НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ
сентябрь–октябрь 2024 Том 24 № 5 http://ntv.ifmo.ru/
SCIENTIFIC AND TECHNICAL JOURNAL OF INFORMATION TECHNOLOGIES, MECHANICS AND OPTICS
September–October 2024 Vol. 24 No 5 http://ntv.ifmo.ru/en/
ISSN 2226-1494 (print) ISSN 2500-0373 (online)

doi: 10.17586/2226-1494-2024-24-5-843-848

Single images 3D reconstruction by a binary classifier

Sallama Adhab Resen✉

Directorate General of Vocational Education, Baghdad, 1001, Iraq
salamaresen@gmail.com✉, https://orcid.org/0009-0007-5044-8857

Abstract
Intelligent systems demand interaction with a variety of complex environments. For example, a robot might need to interact with complicated geometric structures in an environment. Accurate geometric reasoning is required to define the objects navigating the scene properly. 3D reconstruction is a complex problem that requires massive amounts of images. The paper proposes producing intelligent systems for 3D reconstruction from single 2D images. Propose a learnable reconstruction context that uses features to realize the synthesis. Proposed methods produce encoding feature lable input to classification, pulling out that information to make better decisions. Binary Classifier Neural Network (BCNN) classifies whether a point is inside or outside the object. The reconstruction system models an object 3D structure and learns feature filter parameters. The geometry and the corresponding features are implicitly updated based on the loss function. The training doesn't require compressed supervision to visualize the task of reconstructed shapes and texture transfer. A point-set network flow results in BCNN having a comparable low memory footprint and is not restricted to specific classes for which templates are available. Accuracy measurements show that the model can extend the occupancy encoder by the generative model, which doesn't request an image condition but can be trained unconditionally. The time required to train the model will have more neurons and weight parameters overfitting.

Keywords
intelligent systems, 3D reconstruction, features filter, convolution neural networks, Binary Classifier Neural Network (BCNN)

For citation: Resen S.A. Single images 3D reconstruction by a binary classifier. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2024, vol. 24, no. 5, pp. 843–848. doi: 10.17586/2226-1494-2024-24-5-843-848

УДК 004.855

Трехмерная реконструкция отдельных изображений с помощью бинарного классификатора

Саллама Адхаб Ресен✉

Главное управление профессионального образования, Багдад, 1001, Ирак
salamaresen@gmail.com✉, https://orcid.org/0009-0007-5044-8857

Аннотация
Интеллектуальные системы требуют взаимодействия с различными сложными окружающими средами. Например, роботу может потребоваться взаимодействовать в обстановке со сложными геометрическими структурами. Для правильного определения объектов, перемещающихся в пространстве, требуется точное геометрическое обоснование. 3D-реконструкция — сложная задача, требующая большого количества изображений. В работе предлагается создание интеллектуальных систем для 3D-реконструкции из отдельных 2D-изображений. Разработан обучаемый контекст реконструкции, который для реализации синтеза использует определенные признаки. Используемые методы осуществляют кодирование признаков метки входных данных для классификации, извлекая эту информацию для принятия более обоснованных решений. Бинарная сверточная нейронная сеть (Binary Classifier Neural Network, BCNN) классифицирует, находится ли точка внутри или снаружи объекта. Система реконструкции моделирует 3D-структуру объекта и изучает параметры фильтра признаков. Геометрия и соответствующие признаки обновляются на основе функции потерь. Обучение модели не требует сжатого наблюдения для визуализации задачи реконструированных форм и переноса текстуры. Поток сети с множеством точек приводит к тому, что BCNN занимает сравнительно малый объем памяти

© Resen S.A., 2024

и не ограничивается определенными классами, для которых доступны шаблоны. Исследование точности метрики показали, что модель может расширить кодировщик занятости с помощью генеративной модели, которая не запрашивает условие получения изображения и может быть обучена безусловно. Таким образом, за время, необходимое для обучения модели, создается большее количество нейронов и весовых переобученных параметров.

Ключевые слова
интеллектуальные системы, 3D-реконструкция, фильтр признаков, сверточные нейронные сети, двоичный классификатор нейронных сетей

Ссылка для цитирования: Ресен С.А. Трехмерная реконструкция отдельных изображений с помощью бинарного классификатора // Научно-технический вестник информационных технологий, механики и оптики. 2024. Т. 24, № 5. С. 843–848 (на англ. яз.). doi: 10.17586/2226-1494-2024-24-5-843-848

Introduction

3D reconstruction is a challenging problem and has been discussed by many researchers and articles. A traditional 3D reconstruction pipeline takes input images. First, the camera poses for each image are computed using structure from motion or bundle adjustment, and then acute correspondences are computed across [1]. Many techniques represent output predicted by a deep neural network. 3D reconstruction for further optimization uses point representation, mesh refinement, or volumetric techniques. Geometric 3D representation illustrated in Fig. 1.

Voxel representations have been proposed for the reconstruction task as they are easy to process with neural networks. With voxel representation limitations, memory grows cubically in three dimensions. Outputs process at limited spatial resolution; the chosen coordinate system defines the voxels [2]. Voxel representations discretize the 3D space into regular grid cells, 2D and 3D. Mesh output representations for 3D reconstruction use deep neural networks. Meshes discretize space into vertices and faces but are still limited in the number of vertices. Meshes are patch-based approaches to self-intersections and non-watertight meshes. Mesh predictions by the neural network need help for output representation, where mesh needs a specific template to deform a mesh model instead of an explicit output representation. Point sets are another representation recently considered as output for neural networks. Point sets discretize the object's surface into 3D points to model the connectivity. Existing approaches could be more extensive in several points that can be processed. Local shape information is hard to encode; thus, point set-generating networks involve global shape encoding. The

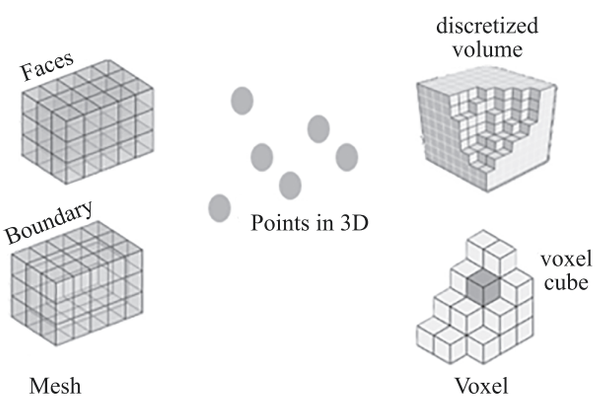

Fig. 1. Geometric 3D representation

images have local and global information, but the essential local features are accurate. To examine further, add more details to the regional features. The idea of using the local and global features was investigated. The paper proposes an implicit output representation that does not require any discretization. It can parallel model arbitrary topology and arbitrary resolution on the Graphics Processing Units (GPU) and input as many points as fit into GPU memory. The method is not restricted to specific categories for available templates with comparable low memory; instead of having a deeper, fully connected network, proposed 3D convolutions are similar to convolutions on the 2D image space. The 2D convolutional network operates on the image domain. The 3D points query the image depending on the projection of the 3D point that falls into that image. Reconstructions were obtained for both geometry and feature representation. Models produce accurate reconstructions for huge spaces.

Literature review

Remarkable progress has been made in 3D reconstruction, mainly credited to the rise of neural implicit modeling and advancements in differentiable rendering. Learning systems can learn dense, high-fidelity 3D from multi-view camera images. 3D reconstruction from a single image is challenging because it involves exploitation, visualization, and extracting information from images. The classical explicit representations differentiable rendering algorithms allow us to learn representations from 2D supervision only from RGB images [3]. Define a differentiable rendering algorithm architecture that combines geometry and appearance prediction [4]. The image encoder takes a 2D image and produces a global latent code for that image [5]. In [6], the forward and backward paths of the network are defined. The forward path corresponds to the rendering path [7]. The model can handle geometric details for entire scenes for single objects. The model leads to fewer occlusion artifacts than previous novel view synthesis baselines [8] which presented representations from 2D images, encode the input 3D shape using a point encoding into a global shape encoding and render the 3D shape into a depth map. 3D surface un-projection through the camera matrix was made for every pixel corresponding to the depth map [9]. Network architecture for predicting occupants has an encoder [10, 11]. The encoder depends on the condition of either a 2D image or a 3D point and on rough voxelization [12]. The encoder produces a set of conditioned layers.

These layers are conditioned on encoding and finding the 3D point location [13]. The shape encodes and predicts an RGB color value. Build generative versions of the model using adversarial loss encoder images. Single image texture reconstruction has the 2D input image and a 2D novel view synthesis baseline results. Synthesis baseline is hard in 2D to 3D translation. Building a model has investigated the representation power of the model reconstruction loss gain encoder [14, 15].

The predicted Neural Network (NN) model takes a point and passes it to a given image condition where the dots are the centers of the voxels [16]. 3D point queries the features using the neighbor interpolation technique. Integrate multiple views and get more precise results on the back of the model [17]. The work [18] combined the idea of implicit representations with primitive-based shape models. From [19] they proposed the universal differentiable renderer for NN representations. The work is considered the normal light location, and the 3D point matches the texture fields for the color value. Pixel- and feature-based reconstructions are performed in low and high-frequency domains. The fused view produced high-quality sentences. The paper found an analytic solution in comparison to a connected network. Instead of representing the 3D shape explicitly, consider the object implicitly as the decision boundary of a Binary Classifier Neural Network (BCNN) classifier. First, the NN geometry must be extracted in a post-processing step which consumes time. The extension to 3D is complex due to dimensionality. The 3D network operates on a fully connected network that has interpolated the features from the 2D image. 3D convolutions have to generate a 3D feature volume. The feature gets as input and predicts 3D object location based on authentic images.

Single image 3D reconstruction

The framework covers two main parts: the learned feature module and 3D pixel reconstruction. The points set the generation network and detect occupancy results. Fig. 2 shows the framework stages. We want to know how to extract information from collected images over time and how to represent the surface and change in 3D reconstruction. The classifier network loses a lot of geometrical detail through occupancy flow, suffering from dimensionality. Areas outlined in the boundary contained

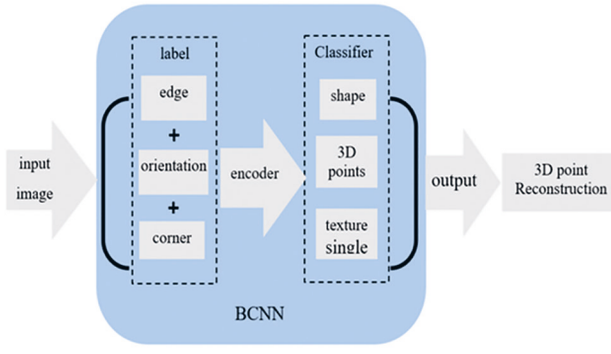

Fig. 2. Proposed stages

on the surface have a significant effect on allocations. The method assumes two hypotheses to produce more autonomous, robust, and safe results.

Stage 1. Relying on local information puts certain constraints on the encoding image. Investigating regional and global features to label them and the global context assisted in representing 3D.

Stage 2. The data set includes different variations in the single image where a features filter is applied to the labeled image. The BCNN network was tiny to determine whether this small patch shows evidence of possibility. Fig. 2 shows the proposed stages. Stage 1 learns features to predicate the shape of an object from a given input image. In stage 2, the establishment field is discovered using the already trained modules in this stage.

In training, the system doesn't require any form of compressed supervision. It learns only from unpaired single-image collections and corresponding feature masses. Once discovered, it reconstructs 3D geometry and models dense correspondences from a single image. These features guide the established reconstruction of BCNN in input geometry. The learned point features are concatenated with the 3D point deformations. 3D shapes are rendered into a depth map by projection through the camera matrix to query 3D points on the surface. Every pixel corresponds to that depth map to get the color value 3D point and link texture. We output the texture in 3D location on the shape encoding, and the image encoding to reconstruct 3D.

Features labels

Image features are detected based on local and global spatial geometry. Many criteria are used to identify features, and there are different ways of using those criteria. Compute image information which distributes first-order spatial derivatives based on orientation. The edges show up as sharp changes in pixel values in the derivative. The edge locations correspond to the minimum and maximum derivatives. Measure the edge sharpness minima and maximum derivatives which are helpful in obtaining magnitudes. The edge locations can be further isolated using local intensity. Corners are detected by first noticing edges where those edges intersect. The corner filter function is detected based on gradients. The features are a number list showing each point position and orientation. A vital feature characteristic is that they are relatively unaffected by scale translation, rotation, and brightness changes. The main idea behind a convolutional neural network is using filters. These filters are responsible for detecting the features or patterns in the image. Many filters pass on the image individually to generate a label. Labeled images are generated by using a filter of 3×3 pixels. The total number of parameters in this is only 9, significantly reducing the number of parameters to train. A single layer of a convolutional neural network will use many filters which might detect the corners, edges, and orientation. These filter results will be passed in parallel onto the further layers which label the features associated with the image. Binary classification is a supervised learning method; a training sample includes input and corresponding labels. Data is labeled before training.

A single layer of a convolutional neural network uses many filters. Three filters are used to get three different images. The entire convolution operation gives small values as parameters to train. After model training, these parameter values take a specific value which can detect the associated features from the images.

BCNN classifier

The classifier considers the object surface as the decision boundary instead of explicitly the total 3D shape representing. The BCNN classifies a point as inside or outside the object depending on classifier function $f_Ø$ in the following equation:

$$f_Ø(p, x) = R^3x \rightarrow [0, 1],$$

where R^3 is 3D location; $R^3 \in (p, x)$; p is input point; x is encoding image conditions; [0,1] is occupancy probability; Ø is neural network parameter.

The occupancy networks update filter feature parameters followed by reconstruction loss. Labels used by the BCNN to detect occupancy points can refine the boundaries. BCNN implicit features models have effective output representations for shape, appearance, materials, and location. Assume the four points are inside, and the rest are outside. Then, iterate at times until the desired resolution is obtained. Fig. 3 shows that the red points are classified as inside, and the blue points are classified as outside. All the blue points have a low occupancy probability close to zero. The red points should be classified as outside and have a high occupancy probability close to one.

The input points network uses the number of points K equals three times the 3D coordinates. The output for each K point is the involved probability value. The classification loss function assumes the ground truth occupies the label for each 3D point that trains the equation below. The loss function updates the prediction parameter Ø while training.

$$Loss(Ø)_o = \sum_{j=1}^{K} Loss(f_Ø(p_{ij,xi}), o_{ij}),$$

where K randomly sampled 3D points p_{ij}; o_{ij} observation dataset.

Parameters Ø takes input 3D location, three coordinates, and a condition vector. The occupancy value of the reconstruction model at time equals zero. Compare this occupancy value to the actual occupancy in observations of the corresponding point location given the image condition corresponding to this label. Fig. 4 represents the shape only at a single point using a 3D occupancy network that can recover geometric detail.

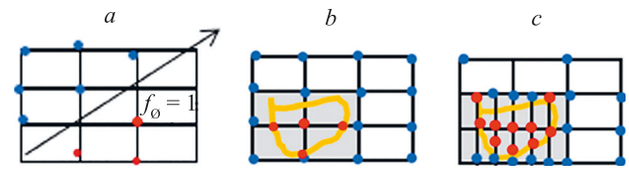

Fig. 3. BCNN occupies probability (*a*) detect grid points, (*b*) evaluate voxels and (*c*) mark surface points

The mapping between the object space and the 3D space points is based on a signed distance-driven functional mapping. A sign-distance generator on the corresponding encodes is used to learn the 3D shape and correspondence field jointly. Signed-distance evaluates the texture field at this predicted surface point and inserts the color value at that corresponding pixel location.

The time limitation of the traditional simple neural network, while dealing with the images, can affect the model performance. The number of operations performed will be significant, and we might need help handling them correctly. Identifying the features from the images permits us to reconstruct the presence of the object automatically in 3D. The BCNN allows the processing of many filters in parallel, significantly reducing time consumption. The only additional time is taken in the filter slice which is still much less than time would have taken by training a simple neural network. Fig. 5 shows the classification that produces labels for an image dataset containing single images. The dataset involves online images of radio, telephone, stairs, sink, car, bin, cabinet, and balcony stone. The results were obtained with a proposed model compared to novel view synthesis baselines.

Fig. 5 shows results that investigated the model power of representation 3D. The first raw is the ground truth example while the second shows label-encoded results. The third raw is to display the overfitting of the texture and object.

The accuracy measurement of the learning model calculated the number of data points detected correctly by comparing the image containing evidence with ground truth. These results are predicted from a single image by combining the implicit representations with primitive-

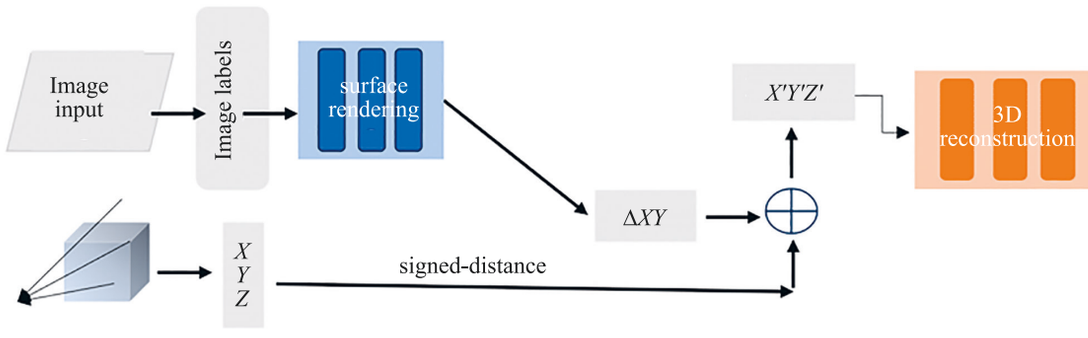

Fig. 4. Mapping 3D point

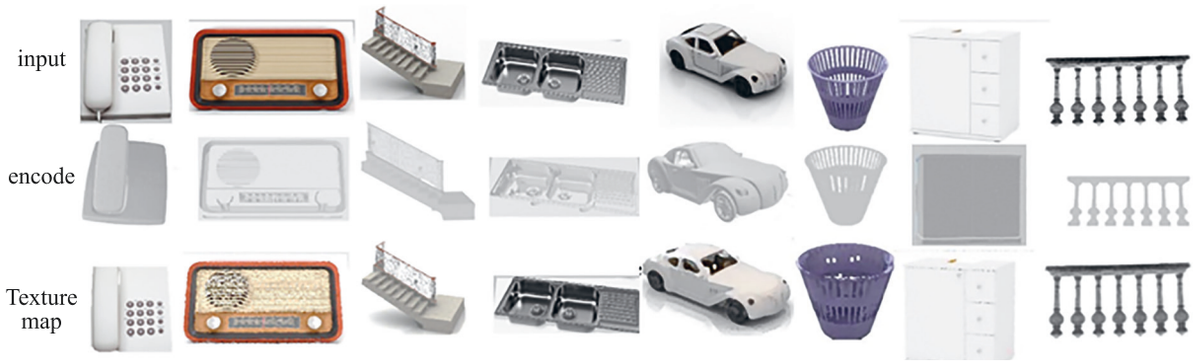

Fig. 5. Reconstruct 3D

based shape models. The high accuracy of the model needs to be corrected in labeling and distinguishing between relevant and irrelevant data. Four measurements are applied to examine the results: precision, recall, chamfer-L1, and Intersection over Union (IoU). Precision is calculated by dividing the predicted label from the overall accurate label. Recall measures the number of relevant elements detected. The recall percentage gives the probability that a randomly selected relevant item from the data set will be detected. Distance chamfer-L1 calculates similarity between set point distance. IoU quantifies the similarity between the predicted bounding and the ground truth. Table displays the model evaluated by accurate measurements.

Precision and recall are calculated to select the suitable model. Recall quantifies the number of positive predictions from all positive examples in the data set. Find the intersection between the two boundaries and the second portion, the union between the ground truth and the predicted. IoU is used for evaluation in the field of object detection. Table displays the results of the generative model showing that latent space interpolations have the same appearance code for different geometry encoding. Smoothing applies acceptable geometric detail structures

Table. Accuracy measurements

Object	Recall	Precision	Chamfer-L1	IoU
car	0.892	0.874	0.291	0.781
recycle bin	0.901	0.910	0.109	0.690
cabinet	0.868	0.849	0.251	0.589
balcony stone	0.859	0.861	0.194	0.609
telephone	0.871	0.880	0.201	0.698
radio	0.909	0.893	0.099	0.794
stair	0.790	0.802	0.301	0.805
sink	0.793	0.799	0.273	0.789

to overfit each object individually. In general, the approach time equals the ground truth object and the object deformed in the model; furthermore, the model can overfit multiple objects simultaneously.

Conclusion

Proposed approaches generally optimize single images to transfer 3D knowledge. BCNN is a simple, fully connected architecture that merges local and global conditioning for real data. It is necessary to input a single image to learn features filter parameters, and then image maps in 3D points. Features are dependent on the 3D location of the points and the viewing direction. Train filters feature parameters using a reconstruction loss function on a data set that has realistic materials. The loss function takes the actual occupancy value and the classifier decision boundary. The loss function is learning weight correspondence information that enhances representing spatially encoded. The approach predicts a latent code for the image that gets reconstructed and overfitted results. Approach characteristics depend on the surface implicitly considered instead of the shape explicitly. 3D convolutions generate a 3D feature volume. The convolution train inputs filter parameters that will detect the associated features from the images. The texture produces field and voxelization of the object. BCNN then queries labels by signed distance from 3D points. The last layers detect the texture and associate it with a particular label in the image. Representing scenes as neural radiance fields uses volume rendering points behind and in front of the surface. The limitation of a fully connected network is that it is sensitive to a large data size, but it is suitable for a single image. The model doesn't require discretization, and arbitrary topology is learned using 2D supervision BCNN. The BCNN is a powerful image model because it performs exceptionally well, and the reconstructions are precise and accurate.

References

1. Häming K., Peters G. The structure from-motion reconstruction pipeline — A survey with focus on short image sequences. *Kybernetika*, 2010, vol. 46, no. 5, pp. 926–937.
2. Molenaar M., Eisemann E. Editing compressed high-resolution voxel scenes with attributes. *Computer Graphics Forum*, 2023, vol. 42, no. 2, pp. 235–243. https://doi.org/10.1111/cgf.14757

Литература

1. Häming K., Peters G. The structure from-motion reconstruction pipeline — A survey with focus on short image sequences // Kybernetika. 2010. V. 46. N 5. P. 926–937.
2. Molenaar M., Eisemann E. Editing compressed high-resolution voxel scenes with attributes // Computer Graphics Forum. 2023. V. 42. N 2. P. 235–243. https://doi.org/10.1111/cgf.14757

3. Oechsle M., Peng S., Geiger A. UNISURF: Unifying neural implicit surfaces and radiance fields for multi-view reconstruction. *Proc. of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2021, pp. 5569–5579. https://doi.org/10.1109/iccv48922.2021.00554
4. Petersen F., Goldluecke B., Borgelt C., Deussen O. GenDR: A generalized differentiable renderer. *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022, pp. 3992–4001. https://doi.org/10.1109/cvpr52688.2022.00397
5. Gromniak M., Magg S., Wermter S. Neural field conditioning strategies for 2D semantic segmentation. *Lecture Notes in Computer Science*, 2023, vol. 14255, pp. 520–532. https://doi.org/10.1007/978-3-031-44210-0_42
6. Zhao Z., Liu W., Chen X., Zeng X., Wang R., Cheng P., Fu B., Chen T., Yu G., Gao S. Michelangelo: Conditional 3D shape generation based on shape-image-text aligned latent representation. *Advances in Neural Information Processing Systems*, 2023.
7. Su G.-M. Joint forward and backward neural network optimization in image processing. *Patent US20230084705A1*, 2023.
8. Greff K., Kaufman R.L., Kabra R., Watters N., Burgess C., Zoran D., Matthey L., Botvinick M., Lerchner A. Multi-object representation learning with iterative variational inference. *Proc. of the 36th International Conference on Machine Learning*, 2019, pp. 4317–4343.
9. Zanuttigh P., Minto L. Deep learning for 3D shape classification from multiple depth maps. *Proc. of the IEEE International Conference on Image Processing (ICIP)*, 2017, pp. 3615–3619. https://doi.org/10.1109/icip.2017.8296956
10. Cheng F., Xiao J., Tillo T., Zhao Y. Global motion information based depth map sequence coding. *Lecture Notes in Computer Science*, 2015, vol. 9314, pp. 721–729. https://doi.org/10.1007/978-3-319-24075-6_69
11. Yuan Z., Zhu Y., Li Y., Liu H., Yuan C. Make encoder great again in 3D GAN inversion through geometry and occlusion-aware encoding. *Proc. of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2023, pp. 2437–2447. https://doi.org/10.1109/iccv51070.2023.00231
12. Liu F., Huang T., Zhang Q., Yao H., Zhang C., Wan F., Ye Q., Zhou Y. BEAM: Beta distribution ray denoising for multi-view 3D object detection. *arXiv*, 2024, arXiv:2402.03634v1. https://doi.org/10.48550/arXiv.2402.03634
13. Wang X., Gupta A. Generative image modeling using style and structure adversarial networks. *Lecture Notes in Computer Science*, 2016, vol. 9908, pp. 318–335. https://doi.org/10.1007/978-3-319-46493-0_20
14. Shu C., Deng J., Yu F., Liu Y. 3DPPE: 3D point positional encoding for transformer-based multi-camera 3D object detection. *Proc. of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2023, pp. 3557–3566. https://doi.org/10.1109/iccv51070.2023.00331
15. Naselaris T., Olman D., Stansbury K., Ugurbil J., Gallant J.L. A voxel-wise encoding model for early visual areas decodes mental images of remembered scenes. *NeuroImage*, 2015, vol. 105, pp. 215–228. https://doi.org/10.1016/j.neuroimage.2014.10.018
16. Du Y.P., Chu R., Tregellas J.R. Enhancing the detection of BOLD signal in fMRI by reducing the partial volume effect. *Computational and Mathematical Methods in Medicine*, 2014. https://doi.org/10.1155/2014/973972
17. Wen X., Zhou J., Liu Y.-S., Su H., Dong Z., Han Z. 3D shape reconstruction from 2D images with disentangled attribute flow. *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022, pp. 3793–3803. https://doi.org/10.1109/cvpr52688.2022.00378
18. Oechsle M., Mescheder L., Niemeyer M., Strauss T., Geiger A. Texture fields: Learning texture representations in function space. *Proc. of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019, pp. 4530–4539. https://doi.org/10.1109/iccv.2019.00463
19. Giannis K., Thon C., Yang G., Kwade A., Schilde C. Predicting 3D particles shapes based on 2D images by using convolutional neural network. *Powder Technology*, 2024, vol. 432, pp. 119122. https://doi.org/10.1016/j.powtec.2023.119122

3. Oechsle M., Peng S., Geiger A. UNISURF: Unifying neural implicit surfaces and radiance fields for multi-view reconstruction // Proc. of the IEEE/CVF International Conference on Computer Vision (ICCV). 2021. P. 5569–5579. https://doi.org/10.1109/iccv48922.2021.00554
4. Petersen F., Goldluecke B., Borgelt C., Deussen O. GenDR: A generalized differentiable renderer // Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022. P. 3992–4001. https://doi.org/10.1109/cvpr52688.2022.00397
5. Gromniak M., Magg S., Wermter S. Neural field conditioning strategies for 2D semantic segmentation // Lecture Notes in Computer Science. 2023. V. 14255. P. 520–532. https://doi.org/10.1007/978-3-031-44210-0_42
6. Zhao Z., Liu W., Chen X., Zeng X., Wang R., Cheng P., Fu B., Chen T., Yu G., Gao S. Michelangelo: Conditional 3D shape generation based on shape-image-text aligned latent representation // Advances in Neural Information Processing Systems. 2023.
7. Su G.-M. Joint forward and backward neural network optimization in image processing. Patent US20230084705A1. 2023.
8. Greff K., Kaufman R.L., Kabra R., Watters N., Burgess C., Zoran D., Matthey L., Botvinick M., Lerchner A. Multi-object representation learning with iterative variational inference // Proc. of the 36th International Conference on Machine Learning. 2019. P. 4317–4343.
9. Zanuttigh P., Minto L. Deep learning for 3D shape classification from multiple depth maps // Proc. of the IEEE International Conference on Image Processing (ICIP). 2017. P. 3615–3619. https://doi.org/10.1109/icip.2017.8296956
10. Cheng F., Xiao J., Tillo T., Zhao Y. Global motion information based depth map sequence coding // Lecture Notes in Computer Science. 2015. V. 9314. P. 721–729. https://doi.org/10.1007/978-3-319-24075-6_69
11. Yuan Z., Zhu Y., Li Y., Liu H., Yuan C. Make encoder great again in 3D GAN inversion through geometry and occlusion-aware encoding // Proc. of the IEEE/CVF International Conference on Computer Vision (ICCV). 2023. P. 2437–2447. https://doi.org/10.1109/iccv51070.2023.00231
12. Liu F., Huang T., Zhang Q., Yao H., Zhang C., Wan F., Ye Q., Zhou Y. BEAM: Beta distribution ray denoising for multi-view 3D object detection // arXiv. 2024. arXiv:2402.03634v1. https://doi.org/10.48550/arXiv.2402.03634
13. Wang X., Gupta A. Generative image modeling using style and structure adversarial networks // Lecture Notes in Computer Science. 2016. V. 9908. P. 318–335. https://doi.org/10.1007/978-3-319-46493-0_20
14. Shu C., Deng J., Yu F., Liu Y. 3DPPE: 3D point positional encoding for transformer-based multi-camera 3D object detection // Proc. of the IEEE/CVF International Conference on Computer Vision (ICCV). 2023. P. 3557–3566. https://doi.org/10.1109/iccv51070.2023.00331
15. Naselaris T., Olman D., Stansbury K., Ugurbil J., Gallant J.L. A voxel-wise encoding model for early visual areas decodes mental images of remembered scenes // NeuroImage. 2015. V. 105. P. 215–228. https://doi.org/10.1016/j.neuroimage.2014.10.018
16. Du Y.P., Chu R., Tregellas J.R. Enhancing the detection of BOLD signal in fMRI by reducing the partial volume effect // Computational and Mathematical Methods in Medicine. 2014. https://doi.org/10.1155/2014/973972
17. Wen X., Zhou J., Liu Y.-S., Su H., Dong Z., Han Z. 3D shape reconstruction from 2D images with disentangled attribute flow // Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022. P. 3793–3803. https://doi.org/10.1109/cvpr52688.2022.00378
18. Oechsle M., Mescheder L., Niemeyer M., Strauss T., Geiger A. Texture fields: Learning texture representations in function space // Proc. of the IEEE/CVF International Conference on Computer Vision (ICCV). 2019. P. 4530–4539. https://doi.org/10.1109/iccv.2019.00463
19. Giannis K., Thon C., Yang G., Kwade A., Schilde C. Predicting 3D particles shapes based on 2D images by using convolutional neural network // Powder Technology. 2024. V. 432. P. 119122. https://doi.org/10.1016/j.powtec.2023.119122

Author

Sallama Adhab Resen — PhD, Lecturer, Directorate General of Vocational Education, Baghdad,1001, Iraq, SC 57204833765, https://orcid.org/0009-0007-5044-8857, salamaresen@gmail.com

Автор

Ресен Саллама Адхаб — PhD, преподаватель, Главное управление профессионального образования, Багдад, 1001, Ирак, SC 57204833765, https://orcid.org/0009-0007-5044-8857, salamaresen@gmail.com

Received 13.05.2024
Approved after reviewing 29.08.2024
Accepted 26.09.2024

Статья поступила в редакцию 13.05.2024
Одобрена после рецензирования 29.08.2024
Принята к печати 26.09.2024