well understood and produce best quality of images but estimation of depth from image is a difficult task. Depth estimation plays an important role in 2D-TO-3D conversion. Depth of visible surface is the distance between surface and camera. The various algorithms are available to estimate depth. They can be classified into two groups. The first groups called Monocular uses single image. The second group called Binocular uses two images. The various monocular cue used are as defocus, motion, relative size, familiar size, absolute size, aerial perspectives, occlusion, texture, defocus cue, height etc. The estimation of depth from monocular cue is difficult task. The binocular cues used are binocular disparity, motion, focus, defocus, and silhouette. Monocular cue provides depth information when viewing a scene with one eye. Binocular cues provide depth information when looking a scene with both eyes. The classification of depth cues can be summarized in Fig.1

The Paper is organized as follows: In section II, we review various algorithms used for depth estimation. In section III, Defocus monocular based depth estimation in detail and effect of various edge

![Fig.2 various monocular cues perception in human](image-url)
detectors. We conclude the paper in section IV.

Fig. 1 Depth Cues Classification

The human eye can perceive monocular cue as shown Fig.2

2. RELATED WORK

Recently, the various approaches are used to estimate depth from monocular cue. They can be summarized in table as [6]. Table I shows comparisons study of different monocular based depth estimation algorithms. The abbreviation used in Table I is as: Occlusion (O), LP (Linear Perspectives), RFS (Relative familiar size), RH (Relative Height), TG (Texture gradients), AP (Aerial Perspective), and S (Shadow).

Elder and Zucker [3] used first and second derivatives of input image to find pint location and defocus in an image. Bae et al. [4] extended this work to obtain defocus map from single image. Tai et al. [5] finds defocus at each pixel and then uses MRF to obtain defocus map.

Table I: Comparisons Study Of Monocular Based Depth Estimation Algorithms [6]

<table>
<thead>
<tr>
<th>Paper</th>
<th>Monocular Cues</th>
<th>Type</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dimicolli et al.[10]</td>
<td>Yes</td>
<td>I/O</td>
<td>Object order</td>
</tr>
<tr>
<td>Criminisi et al.[11]</td>
<td>Yes</td>
<td>I/O</td>
<td>3D model</td>
</tr>
<tr>
<td>Battagio et al.[12]</td>
<td>Yes</td>
<td>I/O</td>
<td>Depth map</td>
</tr>
<tr>
<td>Wang et al.[13]</td>
<td>Yes</td>
<td>I/O</td>
<td>Measurement information</td>
</tr>
<tr>
<td>Huang et al.[14]</td>
<td>Yes</td>
<td>I</td>
<td>3D model</td>
</tr>
<tr>
<td>Chen et al. [15]</td>
<td>Yes</td>
<td>I/O</td>
<td>Object depth</td>
</tr>
<tr>
<td>Torralba et al.[16]</td>
<td>Yes</td>
<td>I/O</td>
<td>Scene mean depth</td>
</tr>
<tr>
<td>He et al.[17]</td>
<td>Yes</td>
<td>O</td>
<td>Depth map</td>
</tr>
<tr>
<td>Hoiem et al.[18]</td>
<td>Yes Yes Yes</td>
<td>O</td>
<td>3D model</td>
</tr>
<tr>
<td>Saxena et al.[19]</td>
<td>Yes Yes</td>
<td>O</td>
<td>3D model+ depth map</td>
</tr>
<tr>
<td>Nedovic et al.[20]</td>
<td>Yes Yes Yes</td>
<td>I/O</td>
<td>3d model</td>
</tr>
<tr>
<td>Liu et al.[21]</td>
<td>Yes Yes Yes</td>
<td>O</td>
<td>Depth map</td>
</tr>
</tbody>
</table>

3. MONOCULAR CUE BASED DEPTH ESTIMATION.

In order to generate depth map, this paper uses cues as defocus. In this paper we focus on depth estimation from single image using defocus monocular cue.

Defocus Monocular Cue:

The defocus in an image acts as important monocular cue for construction of 3D view from 2D images. The different causes for defocus cues are as-1. Improper opening and closing of shutter (2. Atmospheric effect
3. Misfocus of lens
4. Relative motion between camera & object which cause motion blur.

To find the depth from defocus cue edge detection plays important role. Edge detection is the process of finding meaningful transition in an image. In this, points are find out where sharp changes in brightness occur. These points can be detecting blur computing intensities differences in local image regions. Edge detection is an approach used most frequently for segmenting image based on changes in brightness intensity. Detecting change in intensity.
for purpose of finding edge can be achieved by using first and second order derivatives.

**Image Gradient:**

The tool of choice for finding edge strength and direction at location \((x, y)\) of image \(f\) is known as gradient \(\nabla f\) and given as-

\[
\nabla f = \text{grad}(f) = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}
\]

The strength (magnitude) is given as-

\[
M(x, y) = \text{magnitude}(\nabla f) = \sqrt{g_x^2 + g_y^2}
\]  
(1)

The direction of gradients is given as-

\[
\alpha(x, y) = \tan^{-1}\frac{g_y}{g_x}
\]  
(2)

where \(g_x = \frac{\partial f}{\partial x}\) and \(g_y = \frac{\partial f}{\partial y}\)

The gradient of image is obtained by calculating partial derivatives \(\frac{\partial f}{\partial x}\) and \(\frac{\partial f}{\partial y}\) at every pixel location in an image

\[
G_x = \frac{\partial f(x,y)}{\partial x} = f(x+1,y) - f(x,y)
\]

\[
G_y = \frac{\partial f(x,y)}{\partial y} = f(x,y+1) - f(x,y)
\]

The different algorithms used for edge detections as follows:-

1. Roberts Kernel
2. Prewitt Kernel
3. Sobel Kernel
4. Frein-chien edge detection algorithm
5. Marr-Hildreth edge detector

The performance of various edge detectors can be characterized as in Table II.

<table>
<thead>
<tr>
<th>Sr. No</th>
<th>Algorithm</th>
<th>Time</th>
<th>Noise sensitivity</th>
<th>False edges</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Sobel</td>
<td>Lower</td>
<td>Less</td>
<td>More</td>
</tr>
<tr>
<td>2</td>
<td>Canny</td>
<td>High</td>
<td>Least</td>
<td>Least</td>
</tr>
</tbody>
</table>

The performance of canny edge detector is superior to other. The canny edge detector defines edges as zero-crossing of second derivatives in the direction of greatest first derivatives. The canny detector works as multi-stage process: a) Image is smoothed by Gaussian convolution b) Then 2D first derivative operator is applied to smoothed images to highlight region of the image with high spatial derivatives. Edges give rise to mark in the gradient magnitude image

Let \(f(x, y)\) be input image, \(G(x, y)\) be Gaussian function

\[
G(x, y) = e^{-\frac{(x^2+y^2)}{2\sigma^2}}
\]

The smoothed image is obtained as follows by convolving \(G\) & \(f\)

\[
F_s(x, y) = G(x, y) * f(x, y)
\]

It will determines gradient magnitudes and directions

\[
M(x, y) = \sqrt{g_x^2 + g_y^2}
\]

\[
\alpha(x, y) = \tan^{-1}\frac{g_y}{g_x}
\]

Detecting change in intensity for purpose of finding edge can be achieved by using first and second order derivatives.

**Calculation of Gaussian Gradient Magnitude of Origination & Reblur Images:**

The Gaussian gradient magnitude of original & reblur image is calculated along \(x\) & \(y\) direction for 2D coordinate (2D image) as-

\[
\nabla x^2 = \frac{-x}{2\sigma^2} \times e^{-\frac{(x^2+y^2)}{2\sigma^2}}
\]  
(3)

\[
\nabla y^2 = \frac{-y}{2\sigma^2} \times e^{-\frac{(x^2+y^2)}{2\sigma^2}}
\]  
(4)

In magnitude,

\[
|\nabla f(x, y)| = \sqrt{\nabla x^2 + \nabla y^2}
\]  
(5)

The blur scales are estimated at each edge location gives sparse depth map. However quantization error at weak edges, noise may cause inaccurate blur
estimate at some edge location. To reduce this noise effect sparse bilateral filter is applied. Thus error can be minimized.

4. **EXPERIMENTAL RESULTS**

We estimated depth map of input image using defocus cue. We compare various edge detector output and analyze effect of various edge detector on depth map. We studied four edge detector as Robert, Prewitt, Sobel and canny edge detector.

The method can be summarized as follows:

1. Take the input image
2. Find the edge map using different detector algorithm such as Roberts Kernel, Prewitt Kernel, Sobel Kernel, Canny edge detector.
3. Calculate Gaussian gradient magnitude of the original and reblur images
4. Calculate the defocus map from the gradient magnitude ratio.
5. Apply the Bilateral filter to obtain smooth map

In this bilateral filter is used to have smooth map. The numbers of filters are available such as median filter low pass filter, Gaussian filter. The bilateral filter is more accurate in proportional to the distance of relative pixels.

The mathematical formula for a bilateral filter is given as-

\[
Depth = \frac{1}{n(x)} \sum_{i} X_i e^{-0.5 \left( \frac{|f_i - x_j|}{\sigma^2} \right) \left( \frac{|u(f_i) - u(x_j)|}{\sigma^2} \right)^2} \quad (6)
\]

![Fig. 3. Generation of Depth Map using Defocus Cue](image1)

![Fig. 4. Effect of different edge detection algorithms on generation of depth map from defocus cue](image2)
where $u(x_i)$ denotes the intensity value of the pixel $x_i$, $n(x_i)$ represents the normalization factor of the filter coefficients.

**Sparse Depth:**

Fig.3 shows depth map estimation from single image based on Defocus cue as monocular cue.

The visual comparison of above figure gives conclusion that canny detector exhibits better performance but it requires more computational time as it smooth image with Gaussian function to reduce noise and then determines gradient to find edges in an image.

From the experimental result elapsed time to estimate depth from defocus from image based on various edge detectors are as follows:

Table III: Computation Analysis of Edge Detectors

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Edge detector used to estimate depth</th>
<th>Total time elapsed for depth (in sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>Sobel</td>
<td>68.17380</td>
</tr>
<tr>
<td>02</td>
<td>Robert</td>
<td>71.36821</td>
</tr>
<tr>
<td>03</td>
<td>Prewitt</td>
<td>69.33557</td>
</tr>
<tr>
<td>04</td>
<td>Canny</td>
<td>71.58488</td>
</tr>
</tbody>
</table>

5. **Conclusion**

In this paper, we summarized edge detector algorithm Roberts Kernel, Prewitt Kernel, Sobel Kernel, Canny edge detector used in digital image processing. In order to generate depth map, this paper uses depth cue for depth estimation as defocus. Also analyze the effect of these detector to estimate the depth from given image. From the experimental results, canny detector is the best detector to estimate depth. The canny detector exhibits better performance but it requires more computational time as it smooth image with Gaussian function to reduce noise and then determines gradient to find edges in an image.

**REFERENCES**


Mr. R. M. Mulajkar received his BE & ME degree in Electronics Engineering in 2006 & 2011 resp. He is currently Ph.D. student in Electronics Engineering at Sant Gadge Baba Amravati University, Amravati, India. His research interests include Image and Video Processing.

Dr.(Mrs.) V. V. Gohokar is working as Associate Professor in Department of E&TC, M.I.T., Pune, India. She has teaching experience of more than 24 years in Engineering colleges. Her areas of interest include Image and Video Processing, Microwave Communication and Internet of Things.