Journal of Theoretical and Applied Information Technology
31st July 2019. Vol.97. No 14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

# GENERATING 3D IMAGE BASED ON DEPTH MAP ESTIMATION USING HYBRID CLUSTERING AND MERGING ALGORITHM

**SREELETHA S H[1*], ABDUL REHMAN M[2]**

[1*]Karpagam Academy Of Higher Education (Kahe), Coimbatore, Tamilnadu, India

[2]Karpagam Academy Of Higher Education (Kahe), Coimbatore, Tamilnadu, India

[1*]sreelatha.lbs@gmail.com

**ABSTRACT**

Nowadays, all the portable devices including camera, mobile phones etc., are being worked based on the multimedia applications. Along with multimedia, 3d technologies are in peak demand. Yet, the hurdles such as poor image quality and increased time complexity makes the prior 2D to 3D transformation technologies weaker as days goes by. So as to overwhelm this problem, our proposed work initiates an 3D Agglomeration framework, which is an integration of an Improved Simple Linear Iterative Clustering (ISLIC) and a Statistical Region Merging (SRM). With our proposed framework, better quality of image is retained thereby abet of Gaussian smoothing with color uniformity principle. Moreover to tackle with time overhead that are experienced with the prior methodologies, a Depth Image Based Rendering (DIBR) method is utilized in our proposed framework. This in turn makes use of an input 2D texture image to construct the required 3D output image. Thus our 3D agglomeration framework generates a high quality 3D texture image with reduced time complexity. Finally, the performance of the proposed framework work is analyzed and is compared with the other existing methodologies.

**Keywords:** *ISLIC-Improved Simple Linear Iterative Clustering; SRM- Statistical Region Merging; 3D image- three dimensional images.*

## 1. INTRODUCTION

The recent inventions [1] in orthodontic and dento facial orthopedic diagnosis with treatment arranging have depended basically upon innovative and mechanical backings, for example, imaging, jaw checking, and useful investigations. The objectives of these procedures are to reproduce or depict the anatomic and physiological realities precisely and to show the three-dimensional (3D) view decisively. Imaging [2] is a standout amongst the most essential instruments for orthodontists to assess and record size and type of craniofacial structures. Orthodontists routinely utilize 2_dimensional (2D) static imaging strategies to record the craniofacial life systems, yet profundity of structures can't be acquired and restricted with 2D imaging.

3D imaging [3] has been created in the right on time of 1990's and has picked up a valuable place in dentistry, particularly in orthodontics, and furthermore in orofacial careful applications. In 3D symptomatic imaging, a progression of anatomical information is assembled utilizing certain mechanical gear, handled by a PC and later appeared on a 2D screen to show the dream of profundity. The fundamental motivation [4] behind three dimensional imaging is to give both subjective and quantitative data around an object or an object system framework from pictures acquired with different modalities including computerized radiography, processed tomography, positron discharge tomography, attractive reverberation imaging and single photon outflow figured tomography, ultrasonically.

Here objects type [5] would be rigidity for example bones, deformable for example muscles, static for example skull, dynamic for example (heart, joints). A key segment for 3D picture creation is depth map estimation from 2D picture. The objective of 3D vision procedure is to produce the depth information which is known as depth map estimation. Profundity delineates [6] a dim scale picture that contains the separation of the items from a perspective. A significant number of traditional 2D-to-3D transformations couldn't have fulfilled the cost/idleness necessity because of the need of substantial memory or confounded calculations which incorporate seeking forms. One of the creators had exhibited a calculation that fulfills such necessities, utilizing straightforward

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

profundity models and warm/cool shading hypothesis.

A classic 2D-to-3D conversion process [7] involves of two steps: depth estimation for a given 2D image and depth-based rendering of a new image in order to form a stereo pair. 3D modeling related to computer technology, graphics on super computer, image processing, virtual reality and other research areas, is a research hotspot in recent years. According to differences in the way of their research, it might be classified into the three types: geometry-based modeling, image-based modeling, hybrid modeling method based on geometry and image. While changing over 2D to 3D [8] a portion of the key highlights must be viewed as like (I) parameters like shape, movement, shading, surface, edges, (ii) Depth signs must be found to create profundity maps, (iii) Fully programmed or semi – programmed technique include basic leadership by human administrators and have been much fruitful in giving expected outcomes and furthermore tedious. Completely programmed techniques nearly make solid assumption about the 3D scene.

Accordingly it turns out to be exceptionally troublesome assignment [9] to build a deterministic scene demonstrate that covers all conceivable foundation and closer view mixes. Keeping in mind the end goal to acquire the profundity outline, picture preprocessing is required. It is made out of the accompanying advances: 1) Image characterization; 2) Vanishing lines and vanishing point extraction.

The stereoscopic combine picture [10] is then produced by computing the parallax estimation of each pixel in the picture extricating data just from the dim level profundity outline. Keeping in mind the end goal to accomplish profundity control, vigorous and ongoing, a solitary opening holoscopic 3D imaging camera is utilized for recording holoscopic 3D picture utilizing a routinely dispersed exhibit of microlens clusters, which see the scene at a marginally extraordinary edge to its neighbor. In any case, the principle issue [11] is that the microlens cluster presents a dim fringes in the recorded picture and this causes mistakes at playback on the holoscopic 3D Display. After the pre-preparing, the 3D picture is prepared for division. The objective of this procedure is to part the picture into continuous parts. These parts can be covering and together can cover the entire picture. Highlights are computed for every one of these fragments.

Highlight extraction is done [12] after the preprocessing stage in pixel acknowledgment framework. The essential undertaking of example acknowledgment is to take an information design and accurately allocate it as one of the conceivable yield classes. Highlight extraction techniques are arranged into three noteworthy gatherings as: Statistical Features, Global Transformation and Series Expansion Features.

Even though the projection of 3D image has played an important role in fields, especially in medical field the enhancement in 3D imaging is very important one. 3D imaging in medical filed such as Ultrasounds, X-ray, and MRI is used to create visual depictions of the internal of the body for clinical analysis and medical involvement of compound diseases in a small period of time. Traditional medical imaging systems afford 2D visual representations of human organs while more advanced digital medical imaging systems (e.g. X-ray CT) can create both 2D and in many cases 3D images of human organs. These requirements brought much attention on 3D image processing.

With increasing demands of 3D contents, conversion of many existing two-dimensional contents to three-dimensional contents has gained wide interest in 3D image processing. It is important to estimate the relative depth map in an image for the 2D-To-3D conversion technique.

While on dealing with the conventional 2D imaging strategies, it is difficult to understand the profundity of image structure or the texture. But with recent days, depth map is analyzed during 2D to 3D transformation. Even though, the traditional 2D to 3D image conversion strategies fails, due to the huge requirement of memory. Since the transformation results with confounded calculations, also high cost is considered as another major hurdle when on implementing the traditional 2D to 3D conversion systems. Simultaneously in some approaches, when trying to determine the depth map, an error might be occurred due to the presence of hole regions i.e., the region with no edge information or image boundary. Furthermore, lack of image quality, clarity etc., are considered as several other problems exist with the existing system. As well the time taken for converting an image from 2D to 3D is also high, i.e., the processing time is high, which leads to increased time complexity.

In this current era, each and every multimedia device deals with 3D technology. But once a period it's difficult to convert a 2-dimensional image into

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

a 3D image. It becomes possible with the advanced technologies, where the researchers convert 2D images into 3D images with some approaches using digital image processing. Though conversion or transformation is achieved, the time consumption is high for this process, as well the image quality is too poor. In addition, due to the confounded calculations, the memory requirement is also high. These are some common hurdles that made the existing approaches undeserving.

In order to deluge these muddles, the proposed work designs a 3D agglomeration framework. Here, the framework combines the idea of clustering and merging mechanisms with the aid of Improved Simple Linear Iterative Clustering (ISLIC) and Statistical Region Merging (SRM) methods. Moreover, a Depth Image Based Rendering (DIBR) method is initiated here, so that the poor quality input 2D texture image can be successfully reconstructed as a high quality3D texture image during less time span results with minimized cost. Accordingly, our proposed framework achieves improved image quality and thereby reduces the time complexity issues that faced by the prior methodologies.

This paper is organized as follows: Section II reviews the conventional researches of this process; Section III explains the proposed methodology and its execution in detail. Section IV discussed the results obtained from the proposed methodology and a brief discussion over the conventional techniques. Conclusions are made in Section V.

## 2. LITERATURE SURVEY

[13] KeGu, VinitJakhetiya et al Introduced a novel reference less quality metric of DIBR-synthesized images using the auto regression (AR)-based local image description. It was discovered that, after the AR expectation, the reproduced mistake between a DIBR-integrated picture and its AR-anticipated picture can precisely catch the geometry mutilation. The visual saliency is then utilized to change the proposed daze quality metric to a sizable edge. This takes reasonable time to finish the process but lacks in an image quality.

[14] Jiasong Zhu, Jie Hu, proposed a different 3-D feature combination system (M3DF3) for hyperspectral picture characterization. To start with, the 2-D Gabor surface feature had been extended into 3-D (3DSF) domains to agree to the spatial– otherworldly structure of the hyperspectral picture, which is directly applied on the first hyperspectral picture instead of the Gabor features. The proposed 3DSF separately describe the hyperspectral picture from three different edges, i.e., morphology, nearby dependence, and shape smoothness. This system maintains the quality of an input image but time complexity is high.

[15] Xiaozhi Chen, KaustavKundu Introduced a new method which goes for producing an arrangement of amazing 3D object recommendations by exploring stereo symbolism. Utilize a convolutional neural net (CNN) that adventures setting and depth data to together relapse to 3D bounding box coordinates and protest posture. This analysis indicate critical execution increases over existing RGB and RGB-D object proposition methods on the testing KITTI benchmark. This system achieves better 3D view of an input image but lacks in clarity.

[16] Ming Xu and Zeyunyu introduced a tetahedralwork based approach for 3D picture division on a given picture volume. The 3d vigilant edge indicator is used to protect critical component limits in the produced tetrahedral work. Each bunch of voxels inside a tetrahedron is known as a supervoxel, which is dealt with as a hub in a chart and weighted edges are framed between neighboring tetrahedra. At that point a diagram cut technique is connected on the weighted chart to finish the tetrahedral work. This system achieved better image quality but processing time is high.

[17] Yang et al. presented a novel algorithm to estimate depth information from a video via scene classification algorithm. In order to obtain perceptually reliable depth information for viewers, the algorithm classifies them into three categories: landscape type, close-up type, linear perspective type firstly. Then employed aspecific algorithm to divide the landscape type image into many blocks, and assign depth value by similar relative height cue with the image. As to the close-up type image, a saliency-based method is adopted to enhance the foreground in the image and the method combine it with the global depth gradient to generate final depth map. By vanishing line detection, the calculated vanishing point which is regarded as the farthest point to the viewer is assigned with deepest depth value. Finally, depth image-based rendering is employed to generate stereoscopic virtual views after bilateral filter. Experiments show that the proposed algorithm can achieve realistic 3D effects and yield satisfactory results, while the perception scores of anaglyph images lie between 6.8 and 7.8. This system has taken less time but it lacks in an image quality.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

In reference [13], this system takes reasonable time to finish the process but lacks in an image quality. In reference [14], this system maintains the quality of an input image but time complexity is high. In reference [15], this system achieves better 3D view of an input image but lacks in clarity. In reference [16], this system achieved better image quality but processing time is high. In reference [17], this system has taken less time but it lacks in an image quality.

## 3. 3D AGGLOMERATION FRAMEWORK

While on dealing with the conventional 2D imaging strategies, it is difficult to understand the profundity of image structure or the texture. But with recent days, depth map is analyzed during 2D to 3D transformation. Even though, the traditional 2D to 3D image conversion strategies fails, due to the huge requirement of memory. Since the transformation results with confounded calculations, also high cost is considered as another major hurdle when on implementing the traditional 2D to 3D conversion systems. Simultaneously in some approaches, when trying to determine the depth map, an error might be occurred due to the presence of hole regions i.e., the region with no edge information or image boundary. Furthermore lack of image quality, clarity etc., are considered as several other problems exist with the existing system. As well the time taken to convert a 2D image into a 3D image is also high, i.e., the processing time is high, which leads to increased time complexity. These are the major hurdles that makes the existing approaches undeserving.

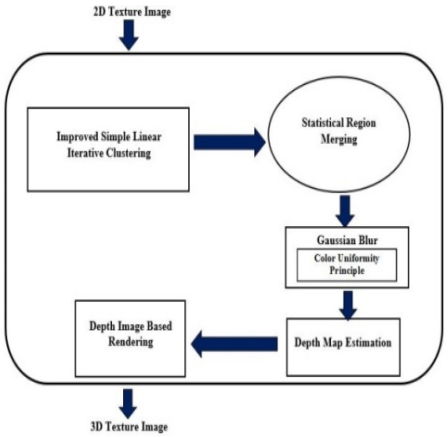


*Figure 1: Proposed system architecture*

Thus to deal with all such issues, our proposed 3D agglomeration framework integrates Improved Simple Linear Iterative Clustering (ISLIC) with Statistical Region Merging (SRM) methods for the effective construction of 3D texture image. With improved SLIC (ISLIC), super pixels are generated based on the color similarities. Moreover, ISLIC provides the information about the edges and the boundary pixels, so that the quality of the resultant 3D image has been improved. It is then followed with Gaussian smoothing in order to analyze the initial depth of input image. Furthermore, the occurrences of noise will be detected and removed with this process. Incase, if any hole region is detected, it is then corrected with the color uniformity principle. Finally, depth map is estimated using the bottom-to-up model with the utilization of Depth Image based Rendering (DIBR) with which, the left and the right views of the image is retained with the integration of the depth map and the original 2D texture image. Thus the resultant 3D texture will be obtained with high quality and at reduced time complexity. The proposed system architecture is described by the figure 1.

### 3.1. Improved SLIC

In this phase, the pixels of an image o would be partitioned into classes that efficiently compare with the objects in an image. In this work ISLIC is proposed to effectively carry out the segmentation process.

While using Conventional SLIC algorithm to generate the super pixels, more number of small areas will be formed in the image surface that would significantly affect the segmentation result so we have under went image normalization to make an improved SLIC

The normalization process transforms an n-dimensional gray scaleimage with the intensity values in the range (Min, Max) into a new image.

$$I : \{X \le R^n\} \rightarrow \{Min, ....., Max\}$$

The linear normalization of a gray scale digital image is performed according to the formula

$$I_N = (I - Min)\frac{newMax - newMin}{Max - Min} + newMin$$

Normalization would be of non-linear if there is not a linear relationship between I and $I_N$.

$$I_N = (newMax - newMin)\frac{1}{1 + e^{-\frac{1-\beta}{\alpha}}} + newMin$$

Where α represents the width of the input intensity range, β represents the intensity around which the range is centered.

Journal of Theoretical and Applied Information Technology
31st July 2019. Vol.97. No 14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Simple linear iterative clustering has only one parameter, k, which is the desired number of equally sized super pixels to be generated. The image is first separated into a grid. The center of each grid tile is then used to initialize a corresponding k-means (up to a small shift to avoid image edges).

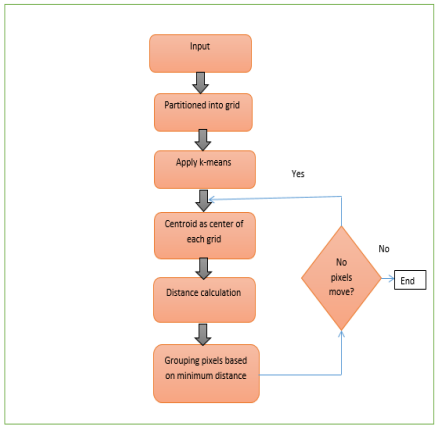


*Figure 2: Flow diagram of ISLIC*

SLIC would not compare each pixel with all pixels in an image. For a region of approximate size S × S, the distance D is computed in a region 2S × 2S around the super pixel center, reducing the number of D calculations. After each pixel is assigned to the nearest cluster center, the new cluster centers are updated for the newly generated super pixels, representing the mean vector of all the pixels belonging to the super pixel. A residual error E is computed, as the distance between previous center locations and recomputed center locations. The iterative procedure would be terminated when the error converges or when it reaches the threshold of input iteration.

$$D = \sqrt{\left(\frac{d_c}{O}\right)^2 + \left(\frac{d_s}{S}\right)^2} \quad (1)$$

$$d_c = \sqrt{\sum_{s_b \in B} \left(P(u_i, v_i, s_b) - P(u_j, v_j, s_b)\right)^2} \quad (2)$$

$$d_s = \sqrt{(u_j - u_i)^2 + (v_j - v_i)^2} \quad (3)$$

Where indicates the distance measurement which includes both distance of color proximity and distance of spatial proximity. The distance of color proximity would maintain the homogeneity of the superpixel and the distance of spatial proximity would maintain the superpixel compactness.

Indicates the color and spatial distance between pixels $P(u_i, v_i, s_b)$ and $P(u_j, v_j, s_b)$ in the spectral band $S_b$. $B$ Indicates the spectral band and $O$ acts as an optimal parameter for managing the compactness of super pixels.

**Step 1** SLIC would commence by dividing an input image O into a rectangular grid with R*S tiles

$$b|u| = g\sqrt{\frac{1}{2Q|u|}\ln\left(6|O|^2 U_{|U|}\right)}$$

Where $R = \left[\frac{width}{G}\right]$ , $S = \left[\frac{height}{G}\right]$ (4)

**Step 2** K means cluster would be initiated from each grid centre and calculate the average weighted

$$x_i = round \quad i\left[\frac{width}{G}\right] y_i = round \quad j\left[\frac{Height}{G}\right] \quad (5)$$

**Step 3** In order to prevent keeping these centres on worst portion of the image, the centres are moved in a 3*3 neighbourhood to reduce the edge strength

$$edge(x, y) = \{P(x+1, y) - I(x-1, y)\} + \{I(x, y+1) - I(x, y-1)\} \quad (6)$$

**Step 4** Then the regions are obtained by running k-means clustering, started from the centres

$$c = \{\sum(x_i, y_i), i = 0,1,.....,R-1, j = 0,1,.....,S-1\} \quad (7)$$

The output of this phase would be clustered form of an input image O.

### 3.2. Region merging based on statistical data

After k-means had achieved better convergence, remerging is adapted to remove each clustered region whose region is small than the minimum region size.

The functionality is to begin with one locale for each pixel and after that applying a factual test on neighboring areas (in rising request of force contrasts) regardless of whether the mean powers are adequately sufficiently comparative to be combined.

Let $O$ be an observed image that contains $|O|$ pixels, every pixel in that is included of R, G, B color channel values fitting to the set {0, 1, . . . , g −

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

1} (where g = 256 for 8 bit RGB images). $O$ is an observation of a true image $O^*$ in which pixels are flawlessly symbolized by a family of distributions from which each of the observed color channel is sampled. The optimal statistical regions in $O^*$ include a property of homogeneity such that inside any statistical region and given any color channel, the statistical pixels have the same expectation, whereas the expectations of adjacent statistical regions vary in at least one color channel. $O$ is obtained from $O^*$ by sampling each statistical pixel for observed RGB values. The color channel values for every pixel in $O^*$ are replaced by a set of Q independent random variables, which take on values from [0, g/Q]. It is to be renowned that the Q parameter can be used to quantify the statistical complexity of $O^*$, the generality of the model, and the statistical difficulty of the problem. Higher values of Q result in finer segmentation and thus the generation of more regions.

Statistical Region Merging holds two components first one is a merging predicate and the second one is the order in examining the predicate.

The predicate is defined as

$$P(U,U') = \begin{cases} true & if \quad \forall a \in \{R,G,B\}, \left|U'_a - U_a\right| \le \left|\sqrt{b^2(U)} + b^2 U\right| \end{cases} \quad (8)$$

$$b|u| = g\sqrt{\frac{1}{2Q|u|}\ln\left(6|O|^2 U_{|U|}\right)}$$

Here $u$ and $u'$ indicates the regions to be tested. $u_a$ Indicates the color channel average in the location is the set of locations with pixels. The order of the merging region is indicated by invariant A which state that when the test is performed between the two locations which is inside a real location, additional tests are not needed.

$u$ and $u'$ be pixels of an image $M$, and $R(U)$ stand for the current region to which a pixel U belongs. The SRM algorithm first sorts these pairs in increasing order according to a function $f(U,U')$. After the sorting is completed, the order is traversed only once, performing the merging test $P(R(u),R(u'))$ for any pair of pixels $(U,U')$ for which $R(u) \neq R(u')$ and merging $R(U)$ and $R(U')$ if it returns true. Thus with SRM, the boundary pixels are identified to perform modification in the edges so as to generate better convergence.

### 3.3. Gaussian Smoothing or Gaussian Blur

The image O which is effectively and equally clustered would be subjected to depth calculation.

Gaussian function is defined as the density function of normal distribution

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \quad (10)$$

Where $\mu$ indicates the average of $x$ and $\mu = 0$ since center point would be the starting point while calculating the average value

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2} \quad (11)$$

The two dimensional Gaussian function is defined as

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\left(x^2+y^2\right)/2\sigma^2} \quad (12)$$

Representing images as matrices of numbers $R = (r_{i,j})$ each number indicates the intensity of a pixel and Gaussian matrix as $S = (s_{i,j})$

Using defined function

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\left(x^2+y^2\right)/2\sigma^2} \quad (9) \quad (13)$$

$$S = \begin{bmatrix} S_{1,1} & S_{1,2} & S_{1,3} \\ S_{2,1} & S_{2,2} & S_{2,3} \\ S_{3,1} & S_{3,2} & S_{3,3} \end{bmatrix} = \begin{bmatrix} g(-1-1) & g(-1,0) & g(-1,1) \\ g(0,-1) & g(0,0) & g(0,1) \\ g(1,-1) & g(1,0) & g(1,1) \end{bmatrix}$$

Blurring R with S

$$C_{k,l} = \sum_{i=1}^{1} \sum_{j=1}^{1} r_{k+i,l+j} s_{1+i,1+j} \quad (14)$$

$$C = \begin{bmatrix} R_{1,1} & R_{1,2} & R_{1,3} \\ R_{2,1} & R_{2,2} & R_{2,3} \\ R_{3,1} & R_{3,2} & R_{3,3} \end{bmatrix} (\&) \begin{bmatrix} S_{1,1} & S_{1,2} & S_{1,n} \\ S_{2,1} & S_{2,2} & S_{2,3} \\ S_{3,1} & S_{3,2} & S_{3,3} \end{bmatrix}$$

The output of this process would be blur (O). With this method, the presence of noise is detected and is removed. As well along with this Gaussian

Journal of Theoretical and Applied Information Technology
31st July 2019. Vol.97. No 14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

smoothing, the color uniformity principle is used to detect the hole regions. Here by smoothing the image the number of big holes are reduced. The blur image is then assigned to estimate the final depth map.

### 3.4. Depth Estimation

After on determining the initial depth map from the un-sharped image generated by the Gaussian blur, have to calculate the depth value for each and every pixel in the texture. Here in our proposed work, bottom-to-up mode is used to estimate the depth map. The bottom-to-up mode is defined by Eq. (15).

$$Depth = White - i \times \left( \frac{White}{Height} \right) \quad (15)$$

Where, 1 < slope, i=1, 2, 3… height. Height is the said to be the height of the texture image.

### 3.5 3D construction using DIBR

Here to construct a 3Dtexture image, the DIBR method is used. In DIBR method, we have to calculate the Shift_value using Eq. (16).

$$Shift\_value = abs[Depth - Zc] / (255) \quad (16)$$

where, Depth be the Depth value obtained from Eq. (15) and Zc be the convergence Distance. Here the Shift_value is used to shift the original pixels with its depth values to calculate the left and the right image view. The left and the right image view is calculated using the below formulas defined by Eq. (17) and Eq. (18).

$$Xlv(x, y) = pic(Xc - Shift\_value) \quad (17)$$

$$Xrv(x, y) = pic(Xc + Shift\_value) \quad (18)$$

where, $Xlv$ be the left view of image and $Xrv$ be the right view of the image respectively. With this DIBR, the real 3D recording camera setup is formulated with a virtual camera configuration formula. Finally based on that configuration formula, both the view of the images is calculated and are integrated with the depth map. It would resultswithafine 3D texture image retaining high quality and at reduced time complexity. This approach is also implemented on Matlab.

## 4. RESULTS AND DISCUSSIONS

This section shows the results obtained with our proposed work. The whole setup is implemented in Matlab. The simulation results obtained for different texture images with our proposed work is shown below.

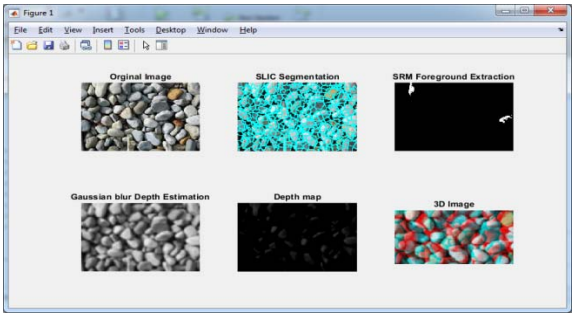


*Figure 3: Experimental result for image 1*

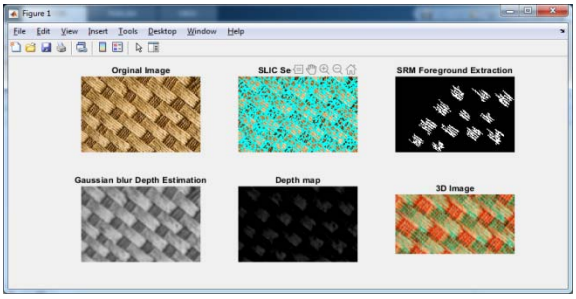


*Figure 4: Experimental result for image 2*

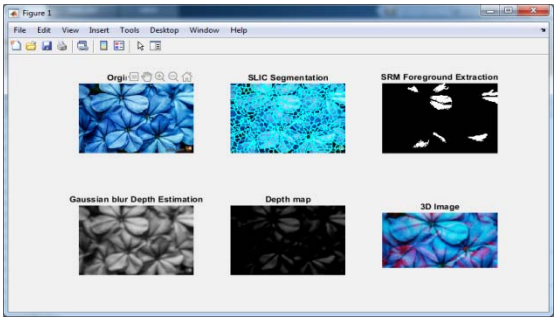


*Figure 5: Experimental result for image 3*

Figure 3 -5represents the experimental results explaining the process of image transformation from 2D to 3D. That the original 2D texture image is given as input to our proposed work. The input image is assigned to the ISLIC method, where the super pixels are generated to find the edge information. It is segmented and the foreground is extracted from the texture image. Further Gaussian smoothing is carried out to blur the image, so as to detect and remove in case of the occurrence of noise is detected. Finally the depth map is determined and is integrated with the original 2D texture image to develop a required 3D texture image as output.

In order to assess the performance of the proposed system, following performance metrics are taken into an account PSNR, structural

Journal of Theoretical and Applied Information Technology
31st July 2019. Vol.97. No 14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

similarity index and virtual information fidelity, which are shown by the performance analysis table 1.

### 4.1. Structural similarity index

SSIM is used for measuring the similarity between two images. The SSIM index is a full reference metric; in other words, the measurement or prediction of image quality is based on an initial uncompressed or distortion-free image as reference.

The SSIM index is calculated on various windows of an image. The measure between two

windows x and y common size N*N is

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

#### 4.1.1. PSNR and MSE

The PSNR block computes the peak signal-to-noise ratio, in decibels, between two images. This ratio is often used as a quality measurement between the original and a compressed image. The higher the PSNR, the better the quality of the compressed, or reconstructed image.

*The Mean Square Error (MSE) and* the *Peak Signal to Noise Ratio (PSNR)* are the two error metrics used to compare image compression quality. The MSE represents the cumulative squared error between the compressed and the original image, whereas PSNR represents a measure of the peak error. The lower the value of MSE, the lower the error.

To compute the PSNR, the block first calculates the mean-squared error using the following equation:

$$MSE = \sum_{M,N} \left[ \frac{I_1(m,n) - I_2(m,n)}{M * N} \right]^2$$

*M* and *N* are the number of rows and columns in the input images, respectively. Then the block computes the PSNR using the following equation

$$PSNR = 10\log_{10}\left(\frac{R^2}{MSE}\right)$$

The performance of the proposed work is examined with the above equation and the resultant performance values are described in terms of graphical representations by figures 6 and 7.

*Table 1: Performance Analysis of Proposed Method*

| | |
|---|---|
| Accuracy | 95.2427 |
| Sensitivity | 98.8732 |
| Specificity | 97.8854 |
| Precision | 96.0603 |
| Recall | 95.5759 |
| F-measure | 96.7439 |
| MSE | 21.0667 |
| PSNR | 17.427 |
| SNR | 12.6339 |
| Noise | 16.6 |
| Execution Time | 5.7204 |
| Bit error Rate | 0.000231 |
| Similarity Structure | 98.36 |

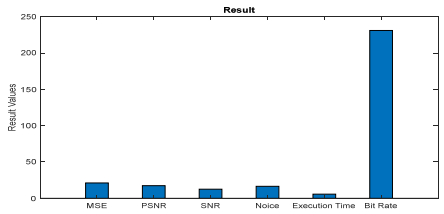


*Figure 6: Performance analysis of proposed method (MSE, PSNR, SNR, noise, execution time and bit rate)*

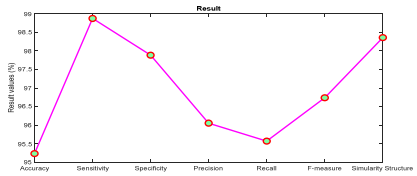


*Figure 7: Performance analysis of proposed method in terms of accuracy, sensitivity, specificity, precision etc.*

Thus the resultant performance analysis shows better outputs for our proposed work in terms of several metrics such as accuracy, sensitivity, specificity, precision etc. here the execution time is low, thus reduced the time complexity face by the prior methodologies.

### 4.2. Performance Comparison

For the purpose of comparison with the proposed system following methodologies are taken into an account such as relative height interface, time coherent depth maps and saliency detection.

*Table 2: PSNR comparison of existing methods with proposed method [18]*

| Methods | Image 1 | Image 2 | Image 3 | Image 4 | Image 5 | Image6 |
|---|---|---|---|---|---|---|
| Relative Height | 10.275 | 12.2076 | 14.8094 | 10.4348 | 10.9573 | 8.5441 |

| Interface | | | | | | |
|---|---|---|---|---|---|---|
| Time-coherent depth maps | 10.5352 | 11.2514 | 11.841 | 12.1487 | 10.358 | 6.5429 |
| Saliency Detection | 15.4871 | 12.4473 | 12.362 | 13.2476 | 10.4552 | 6.5943 |
| Proposed System | 17.427 | 16.257 | 15.325 | 17.255 | 14.36 | 11.25 |

Table 2 describes about the performance comparison in terms of PSNR with our proposed work and the prior technologies such as Relative height Interface, Time coherent depth maps and saliency detection.

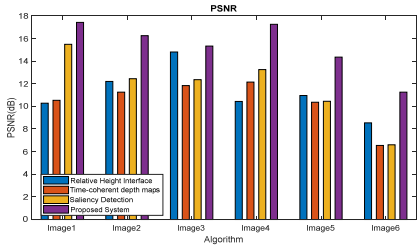


*Figure 8: PSNR comparison of existing methods with proposed method*

Figure 8 shows a better comparison result exhibiting high PSNR rate for our proposed work when compared it with the other prior methodologies. Thus the resultant texture image retains good quality and clarity than the existing methodologies.

*Table 3: SSIM comparison of existing methods with proposed method [18]*

| Methods | Image 1 | Image 2 | Image 3 | Image 4 | Image 5 | Image 6 |
|---|---|---|---|---|---|---|
| Relative Height Interface | 0.5219 | 0.688 | 0.5591 | 0.6171 | 0.6493 | 0.9747 |
| Time-coherent depth maps | 0.4777 | 0.6744 | 0.5693 | 0.6772 | 0.555 | 0.9665 |
| Saliency Detection | 0.5786 | 0.7065 | 0.6586 | 0.7061 | 0.6305 | 0.9667 |
| Proposed system | 0.9438 | 0.9754 | 0.8712 | 0.9741 | 0.8423 | 0.9817 |

Above comparison table 3 for SSIM demonstrates about the comparison of our proposed work with some prior techniques in which the proposed work with ISLIC shows better results, yeild good image quality. As our proposed method utilizes improved SLIC method for segmenting the appropriate features from the image, it provides an output image having fine quality. The diagrametic representation for this comparison table is shown by the figure 9.

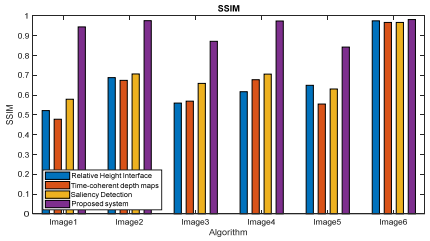


*Figure 9: SSIM comparison of existing methods with proposed method*

Table 4 describes about the VIF comparison with this table and figure 10, it is confirmed that the proposed method results with high VIFs values indicating the presence of strong or high correlation rate. As to the image quality assessment VIF, the average VIF value of our algorithm is higher than the relative height interface, time coherent depth maps and saliency detection.

*Table 4: VIF comparisons of existing methods with proposed method [18]*

| Methods | Image 1 | Image 2 | Image 3 | Image 4 | Image 5 | Image 6 |
|---|---|---|---|---|---|---|
| Relative Height Interface | 0.0504 | 0.1065 | 0.0279 | 0.038 | 0.1847 | 0.2142 |
| Time-coherent depth maps | 0.035 | 0.0554 | 0.032 | 0.0238 | 0.0721 | 0.0991 |
| Saliency Detection | 0.1002 | 0.1287 | 0.0445 | 0.058 | 0.0974 | 0.1185 |
| Proposed System | 0.1358 | 0.1587 | 0.0687 | 0.0874 | 0.1972 | 0.368 |

Journal of Theoretical and Applied Information Technology
31st July 2019. Vol.97. No 14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

*Figure 10: VIF comparison of existing methods with proposed method*

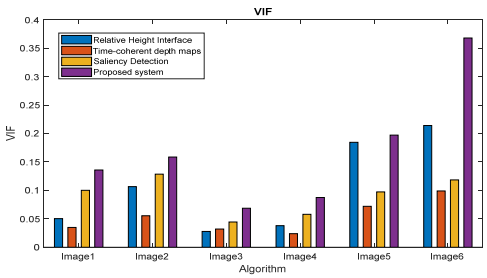


Table 5 and figure 10 shows the experimental results comparison in terms of Feature Similarity Index for Image Quality Assessment(FSIM). Here the proposed work achieves better results when comparing the FSIM values with the existing methods such as Relative Height Interface,Time-coherent Depth maps and Saliency Detection.

*Table 5: FSIM Values For Different Methods*

| Methods | Mask | Home |
|---|---|---|
| Relative Height Interface | 0.9062 | 0.956 |
| Time-coherent Depth maps | 0.9454 | 0.915 |
| Saliency Detection | 0.96 | 0.956 |
| Proposed | 0.982 | 0.978 |

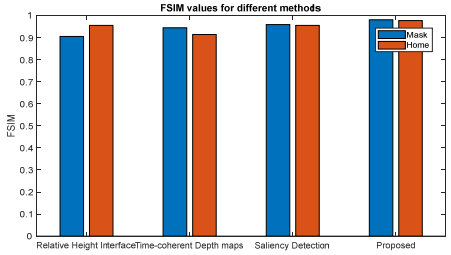


*Figure 11: FSIM values for different methods*

Gradient Magnitude Similarity deviation (GMSD) values in table 6 and figure 12 describes about the similarity verification results of the segmented images with the original image using some quantitative similarity metrics. Here our comparison results shows a better output having good similarities, which in turn represents the fine quality of the output image obtained using our proposed method than the prevailing ones.

*Table 6: GMSD Values For Different Methods*

| Methods | Mask | Home |
|---|---|---|
| Relative Height Interface | 0.8624 | 0.8502 |
| Time-coherent Depth maps | 0.8998 | 0.889 |
| Saliency Detection | 0.9198 | 0.9109 |
| Proposed | 0.9681 | 0.9455 |

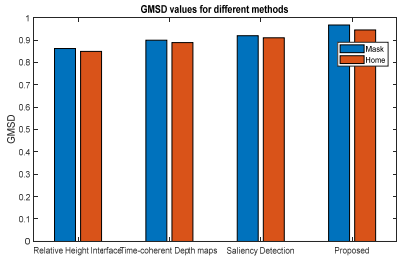


*Figure 12: GMSD values for different methods*

Thus the above tables and figures shows a better result for this proposed work, thereby it concludes that in our proposed framework using ISLIC, the drawbacks faced with the existing methodologies are successfully overwhelmed with reduced time complexities.

## 5. CONCLUSION

The conventional techniques for converting image from two dimensional to three dimensional had availed the parameters such as size of an image, edge, motion and texture for constructing stereo pairs. This leads to various problems such as distortion, lack of image quality, increased time complexity and so on. In order to deal with all those hurdles, an improved SLIC with statistical region merging (SRM) is proposed. With ISLIC, super pixels are generated and the foreground extraction is carried out by SRM. It is then followed by smoothing the image using Gaussian variables, thus the existence of noise will be reduced. Accordingly, the depth map is estimated and finally with DIBR, the 2D texture image is successfully constructed as an effective 3D image with improvised quality by the configuration formula combining the depth map with the left and right views of the texture image. Furthermore, the time complexity issues faced with the prior methods are reduced here with DIBR. Though the speed and power of the transformation system has to be enhanced. Hence in future, the work is extended with some other advanced technologies so as to enhance the speed and power of the transformation system.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

## REFERENCES

[1] Karatas, OrhanHakki, and Ebubekir Toy, "Three-dimensional imaging techniques: A literature review," European journal of dentistry, vol. 8, no. 1, pp. 132, 2014.

[2]Udupa, K. Jayaram, "Three-dimensional visualization and analysis methodologies: a current perspective," Radiographics,vol. 19, no. 3, pp. 783-806, 1999.

[3] Mathai, Shelmy, P. Paul Mathai, and K. A. Divya, "Automatic 2D to 3D video and image conversion based on global depth map,"Computational Intelligence and Computing Research (ICCIC), 2015 IEEE International Conference on, IEEE, 2015.

[4] Yamada, Kunio, and Yasunari Suzuki, "Real-time 2D-to-3D conversion at full HD 1080P resolution, Consumer Electronics, 2009," ISCE'09.IEEE 13th International Symposium on, IEEE, 2009.

[5] Konrad, Janusz, et al., "Learning-based, automatic 2D-to-3D image and video conversion," IEEE Transactions on Image Processing, vol. 22, no. 9, pp. 3485-3496.

[6] 3D Modeling Visualization Algorithm Research and Implementation

[7] D. Vetriselvi and Mr L. Megalan Leo, "2D TO 3D CONVERSION OF DENTAL IMAGES: SURVEY," International Journal of Applied Engineering Research, vol. 10, no. 20, 2015.

[8] Battiato, Sebastiano, et al., "3D stereoscopic image pairs by depth-map generation," 3D Data Processing, Visualization and Transmission, 3DPVT 2004.Proceedings.2nd International Symposium on, IEEE, 2004.

[9] Swash, Mohammad Rafiq, et al., "Pre-processing of holoscopic 3D image for autostereoscopic 3D displays," 3D Imaging (IC3D), 2013 International Conference on, IEEE, 2013.

[10] Uher, Vaclav, and Radim Burget, "Automatic 3D segmentation of human brain images using data-mining techniques," Telecommunications and Signal Processing (TSP), 2012 35th International Conference on, IEEE, 2012.

[11] Kumar, Gaurav, and Pradeep Kumar Bhatia, "A detailed review of feature extraction in image processing systems," Advanced Computing & Communication Technologies (ACCT), 2014 Fourth International Conference on, IEEE,2014.

[12] Haralick, M. Robert and KarthikeyanShanmugam, "Textural features for image classification," IEEE Transactions on systems, man, and cybernetics, vol. 6, pp. 610-621, 1973.

[13] Gu, Ke, et al., "Model-based referenceless quality metric of 3D synthesized images using local image description," IEEE Transactions on Image Processing,vol. 27, no. 1, pp. 394-405, 2018.

[14] Zhu, Jiasong, et al., "Multiple 3-D Feature Fusion Framework for Hyperspectral Image Classification," IEEE Transactions on Geoscience and Remote Sensing, 2018.

[15] Chen, Xiaozhi, et al., "3D object proposals using stereo imagery for accurate object class detection," IEEE transactions on pattern analysis and machine intelligence, vol. 40, no. 5, pp. 1259-1272, 2018.

[16] Xu, Ming, and Zeyun Yu, "3D image segmentation based on feature-sensitive and adaptive tetrahedral meshes, Image Processing (ICIP)," 2016 IEEE International Conference on, IEEE, 2016.

[17] Yang, Yizhong, Xionglou Hu, Nengju Wu, Pengfei Wang, Dong Xu, and Shen Rong. "A depth map generation algorithm based on saliency detection for 2D to 3D conversion." 3D Research 8, no. 3 (2017): 29.