Deep 3D Shape Reconstruction from Single-View Sketch Image

Fei Wang*, Yu Yang †‖, Baoquan Zhao‡, Junkun Jiang §, Teng Zhou *, Dazi Jiang *, and Tie Cai ¶
*Shantou University, Shantou, China
†Shenzhen Securities Information Co., Ltd, Shenzhen, China
‡Guilin University of Electronic Technology, Guilin, China
§Sun Yat-sen University, Guangzhou, China
¶Shenzhen Institute of Information Technology, Shenzhen, China

***Abstract*—In this paper, we introduce a novel 3D shape reconstruction method from single-view sketch image based on deep neural network. The proposed pipeline is mainly composed of three modules. The first module is sketch component segmentation based on multi-modal DNN fusion and is used to segment a given sketch into a series of basic units and build transformation template by the knots between them. The second module is a non-linear transformation network for multifarious sketch generation with the obtained transformation template. It creates the transformation representation of a sketch by extracting the shape features of an input sketch and transformation template samples. The third module is deep 3D shape reconstruction using multifarious sketches, which takes the obtianed sketches as input to reconstruct 3D shapes with a generative model. It fuses and optimizes features of multiple views and thus is more likely to generate high-quality 3D shapes. To evaluate the effectiveness of the proposed method, we conduct extensive experiments on a public 3D reconstruction dataset. The results demonstrate that our model can achieve better reconstruction performance than peer methods.**

***Keywords*-Sketch-based 3D shape reconstruction; deep neural networks; sketch understanding**

I. Introduction

As one of the most fundamental problem in computer graphics, 3D shape reconstruction is playing an increasingly important role in a wide variety of fields such as virtual/augmented reality, computer aided geometric design, gaming, medical imaging and so on. However, manually reconstructing 3D models is not a trivial task. It usually takes designers a lot of time and effort to harvest an exquisite 3D model since the procedure involves intensive interactions. To alleviate this situation, a large body of technologies have been developed to facilitate the 3D shape reconstruction process. Among them, automatically reconstructing 3D models from sketch images is gaining more and more popularity due to its high efficacy and simplicity of interaction. Especially with the recent advance in deep learning technology, sketch-based deep 3D shape reconstruction has achieved remarkable progress. Unfortunately, there are still many challenging issues in this area that have not been effectively addressed, which is seriously hindering the adoption of this technique in many applications.

These issues are mainly featured in the following three aspects. Firstly, there is great semantic gap between sketch image and 3D models. Compared with nature images, sketch is a visual representation form with high-level of abstraction and can easily cause ambiguity. This character could bring great challenges to the understanding of sketches, and will affect the performance of 3D reconstruction. Besides, drawing skills and painting styles of different users vary greatly, which further exacerbate the difficulties of sketch semantic analysis. Secondly, deep learning based 3D reconstruction models are generally highly dependent on sufficient training data. As the diversity of users' painting styles, it is very time-consuming and costly to collect tens of thousands of well-labelled sketch images that can be used to fed into deep neural networks for effective training. Last but not least, the dimensionality and feature of sketches and 3D models are quite different. How to effectively exploit the very limit visual clues in sketch images to reconstruct 3D shape accurately is another challenging task that is still inadequately addressed by existing studies.

To tackle the aforementioned issues, in this paper, we introduce a novel 3D shape reconstruction framework from sketch image using deep neural network. Unlike conventional methods that need several sketch images from multiple-views, the proposed approach only takes single-view sketch image as input. This can significantly reduce the interaction during the reconstruction process, which is a very important factor that is highly concerned by practitioners in real-world applications. However, compared with muti-view based reconstruction methods, extracting meaningful features from a single-view sketch may be insufficient for accurate 3D shape reconstruction. This is because sketches from multiple views are more likely to convey more useful features to infer the underlying structure of a 3D shape. To gain the merit of multi-view sketch based reconstruction frameworks, we propose to first generate several more sketch images according to the information provided by input one, and then use the obtained sketch images to reconstruct 3D shape model.

The rest of this paper is organized as follows. In next section, we briefly introduce existing researches on 3D shape reconstruction from sketch images. The details of the

978-1-7281-9234-5/20/$31.00 ©2020 IEEE
DOI 10.1109/ICDH51081.2020.00039

proposed framework is presented in Sec. III. We conduct an extensive experiment to demonstrate the effectiveness of the proposed method in Sec. IV. Sec. V concludes the paper.

II. Related Work

In this section, we briefly review existing researches that are closely related to our work from two aspects: (1) sketch understanding and (2) sketch-based composition and reconstruction.

A. Sketch Understanding

Sketch understanding is an emerging branch in the field of artificial intelligence, aiming at recognizing and extracting the semantic knowledge from sketch images in a fully-/semi- automatic fashion. It mainly encompasses two parts, i.e., sketch recognition and semantic understanding [5]. A pioneer work on large-scale sketch recognition is ”How do users draw sketches?” [1], in which 20k sketch images in total were collected. It adopts Gaussian derivative to estimate the direction of lines and utilizes bag-of-words model to encode local curve direction as the feature vector of sketch. Then, support vector machine (SVM) is employed to classify sketch images. With the great achievement of deep neural network (DNN) on natural image recognition, convolutional neural network (CNN) has also been applied to sketch recognition. However, taking sketch image as 2D pixel array, CNN-based methods usually need to learn a huge amount of network parameters, which is very inefficient to train the network. Compared with natural images, the features of sketch images are sparse. Such spare data can significantly improve the compactness of networks.

B. Sketch-based Composition and Reconstruction

A large body of studies have been carried out to extract features from sketch images and use them to perform a variety of tasks such as 3D model retrieval, shape reconstruction, and so on. For example, Chen et al. [6] proposed to retrieve 3D model using sketch by constructing two individual deep CNN and metric network. One network is for sketch images and the other one is for 3D models. Interleaved active metric learning (IAML) is used to learn specific features from these two modalities, which is capable of mining important features from samples for training and learning discriminative feature representation effectively. Besides, to reduce the cross-modality difference between sketch features and 3D shapes, it also introduced a modality transformation network to convert sketch feature into the feature space of 3D models, which achieves better retrieval performance.

This technique also paves the way for the development of sketch-based 3D shape reconstruction [2], [3]. Wang et al. [7] developed a label-free sketch neural network 3D-GAN. This model is integrated with an embedding latent vector space to harvest a similar feature vector distribution between sketch image and 2D rendered image. Then, it obtains final results by retrieving the k most similar 3D models with sketch as the prior knowledge. Lun et al. [8] mapped sketch to 3D shape by training a ConvNet to infer the structure and reconstructed 3D model using a multi-view framework. Unlike conventional methods that adopts voxels to represent 3D models, the shape of 3D objects can be represented with surface-based forms (for example, polygon mesh), which can achieve more accurate prediction by combining feedforward frameworks.

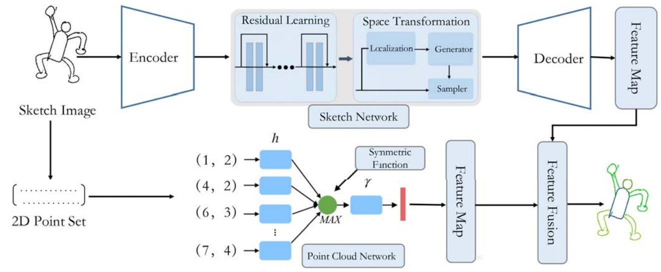

Figure 1. Sketch component segmentation based on multi-modal DNN fusion.

III. The Proposed Deep Neural Network for 3D Shape Reconstruction

A. Overview of the Proposed Method

The proposed pipeline is mainly composed of three components: (1) sketch component segmentation based on multi-modal DNN fusion, (2) multifarious sketch generation based on non-linear transformation network, and (3) deep 3d shape reconstruction using multifarious sketches. The first component is a 2D point cloud based sketch segmentation model, which is used to segment a given sketch into a series of basic units and build transformation template by the knots between them. The advantage of this component is that it only relies on a small amount of sample data to achieve the learning task of transformation network. The second component is developed for generating multifarious sketches with the aforementioned transformation template. It creates the transformation representation of a sketch by extracting the shape features of an input sketch and transformation template samples, which can avoid the problems existing in conventional models such as unitary of transformation structure, distortion, etc. The third component takes the obtianed sketches as input to reconstruct 3D shapes with a generative model. It fuses and optimizes features of multiple views and thus is more likely to generate high-quality 3D shapes. Details of each component will be introduced in the following three subsections.

B. Sketch Component Segmentation Based on Multi-modal DNN Fusion

Sketch Component Segmentation Net is inspired by the work [9]. The sketch network mainly consists of two parts:

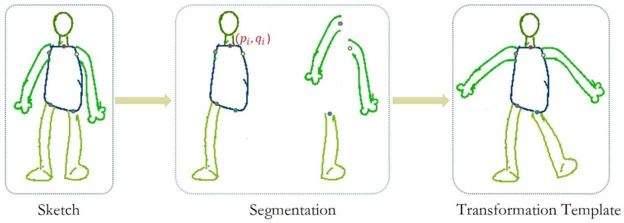

Figure 2. Sketch segmentation, knots and transformation template.

encoder and decoder. On the one hand, the network obtains the global feature of a sketch through the encoder. As shown in Fig. 1, feature representation is learned and extracted using spatial invariance enhanced residual (SIER), which is composed of two modules: residual learning module and spatial transformation module. These features will be combined together in the decoding phrase to generate pixel-level feature segmentation image. On the other hand, the coordinate information of sketch contour is an important geometry structure. Therefore, 2D point cloud network can obtain the feature of sketch by representing each point with 2D coordinates (x, y). Let $P = \{p_i | i = 1, \cdots, n\}$ be the coordinates set of sketch contour, where p_i represents the coordinate of each sample point, 2D point cloud network takes P as input and gets the global features by gathering point features with maximum function MAX. Then the probability of each point in P associated with all semantic units can be obtained by connecting local and global features through segmentation network. More specifically, let $f : X \to R$ be a continuous set function regarding Hausdorff distance $d_H(\cdot, \cdot)$. For $\forall \epsilon > 0$, there exists a continuous function $g(x_1, \cdots, x_n) = \gamma \circ MAX$ such that for arbitrary $x_i \in X$,

$$|f(\{x_i, \cdots, x_n\}) - \gamma(MAX(h(x_q), h(x_2), \cdots, h(x_n)))| < \epsilon \tag{1}$$

where γ and h are continuous functions, and MAX is vector maximum description operator. Therefore, the general function of arbitrary 2D point set can be represented as:

$$f(\{x_1, \cdots, x_n\}) \approx g(h(x_1), \cdots, h(x_n)) \tag{2}$$

As shown in Fig. 1, h can be learned wirth multi-layer perception (MLP), and $g = \gamma \circ MAX$ can be obtained with single variable function and max pooling.

After performing sketch segmentation, we use the transformation templates of the original sketch to train the network. As shown in Fig. 1, a sketch is segmented into a series of semantic units. The knot between units is represented with $\{(p_i, q_i)\}$, which is highlighted with red circle in the figure.

C. Multifarious Sketch Generation Based on Non-linear Transformation Network

In this step, we first carry out transformation feature extraction and representation. For $m(m \geq 2)$ sketch images, each of which has n vertexes, let p_i be the v_i-th vertex of an input sketch and p'_i be the one in the transformation template, the transformation gradient T_i can be obtained by minimizing the energy function below:

$$E(T_i) = \sum_{j \in N_i} c_{ij} \|e'_{ij} - T_i e_{ij}\|^2 \tag{3}$$

where N_i is the neighborhood set of v_i, $e'_{ij} = p'_i - p'_j$, $e_{ij} = p_i - p_j$, and $c_{ij} = cot\alpha_{ij} + cot\beta_{ij}$ is cotangent weight. The affine transformation matrix can be decomposed into rotation part and scaling part $T_i = R_i S_i$. The rotation difference from v_i to v_j is given by:

$$dR_{ij} = R_i^T R_j \tag{4}$$

Thus, the energy function is redefined as:

$$E(T_i) = \sum_{j \in N_i} c_{ij} \sum_{t \in N_i} c'_i \|e'_{ij} - R_i dR_{ti} S_i e_{ij}\|^2 \tag{5}$$

where $c'_i = \frac{1}{|N_i|}$, and $|N_i|$ is the number of neighbourhood of v_i. Finally, the feature of transformation template on v_i can be represented as:

$$f_i^j = \{logdR_{ij}; S_i\}(\forall i, j \in N_i). \tag{6}$$

The Sketch-VAE network aims to find an encoder and a decoder, where the goal of encoder is to map the posterior distribution of x to hidden vector z, while that of decoder is to generate a credible x. The loss function of Sketch-VAE model is defined as:

$$L_{VAE} = \sum_{j=1}^{M} \sum_{i=1}^{K} (\hat{f}_i^j - \tilde{f}_i^j)^2 + D_{KL}(q(z|\tilde{f}) \| p(z)) \tag{7}$$

where $\tilde{f}_i^j)$ is the transformed feature after preprocessing, $\hat{f}_i^j)$ is the output of Sketch-VAE model. z is a hidden vector, $p(z)$ and $q(z|\tilde{f})$ are prior and posterior probability, respectively, and D_{KL} is KL divergence.

To avoid incorrect output caused by the separation between sketch components, we add a constrain condition and define the loss function of knots as:

$$L_{joints} = \sum_{i=1}^{n} \|p_i - q_i\|^2. \tag{8}$$

Besides, we also add a regularization constrain to the network optimization network to avoid distortion and define the loss function as:

$$L_{reg} = \sum_{i=1}^{|V|} \sum_{j \in N_i} w_{ij} \|(\hat{v}_i - \hat{v}_j) - R_i(v_i - v_j)\|^2 \tag{9}$$

where v_i and $\hat{v}_i$ are the vertexes in the original and transformed sketches, respectively. Thus, the final loss function of the transformation network can be formulated as:

$$L_{total} = L_{VAE} + \lambda_1 L_{joints} + \lambda_2 L_{reg}. \tag{10}$$

Table I
MAN-MADE OBJECTS (SYNTHETIC)

	Ours	ShapeMVD	Nearest retrieval	Tatarchenko et al. [13]	[13]+ U-net	Volumetric decoder	R2N2
Hausdorff distance	**0.076**	0.092	0.165	0.142	0.121	0.113	0.144
Chamfer distance	**0.011**	0.015	0.025	0.022	0.017	0.021	0.026
normal distance	**26.45**	30.66	42.57	35.58	32.32	49.40	48.78
depth map error	**0.013**	0.026	0.049	0.039	0.030	0.038	0.045
volumetric distance	**0.276**	0.344	0.501	0.442	0.374	0.432	0.512

Table II
CHARACTER MODELS (SYNTHETIC)

	Ours	ShapeMVD	nearest retrieval	Tatarchenko et al. [13]	[13]+ U-net	volumetric decoder	R2N2
Hausdorff distance	**0.065**	0.089	0.200	0.119	0.092	0.152	0.148
Chamfer distance	**0.010**	0.015	0.036	0.025	0.016	0.026	0.032
normal distance	**26.47**	30.61	44.93	34.98	31.00	53.84	53.13
depth map error	**0.014**	0.018	0.040	0.030	0.019	0.031	0.036
volumetric distance	**0.248**	0.313	0.541	0.428	0.329	0.437	0.493

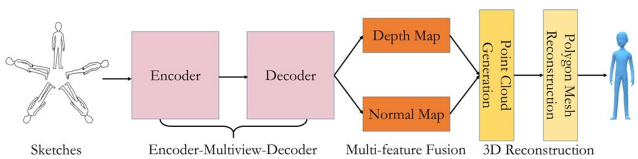

Figure 3. Deep 3D shape reconstruction using multifarious sketches.

D. Deep 3D Shape Reconstruction Using Multifarious S-ketches

The proposed deep 3D shape reconstruction framework is illustrated in Fig. 3. For the sketches obtained in last step, the encoder first extracts the feature of each sketch. It consists of a series of convolution and all layers use ReLU as active function. Then the decoder transforms these features into depth and normal images, which will be fused into 3D point cloud subsequently. The first step of fusion is to map all foreground pixels to 3D points. If the prediction probability of a pixel is lager than 50%, we consider it as a foreground pixel. Let the depth of a foreground pixel p be $p_{p.v}$, the coordinate set of graph space in the i-th sketch image be $\{p_x, p_y\}$, the position of a 3D point $q_{p,i}$ can be calculated as:

$$q_{p,i} = R_v[\kappa p_x \kappa p_y d_{p,i}]^T + e_i, \tag{11}$$

where κ is a scaling coefficient, representing the distance between adjacent pixels and center. Finally, 3D model can be obtained by transforming it into polygon mesh with Poisson surface reconstruction algorithm [10].

IV. EXPERIMENT AND DISCUSSION

To evaluate the effectiveness of the proposed method for 3D shape reconstruction, we conduct a comparative experiment on a public dataset. To train our neural network, we use the dataset presented in [8], which mainly consists of three different types of 3D models, i.e., human/humanoid, airplanes and chairs. Among them, human/humanoid involves human models, aliens and virtual cartoon characters, which come from *The Models Resource* dataset [11]. While airplane and chair models are mainly from *3D ShapeNet* [12], which has a large variety in shape geometry. There are 120 sketch images in total in the test dataset. Among them, 90 are synthetic sketches, which are generated from test images with line painting techniques; while the rest 30 sketch images are drawn by two professional artists. They were asked to draw 10 sketch images for each category.

We tested the reconstructions produced by our method (called ShapeMVD) versus the following methods: (a) a network based on the same encoder as ours but using a volumetric decoder baseline instead of our multiview decoder, (b) a network based on the same encoder as ours but with the Tatarchenko et al.s view-based decoder [18] instead of our multi-view decoder, (c) the convolutional

Peer sketch-based 3D shape reconstruction methods s-elected for comparison include ShapeMVD [8], Nearest retrieval, Tatarchenko et al. [13], U-net [15], volumetric decoder and R2N2 [14], which are state-of-the-art models and widely used by existing studies for performance evaluation. For the nearest-neighbor baseline, we extract the representation of the input test sketches based on our encoder. This is used as a query representation to retrieve the training shape whose sketches have the nearest encoder representation based on Euclidean distance.We additionally implemented a variant of Tatarchenko et al.'s decoder by adding U-net connections between the encoder and their decoder. The volumetric decoder consisted of five transpose 3D convolutions of stride 2 and kernel size $4 \times 4 \times 4$. The number of filters starts with 512 and is divided by 2 at each

Table III
MAN-MADE OBJECTS (HUMAN DRAWING)

	Ours	ShapeMVD	nearest retrieval	Tatarchenko et al. [13]	[13]+ U-net	volumetric decoder	R2N2
Hausdorff distance	**0.094**	0.116	0.176	0.153	0.153	0.130	0.149
Chamfer distance	**0.011**	0.017	0.031	0.024	0.025	0.022	0.028
normal distance	**21.058**	27.04	40.96	32.40	30.45	48.32	48.12
depth map error	**0.011**	0.021	0.042	0.033	0.032	0.032	0.042
volumetric distance	**0.202**	0.311	0.544	0.405	0.403	0.405	0.500

Table IV
CHARACTER MODELS (HUMAN DRAWING)

	Ours	ShapeMVD	nearest retrieval	Tatarchenko et al. [13]	[13]+ U-net	volumetric decoder	R2N2
Hausdorff distance	**0.102**	0.117	0.188	0.139	0.136	0.178	0.168
Chamfer distance	**0.013**	0.021	0.036	0.025	0.024	0.032	0.036
normal distance	**28.22**	33.44	43.81	36.11	34.74	54.91	54.29
depth map error	**0.012**	0.026	0.040	0.031	0.027	0.037	0.040
volumetric distance	**0.217**	0.298	0.458	0.342	0.307	0.420	0.436

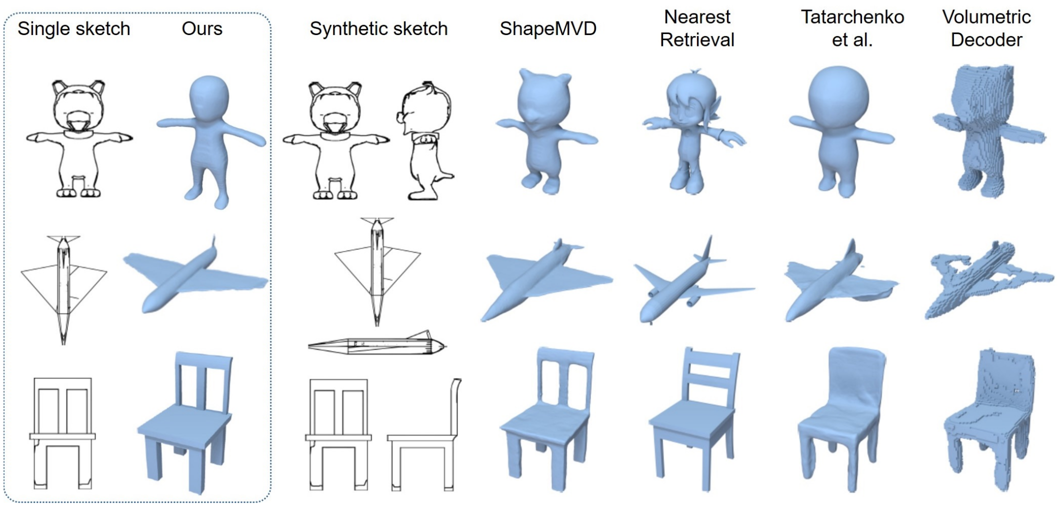

Figure 4. Comparison of 3D models reconstructed with different methods.

layer. Leaky ReLU functions and batch normalization were used after each layer. We note that we did not use skip-connections (U-net architecture) in the volumetric decoder because the size of the feature representations produced in the sketch image-based encoder is incompatible with the ones produced in the decoder.

Following the common practice, we also compare the similarity between the reconstructed 3D models and input sketches using the following five distance metrics: Chamfer distance, Hausdorff distance, surface normal distance, depth map error and volumetric Jaccard distance. The comparison results are illustrated in Table I IV, where Table I and II are the results of synthetic sketches while Table III and IV those of human drawing sketches. Bold values in the tables indicate the best results among all methods. From Table I and II, we can see that the proposed method achieves the smaller distances than peer ones in terms of the five evaluation metrics on both man-made objects and character models, which demonstrates that our method can generate more accurate 3D models. We can find that ShapeMVD performs better than the conventional methods like U-net. However, its performance gain is just marginally higher. Our model outperforms all of these method and performance gain is significant. The superiority of the proposed method can also be reflected on the human drawing sketches dataset, as shown in Table III and IV. Likewise, the distance reduction compared with other methods is also prominent.

Beyond the aforementioned quantitative comparison, we also demonstrate the effectiveness of the proposed method by visually comparing the 3D shape reconstruction results.

As shown in Fig. 4, the proposed framework can obtain comparable reconstruction results (highlighted with blue bounding box in the figure) that are in line with those generated by peer methods. It is worth noting that the 3D model of chair reconstructed with our method is pretty good, compared with ShapeMVD, Tatarchenko et al. and volumetric decoder.

V. Conclusion

In this paper, we have introduced a novel 3D shape reconstruction method from single-view sketch image using deep neural network. Our model is general, and can be easily extended to other applications, such as biomedical, intelligent computing [4], [16]–[23], [25]–[29]. The proposed method first generates a series of sketch images from differ view points by analysing the semantic information of the input sketch image. Then, the obtained sketch images are fed into deep neural network to reconstruct the 3D shapes. Compared with multi-view based approaches, the proposed method only takes a single sketch image as input, which can significantly reduce time used for drawing sketches and remarkably improve the reconstruction efficiency. Besides, using the input sketch image as visual clues to generate multi-view sketch images is helpful to reconstruct 3D shapes more accurately, which is superior then conventional single sketch image based 3D shape reconstruction methods. Extensive experiments on a public 3D shape reconstruction dataset have demonstrated the efficacy of the proposed model.

Acknowledgment

This research is supported by the Scientific Research Start-up Fund of Shantou University (09420021), the General Universities and Colleges Young Innovative Talents Project of Guangdong Province, the National Natural Science Foundation of China (61902087,61902232, 61902231) and in part by the Natural Science Foundation of Guangxi (2018GXNSFAA294127), the Natural Science Foundation of Guangdong Province (No. 2018A030313291, 2019A1515010943). The corresponding author is Yu Yang (ahhffdyy@163.com).

References

[1] Eitz M, Hays J, Alexa M. How do humans sketch objects?[J]. ACM Trans. Graph., 2012, 31(4) : 44 C 1.

[2] Fei Wang, Shujin Lin, Hanhui Li, Hefeng Wu, Tie Cai, Xiaonan Luo, Ruomei Wang: Multi-column point-CNN for sketch segmentation. Neurocomputing 392: 50-59 (2020).

[3] Fei Wang, Shujin Lin, Xiaonan Luo, Baoquan Zhao, Ruomei Wang: Query-by-sketch image retrieval using homogeneous painting style characterization. J. Electronic Imaging 28(02): 023037 (2019).

[4] Fei Wang, Shujin Lin, Xiaonan Luo, Ruomei Wang: Coupling Computation of Density-Invariant and Divergence-Free for Improving Incompressible SPH Efficiency. IEEE Access 8: 135912-135919 (2020).

[5] Fei Wang, Shujin Lin, Hefeng Wu, Hanhui Li, Ruomei Wang, Xiaonan Luo, Xiangjian He: SPFusionNet: Sketch Segmentation Using Multi-modal Data Fusion. ICME 2019: 1654-1659.

[6] Chen J, Fang Y. Deep cross-modality adaptation via semantics preserving adversarial learning for sketch-based 3d shape retrieval[C] // Proceedings of the European Conference on Computer Vision (ECCV). 2018 : 605 C 620.

[7] Wang L, Qian C, Wang J, et al. Unsupervised learning of 3D model reconstruction from handdrawn sketches[C] // 2018 ACM Multimedia Conference on Multimedia Conference. 2018 : 1820 C 1828.

[8] Lun Z, Gadelha M, Kalogerakis E, et al. 3d shape reconstruction from sketches via multiview convolutional networks[C] // 2017 International Conference on 3D Vision (3DV). 2017 : 67 C 77.

[9] Wang F ,Lin S., Wu H., Li H., Luo X. , He X. (2019). SPFusionNet: Sketch Segmentation Using Multi-modal Data Fusion. 1654-1659.

[10] Kazhdan M, Hoppe H. Screened poisson surface reconstruction[J]. ACM Transactions on Graphics (ToG), 2013, 32(3) : 29.

[11] T. M. Resource, “https://www.models-resource.com/,” 2017.

[12] Chang A X , Funkhouser T , Guibas L , et al. ShapeNet: An Information-Rich 3D Model Repository[J]. Computer Science, 2015.

[13] Tatarchenko M , Dosovitskiy A , Brox T . Multi-view 3D Models from Single Images with a Convolutional Network[C]// European Conference on Computer Vision. Springer International Publishing, 2016.

[14] C. B. Choy, D. Xu, J. Gwak, K. Chen, and S. Savarese, 3dr2n2: A unified approach for single and multi-view 3d object reconstruction, in Proc. ECCV, 2016.

[15] Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Lecture Notes in Computer Science, vol 9351. Springer, Cham, 2015.

[16] Lu, H.; Huang, D.; Song, Y.; Jiang, D.; Zhou, T.; Qin, J. ST-TrafficNet: A Spatial-Temporal Deep Learning Network for Traffic Forecasting. Electronics 2020, 9, 1474.

[17] Xiao, Guorong and Tu, Geng and Zheng, Lin and Zhou, Teng and Li, Xin and Ahmed, Syed and Jiang, Dazhi. (2020). Multi-modality Sentiment Analysis in Social Internet of Things based on Hierarchical Attentions and CSATTCN with MBM Network. IEEE Internet of Things Journal. PP. 1-1.

[18] Cai L , Yu Y , Zhang S , et al. A Sample-rebalanced Outlier-rejected k-nearest Neighbour Regression Model for Short-Term Traffic Flow Forecasting[J]. IEEE Access, 2020:1-11.

[19] Dazhi Jiang, Kaichao Wu, Dicheng Chen, Geng Tu, Teng Zhou, Akhil Garg and Liang Gao. A probability and integrated learning based classification algorithm for high-level human emotion recognition problems, Measurements, 150(2020) 107049.

[20] Dazhi Jiang, Z. Liu, L. Zheng and J. Chen, "Factorization Meets Neural Networks: A Scalable and Efficient Recommender for Solving the New User Problem," in IEEE Access, vol. 8, pp. 18350-18361, 2020.

[21] Dazhi Jiang, Geng Tu, Donghui Jin, Kaichao Wu, Cheng Liu, Lin Zheng, Teng Zhou. A Hybrid Intelligent Model for Acute Hypotensive Episode Prediction with Large-scale Data, Information Sciences, 2020.

[22] Guorong Xiao,Yunju Ma, Cheng Liu, Dazhi Jiang. A Machine Emotion Transfer Model for Intelligent Human-machine Interaction based on Group Division. Mechanical Systems and Signal Processing, 142(2020) 106736.

[23] Guorong Xiao; Tu, Geng; Zheng, Lin; Zhou, Teng; Li, Xin; Ahmed, Syed Hassan; Jiang, Dazhi. Multi-modality Sentimental Analysis in Social Internet of Things based on Hierarchical Attentions and CSAT-TCN with MBM Network, IEEE Internet of Things Journal, 2020.

[24] Lin Zheng, Naicheng Guo, Jin Yu, Dazhi Jiang:Memory Reorganization: A Symmetric Memory Network for Reorganizing Neighbors and Topics to Complete Rating Prediction. IEEE Access 8: 81876-81886 (2020).

[25] H. Lu, D. Huang, S. Youyi, D. Jiang, T. Zhou, and J. Qin, "St-trafficnet: A spatial-temporal deep learning network for traffic forecasting," *Electronics*, pp. 1–17, 2020.

[26] H. Lu, Z. Ge, S. Youyi, D. Jiang, T. Zhou, and J. Qin, "A temporal-aware lstm enhanced by loss-switch mechanism for traffic flow forecasting," *Neurocomputing*, 2020.

[27] C. Li, S. Tang, H. K. Kwan, J. Yan, and T. Zhou, "Color correction based on cfa and enhancement based on retinex with dense pixels for underwater images," *IEEE Access*, 2020.

[28] C. Li, S. Tang, J. Yan, and T. Zhou, "Low-light image enhancement via pair of complementary gamma functions by fusion," *IEEE Access*, 2020.

[29] C. Li, S. Tang, J. Yan, and T. Zhou, "Low-light image enhancement based on quasi-symmetric correction functions by fusion," *Symmetry*, 2020.