Authoring 2.5D Designs with Depth Estimation

Xia Su[*]
Paul G. Allen School of Computer Science and Engineering
University of Washington
Seattle, Washington, USA
xiasu@cs.washington.edu

Cuong Nguyen
Adobe Research
San Francisco, California, USA
cunguyen@adobe.com

Matheus A Gadelha
Adobe Research
San Jose, California, USA
gadelha@adobe.com

Yu Shen
Adobe Research
San Jose, California, USA
shenyu@adobe.com

Stefano Petrangeli
Adobe
San Jose, California, USA
petrange@adobe.com

Jon E. Froehlich
Paul G. Allen School of Computer Science & Engineering
University of Washington
Seattle, Washington, USA
jonf@cs.uw.edu

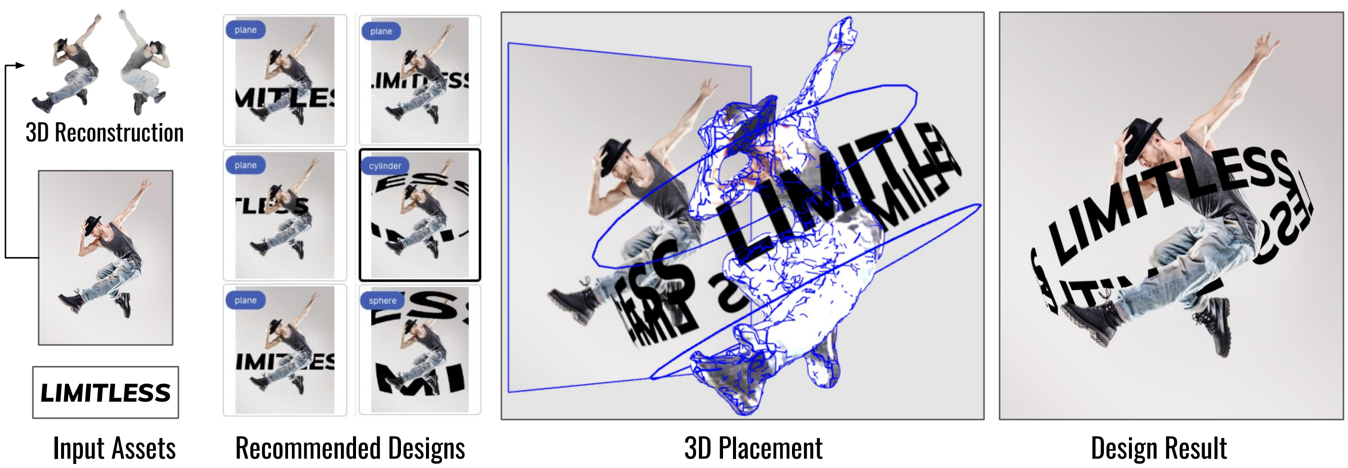

Figure 1: We introduce *DepthScape*, a Human-AI collaborative authoring system for 2.5D visual designs. DepthScape takes input images and use 3D reconstruction to estimate its inherent depth information. With AI-assisted design recommendation, users can quickly layout design elements in the implicit 3D space. The output is a visual design with realistic perspective and occlusion effects following depth cues in the input image.

Abstract

2.5D effects like occlusion and surrounding, which incorporate 3D blending in 2D designs, enhance visual dynamics and realism. However, it is hard to create 2.5D effects due to complexity of human depth perception and the labor intensive task of realistically conforming design elements to real-world perspective and occlusion. We introduce *DepthScape*, a Human-AI collaborative authoring pipeline that creates 2.5D visual designs by placing design elements in an estimated depth space. Leveraging a single-view 3D reconstruction model, DepthScape processes input images into 3D scenes, enabling realistic 2.5D effects by positioning and rendering visual elements in the 3D scene. DepthScape also suggests element placement in the 3D space based on examples from our design library, and provides interface for users to further fine tune the design. Our user study among 9 users shows that DepthScape is capable of creating sophisticated visual effects, accelerating the authoring process, and also inspiring new ideas.

[*]Work completed during an internship at Adobe Research.

CHI EA '25, Yokohama, Japan

ACM ISBN 979-8-4007-1395-8/25/04
https://doi.org/10.1145/3706599.3719727

CCS Concepts

• **Human-centered computing** → **Interactive systems and tools**.

Keywords

2.5D Design, Depth Estimation, Design Tool, Creativity Support

ACM Reference Format:
Xia Su, Cuong Nguyen, Matheus A Gadelha, Yu Shen, Stefano Petrangeli, and Jon E. Froehlich. 2025. Authoring 2.5D Designs with Depth Estimation. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25), April 26–May 01, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3706599.3719727

1 Introduction

Humans gain 3D perception from 2D images with depth cues like occlusion and shading [3, 7, 15]. Leveraging this visual ability, creators have long been applying depth-enhancing techniques such as layering, projecting, and shading, to further improve visual realism in 2D designs. Borrowing from graphics research [13, 19], we call this type of visual design "*2.5D design*". However, creating such powerful visual designs can be challenging. For instance, when a designer wants to wrap text around a human figure, they could use a 2D design tool to apply transformations to the text pixels to create perspective and occlusion effects, which is tedious and error-prone. There are also alternative 3D workflows of modeling and rendering, which has a steep learning curve for 2D designers. In both cases, exploring the design space and iterating on the design is very time-consuming, leaving much of the potential design space unexplored and limiting creativity.

How do we make it easier for 2D designers to craft these effects? Inspired by the aforementioned visual perception theories and further validated by hundreds of professional design examples, we summarize a key design rule of 2.5D visual designs: creating realistic perspective and occlusion effects to integrate 2D elements seamlessly into 3D scenes. For instance, a text element partially occluded by a human figure, or a circular text with a surrounding effect that goes behind the figure while other parts stay in front. Building on these insights, we propose a simplified 2.5D design paradigm: inferring depth from user-provided input images to create an *implicit 3D space*, where users can place design elements to achieve perspective and occlusion. This approach reduces the complexity of creating realistic effects, allowing designers to focus on creativity rather than manual adjustments.

We introduce DepthScape, a graphic authoring system that leverages depth estimation and 3D rendering to create 2.5D designs. We utilize a single-view-to-3D reconstruction model [18] to transform input images into 3D scenes, creating an *implicit 3D space* where other 2D elements, like text, icons, and shapes, can be placed. Here we define this *implicit 3D space* as a 3D reconstructed space that is not directly interactive with user but encodes critical depth cues. We tune the rendering pipeline to make the implicit 3D scene and other design elements occlude each other following their 3D placement, while the 3D scene itself remains invisible and reveals the raw pixels of the input image. In this case, we avoid distorting, layering, and masking workload and create an enlarged design space where novel 2.5D designs can be easily explored.

To evaluate DepthScape, we conducted a user study with nine participants, guiding them to first replicate example 2.5D designs and then freely explore with the DepthScape system. Our interview about the user experience shows positive overall rating of DepthScape's usability, especially in terms of creativity. Suggestions regarding the interface and the technical pipeline are also collected to support future improvements.

In summary, this work makes three primary contributions: (1). A novel pipeline for creating 2.5D visual design leveraging depth estimation. (2). An authoring interface with AI suggestion. (3). An exploratory system evaluation and findings that supports future improvements.

2 Related Works

Humans gain 3D perception from not only binocular vision, but also monocular cues to depth, such as linear perspective, interposition, Gestalt principles, and shadows [3, 7, 15]. Techniques based on this instinctive human ability are widely applied in graphics research and engineering. From 1970s, developers have been using methods like scaling sprites [2] and parallax scrolling [1] to create pseudo-3D effects in arcade games. Similar techniques are also widely applied in visual arts from paintings to modern designs of posters, illustrations, websites, *etc.*

In addition to using design techniques and graphic technologies to achieve depth perception, researchers have increasingly leveraged real estimated depth to assist creative tasks. With recent advancements in computer vision, the depth information of a single image can now be easily obtained, whether as a depth map [16, 22, 23, 27] or a 3D model [18, 20, 21, 25]. These output results can be applied in various creative applications, like *ZoomShop* [10], which uses image depth information to edit image composition, enlarge distant objects, and adjust relative size and positions of objects. Similarly, *VideoDoodle* [24] uses depth information of video scenes to enable hand-drawn animations into the video. Lu et al.[11] uses depth map to better capture key contours in the input image during image vectorization.

Following this thread of work, we gain insights on how depth information can be helpful to creation tasks: depth information provides spatial details that ensure content edits true to reality. Since 2.5D designs are also heavily dependent on depth perception and aims for visual realism, utilizing the estimated depth information to simplify the image editing workload for 2.5D effects creation becomes a natural solution.

3 The Formulation of DepthScape

To build a system for 2.5D design creation, three key questions must be addressed:

What are the target 2.5D effects? By analyzing over 200 professional 2.5D poster designs from Pinterest, we identified common patterns: Most designs feature one main object and multiple text or shape elements positioned around the main object. Meanwhile, element placement typically falls into three categories: *Plane* (flat surface), *Surrounding* (warped around the main object), and *Decaling* (rendered onto 3D surfaces). See Figure 2.

What are the key barriers to creating these 2.5D effects? Designers often use image editing tools to segment, mask, arrange, and warp elements, which is time-consuming and requires precise pixel transformations for 3D perspectives and occlusion effects. Alternatively, 3D workflows (e.g., photogrammetry, modeling) ensure realistic results but demand significant expertise and effort to switch between 2D and 3D tools [5, 6, 26], complicating spatial reasoning. These challenges make design iterations tedious and limit exploration of the design space.

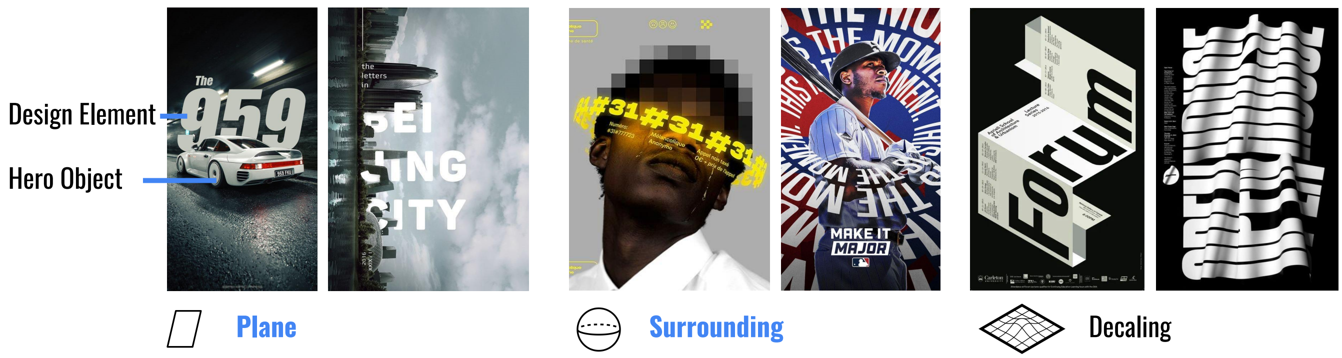

Figure 2: Three common types of design observed from design library. We focus on plane and surrounding for this current stage.

How can we resolve these barriers and accelerate the design process? Inspired by prior work that utilize depth information to edit images and videos [10, 24], leveraging depth information from input images offers a promising solution. By estimating depth, we can construct an implicit 3D space to place design elements, providing three key advantages: (1) Occlusion effects are achieved through depth-axis placement, eliminating the need for layering and masking. (2) Realistic perspectives are achieved via 3D positioning and rotation, avoiding manual distortion. (3) Parameterized placement simplifies operation, automation, and analysis.

4 The DepthScape System

Building on the above discussion, we introduce *DepthScape*, a human-AI collaborative authoring pipeline that simplifies 2.5D graphic design through depth estimation. DepthScape constructs implicit 3D spaces based on visual depth cues in input images, enabling the placement and 3D deformation of 2D design elements to achieve realistic occlusion and perspective effects. The system accelerates 2.5D design with design templates and AI-driven recommendations. Additionally, DepthScape offers a tuning and visualization interface for fine adjustments, supporting iterative exploration of the design space.

4.1 Creating an implicit 3D space

We use depth estimation to create an implicit 3D space that reflects visual depth cues in a given image. We adopt a state-of-the-art single image 3D reconstruction model *CRM* [18] to reconstruct input images into high quality 3D mesh models, which also aligns with the input image in a front orthographic view. This created *implicit 3D space*, which isn't directly manipulated by users but encodes critical depth cues, can host other visual elements to render realistic occlusions and natural perspectives (Figure 4A), thus creating 2.5D effects with ease.

However, standard 3D rendering pipelines reveals the generated 3D mesh (Figure 4C), which has lower graphic quality compared to the original input image. Thus, we engineer on the rendering pipeline to conceal the reconstructed 3D model while still enabling it to occlude other design elements. (See Figure 4B) Specifically, we first render the input image as a background of the canvas, then render the entire implicit 3D space while disabling color pixel writing of the reconstructed 3D models. This creates the illusion that objects in the original input images are occluding other design elements. See Figure 4D for an example result.

4.2 Placing design elements

In the implicit 3D space, designers can place visual elements in two steps: (1) **Position primitive surfaces** using 10 parameters, including the surface type (plane, cylinder, or sphere), and its position, rotation, and scale along the x, y, and z axes. (2) **Place design elements** (PNG images) on the primitive surfaces using 5 parameters: scale (u, v), offset (u, v), and rotation.

This results in 15 parameters per element, creating a vast design space but also posing challenges for users unfamiliar with 3D editing. To simplify, DepthScape employs parameter fitting and normalization. Primitive surfaces are fitted to the reconstructed 3D object, and all position and scale parameters are normalized to the bounding box. This provides a default set of parameters, offering a baseline design to streamline the process.

4.3 AI-recommendation of design

With normalized parameters, each element placement can be represented by the aforementioned 15 parameters, forming a reusable design template. To facilitate quick exploration of the design space, we built a design library of example templates and further implemented an AI recommendation system using CLIP embeddings [17]. When a user provides an input image, DepthScape compares its CLIP embedding with those of images in the design library, retrieving designs with the most similar input images. These recommended designs, as parameterized design templates, are rendered with user provided image assets and displayed as thumbnails, accelerating design exploration and inspiring creativity.

4.4 The DepthScape Interface

The DepthScape interface is currently implemented as a web demo (Figure 3) using React [4] and Babylon.js [14]. The interface has three main parts: (1) the Asset Panel manages imported assets and shows suggested design thumbnails; (2) the canvas renders the 2.5D effects, (3) the Edit Panel enables fine-tuning of the design via a series of categorized sliders.

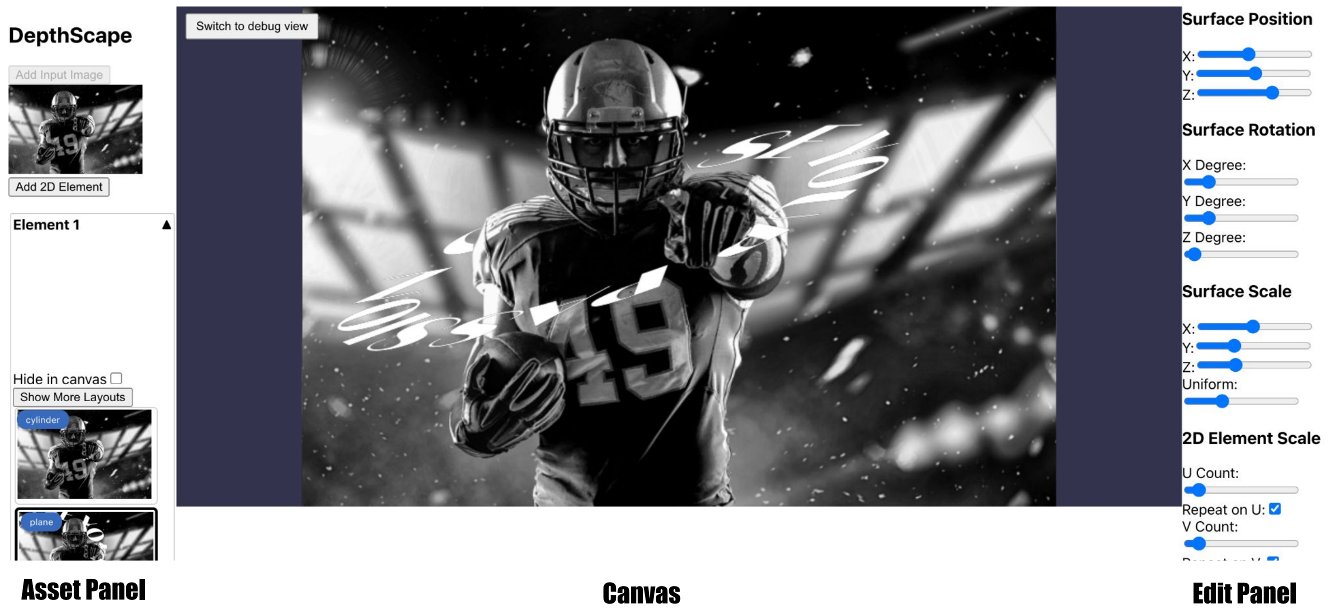

Figure 3: DepthScape's User Interface

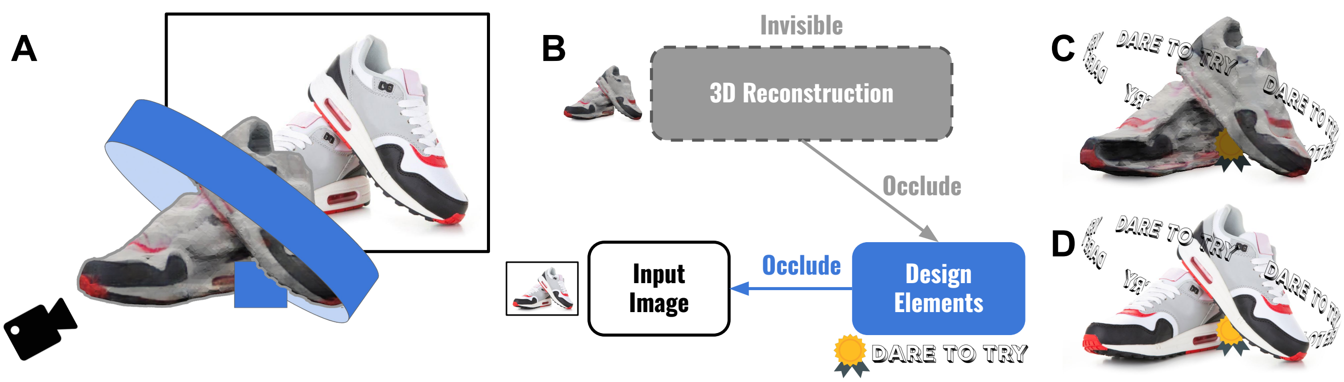

Figure 4: DepthScape's implicit 3D space and rendering setting. (A) A diagram showing the composition of the implicit 3D space, including the raw input image, the 3D reconstruction model, design elements, and an orthographic camera generating design results. (B) The occlusion relations in the rendering pipeline. (C) A design result in default render setting. (D) A design result in engineered render setting.

To use DepthScape, the user starts by importing an input image, which is processed into a 3D scene in real time. Next, they can add 2D assets like text, icons, or shapes to the 2.5D design. For each asset, DepthScape generates a series of AI-recommended layouts, displayed as thumbnails in the Asset Panel. Users can click these thumbnails to apply a selected design. Further adjustments to element placement—such as orientation, size, and repetition—can be made using the Edit Panel. To aid placement, a debug view is available, showing the canvas with a perspective camera to reveal the implicit 3D space. See the supplementary video for a demonstration.

5 User Study

To examine the usability and authoring performance of the DepthScape system, we conducted a user study across nine participants (eight male and one female, aged 20-60), who tried the DepthScape

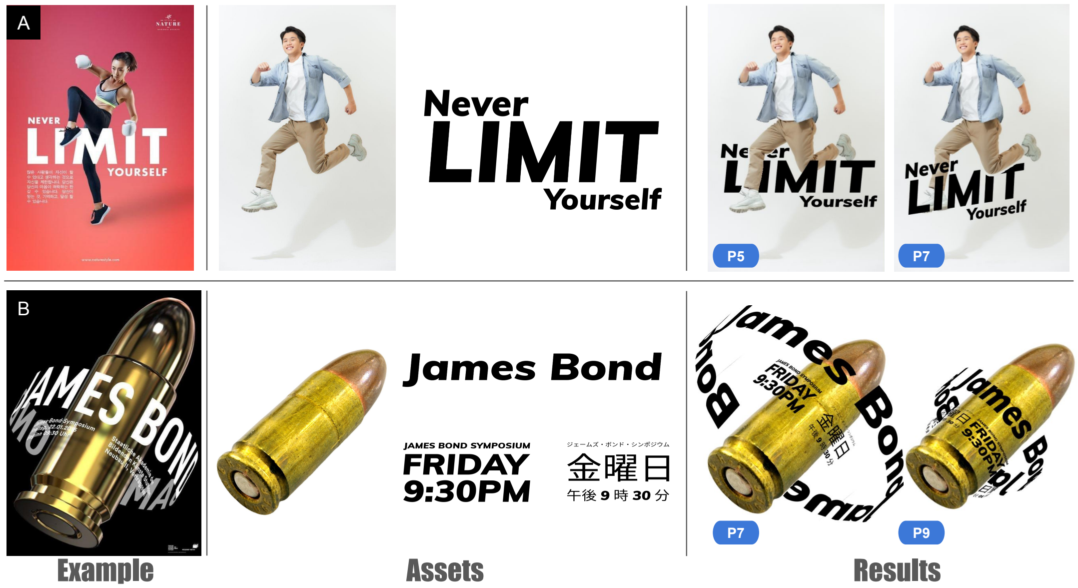

Figure 5: Replication tasks and results of the user study. Left: replication target; Middle: image assets for task A and B; Right: selected design results.

system and evaluated their experiences. Among them four have professional design experience in visual and graphics. All participants self-reported high 2D visual design expertise, while six have moderate to professional 3D design skills.

We first introduced the project motivation and the DepthScape interface, then guided participants to recreate two example designs with provided image assets (Figure 5). After they finish, we asked follow-up questions about their experiences. We then encouraged participants to explore the DepthScape system freely and create designs as they wish. Lastly, we followed-up the design experience with more usability questions and Likert-scale ratings.

On average, the study lasted 61 minutes and 30 seconds. All participants successfully utilized DepthScape to create multiple 2.5D designs. In total, participants created 18 open-ended designs, with an average of 2 designs per participant. Most participants found the first replication task (Figure 5A) fairly simple and straightforward to finish with the AI recommended layout and the edit panel. The second replication task (Figure 5A) was more challenging and time-consuming compared to the first one, especially for the alignment of cylinder surfaces to the bullet shape. Many participants used the debug view to support 3D alignment. P8 commented *"I like the debug view because you could actually see how curved (the surface) is or if it's pointing downwards, but in 2D I can only see the final render."*

Participants produced diverse and visually compelling designs in the open-ended phase (Figure 6). Designs, like P1-2, P3, and P4, though created quickly and aimed for relatively simple depth cues, achieved realistic partial occlusion effects due to the accurate depth information in the *implicit 3D space*. More complex designs, like P1-1, P2, P6, and P8, were also quickly created leveraging repetition and surrounding effects. Some designs, such as the helix design in P2 and the purple haze in P8, were created beyond expectations when exploring the design space. By accidentally editing the element repetition and rotation, P8 unexpectedly repurposed a rectangle frame into a haze effect behind the car.

From these design explorations, we observe three ways that DepthScape supports creativity: (1) the accurate depth information in the *implicit 3D space* simplifies the creation of realistic occlusion effects; (2) the capability of rendering complex deformation and repetition effects in 3D space supports the creation of complex but orderly perspective effects. (3) the parameters of design element placement creates a design space that is big enough to contain serendipitous designs, while still being easy to explore.

We received a positive overall rating for usability, especially for the creativity support of the system. Participants highly agree that *"I felt that the prototype supported my creativity."* (avg = 4.67/5). P8 commented *"I can do anything I want and try out things"*. Participants also commonly agreed that DepthScape makes occlusion (avg = 4.56/5) and surrounding effects (avg=4.56/5) easy to achieve. P2 said *" (In existing editing tools) If I want a slightly different rotation, I had to redo everything from scratch. But in DepthScape I can just change the parameter"*.

Besides the positive results, there were also aspects that can be improved. One main improvement is in the result quality (avg=2.22/5), which was mainly due to misalignments of 3D models with input images and the rendering quality of elements. Another often wished improvement was in the the general usability of interface (avg=3.56/5), mainly due to the challenge in finding the right sliders for the desired adjustments, and quality of AI recommended designs (avg= 3.67/5). P2 commented: *"The sliders creates cool effects, but they are sometimes unintuitive"*. Participants also suggested new features that can improve usability, like enabling direct manipulation in the canvas, having an undo function, enabling parameter

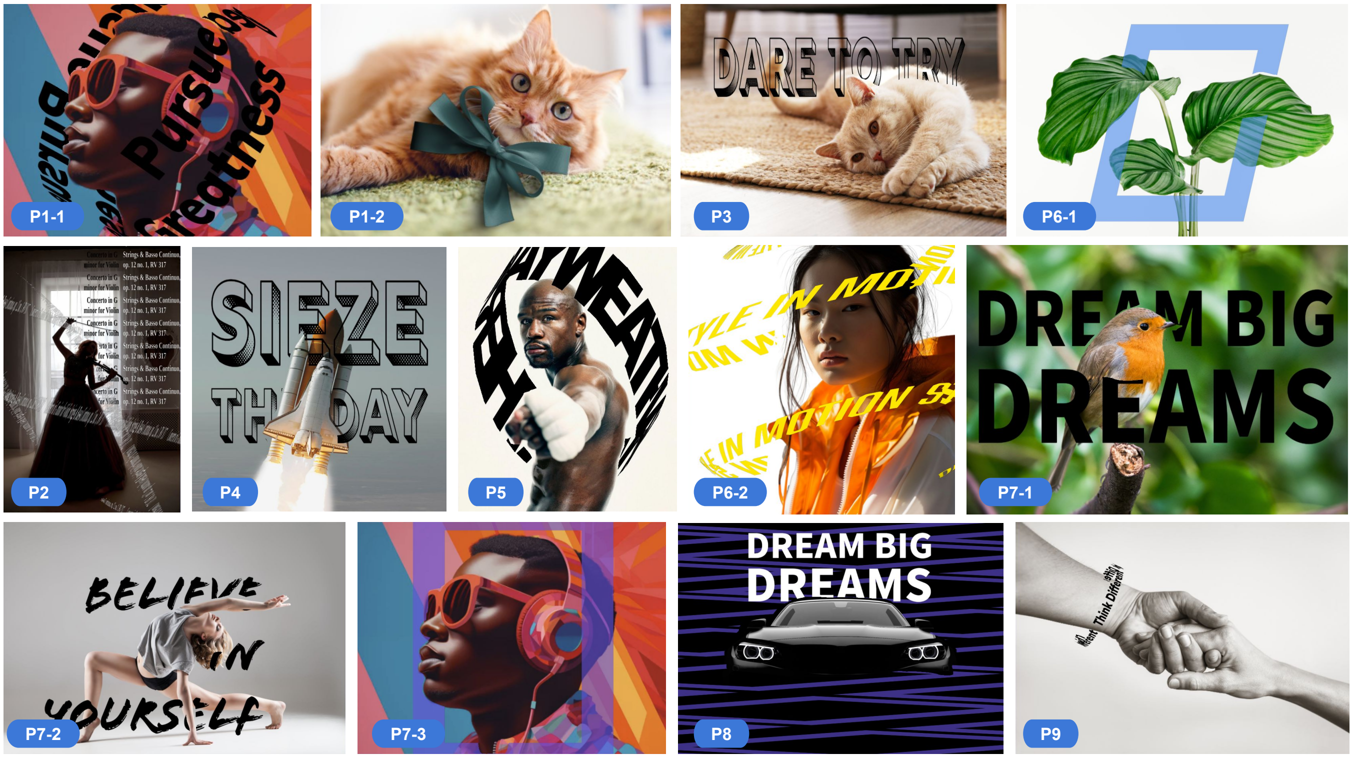

Figure 6: Open-ended designs from study participants.

syncing to align multiple elements, enabling perspective rendering for better 3D effects, etc.

6 Discussion

Application Scenarios: As suggested by eight of our nine user study participants, DepthScape's 2.5D design pipeline could be integrated into existing image editing tools like Adobe Photoshop, Adobe Express, and Figma. Its depth-based editing capabilities would introduce an inspiring new functionality to these platforms, while the advanced image and text editing features of these platforms could complement current limitations of the DepthScape prototype.

Beyond static image design, we also envision DepthScape being applied to animation creation. First, the 3D placement parameters of design elements in the implicit 3D space—such as rotation and position—could be tied to time to create simple motion animations. Additionally, recent advancements in video depth estimation [8, 9, 12] open the possibility of extending DepthScape's depth-based 2.5D creation pipeline into video, enabling the placement of visual elements within a 3D video space. This could lead to a powerful tool for dynamic and immersive content creation.

Limitations and Failure Cases: While DepthScape was positively evaluated in our user study and generated many visually compelling designs, we observed limitations in system performance and adaptability.

Firstly, as reflected in user study, professional designers wished for higher rendering quality. Currently, the 3D reconstruction of input image generated by CRM [18] sometimes have small misalignments with the input image, leading to offsets in the occlusion effects. Also, the alignment of 3D reconstruction from CRM to the original input image requires orthographic rendering, which limits the foreshortening in element rendering.

Secondly, the current depth estimation and alignment pipeline fails on complex images like urban, forest, and indoor scenes, or images without clear main objects. The depth quality is also impacted at unclear edges (*e.g.* furry, steamy, half-transparent, and unfocused object edges).

Thirdly, the current CLIP-based design recommendation only focuses on semantic information of the input image, neglecting the semantics in the design element image and user design goal. In this case, some recommended designs may mismatch the overall design semantics and user intention.

Future Work: We aim to explore better depth estimation methods that support more complex scenes and improve the alignment between input images and their corresponding 3D scenes. Additionally, we plan to enhance perspective rendering in the *implicit 3D space* to better showcase foreshortening and other perspective depth cues. We **do not** plan to enable users to edit the depth reconstruction directly.

To further advance the system, we intend to implement more sophisticated AI-driven design support by leveraging the latest vision-language models. By semantically understanding and extracting key objects from input images and locating their positions in the depth space, we can enable more context-aware placements of visual elements. For example, text layers could be aligned parallel to a building facade or oriented to face the forward direction of a human figure. These improvements will enhance the system's ability to create intuitive and visually coherent 2.5D designs.

7 Conclusion

Leveraging the latest advancements in depth estimation, we propose a novel 2.5D design pipeline that transforms depth cues in input images into implicit 3D spaces, enabling easy creation of realistic occlusion and surrounding effects. By offering AI-assisted element placement recommendations and a fine-tuning interface, users can efficiently create visually compelling 2.5D designs. A user study with 9 participants demonstrated strong creativity support, with participants praising the system's ease of use and creative potential. We believe DepthScape introduces a new design paradigm and has the potential to be integrated into professional image editing workflows.

References

[1] Giant Bomb. n.d.. Parallax Scrolling. https://www.giantbomb.com/parallax-scrolling/3015-2915/ Accessed: 2024-09-11.

[2] Giant Bomb. n.d.. Sprite Scaling. https://www.giantbomb.com/sprite-scaling/3015-7122/ Accessed: 2024-09-11.

[3] Eli Brenner and Jeroen B Smeets. 2018. Depth perception. In *Stevens' handbook of experimental psychology and cognitive neuroscience: Sensation, perception, and attention*. Wiley, 385–414.

[4] Artemij Fedosejev. 2015. *React. js essentials*. Packt Publishing Ltd.

[5] Taro Fujita, Yutaka Kondo, Hiroyuki Kumakura, Susumu Kunimune, and Keith Jones. 2020. Spatial reasoning skills about 2D representations of 3D geometrical shapes in grades 4 to 9. *Mathematics Education Research Journal* 32 (2020), 235–255.

[6] Kristin M Gagnier, Kinnari Atit, Carol J Ormand, and Thomas F Shipley. 2017. Comprehending 3D diagrams: Sketching to support spatial reasoning. *Topics in cognitive science* 9, 4 (2017), 883–901.

[7] Ian P Howard and BJ Rogers. 2002. Depth perception. *Stevens' handbook of experimental psychology Sensation and Perception* 1 (2002), 77–120.

[8] Wenbo Hu, Xiangjun Gao, Xiaoyu Li, Sijie Zhao, Xiaodong Cun, Yong Zhang, Long Quan, and Ying Shan. 2024. DepthCrafter: Generating Consistent Long Depth Sequences for Open-world Videos. arXiv:2409.02095 [cs.CV] https://arxiv.org/abs/2409.02095

[9] Johannes Kopf, Xuejian Rong, and Jia-Bin Huang. 2021. Robust consistent video depth estimation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 1611–1621.

[10] Sean J Liu, Maneesh Agrawala, Stephen DiVerdi, and Aaron Hertzmann. 2022. ZoomShop: Depth-Aware Editing of Photographic Composition. In *Computer Graphics Forum*, Vol. 41. Wiley Online Library, 57–70.

[11] Shufang Lu, Wei Jiang, Xuefeng Ding, Craig S Kaplan, Xiaogang Jin, Fei Gao, and Jiazhou Chen. 2019. Depth-aware image vectorization and editing. *The Visual Computer* 35 (2019), 1027–1039.

[12] Xuan Luo, Jia-Bin Huang, Richard Szeliski, Kevin Matzen, and Johannes Kopf. 2020. Consistent video depth estimation. *ACM Transactions on Graphics (ToG)* 39, 4 (2020), 71–1.

[13] D Man and A Vision. 1982. A computational investigation into the human representation and processing of visual information. *WH San Francisco: Freeman and Company, San Francisco* 1 (1982), 1.

[14] Julien Moreau-Mathis. 2016. *Babylon. js Essentials*. Packt Publishing Ltd.

[15] John Murray. 1994. Some perspectives on visual depth perception. *ACM SIGGRAPH Computer Graphics* 28, 2 (1994), 155–157.

[16] Suraj Patni, Aradhye Agarwal, and Chetan Arora. 2024. ECoDepth: Effective Conditioning of Diffusion Models for Monocular Depth Estimation. *ArXiv* abs/2403.18807 (2024). https://api.semanticscholar.org/CorpusID:268723777

[17] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, and Ilya Sutskever. 2021. Learning Transferable Visual Models From Natural Language Supervision. arXiv:2103.00020 [cs.CV] https://arxiv.org/abs/2103.00020

[18] Zhengyi Wang, Yikai Wang, Yifei Chen, Chendong Xiang, Shuo Chen, Dajiang Yu, Chongxuan Li, Hang Su, and Jun Zhu. 2024. Crm: Single image to 3d textured mesh with convolutional reconstruction model. *arXiv preprint arXiv:2403.05034* (2024).

[19] Jiajun Wu, Yifan Wang, Tianfan Xue, Xingyuan Sun, Bill Freeman, and Josh Tenenbaum. 2017. Marrnet: 3d shape reconstruction via 2.5 d sketches. *Advances in neural information processing systems* 30 (2017).

[20] Jiale Xu, Weihao Cheng, Yiming Gao, Xintao Wang, Shenghua Gao, and Ying Shan. 2024. Instantmesh: Efficient 3d mesh generation from a single image with sparse-view large reconstruction models. *arXiv preprint arXiv:2404.07191* (2024).

[21] Yinghao Xu, Zifan Shi, Wang Yifan, Hansheng Chen, Ceyuan Yang, Sida Peng, Yujun Shen, and Gordon Wetzstein. 2024. Grm: Large gaussian reconstruction model for efficient 3d reconstruction and generation. *arXiv preprint arXiv:2403.14621* (2024).

[22] Lihe Yang, Bingyi Kang, Zilong Huang, Xiaogang Xu, Jiashi Feng, and Hengshuang Zhao. 2024. Depth anything: Unleashing the power of large-scale unlabeled data. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 10371–10381.

[23] Lihe Yang, Bingyi Kang, Zilong Huang, Zhen Zhao, Xiaogang Xu, Jiashi Feng, and Hengshuang Zhao. 2024. Depth Anything V2. *arXiv preprint arXiv:2406.09414* (2024).

[24] Emilie Yu, Kevin Blackburn-Matzen, Cuong Nguyen, Oliver Wang, Rubaiat Habib Kazi, and Adrien Bousseau. 2023. VideoDoodles: Hand-Drawn Animations on Videos with Scene-Aware Canvases. *ACM Trans. Graph.* 42, 4, Article 54 (jul 2023), 12 pages. https://doi.org/10.1145/3592413

[25] Ruowen Zhao, Zhengyi Wang, Yikai Wang, Zihan Zhou, and Jun Zhu. 2024. FlexiDreamer: Single Image-to-3D Generation with FlexiCubes. *arXiv preprint arXiv:2404.00987* (2024).

[26] Tongyu Zhou, Joshua Kong Yang, Vivian Hsinyueh Chan, Ji Won Chung, and Jeff Huang. 2024. PortalInk: 2.5D Visual Storytelling with SVG Parallax and Waypoint Transitions. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology (UIST '24)* (Pittsburgh, PA, USA). Association for Computing Machinery (ACM), New York, NY, USA. https://doi.org/10.1145/3654777.3676376

[27] Ruijie Zhu, Ziyang Song, Li Liu, Jianfeng He, Tianzhu Zhang, and Yongdong Zhang. 2024. HA-Bins: Hierarchical Adaptive Bins for Robust Monocular Depth Estimation Across Multiple Datasets. *IEEE Transactions on Circuits and Systems for Video Technology* 34 (2024), 4354–4366. https://api.semanticscholar.org/CorpusID:265359133