DrawingSpinUp: 3D Animation from Single Character Drawings Supplemental Materials

JIE ZHOU, City University of Hong Kong
CHUFENG XIAO, Hong Kong University of Science and Technology
MIU-LING LAM, City University of Hong Kong
HONGBO FU*, Hong Kong University of Science and Technology

ACM Reference Format:
Jie Zhou, Chufeng Xiao, Miu-Ling Lam, and Hongbo Fu*. 2024. DrawingSpinUp: 3D Animation from Single Character Drawings Supplemental Materials. In *SIGGRAPH Asia 2024 Conference Papers (SA Conference Papers '24), December 3–6, 2024, Tokyo, Japan.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3680528.3687593

1 CONTOUR REMOVAL

1.1 Training Data Preparation

We train our contour removal network on the 3DBiCar dataset [Luo et al. 2023]. We render front-view images and contour lines of different thicknesses from 1,500 textured 3D biped cartoon characters in 3DBiCar with Blender [Blender Development Team. 2022]. In order to increase data diversity, we render front-view images with random rotation and render contour lines with random stroke widths (ranging from 1 to 30) and random colors. We train our removal network for only 4 epochs since too many training epochs might lead to over-fitting.

1.2 Comparison of Different Methods

Fig. 1 shows the results of contour removal based on different backbones, including a DoG edge extractor mentioned by Panic-3D [Chen et al. 2023], Pix2PixHD [Wang et al. 2018], and FFC-ResNet [Suvorov et al. 2022]. The DoG-based method treats all lines indiscriminately and is unable to differentiate between external contour lines and internal texture lines, while Pix2pixHD tends to give inaccurate predictions due to a limited receptive field.

2 SHAPE REFINEMENT

2.1 Intersection Removal

Considering that the thinning deformation might affect adjacent areas (for example, the face might also be deformed when we thin the top of the neck), we need to ignore these intersections, i.e., pixels within the intersecting areas. We design an intersection removal algorithm to locate these intersections, as described in Algorithm 1.

ALGORITHM 1: IntersectionRemoval

```
Input: S_mov, S, r
endpoints1 = GetEndPoints (S_mov);
endpoints2 = GetEndPoints (S);
for each point p in endpoints1 do
    if p not in endpoints2 then
        for each pixel in a circle of radius r around p do
            set pixel value to 0 in S_mov;
        end
    end
end
Function GetEndPoints(skeleton):
    endpoints ← Empty list;
    for each pixel (x, y) in skeleton do
        if skeleton[x, y] = 1 then
            count ← the sum of 3 × 3 neighborhood of (x, y);
            if count = 2 then
                Add (x, y) to endpoints;
            end
        end
    end
    return endpoints;
Output: S_mov with removed interseactions
```

2.2 Shearing Transformation

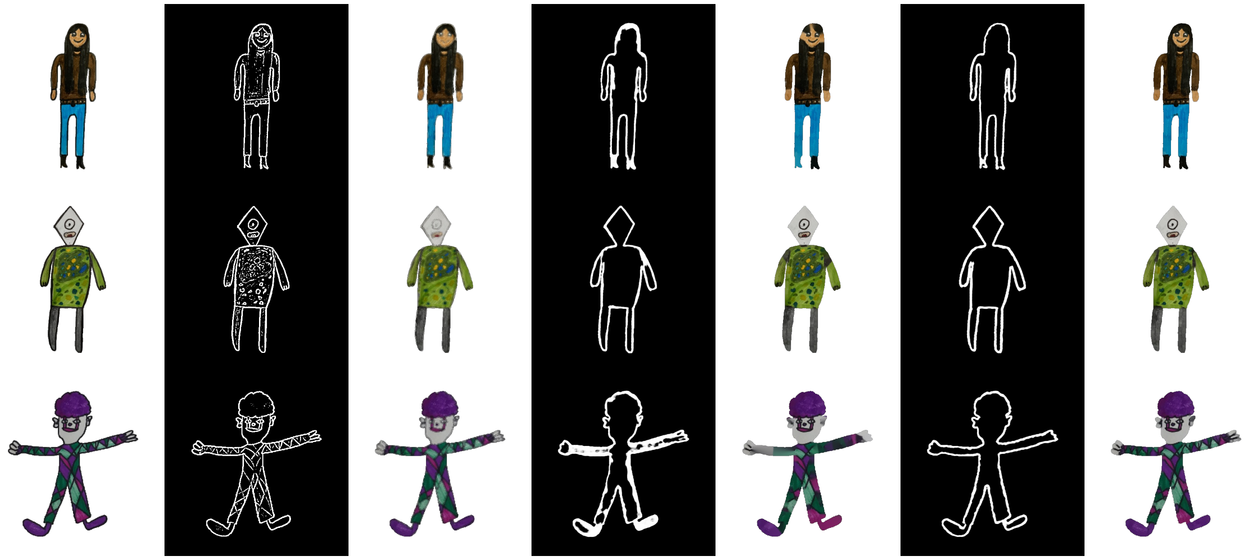
Since Wonder3D [2024] adopts an independent coordinate system for each object that is related to the input view, the posture of the reconstructed character may be slightly tilted, e.g., making the body lean forward or back. We mitigate this problem with a simple shearing transformation, as illustrated in Fig. 2. We adopt shearing instead of rotation because we want to keep the front-view boundary consistent with the input image. We project all vertices onto the YOZ plane. Then, we calculate the eigenvectors with the PCA algorithm and get the incline angle $theta$. So we can know the displacement in the z-direction of each vertex and update the z value of each vertex:

$$z_{shear} = z - y \cdot \tan\theta. \quad (1)$$

*Corresponding author.
Authors' addresses: Jie Zhou, jzhou67-c@my.cityu.edu.hk, City University of Hong Kong; Chufeng Xiao, chufengxiao@outlook.com, Hong Kong University of Science and Technology; Miu-Ling Lam, miu.lam@cityu.edu.hk, City University of Hong Kong; Hongbo Fu*, fuplus@gmail.com, Hong Kong University of Science and Technology.

SA Conference Papers '24, December 3–6, 2024, Tokyo, Japan

ACM ISBN 979-8-4007-1131-2/24/12
https://doi.org/10.1145/3680528.3687593



(a) Input (b) DoG + FastMarching (c) Pix2PixHD + FastMarching (d) FFC-ResNet + FastMarching

Fig. 1. Comparison of different backbones for contour removal.

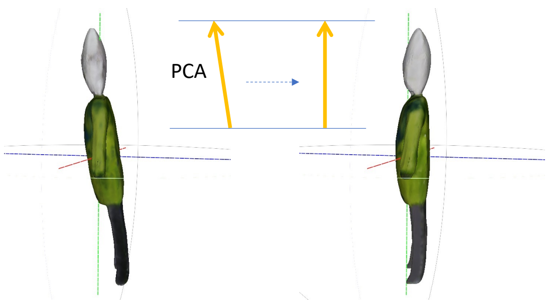


Fig. 2. An illustration of shear transformation.

2.3 Different Deformation Strategies

In addition to the bilateral deformation approach we employed, we also experimented with single-side deformation strategies, as shown in Fig. 3. It can be observed that single-side deformation tends to generate distorted surfaces.

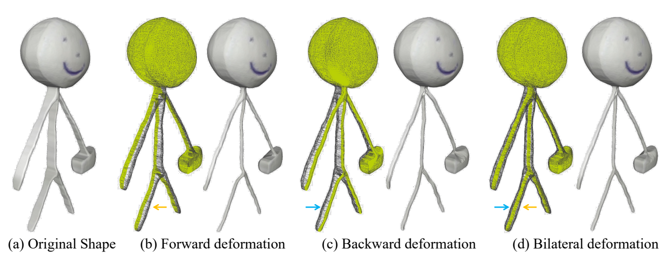


Fig. 3. Different deformation strategies.

REFERENCES

Blender Development Team. 2022. Blender (Version 3.1.0) [Computer software]. https://www.blender.org

Shuhong Chen, Kevin Zhang, Yichun Shi, Heng Wang, Yiheng Zhu, Guoxian Song, Sizhe An, Janus Kristjansson, Xiao Yang, and Matthias Zwicker. 2023. PAniC-3D: stylized single-view 3D reconstruction from portraits of anime characters. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 21068–21077.

Xiaoxiao Long, Yuan-Chen Guo, Cheng Lin, Yuan Liu, Zhiyang Dou, Lingjie Liu, Yuexin Ma, Song-Hai Zhang, Marc Habermann, Christian Theobalt, et al. 2024. Wonder3d: Single image to 3d using cross-domain diffusion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 9970–9980.

Zhongjin Luo, Shengcai Cai, Jinguo Dong, Ruibo Ming, Liangdong Qiu, Xiaohang Zhan, and Xiaoguang Han. 2023. RaBit: Parametric Modeling of 3D Biped Cartoon Characters with a Topological-consistent Dataset. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 12825–12835.

Roman Suvorov, Elizaveta Logacheva, Anton Mashikhin, Anastasia Remizova, Arsenii Ashukha, Aleksei Silvestrov, Naejin Kong, Harshith Goka, Kiwoong Park, and Victor Lempitsky. 2022. Resolution-robust large mask inpainting with fourier convolutions. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*. 2149–2159.

Ting-Chun Wang, Ming-Yu Liu, Jun-Yan Zhu, Andrew Tao, Jan Kautz, and Bryan Catanzaro. 2018. High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.