Journal of Applied Information Science
6 (2), December 2018, 06-13
http://www.publishingindia.com/jais

2D to 3D Conversion of Images Using Defocus Method Along with Laplacian Matting for Improved Medical Diagnosis

Manonmani S.[1*], Shanta Rangaswamy[2], Dhanush Kumar K.[3], Ishan Srivastava[4], Jonnalagadda Akhilesh[5] and Joshua Issac[6]
[1]Rashtreeya Vidyalaya College of Engineering, Bengaluru, Karnataka, India.
Email: manonmanis@rvce.edu.in
[2]Rashtreeya Vidyalaya College of Engineering, Bengaluru, Karnataka, India.
Email: shantharangaswamy@rvce.edu.in
[3]Rashtreeya Vidyalaya College of Engineering, Bengaluru, Karnataka, India.
Email: dhanushkr16@gmail.com
[4]Rashtreeya Vidyalaya College of Engineering, Bengaluru, Karnataka, India.
Email: ishansrivastava104@gmail.com
[5]Rashtreeya Vidyalaya College of Engineering, Bengaluru, Karnataka, India.
Email: jonnalagaddaakhilesh@gmail.com
[6]Rashtreeya Vidyalaya College of Engineering, Bengaluru, Karnataka, India.
Email: joshuaissacis@gmail.com
*Corresponding Author

Abstract: In today's medicine field, there is an increasing need to be more efficient and to be able to develop new techniques to diagnose and cure different diseases and ailments. From the wide range of medical images, there has been an initiative to concentrate or restrict the implementation to X-rays, since X-ray images are the only valid analogies that can be compared to vision from a camera with a perspective. Due to the older diagnosis methods implemented, there is an excess of 2D data when compared to 3D data. Therefore, conversion of 2D data to 3D plays an important factor in arriving at the required result with higher efficiency since it is more cost effective to convert 2D images to 3D rather than create a 3D image from scratch. In this work, the first concept used is of defocus method to perform the depth estimation of the given X-ray image, with the help of edge detection and the second concept used is, gradient magnitude calculation along with image matting using Laplacian process to produce a 3D structure of the 2D image.

ACM CCS (2012) Classification:

Computing methodologies → Image and video acquisition → 3D imaging.

Keywords: ARE and RMSE, Defocus method, Image quality metrics, Laplacian matting.

I. Introduction

Medical imaging has come a long way since its initial inception of applications and today it has become a major tool or facility that can be used by the doctors to diagnose their patients about the kind and intensity of any problem in their body. In fact, the range of medical imaging has spread so vast that there are specific treatment processes and techniques to diagnose and evaluate different types of regions of interest in their body depending on the disease or injury or state of the medical condition in which the patient is. Some of the common types of medical imaging include X-ray, MRI (Magnetic Resonance Imaging), CT scan (Computer Tomography scan), Ultrasound and nuclear medicine imaging including PET (Positron-Emission Tomography). The method of using X-rays is prominently the cheapest available and often serves as a best answer to diagnose the structure of the body, mainly on the bones in case of an external injury that can possibly cause damage to the integrity of the bone.

Although X-ray imaging is the cheapest solution compared others it isn't very cheap looking at it from an absolute value perspective and for some specific cases there might be a need of more than one X-ray to diagnose a medical condition. It also poses a increasing chance (very less) of occurrence of cancer since it involves the emission of ionization particles that might trigger certain cells in the body to develop a tumor which might ultimately lead to cancer lately.

So, an attempt have been made to look at different possible techniques that can be applied on a single X-ray image to extract more information from it thus possibly reducing the necessity of getting exposed to the radiation multiple times.

Related Work

The paper [1], the technique mentioned for measuring the depth of a query image involves the usage of a pre-existing image (color and depth) database. The input image is taken and the database is scanned for photometric similarity. This similarity is detected by comparing their gist (structure of the image) descriptors.

A depth map [2] is estimated by fusing the two maps. A clustering-based hierarchical search using compact SURF (Speeded-Up Robust Features) descriptors to characterize images is proposed.

In [3], the method proposed involves searching for similarities between a query image and the images from the depth of an image in the database based on the nearest neighbor search. In this approach, only k images that are most similar to the query images in terms of image functions like color and edges.

In [4], the proposed method utilizes the learned distance of an image rather than the Euclidian distance. It says that the distance metric is learned based on the fact that similar and dissimilar 3D image structures are greatly different. There is also a method proposed that uses the Gaussian weighted function for the same output.

In [5], the technique mentioned for measuring the depth of a query image involves the usage of an image (color and depth) database. The input image is taken and the database is scanned for structural similarity. The similarity is detected by using Local Binary Pattern which are used to characterize the structure color image.

In [6], focus has been on the creation of a depth map from a single RGB image which is a challenging task. A regression model is proposed with the help of Convolutional Neural Networks. This model is optimized using the Tukey's Bi-weight Loss Function (To find difference between estimation and ground truth) which is an M-estimator. In [7], a model is proposed with the help of Convolutional Residual Networks. This model is optimized using the Huber Loss Function (To find difference between estimation and ground truth).

In [8], focuses on depth map and the depth is found through discrete-continuous variables. The continuous variables find the depth of the pixels and the discrete variables represent the relationship between the neighboring pixels.

The main focus is on Image Matting which focuses on foreground and background extraction. Given an Image to produce a matte (a transparency value α at each pixel) such that each pixel can be deconstruct the color value I(i) into a sum of two samples, one from F(i) i.e. foreground color and one from B(i) i.e. from background color. In [10], uses the concept of stereovision as it gives a high spatial resolution. It mainly uses image pairs taken from two different cameras with different perspective to obtain depth values. Factors which affect this approach are computational efficiency, depth discontinuities.

In [11], the depth information is extracted from the single image in order to achieve this scene classification and object detection is applied. In this image is divided into different image sets based on the different scene structure.

II. Theories and Concepts

In 2D to 3D conversion of X-ray images the first stage is conversion of image to Grey scale which performs the conversion of multiband image to binary image which is followed by the Noise Removal and Edge detection. It gives the outline of X-ray image. Feature extraction stage includes calculation of Gaussian gradient magnitude, defocus map and Laplacian matting.

Conversion of the Image to Grayscale

This module deals with the conversion of a multiband image into a single band image-:

Multiband Image - A Multiband image is an RGB color image in which each component consists of 8 bits, hence, the entire multiband image occupies 24 bits.

Single Band Image - A Single band Image is a binary image in which the components occupy a total of 8 bits.

The multiband image is converted into a single band image so as to reduce the overhead of time in the overall performance.

The Conversion Formula:

$$G = \frac{R+G+B}{3} \tag{1}$$

Noise Removal with the Use of Median Filter

The original image is smoothened initially by using Median filter [9] in order to increase the performance of processing.

$$B = \mathrm{MEDFILT2}(A, [M\ N]) \tag{2}$$

Apply Median filtering on two dimensional matrix A [14] [19] and output pixel around the corresponding pixel in the input image has the median value in M by N neighborhood in the input image. Apply MEDFILT2 on the image is used to pad the image with zero on the edge then the median values for the points within [M N]/2 of the edges may appear distorted.

Canny Edge Detection

Canny's intentions involved enhancing the many existing edge detectors comparatively. The ideas and methods are proven to be very successful. A list of criteria has been taken into consideration to improve the contemporary methods of edge detection. The low error rate is taken into consideration as the initial and most important criteria. It is of great significance that the various edges that occur in images should not be missed and non-edges should not be responded too. The second criteria that should be taken into consideration is the edge points be well localized. "The edge point be well localized" indicates that the distance found by the detector between the actual and the edge pixel is to be minimum. Only one response is given to a single edge is taken into consideration for third criteria. The first two criteria is more enough to eliminate the possibility of getting multiple responses to an edge. This was implemented already. In canny edge detector, these criteria are used for smoothening the image to eliminate noise. Next step is to find image gradient in order to height light region with high spatial derivatives. Then, suppress any pixels that are not at maximum (i.e. non-maximum suppression) by tracking along those regions with the help of algorithm. The canny operator uses a method called Hysteresis that is used to track remaining pixels that have not been suppressed. The gradient array is reduced further more with the help of Hysteresis.

Gaussian Filter

Several steps to be followed to implement the Canny edge detector algorithm. Before trying to locate and to detect edges present in the image, first apply filtering algorithm to filter out noise from the original image. Gaussian filter is used exclusively in the Canny algorithm to remove noise from the image. The Gaussian filter can be calculated with the use of a simple mask before the Standard convolution method can be used to perform the Gaussian smoothing [21]. A convolution mask is generally much smaller than the image itself. As a result, the mask is slid over the image, manipulating a square of pixels at a time. The possibility of detecting noise decreases gradually by increasing the width of the Gaussian mask. As a result, the localization error will also be increased slightly by increasing Gaussian width.

Sobel Operator Convolution Masks

After eliminating and smoothing the image, next is to find the edge strength by considering the gradient of the image. The Sobel operator is used to find a 2D spatial gradient measurement on an image [20]. Then the edge strength (approximate absolute magnitude) at each point can be calculated. Then, the approximate absolute gradient magnitude (edge strength) at each point can be found. A pair of 3x3 convolution mask used by Sobel operator to estimate the gradient in x-direction (columns) as well as in y-direction (row) as in Fig. 1 as shown below:

-1	0	+1
-2	0	+2
-1	0	+1

+1	+2	+1
0	0	0
-1	-2	-1

Fig. 1: Sobel Operator Convolution Masks

The formula for approximation of the magnitude or edge strength of the gradient is given below:

$$G = |GX| + |GY| \tag{3}$$

Finding Edge Direction

After calculating the gradient in both the direction (x and y), It is very easy to find the edge direction. The code should enforce the restriction whenever the error will be generated (error will be generated when sumX=0). However, and error will be generated when sumX is equal to zero. Whenever the sumX=0 (gradient in x direction), set the edge direction as 0 degrees or 90 degrees, depending on the value of the gradient in y direction. If the gradient in y direction (sumY=0) is equal to zero then the edge direction will be 0 degree else will be 90 degree [16].

The formula to find the edge direction is as follows:

$$\text{theta} = \text{invtan}\ (Gy\ /\ Gx) \tag{4}$$

After finding the edge direction, the next step is relating the direction that should be traced in an image to edge direction. The alignment of pixels of 3x3 image is as follows:

x x x
x a x
x x x

By considering the pixel "a" in the above alignment, it is clearly stating that only four possible direction are there while considering the surrounding pixels. The four possible directions are given below:

1. Horizontal direction-0 degrees,
2. Along the positive diagonal-45 degrees,
3. Vertical direction-90 degree and
4. Along the negative diagonal-135 degrees.

Now edge orientation should be mapped with any one of the direction described above. For example if the edge orientation = 5 degrees then edge direction = 0 degrees. By considering the above calculation, the semicircle is divided into 5 regions as shown in Fig. 2 and edge direction is given in the Table I.

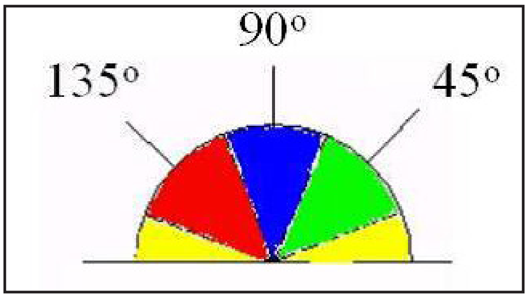

Fig. 2: Classification of Edge Direction (source:www.uta.edu/faculty/krrao/dip/Courses/EE5356/Edge%20detection%20project.doc)

TABLE I: VALUES OF EDGE DIRECTION

Sl. No	Color Region	Range in Degrees	Direction in Degrees
1	Yellow	0-22.5 &157.5-180 degrees	0 degrees
2	Green	22.5-67.5 degrees	45
3	Blue	67.5-112.5 degrees	90
4	Red	112.5-157.5	135

Non-Maximum Suppression

After finding the edge direction, the next step is to suppress any pixel (set pixel value is equal to zero) value that is not considered to be an edge by tracing along the edge in the edge direction. This process is called Non-Maximum Suppression. This suppression will produce a thin layer in the output image.

Double Threshold Edge Classification

Hysteresis is used as a means to eliminate streaking. Streaking is essentially the breaking up of an edge contour. It is caused by the operator output fluctuating above and below the threshold. Consider a single threshold, T1 is applied to an image, and an edge has an average strength equal to T1, then due to noise, there will be various instances in which the edge may dip below the threshold. Similarly, it can also extend above the given threshold, resulting in an edge looking like a dashed line. In order to avoid this, hysteresis uses two thresholds, a high and a low. Any pixel in the image that has a greater value than that of T1 is presumed to be an edge pixel, and is immediately marked as such. Then, any pixels that are connected to this edge pixel and that have a greater value than that of T2 are also selected as edge pixels. If you think of following an edge, you need a gradient of T2 to start but you don't stop till you hit a gradient below T1.

Using Defocus for Depth Estimation

In application like computer graphics and computer vision includes depth estimation, image de-blurring, image quality assessment and refocusing, Defocus for Depth Estimation plays an important role. In conventional camera, Focusing is more challenging one. Image captured by conventional camera that is not calibrated is facing more challenging problem of recovering the defocus map from a single image. By using the first and second order derivatives of the input image, Elder and Zucker [12] found the blur amount and locations of the edges. The obtained defocus map is sparse.

Calculation of the Gaussian Gradient Magnitude

The Gaussian kernel is used to re-blur edge in an image. After finding out the re-blur an edge in an image, then calculate the ratio between the gradient magnitude of the step edge and the re-blurred edge is calculated. The ratio is maximum at the edge location. With the help of obtained maximum value, compute the amount of the defocus blur at the edge location. Using the formula given below in equation 5, calculate the gradient of the re-blurred edge. Here, in the re-blur Gaussian kernel, σ_o is the standard deviation [12].

Gaussian Gradient Magnitude,

$$G = \frac{A}{\sqrt{2\pi(\sigma^2+\sigma_0^2)}} \exp\left(\frac{x^2}{2(\sigma^2+\sigma_0^2)}\right) \tag{5}$$

It is called the re-blur scale. The gradient magnitude ratio between the original and re-blurred edges is given by the formula shown in equation 6.

Ratio of Gradient Magnitude,

$$R = \frac{|\nabla i(0)|}{|\nabla i_1(0)|} = \sqrt{\frac{\sigma^2+\sigma_0^2}{\sigma^2}} \tag{6}$$

Giving an insight on both the above equations, it is observed that the gradient edge depends on both σ (blur amount) and the edge amplitude A. So, maximum gradient magnitude ratio R eliminates the effect of edge amplitude A. Then maximum gradient magnitude R depends only on $\sigma_{o \text{ and }} \sigma$.

Removal of Noise (Gradient Image)

IMFILTER is used for removing noise in the gradient image [18].

$$\mathrm{B} = \mathrm{IMFILTER}(\mathrm{A}, \mathrm{H}) \quad (7)$$

Here, multidimensional filter (H) is used to filter multidimensional array (A). Multidimensional Array A can be of any dimension and class [18]. Array A can be either a non-sparse numeric array or logical array. The size and class of output array B is same that of array A. Floating point mathematics is used to compute each element in an output array B. Fractional values are rounded off in an array A, when array A is an integer or a logical array.

Generation of the Defocus Map

The output of the defocus blur estimation is a sparse defocus map d0(x). So the next step is to obtain a full depth map d(x) by propagating the defocus blur estimation from edge location to the entire image. To obtain full depth map, d(x) (defocus map) is required which is close to the d0(x) (sparse defocus map) at each edge location. It is preferably ideal that the defocus blur discontinuities are to be aligned with the image edges [15]. Edge-Aware interpolation methods are generally used for tasks such as these. Here, Bilateral filter [13] is applied to the existing defocus map which smoothes while preserving the edges and thus the final thin focus depth map is obtained. The Laplacian matting is applied to perform the defocus map interpolation.

Using Laplacian Process for Depth Estimation

When dealing with image matting, a graph can be constructed, in such a way that each pixel is a vertex and, adjacent pixels share an edge. A graph such as this is naturally non looping and undirected. A large part of image matting, however, is determining appropriate weights for the edges, so it cannot immediately construct Graph Laplacian. Instead, a variant of the Graph Laplacian called the Matting Laplacian is used. Although identical in purpose, D and W must be produced differently from simple adjacency. The distance between pixels must be taken into account, but will be weighted by differences in color. For example, a red and a blue pixel directly next to each other would likely be less "adjacent" than two red pixels a slight distance apart. The way these numbers are produced is called the affinity function, and it has been defined in various ways.

III. Evaluation Metric

The evaluation metrics are the criteria used for testing the efficiency of the developed system. They are the set of predefined rules that analyze the results obtained from the various test cases to get the quantitative measure of the accuracy of the system.

Given a predicted depth image and the corresponding ground truth, with ˆdp and dp denoting the estimated and ground-truth depths respectively at pixel p, and T being the total number of pixels for which there exist both valid ground truth and predicted depth, the following metrics have been reported in literature [22]:

Absolute Relative Error =

$$\frac{1}{T}\sum_p \frac{|d_p-\widehat{d_p}|}{d_p} \quad (8)$$

Linear Root Mean Square Error (RMSE) =

$$\sqrt{\frac{1}{T}\sum_p (d_p-\widehat{d_p})^2} \quad (9)$$

Log scale invariant RMSE =

$$\frac{1}{T}\sum_p (\log\widehat{d_p} - \log d_p + \alpha(\widehat{d_p}, d_p))^2\widehat{d_p} \quad (10)$$

Where, α(ˆdp, dp), addresses scale alignment.

We consider Absolute Relative Error Linear and Root Mean Square Error (RMSE) [17] during this metric evaluation since log scale invariant RMSE [21] is just another variant of RMSE and the results are proportional to each other.

IV. Proposed Work and Discussion

The objective of estimating the depth in X-ray images has been achieved with this developed model using "Defocus method" and "Laplacian Matting" in which the depth of a map is computed based on color similarities. With the metrics derived from Absolute Relative Error (ARE) and Relative Mean Square Error (RMSE) an average accuracy of up to 93% has been achieved which makes this model reliable and will be useful to reduce both the amount of radiation to which the patient is exposed and the expenses up to 50%.

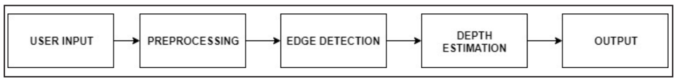

Fig. 3: Block Diagram of Depth Estimation Model

- User feeds the medical image as an input.
- The system then removes the noise from the image.
- The system then identifies significant edges in image based on the threshold set.
- The output is received as a depth image of given 2D-image.

V. Experimental Results and Analysis of the Depth Estimation in X-Ray Images

The analysis of the system is performed to determine which inputs provided the best results and how these inputs affect the overall efficiency of the system. In the testing phase, various modules of the system are tested individually by providing sample inputs and observing the outputs they produce. In this chapter the evaluation metrics that are available to estimate the accuracy and reliability of the results obtained from input images so that it is possible to determine for which kind of input the system functions more efficiently [22].

Experimental Dataset

The Experimental Dataset used for the project has been accumulated from multiple medical resources i.e. medical institutions and diagnosis centers for X-ray cases. For the project to get effective output there is a need for datasets that involves two views of an object mandatorily and it is seldom found in cases of global directories of datasets. The dataset mainly focuses on the images which are in jpeg format. The main reason for using jpeg format is that the images can be compressed easily. They take lesser file size. The picture quality is high with little level of compression.

Performance Analysis

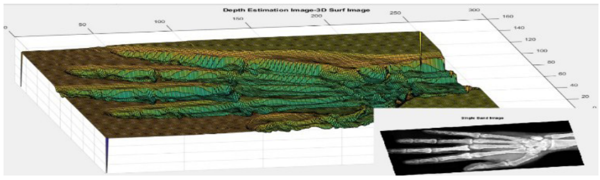

Fig. 4: X-ray Projection of a Hand

Fig. 4 shows the surf image of the input image that shows the estimated 3D structure of the given input image. Since the model produces the output images now, there is a need to justify that the output that has been displayed is accurate enough to be relied on. For that, evaluation metric is applied to calculate the relative errors and estimate the accuracy of the output.

One of the main inferences that are to be taken from the 3D plot obtained is that the scale used to reflect the estimated depth value is unique and there is a need to convert the value into useful units i.e. inches or centimeters. So, these values are referred on the vertical axis as Depth units denoted by Z. With the help of multiple trials, a conversion factor of '3.1' was calculated which when multiplied with the obtained z values, gives the estimated depth values in terms of inches and this when further multiplied by 2.54 gives the same values in terms of centimeters. But, one of the problems that has been encountered is that, irrespective of the size of the object in the image, if the gradients obtained from the input are similar in nature, the conversion estimates the depth value of the image always scaling up to a certain size because every input image will be compressed to a fixed size and the conversion factor scales it up back to a fixed dimension. So, this might result in getting the same estimated depth for two different sizes of input if they have similar gradient values which are a not to be overlooked at all. To overcome this limitation, there is a need to include at least one size-parameter from the actual object while rendering this process so that it can help us to scale up back to the desired dimension which is of life size.

$$Depth\ x = \frac{Z*3.1*2.54*\text{optimal width}}{\text{standardised width}}\ \text{cm} \qquad (11)$$

Where, Z i.e. depth value is obtained by taking the difference between upper bound and lower bound of the intensity values in the 3D map, 3.1 is the conversion factor, 2.54 to convert inches to cm, Optimal width is the width of the object in the image which is given as an input to the size-parameter that helps in scaling up the image to real size. Standardized width-this value is fixed and is a consequence of the conversion factor.

The above formula will be the solution for the existing problem. The standardized width is found out to be 9.5cm after testing with the sample cases that resulted in scaling up back to more or less an equal value of width. So, the formula now becomes simpler and can be written as:

$$Depth\ x = \frac{Z*3.1*2.54*\text{optimal width}}{9.5}\ \text{cm} \qquad (12)$$

This enables us to take any two values in the 3D surf image that has been created and find the estimated difference in depth between them in centimeters (up to 4 decimal points).

The below comparison shows us a form of a histogram based values taken from the actual and the estimated images at regular intervals (in centimeters and real size) and these values will be substituted into the formulae in the evaluation metrics and thus can come to a decision on the correctness of the model and on the amount of error that can be tolerated during this process of getting the estimated output.

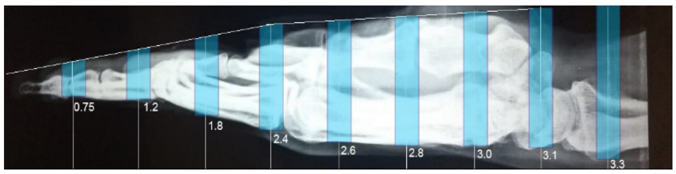

Fig. 5: Histogram Values on the Ground Truth Image1

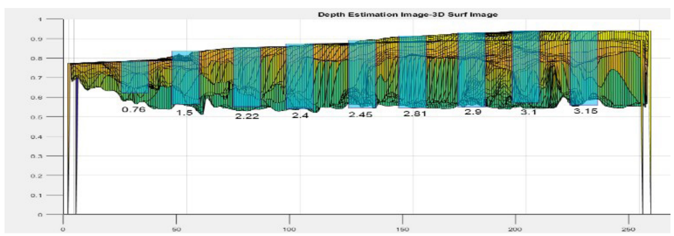

Fig. 6: Histogram Values on the Estimated Surf Image1

Table II: Values of Relative Error for Image1 (ARE and RMSE) Image 1:Absolute Relative Error: 0.026 RMSE: 0.2264

Estimated Depth (Z1) (in cm)	0.76	1.5	2.22	2.4	2.45	2.65	2.81	2.9	3.05	3.15
Actual Depth (Z1) (in cm)	0.75	1.2	1.8	2.4	2.6	2.8	3	3.1	3.3	3.4
Relative Error (RMSE)	0.0001	0.09	0.1764	0.0225	0.0225	0.0361	0.04	0.0625	0.0626	0.0625
Relative Error (Absolute)	0.013	0.25	0.23	0	0.05	0.053	0.061	0.064	0.075	0.073

The Table II above shows us the metrics obtained from the Absolute relative error and Linear Root Mean Square Error (RMSE) for which inputs have been taken from multiple histogram values taken from each image and here, the table specifically shows metrics from the values of above shown input image1.

Inference

From the results that have been obtained, an estimate can be given that the range of the error through the metric of Absolute Relative error has been from 0.5-3% from the data set that have been used and can possibly have a worst case of 6% for any unusual case. So talking in terms of accuracy [24], it can be said that the model can give an average accuracy of 96%. In contrast, the RMSE values [23] are also promising with the tolerable deviations of 2-5% on various test cases and can expect a worst case of 10% deviation from the original which leads to come to a concluding fact that through RMSE, there is an accuracy range 95-98% and 90% at the worst.

VI. Conclusion

There has been a success in developing an algorithm that performs 2D to 3D conversion of X-ray images that estimates the 3D structure for a given input with a fair accuracy. This project covered methods used to estimate depth in X-ray images and classified these methods based on the algorithms used for depth estimation and cost effectiveness. The objective of estimating the depth in X-ray images has been achieved with this developed model using “Defocus method” and “Laplacian Matting” in which the depth of a map is computed based on color similarities. This model will be useful to reduce the amount of radiation to which the patient is exposed and the expenses up to 50%.

Limitations of the Project

Though, good results have been obtained on the working cases of input, the actually scope of choosing the kind of input is restricted to mostly the X-rays of hands and legs only. So this model is not so suitable to be applied on other kind of X-rays for example, chest or spine or skull etc. This model cannot estimate the 3D structure of a given input image if the orientation of the object is not in the format of a proper front view, which means other X-rays that have a different orientation of profile, cannot be used for the estimation of depth.

Future Enhancement

- Extending the proposed method to perform clinical task like analyzing the depth of bone cracks and be able to locate it with greater accuracy in a given X-ray.
- Integrating multiple perspectives of a medical object (input object) to generate a 3D model.
- Extending the proposed solution to perform depth estimation for other categories of X-rays (Legs, Knee, Elbow, Chest, etc).

References

[1] J. L. Herrera, J. Konrad, C. R. del-Blanco, and N. García, “Learning-based depth estimation from 2D images using gist and saliency,” *2015 IEEE International Conference on Image Processing (ICIP),* IEEE, pp. 4753-4757, Quebec City, QC, Canada, 27-30 September 2015.

[2] J. L. Herrera, C. R. del-Blanco, and N. García, “Fast 2D to 3D conversion using a clustering-based hierarchical search in a machine learning framework,” *2014 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON)*, IEEE, pp. 384-387, Budapest, Hungary, 2-4 July 2014.

[3] P. R. Kapse, and S. S. Hippargi, “Automatic 2D-to-3D image conversion using KNN search,” *2016 IEEE International Conference on Advances in Electronics, Communication and Computer Technology (ICAECCT),* IEEE, pp. 210-215, Pune, India, 2-3 December 2016.

[4] H. Tian, B. Zhuang, Y. Hua, and A. Cai, “Depth extraction from a single image by sampling based on distance metric learning,” *2014 IEEE International Conference on Image Processing (ICIP)*, IEEE, pp. 2017-2021, Paris, France, 27-30 October 2014.

[5] J. L. Herrera, C. R. del-Blanco, and N. García, “Learning 3D structure from 2D images using LBP features,” *2014 IEEE International Conference on Image Processing (ICIP),* IEEE, pp. 2022-2025, Paris, France, 27-30 October 2014.

[6] A. J. Afifi, and O. Hellwich, “Object depth estimation from a single image using fully convolutional neural network,” *2016 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, IEEE, pp. 582-588, Gold Coast, QLD, Australia, 30 November - 2 December 2016.

[7] I. Laina, C. Rupprecht, V. Belagiannis, F. Tombari, and N. Navab, “Deeper depth prediction with fully convolutional residual networks,” *2016 Fourth International Conference on 3D Vision*, IEEE, pp. 239-248, 2016.

[8] M. Liu, M. Salzmann, and X. He, “Discrete-continuous depth estimation from a single image,” *2014 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, pp. 716-723, Columbus, OH, USA, 23-28 June 2014.

[9] S. Manonmani, V. P. Lalitha, and S. Rangaswamy, “Survey on image denoising techniques,” *International Journal of Science, Engineering and Technology Research (IJSETR),* vol. 5, no. 9, pp. 2824-2827, September 2016.

[10] A. L. Wicks, A. T. Asbeck, and S. C. Southward, "Using texture features to perform depth estimation," Bhavi Bharat Kotha, Blacksburg, Virginia, 15 December 2017.

[11] Y. Ren, and J. Liu, "Single outdoor image depth map generation based on scene classification and object detection," *2014 7th International Conference on Ubi-Media Computing and Workshops*, IEEE, pp. 91-94, Ulaanbaatar, Mongolia, 12-14 July 2014.

[12] S. Zhuo, and T. Sim, "Defocus map estimation from a single image," *Pattern Recognition*, vol. 44, pp. 1852-1858, Elsevier, 2011.

[13] S. Paris, P. Kornprobst, J. Tumblin, and F. Durand, "Bilateral filtering: Theory and applications," *Foundations and Trends® in Computer Graphics and Vision*, vol. 4, no. 1, pp. 1-73, 2009.

[14] S. H. Teoh, and H. Ibrahim, "Median filtering frameworks for reducing impulse noise from grayscale digital images: A literature survey," *International Journal of Future Computer and Communication*, vol. 1, no. 4, pp. 323-326, December 2012.

[15] J. Konrad, M. Wang, and P. Ishwar, "2D-to-3D image conversion by learning depth from examples," *2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, IEEE, pp. 16-22, Providence, RI, USA 16-21 June 2012.

[16] S. Rangarajan, "Algorithms for edge detection,". Available at http://www.uta.edu/faculty/krrao/dip/EE Courses/5356/Edge%20detection%20project.doc [Accessed on January 25th 2018]

[17] A. Kulkarni, S. Rangaswamy, and Manonmani S., "Comparative study of major image enhancement algorithms," *European Journal of Engineering Research and Science*, vol. 2, no. 7, July 2017.

[18] J. Canny, "A computational approach to edge detection," *Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679-698, November 1986.

[19] Y. L. Wei, B. Li, J. D. Li, and Y. Y. Fan, "Impulse noise filtering algorithm based on dual threshold criterion," *Applied Mechanics and Materials*, vol. 437, pp. 849-852, 2013.

[20] T. Sun, M. Gabbouj, and Y. Neuvo, "Center weighted median filters: Some properties and their applications in image processing," *Signal Processing*, Elsevier, vol. 35, no. 3, pp. 213-229, 1994.

[21] R. Chandel, and G. Gupta, "Image filtering algorithms and techniques: A review," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 3, no. 10, pp. 198-202, 2013.

[22] M. Basu, "Gaussian-based edge-detection methods - A survey," *IEEE Transactions on Systems, Man and Cybernetics, Part C: Applications and Reviews*, vol. 32, no. 3, pp. 252-260, 2002.

[23] C. Cadena, Y. Latif, and I. D. Reid, "Measuring the performance of single image depth estimation methods," *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS),* 9-14 October 2016.

[24] C. Chen, J. Twycross, and J. M. Garibaldi, "A new accuracy measure based on bounded relative error for time series forecasting," Published, 24 March 2017.