



Reconstruction of Simple and Complex Three Dimensional Images Using Pattern Recognition Algorithm

Shilpa Rani

Ph.D. Scholar, Lovely Professional University, Punjab; Assistant Professor, Department of CSE, Neil Gogte Institute of Technology, Hyderabad, Telangana, India. E-mail: shilpachoudhary2020@gmail.com

Deepika Ghai

Assistant Professor, Lovely Professional University, Punjab, India. E-mail: deepika.21507@lpu.co.in

Sandeep Kumar

Professor, Department of CSE, Koneru Lakshmaiah Education Foundation, A.P., India. E-mail: er.sandeepsahratia@gmail.com

Abstract

The reconstruction of 3D images is always a difficult task for the researchers. The 3D reconstruction of the image is a core technique of various fields such as Computer graphics, computer vision, CAD systems, medical science, computer application, etc. Reconstruction of the 3D image allows us to gather the quantitative features of the objects such as the shape, size, and volume of the objects. The existing computer algorithms need spatial dimension information to make the distinguished inference from the given 3D image which is not always possible. This paper simplifies the 3D reconstruction of the image. This research paper introduced a novel algorithm for the representation of the Three Dimensional images into a textual form. The syntactic approach is used for the extraction of the features of the image and these are called knowledge vectors. The knowledge vector consists of the direction information and length information. This a new approach in the field of image processing where images can be represented as a knowledge vector and it could be a great contribution in the field where security is a major concern. Further, the knowledge vector is used for the reconstruction of the 3D image. The performance of the algorithm is evaluated on the PASCAL 3D + and example-based Synthesis of the 3D Object Arrangements dataset. According to the obtained results, the proposed methodology is having better accuracy, and the processing time of reconstruction of the original 3D image is 1.02 Seconds. Single-pass is sufficient for reconstructing the original image.

Keywords: Syntactic approach; Construction; Reconstruction; 3D images.

Introduction

The processing of Three Dimensional images is an important area of research in the field of image processing. Three-dimensional images always give more information compared to two-dimensional images because of depth information. It's always a difficult task to obtain a 3D image of a particular object. Therefore most of the researchers are working on the reconstruction of the 3D model using 2D images. (V. Blanz et al. 1999) proposed a photometric stereo method in which a minimum of two images would be taken under different lighting conditions to get the depth information. This is only possible when a large amount of dataset is available. But the computation cost of these methods is very high in terms of time complexity therefore it is not a popular approach for the reconstruction of 3D images. (B. K. Horn 1989) proposed a method with the name 'shape and shading' which requires only a single image for the reconstruction of a 3D image. (I. Kemelmacher-Shlizerman 2011) worked on the concept that objects of similar types consist of common characteristics and features. In this approach, only reference depth information is required other than the original image to get the depth information from the image. These above-mentioned methods have their own pros and cons and these methods are based on the statistical approach. Statistical approaches are a good alternative if images are blurred or noisy but this approach could fail if the image is more complex and containing complex patterns. Therefore, for complex images, structural-based methods are a good choice. Taking consideration into that, this research paper introduced a novel algorithm that is based on structural pattern recognition techniques (T.Pavlidis 1977, Fu et al. 1974, J. Gips 1975). Most of the pattern recognition techniques will follow some basic steps: preprocessing, feature extraction, and finally classification which is represented in the following Figure. 1.



Figure 1. General steps of Pattern recognition techniques.

Feature extraction is always an important step of any pattern recognition technique and there are different feature extraction techniques available in both statistical and structural pattern recognition. The proposed methodology is based on structural pattern recognition and we worked on representing the features of the image into a textual form which is a novel approach. The syntactic approach is used for the extraction of the features of the image and these are called knowledge vectors. Knowledge vector is nothing but it is a collection of strings and it consists of direction and length information. This is a new approach in the field of image processing where an image can be represented as a knowledge vector and further this knowledge vector is used as an input for the reconstruction of the input 3D image and a detailed description of the proposed method is explained further.

Objective and Motivation of the research

Structural pattern recognition is a less explored by the researchers but this method is more suitable for understanding the complex pattern of the images and able to retrieve more specific information about the objects which are present in to the image. The main objective of this research is to understand the image in textual form and further reconstruct the original image using textual information. The proposed method is also provides the various statistics of the original and reconstructed image and this statistics could be helpful for further findings.

Further, this paper is divided into the following section: section 2 describes the literature survey, section 3 describes the proposed methodology, section 4 describes the Dataset and its parameter and section 4 represents the results and section 5 represents the conclusion and future work.

Literature Survey

The traditional approach of 3D reconstruction requires multiple views of images. Reconstruction of 3D images using a single image is more appropriate comparatively and it is divided into two categories: Voxel-based and point cloud based reconstruction. Both methods are well explored by different researchers. (A. Dai et al. 2017) proposed a shape-based method for reconstruction. The authors had combined the 3d deep learning architecture with the 3D shape synthesis method. The deep learning-based method can infer the global structure of the image. (Tulsiani et al. 2018) proposed a method based on the neural network which considered the multiple views of the same image. (Choy et al. 2016) proposed a 3D reconstruction method using RNN which learned using 2D images and later able to construct the 3D images and experiments were performed on synthesis dataset. (S. Tulsiani et al. 2017) proposed a method based on the projection of 2D images and produced a 3D reconstructed image. (G. Riegler et al. 2017) worked on the two different defects of voxels which are sparse information and complexity of computation. Authors had achieved better results compared to existing methods but still, there could be a scope of improvement. In the same way, many authors explored the reconstruction of a 3D image using point cloud-based methods. (Fan et al. 2017) was the first author who proposed the concept of point cloud-based 3D reconstruction. In this CD (Chamfer distance) and EMD (Earth Mover's distance) was used as a loss function to train the model. (P. Mandikal et al. 2018) proposed a method with the name 3D-PSRNet which propagated the segmentation information to other task and used the Chamfer distance and location-based segmentation as loss function. (L. Jiang et al. 2018) proposed geometric loss and conditional loss components based on geometric adverse loss. The geometric loss component ensures that the shape of the 3D reconstructed image is close to the original image and conditional adversarial is responsible for generating meaningful point clouds. (K. L. Navaneet et al. 2018) proposed 3D LMNet. In this also authors used the Chamfer loss function for the training of encoder and later used diversity loss function and latent matching

loss function for the mapping the vector of an autoencoder which solved the uncertainty of 2D inputs. (Chen Zhang et al. 2019) proposed a volumetric based method in which objects are refereed as a cuboids. This method is able to preserve the smoothness and cleanliness of the objects while reconstruction and able to reconstruct the hidden part which are back to the surface. (Bin Li et al.2020) proposed an algorithm with the name 3D-ReConstnet which creates an end to end 3D reconstruction network. This method used the concept of residual network for extracting the features from the image and Gaussian probability distribution used for the prediction of the point clouds. This method is able to determine the 3D output from the 2D images and also able to reconstruct the 3D image even if the objects are self-occluded. Results of this method are satisfactory compared to existing methods.

Most of the Researchers worked on 3D reconstruction techniques using machine learning-based methods and every method have their pros and cons. considering this; inthe proposed methodology we used the syntactic approach for the reconstruction of the 3D image which could improve the 3D reconstruction based methods.

Methodology

In the above section, we discussed the available approach and their problems for the construction and reconstruction of the 3D images. Considering the problem of existing techniques we introduced a new methodology for the construction & reconstruction of the Three Dimensional images. The proposed methodology is working on the construction and reconstruction of the 3D images using the feature vector. The working diagram of the proposed methodology has given in Figure 2 and basically, it is divided into 3 parts and it is explained further:

Knowledgeacquisition

The first step is knowledge acquisition in which knowledge of extracting the features of a specific domain or application is acquired. This is the basic step of structural pattern recognition. Knowledge techniques are used to gather knowledge about the dataset and the environment in which the dataset is generated and it produces the knowledge base which contains a detailed description of the data. In the second step domain/ application-specific feature extraction method is developed using the dataset and knowledge base.

Feature Extraction

For extracting the features from the image, a syntactic approach is used in which the components of the image can be represented as a string. Features are extracted in such a way that it can reconstruct the original 3D image. For the extraction of the features of the image, different directions are explored in which a pixel could be present. For the scanning of the image for feature extraction, a 3X3X3 size of the sliding window is used and it is represented in Figure 3. The sliding window will start the tracing of the image from the initial pixel and continue until it will not scan the whole image. The preferred direction of scanning the image is the Right

direction(R) and further, the algorithm will scan the image in all possible directions. In a window of 3X3X3, there would be possibly 26 directions where a neighborhood pixel could be present. The scanning of the image continues until it will not scan the entire neighborhood pixel from the current position and at the same time it will remove the pixel which is already scanned. If further no pixel is present then the algorithm will display the knowledge vector of that object and start scanning the other object if it is present in the image. The algorithm for extracting the feature vector from the image is given below.

Algorithm 1

Step 1: Start

Step 2: Initialize a 3X3X3 window.

Step 3: Use the 3X3X3 window for scanning of the image and start tracing the image from the starting position and find the initial pixel of the image (x_i, y_i, z_i) which could be the initial pixel of the first component of the image.

Step 4: Trace the image in all possible directions to find the next connected pixel.

Step 5: Repeat step 4 and simultaneously remove the scanned pixel until there exists any neighborhood pixel in any possible directions.

Step 6: Print the feature vector which consists of the initial pixel position, the respective directions, and their length, the end pixel position.

Step 7: Go back to (x_i, y_{i+1}, z_i) and repeat step 4 to step 6.

Step 8: Display the final feature vector of all the objects which are present in the image.

It is possible that any of the neighborhood pixels is not present in the image. Therefore, it is possible that maybe one or two or more or the number of pixels is not present in the image. Considering this, these possibilities are divided into 8 different groups and this is represented in Table 1.

Table 1. The 8 Possible groups

Groups	Number of pixels Eliminated
Group P	Nil
Group Q	One of the Corner Pixels is eliminated.
Group R	Any two corner pixels are eliminated.
Group S	Any three pixels are eliminated.
Group T	Any four pixels are eliminated.
Group U	Any five pixels are eliminated.
Group V	Any Six pixels are eliminated.
Group W	Any Seven pixels are eliminated.
Group X	All corner pixels are eliminated.

Reconstruction of the original Image using feature vector

In the previous step, the knowledge vector of the image is identified which is nothing but features of the image. In this step, the feature vector is considered as an input and a novel algorithm is written which converts the knowledge vector into the image.

Algorithm 2

Step 1: Read the input knowledge vector of the image.

Step 2: Identify the starting pixel of the first object using the direction and length code of the feature vector.

Step 3: Read the present pixels in all the directions and set the pixel value as one.

Step 4: Repeat step 2 & 3 until it will not read all the objects of the image.

Step 5: Display the original image.

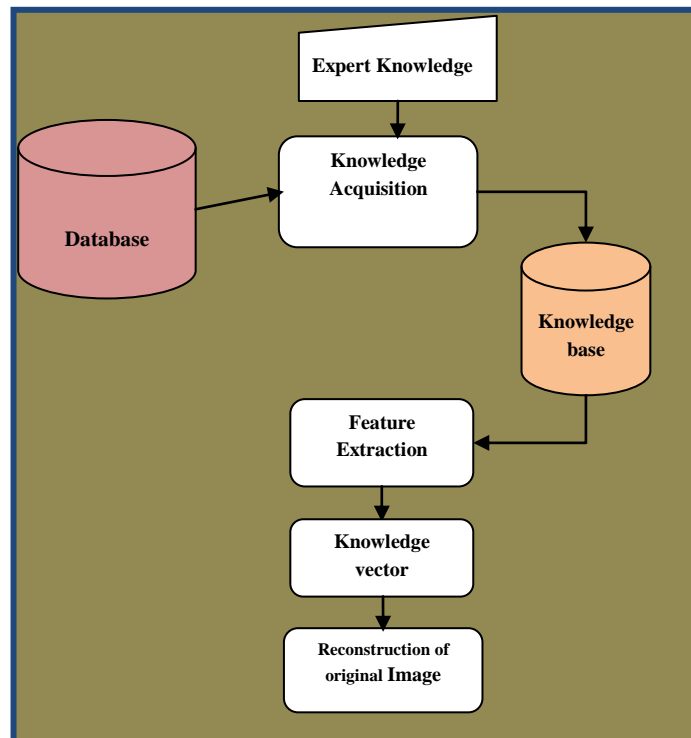


Figure 2. Flow Chart of Proposed Methodology

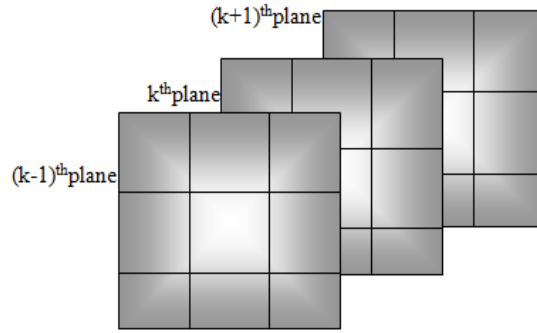


Figure 3. Empty scanning window with size 3X3X3.

Experiments setup and its parameters

Two algorithms are implemented for the proposed work and both algorithms are implemented in python using the Anaconda tool. The experiments are done on an Acer laptop, 4GB RAM, 2GB NVIDIA Graphics card, 5 core processors, Windows operating system, and only one core is used for the computation.

Datasets

The efficiency of the algorithm is tested on two standard datasets. The datasets are the 3D Object Arrangements dataset and the PASCAL 3D+ dataset. An example-based 3D object arrangement dataset is a collection of different scenes, models, and textures. PASCAL 3D+ is a standard dataset for 3D object reconstruction and it consists of more than 3000 instances of each category.

Evaluation Parameters

To evaluate the performance of the proposed methodology some evaluation parameters are considered. The First parameter is Accuracy which means how accurately the algorithm can find the feature vector of the image and at the same time how accurately the algorithm can reconstruct the 3D image using knowledge vector.

The second parameter is Processing time it means the time which is required for the construction of the knowledge vector and reconstruction of the 3D image using the knowledge vector. Processing time can be calculated using the following formula:

$$P.T = \frac{1}{\text{Performance of the processing unit}} \quad (1)$$

The third evaluation parameter is the number of passes which is required for reconstructing the original image using a knowledge vector.

The other evaluation parameters are Mean and standard deviation. Mean deviation is a statistical measurement that is used to find the difference between the actual value and the expected value. The formula for mean deviation is as follows:

$$\mu = \frac{\sum_{i=1}^n x_i}{n} \quad (2)$$

μ = Mean of the Sample

n = number of Samples

x_i = Each value of the Sample

Standard deviation is used to measure the dispersion of the sample dataset from its mean value. The formula for Standard deviation is as follows:

$$S.D = \frac{\sqrt{\sum (x_i - \mu)^2}}{N} \quad (3)$$

Here

S.D = Standard Deviation

x_i = Each value from the Sample

μ = the mean of Sample

N = Size of the Sample

Results

Performance analysis of the proposed algorithm has been done using the mentioned evaluation parameters and the comparative analysis have been done on a collection of simple 3D images, PASCAL 3D+ Dataset, and an Example-based 3D object arrangement dataset and the results are discussed below:

Accuracy

Accuracy is measured in terms of checking that how accurately a knowledge vector can reconstruct the original image and results are displayed in Table 2.

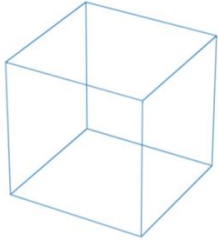
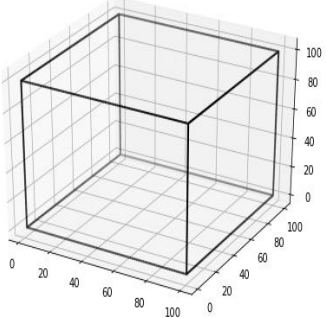
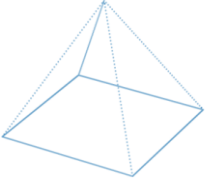
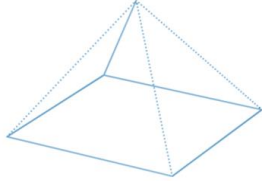
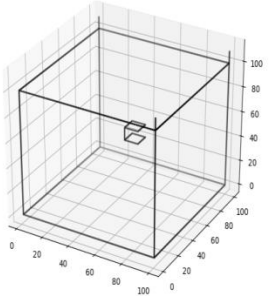
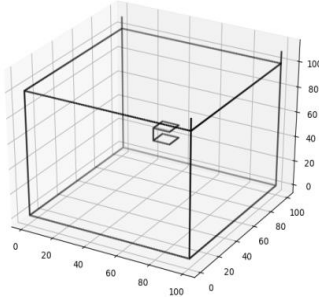
Mean and Standard Deviation

From each dataset randomly 200 images were selected for the testing of the algorithm. Minimum value, Maximum value, Total count, Mean, median, and standard deviation of the RGB components are calculated. Luminance and saturation minimum, maximum, mean, median, and standard deviation is also calculated. All the above statistics for sample images are represented in Figure4, 5,6 &7.

Processing Time and Number of Passes

The processing time of reconstruction of the original 3D image is 1.02 Seconds and the single pass is sufficient for reconstructing the original image.

Table 2: Original Image, Knowledge vector, and reconstructed images.

	<p> $\langle 1,1,1 \rangle R \square D \square L \square U \square BU \square B \square R \square D$ $\square L \square U \square *99 \square 99 \square 99 \square 98 \square 1 \square 98 \square 99$ $\square 99 \square 99 \square 98 \square \langle 2,1,100 \rangle \langle 1,100,2 \rangle$ $B \square *97 \square \langle 1,100,99 \rangle \langle 100,1,2 \rangle B \square *$ $97 \square \langle 100,1,99 \rangle \langle 100,100,2 \rangle B \square *97$ $\square \langle 100,100,99 \rangle$ </p>	
	<p> $\langle 1,1,1 \rangle R \square D \square L \square U \square BR \square BDR \square FD$ $R \square *104 \square 104 \square 104 \square 103 \square 1 \square 51 \square 51$ $\square \langle 104,104,2 \rangle \langle 2,104,2 \rangle BDL \square *50$ $\square \langle 52,54,52 \rangle \langle 104,2,2 \rangle BUR \square *50 \square$ $\langle 54,52,52 \rangle$ </p>	
	<p> $\langle 1,1,1 \rangle R \square D \square L \square U \square BU \square B \square R \square D$ $\square L \square U \square *99 \square 99 \square 99 \square 98 \square 1 \square 98 \square 99$ $\square 99 \square 99 \square 98 \square \langle 2,1,100 \rangle \langle 1,100,2 \rangle$ $B \square *107 \square \langle 1,100,99 \rangle \langle 100,1,2 \rangle B \square *$ $*107 \square \langle 100,1,99 \rangle \langle 100,100,2 \rangle B \square *$ $107 \square \langle 100,100,99 \rangle \langle 50,50,50 \rangle R \square D$ $\square L \square U \square BU \square B \square R \square D \square L \square U \square *10$ $10 \square 10 \square 9 \square 1 \square 9 \square 10 \square 10 \square 9 \square \langle 1$ $00,100,99 \rangle$ </p>	

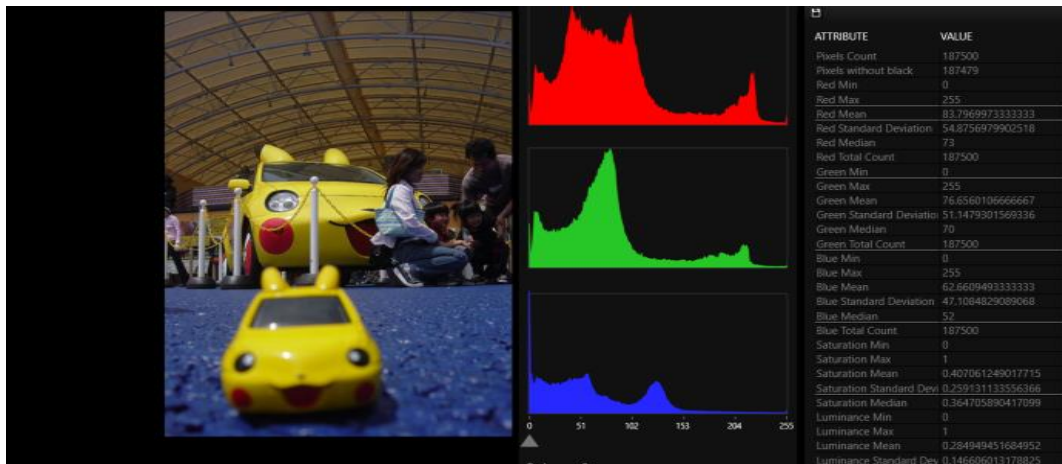


Figure 4. Original image and its corresponding histogram, Image statistics of PASCAL 3D + Dataset.

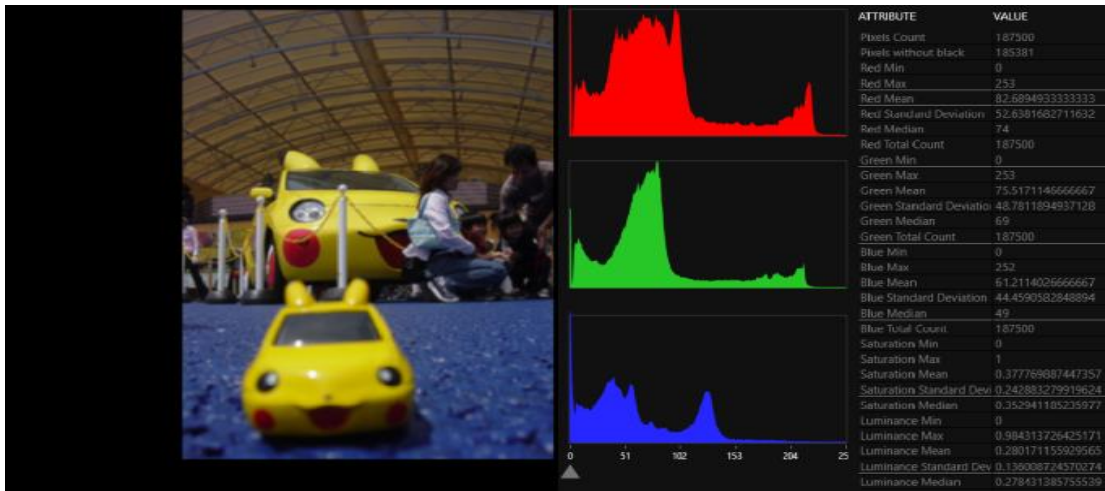


Figure 5. Reconstructed image and its corresponding histogram, Image statistics PASCAL 3D + Dataset.



Figure 6. Original image and its corresponding histogram, Image statistics Example based 3D object arrangement Dataset.



Figure 7: Reconstructed image and its corresponding histogram, Image statistics of Example-based 3D object arrangement Dataset.

Discussion

This research is carried out for adding a new method in the field of structural pattern recognition. This technique allows representing the simple and complex patterns which is not possible in the statistical pattern recognition. Algorithm is working pretty fine even if more than one objects present in the image. For the evaluation of the proposed algorithm various evaluation parameters are used and which is discussed in the above section. The best part of this algorithm is processing time which is 1.02 Sec. Processing time is the time which is required for reconstructing the original image from the knowledge vector and single pass is sufficient for the reconstruction.

Conclusion & Future scope

In this research paper, we introduced a methodology for representing the image into a textual form and the reconstruction of the original image using the knowledge vector. Feature extraction and representation have been done using a syntactic approach and the same string is used as input for the reconstruction of the input image. Experiments are performed on 3D shapes as well as complex 3D images and results are outperformed. The size of the knowledge vector is large comparatively so the knowledge vector of complex images is not displayed in the result section. In the future, we can try to reduce the size of the knowledge vector, and classification techniques can be used for 3D object recognition.

Conflict of interest

The authors declare no potential conflict of interest regarding the publication of this work. In addition, the ethical issues including plagiarism, informed consent, misconduct, data fabrication and, or falsification, double publication and, or submission, and redundancy have been completely witnessed by the authors.

Funding

This research did not receive any specific grant from agencies in the public, commercial or not for profit sectors

References

- Blanz, V., & Vetter, T. (1999, July). A morphable model for the synthesis of 3D faces. *Proceedings of the 26th annual conference on Computer graphics and interactive techniques* (pp. 187-194).
- Horn, B. K. (1989). Obtaining shape from shading information.
- Kemelmacher-Shlizerman, I., & Basri, R. (2010). 3D face reconstruction from a single image using a single reference face shape. *IEEE transactions on pattern analysis and machine intelligence*, 33(2), 394-405.
- T.Pavlidis(1977). Structural Pattern Recognition. New York: Springer-Verlag.
- Fu, King Sun (1974). Syntactic methods in pattern recognition. Elsevier.

- J.Gips(1975). Shape Grammars and their Uses. Birkhauser: Verlag, Basel and Stuttgart.
- Dai, A., Ruizhongtai Qi, C., & Nießner, M. (2017). Shape completion using 3d-encoder-predictor cnns and shape synthesis. *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 5868-5877).
- Tulsiani, S., Efros, A. A., & Malik, J. (2018). Multi-view consistency as supervisory signal for learning shape and pose prediction. *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2897-2905).
- Choy, C. B., Xu, D., Gwak, J., Chen, K., & Savarese, S. (2016, October). 3d-r2n2: A unified approach for single and multi-view 3d object reconstruction. *European conference on computer vision, Springer, Cham* (pp. 628-644).
- Riegler, G., Osman Ulusoy, A., & Geiger, A. (2017). Octnet: Learning deep 3d representations at high resolutions. *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3577-3586).
- Fan, H., Su, H., & Guibas, L. J. (2017). A point set generation network for 3d object reconstruction from a single image,” in *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 605-613).
- Kumar, S., Singh, S., & Kumar, J. (2019). Multiple face detection using hybrid features with SVM classifier. *Data and Communication Networks*, Springer, Singapore (pp. 253-265).
- Kumar, S., Singh, S., & Kumar, J. (2021). Face spoofing detection using improved Seg-Net architecture with a blur estimation technique. *International Journal of Biometrics*, 13(2-3), (pp.131-149).
- Kumar, S., Singh, S., & Kumar, J. (2018). Live detection of face using machine learning with multi-feature method. *Wireless Personal Communications*, 103(3), (pp.2353-2375).
- Kumar, S., Singh, S., & Kumar, J. (2019, January). Gender classification using machine learning with multi-feature method. *IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)* (pp. 0648-0653).
- Mandikal, P., KL, N., & Venkatesh Babu, R. (2018). 3D-psrnet: Part segmented 3d point cloud reconstruction from a single image. *Proceedings of the European Conference on Computer Vision (ECCV) Workshops* (pp. 0-5).
- Rani, S., Ghai, D., & Kumar, S. (2021). Construction and Reconstruction of 3D Facial and Wireframe Model Using Syntactic Pattern Recognition. *Cognitive Behavior and Human Computer Interaction Based on Machine Learning Algorithm*, (pp-137-156).
- Zhang, C. (2019). CuFusion2: Accurate and de-noised volumetric 3D object reconstruction using depth camera. *IEEE Access*, 7, (pp.49882-49893).
- Li, B., Zhang, Y., Zhao, B., & Shao, H. (2020). 3D-ReConstnet: a single-view 3d-object point cloud reconstruction network. *IEEE Access*, 8, (pp-83782-83790).
- Rani, S., Lakhwani, K., & Kumar, S. (2020, December). Three dimensional wireframe model of medical and complex images using cellular logic array processing techniques. *International Conference on Soft Computing and Pattern Recognition, Springer, Cham* (pp. 196-207).
- Jain, A., Gahlot, A.K., Dwivedi, R., Kumar, A. and Sharma, S.K., (2018). Fat Tree NoC Design and Synthesis, in *Intelligent Communication, Control and Devices* (pp. 1749-1756).
- Sarita, K., & Choudhary, S. (2012). An improved BPCS image steganography in integer wavelet transform domain using 4-4 block size. *Int. J. Eng. Res. Technology*, 1, (pp.1-8).
- Jain, A. and Kumar, A., 2021. Desmogging of still smoggy images using a novel channel prior. *Journal of*

Ambient Intelligence and Humanized Computing, 12(1), (pp.1161-1177).

Ghai, D., Gianey, H. K., Jain, A., & Uppal, R. S. (2020). Quantum and dual-tree complex wavelet transform-based image watermarking. *International Journal of Modern Physics B*, 34(04), 2050009.

Choudhary, S., Lakhwani, K., & Agrwal, S. (2012). An efficient hybrid technique of feature extraction for facial expression recognition using AdaBoost Classifier. *International Journal of Engineering Research & Technology*, 8(1).

Kumar, Sukhwinder Singh and Jagdish Kumar(2018). Automatic Live Facial Expression Detection Using Genetic Algorithm with Haar Wavelet Features and SVM. *Wireless Personal Communication*, Vol. 103, No. 3, (pp. 2435-2453).

Rohit Raja, S. Kumar, and Rashid Mahomood(2020). Color Object Detection Based Image Retrieval using ROI Segmentation with Multi-Feature Method. *Wireless Personal Communication*, (pp. 1-24).

Kumar, S., Jain, A., Shukla, A. P., Singh, S., Raja, R., Rani, S., & Masud, M. (2021). A Comparative Analysis of Machine Learning Algorithms for Detection of Organic and Nonorganic Cotton Diseases. *Mathematical Problems in Engineering*.

Raja, R., Kumar, S., Rani, S., & Laxmi, K. R. (2020). Lung Segmentation and Nodule Detection in 3D Medical Images Using Convolution Neural Network. *Artificial Intelligence and Machine Learning in 2D/3D Medical Image Processing*, CRC Press(pp. 179-188).

Kumar S, V.TajKiran, SekuriSwetha, and Prashant Johri(2018). IoT based Smart Home Surveillance and Automation. *IEEE International Conference on Computing, Power and Communication Technology (GUCON)*, (pp. 795-799).

Bibliographic information of this paper for citing:

Rani, Shilpa; Ghai, Deepika & Kumar, Sandeep (2022).Reconstruction of Simple and Complex Three Dimensional Images Using Pattern Recognition Algorithm *Journal of Information Technology Management*, Special Issue, 235-247.

Copyright © 2022, Shilpa Rani, Deepika Ghai and Sandeep Kumar

