Supplemental Material for ClipGen : A Deep Generative Model for Clipart Synthesis

I-Chao Shen Bing-Yu Chen
National Taiwan University

1. Network Architecture

1.1. Continue and location

Figure 18 shows the framework of combining three sources of input (i.e., visual representation of current canvas, target category vector and existing layer count) for predicting whether to continue adding a new layer and where to add it. In the decoder part, we include N deconvolution blocks (each of them composed of 3x3 transpose convolution with stride 2 followed by a Batch Normalization layer and a ReLU layer). The final prediction of probability is obtained using a 1x1 convolution of the upscaled state.

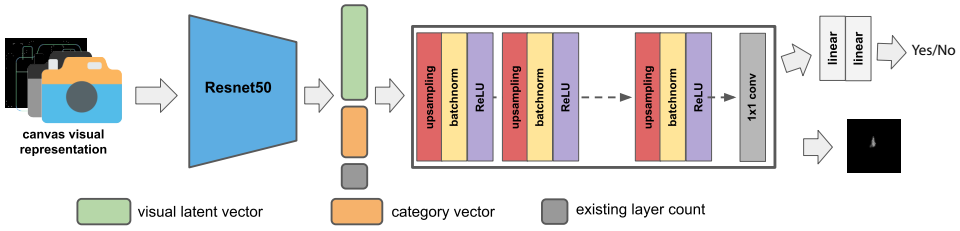

Figure 18. Architecture of the first module of our framework.

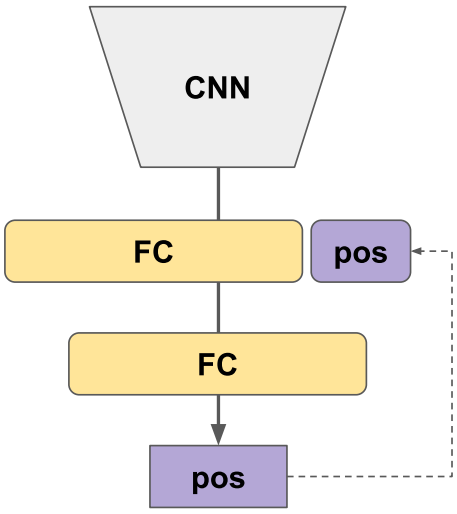

Figure 19. Architecture of the CNN network we used to compare with the RNN recurrent architecture. In each step, we predict an 2D position and feed it back to the network to predict the next 2D position.

2. Dataset

Currently, In *ClipNet*, we collect ten categories of man-made object cliparts. We show the example of each category in Figure 20. We share all the clipart we collected and a viewer at https://drive.google.com/file/d/1-0AM1_dLJ26MHIiD3VZFlxCm_1QN996z/view?usp=sharing.

Figure 20. Example cliparts of each category included in the *ClipNet* dataset.

3. Nearest Neighbor Visualization

We show the nearest neighbor cliparts in the dataset for all the synthesized cliparts shown in the main paper in Figure 21 and Figure 22.

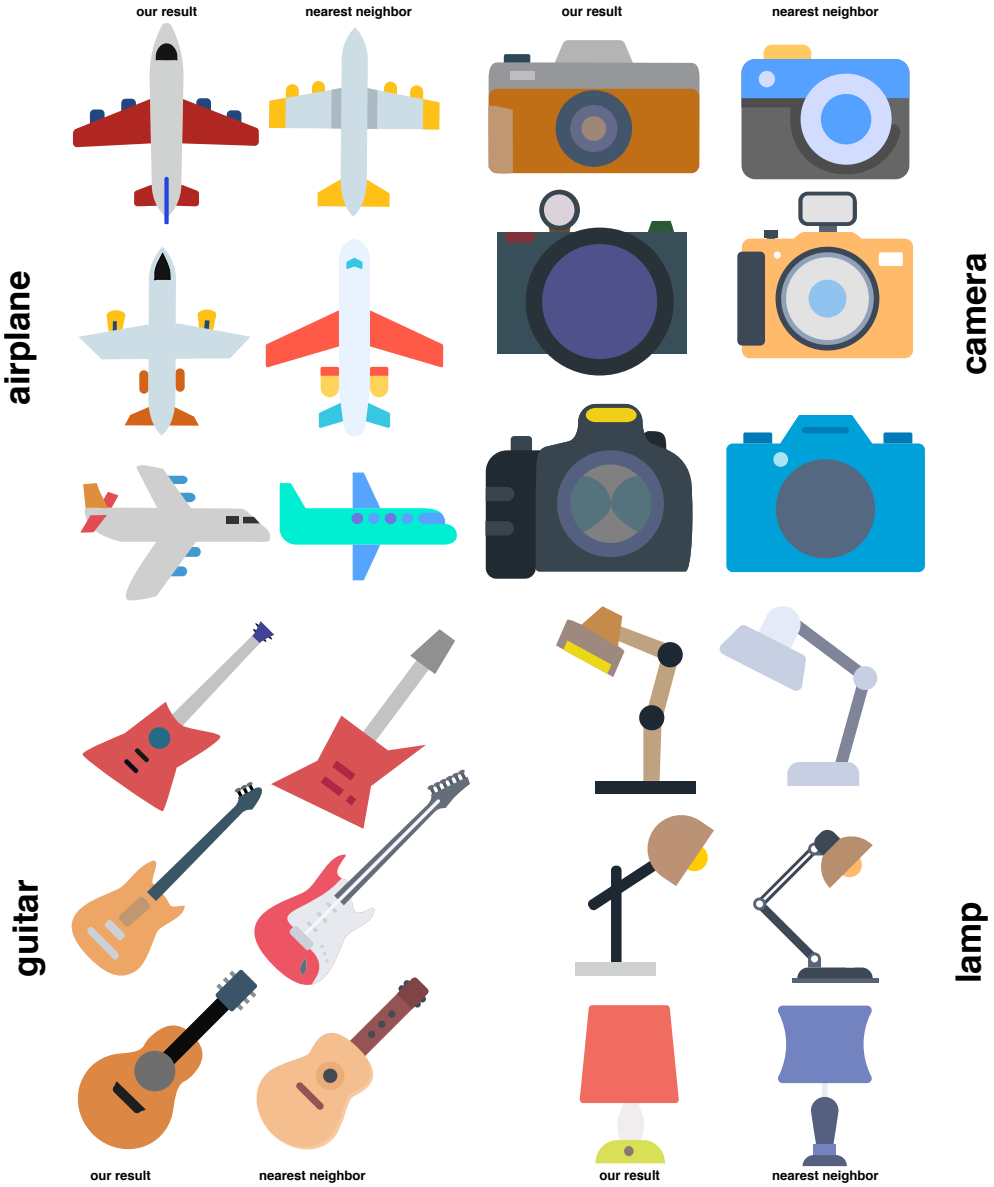

Figure 21. For each synthesized cliparts, we show it's nearest neighbor clipart in the dataset.

Figure 22. For each synthesized cliparts, we show it's nearest neighbor clipart in the dataset.

4. More comparison

We show more comparison with [Hoshyari et al. 2018] for synthesis results from Figure 23 to Figure 30.

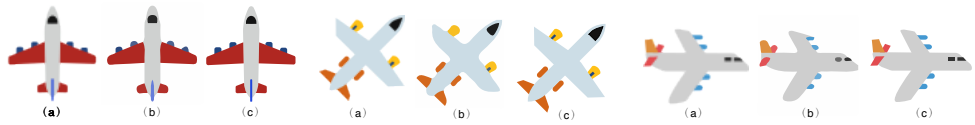

Figure 23. Comparison for airplane cases. (a) Input raster image, (b) result of Hoshyari [Hoshyari et al. 2018], and (c) our result.

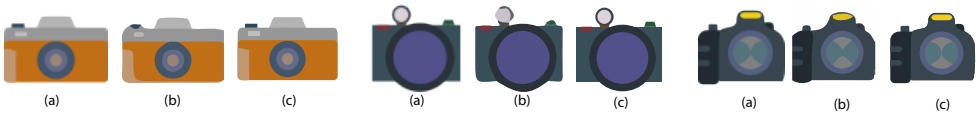

Figure 24. Comparison for camera cases. (a) Input raster image, (b) result of Hoshyari [Hoshyari et al. 2018], and (c) our result.

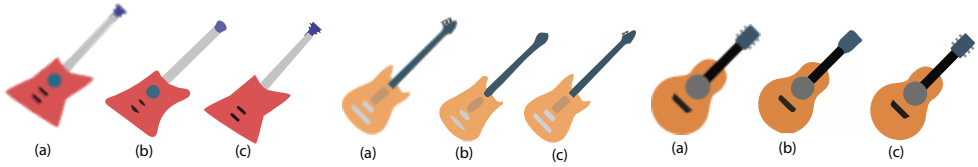

Figure 25. Comparison for guitar cases. (a) Input raster image, (b) result of Hoshyari [Hoshyari et al. 2018], and (c) our result.

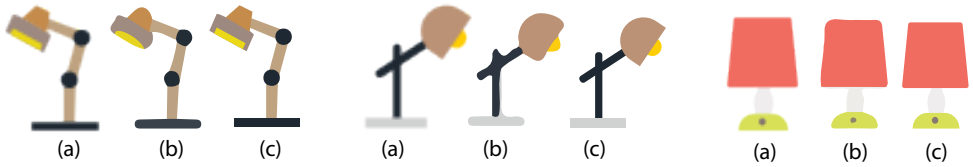

Figure 26. Comparison for lamp cases. (a) Input raster image, (b) result of Hoshyari [Hoshyari et al. 2018], and (c) our result.

5. Layering visualization of synthesis result

We followed the visualization method used in [Favreau et al. 2017].

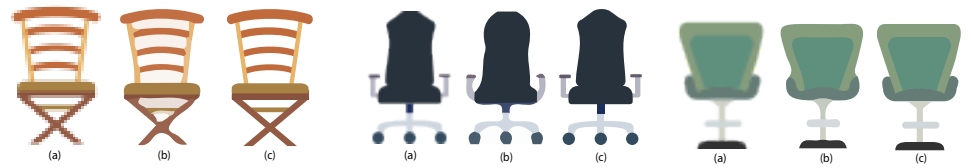

Figure 27. Comparison for chair cases. (a) Input raster image, (b) result of Hoshyari [Hoshyari et al. 2018], and (c) our result.

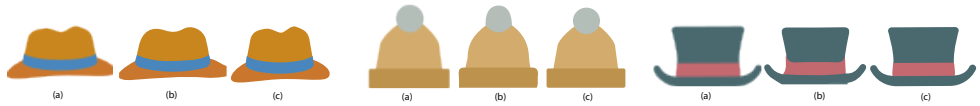

Figure 28. Comparison for hat cases. (a) Input raster image, (b) result of Hoshyari [Hoshyari et al. 2018], and (c) our result.

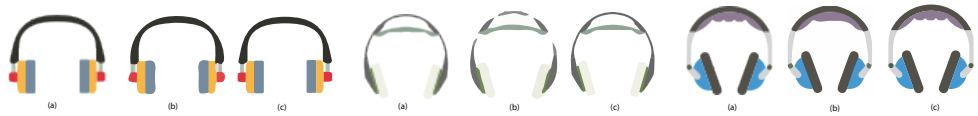

Figure 29. Comparison for headphone cases. (a) Input raster image, (b) result of Hoshyari [Hoshyari et al. 2018], and (c) our result.

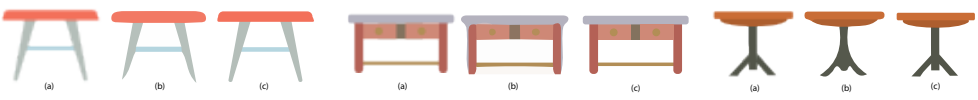

Figure 30. Comparison for table cases. (a) Input raster image, (b) result of Hoshyari [Hoshyari et al. 2018], and (c) our result.

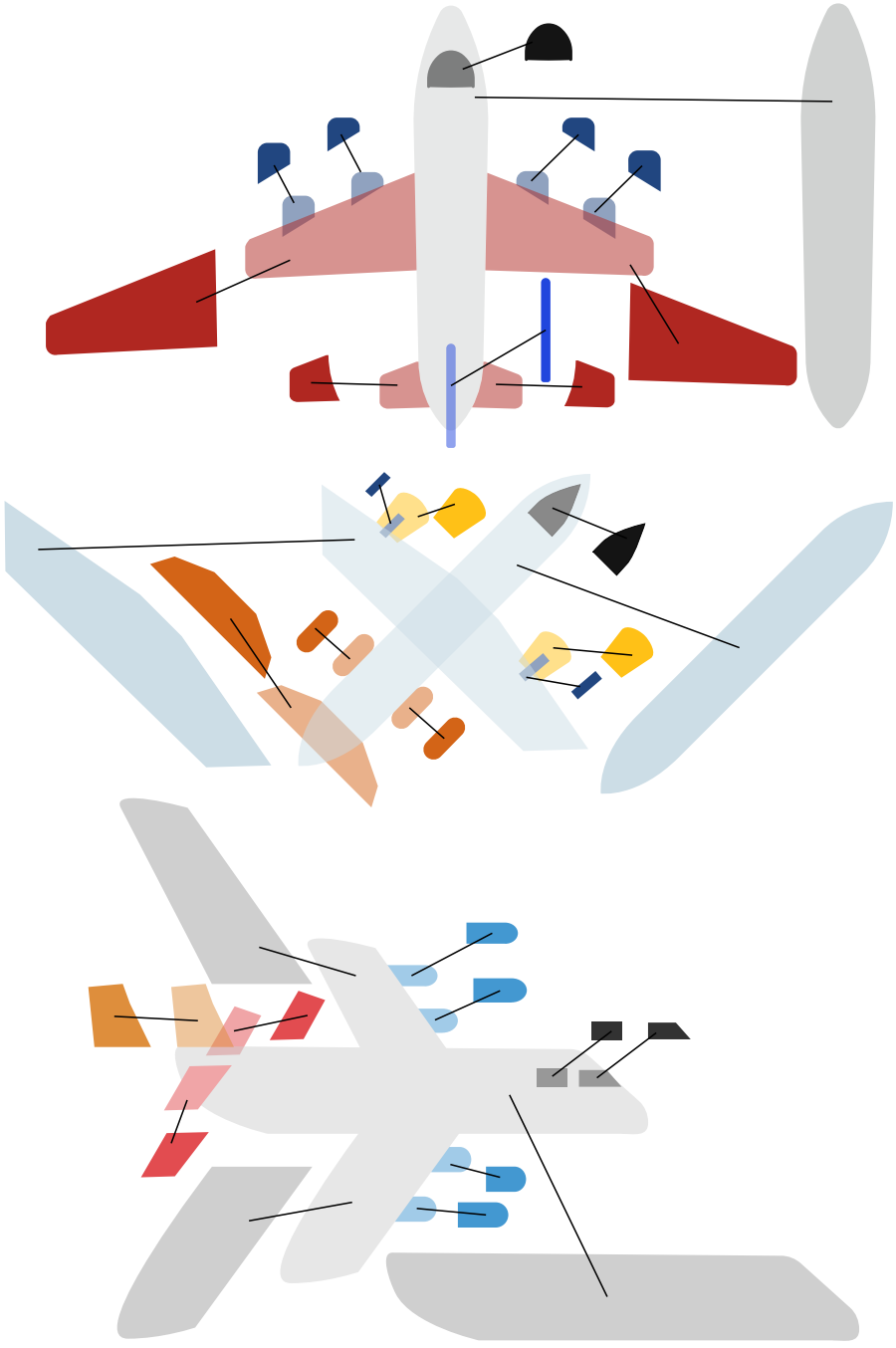

Figure 31. Layering of the synthesized airplane clipart

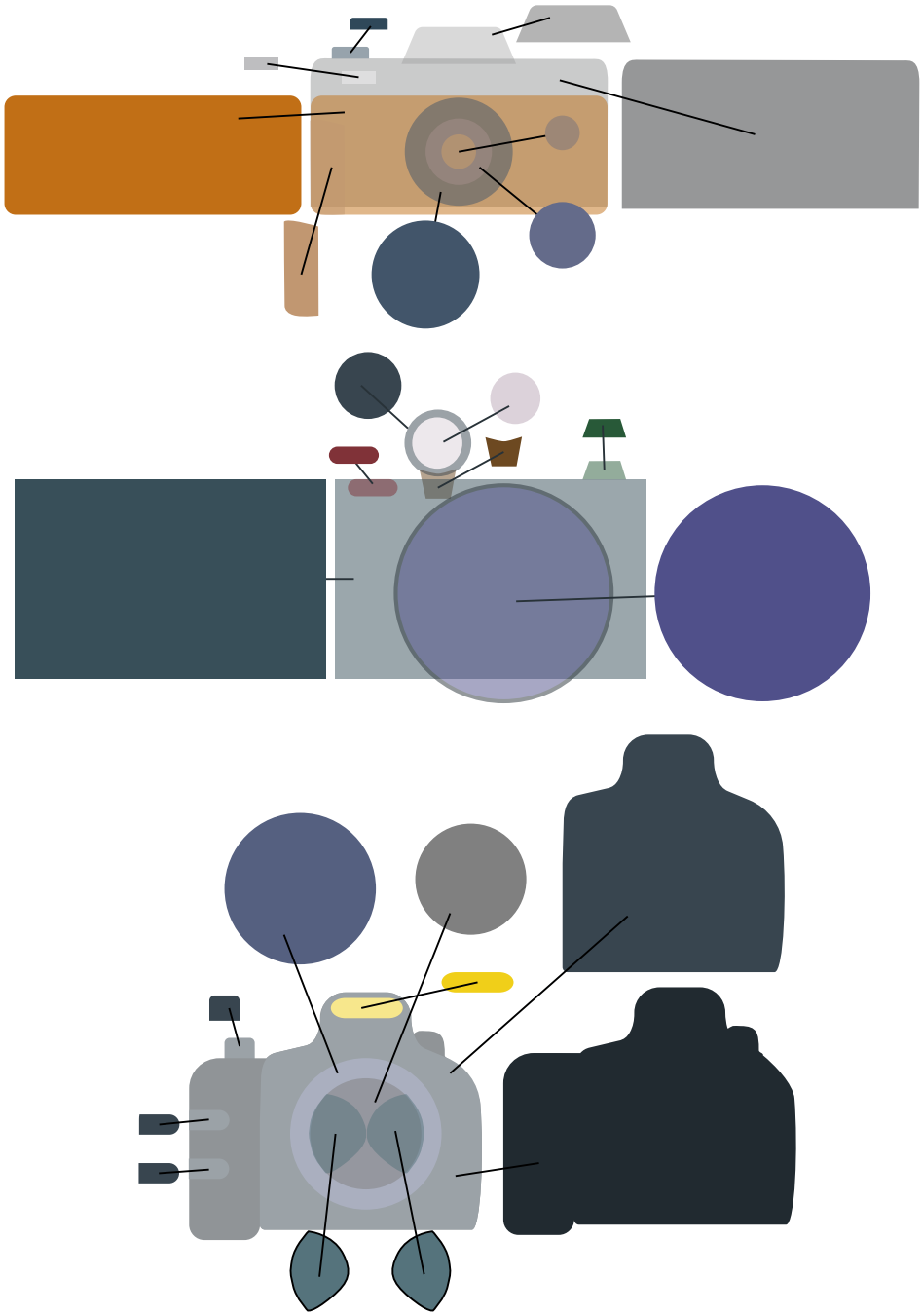

Figure 32. Layering of the synthesized camera clipart.

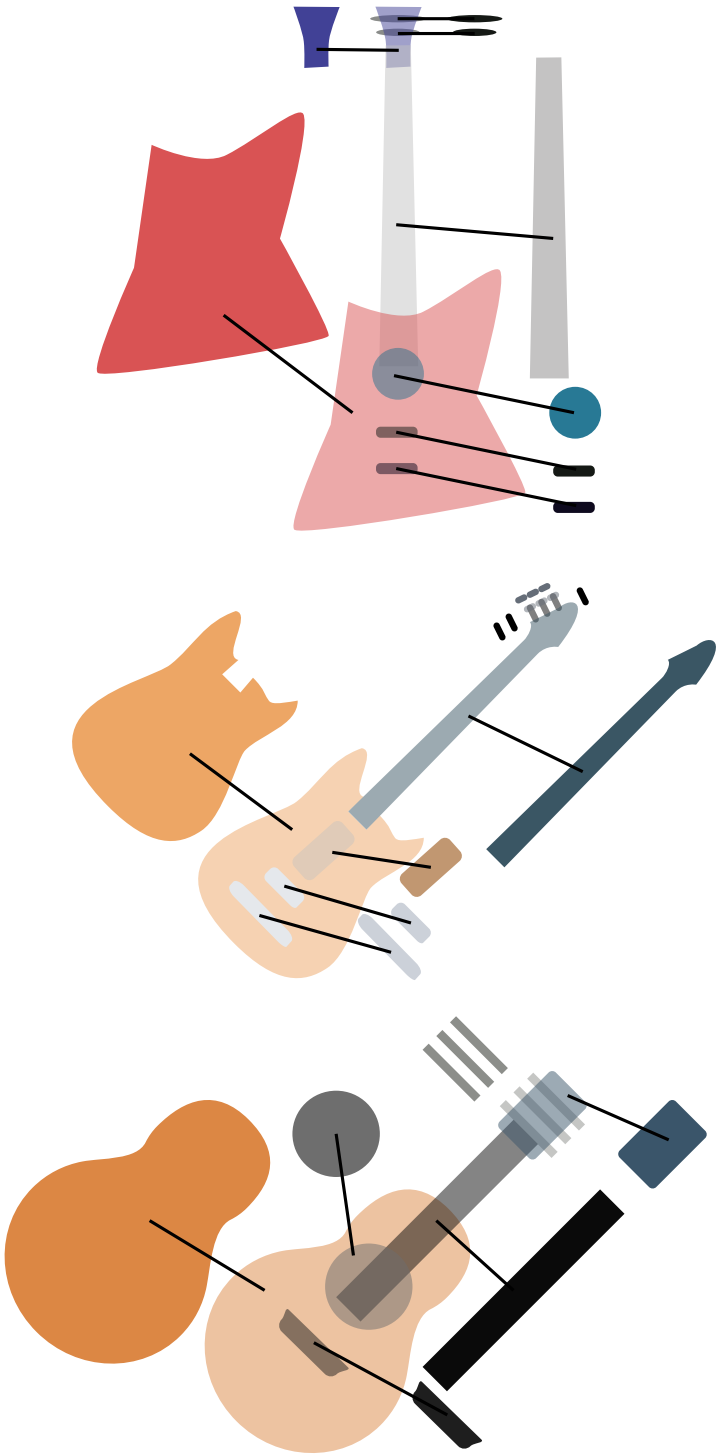

Figure 33. Layering of the synthesized guitar clipart.

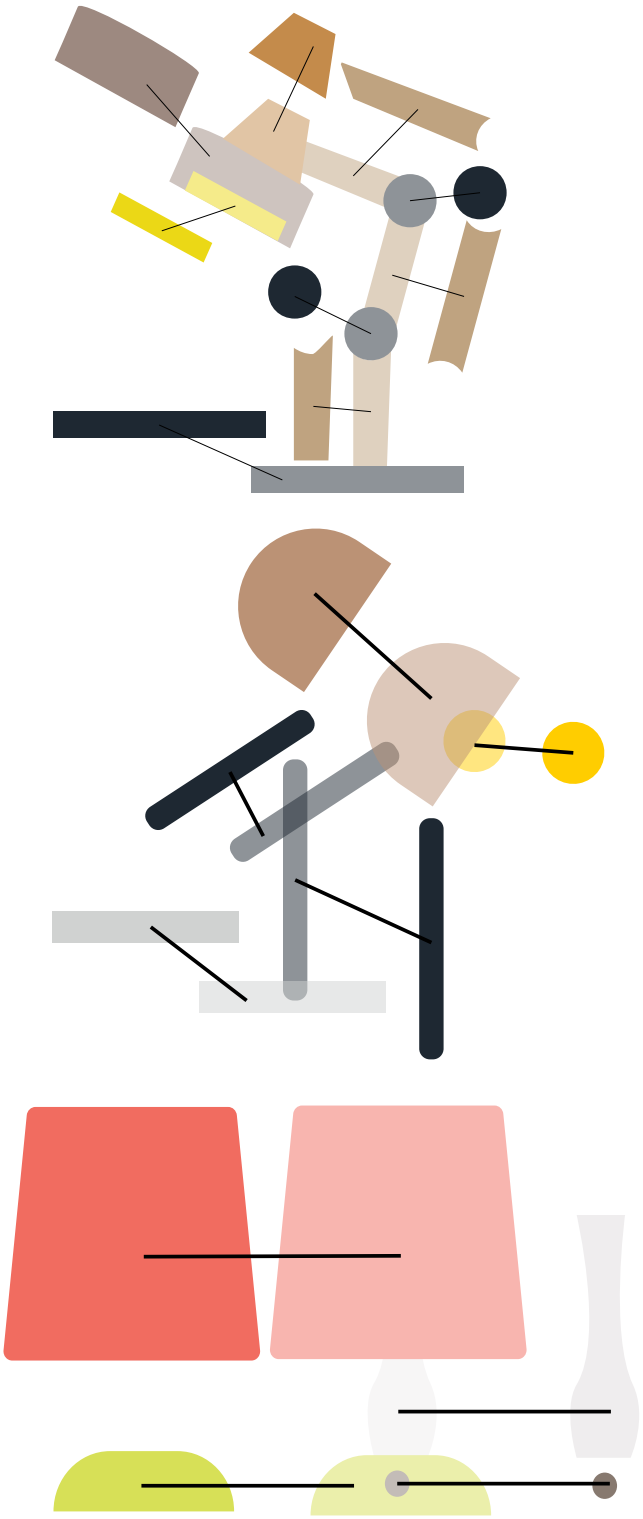

Figure 34. Layering of the synthesized lamp clipart.

Figure 35. Layering of the synthesized chair clipart.

Figure 36. Layering of the synthesized headphone clipart.

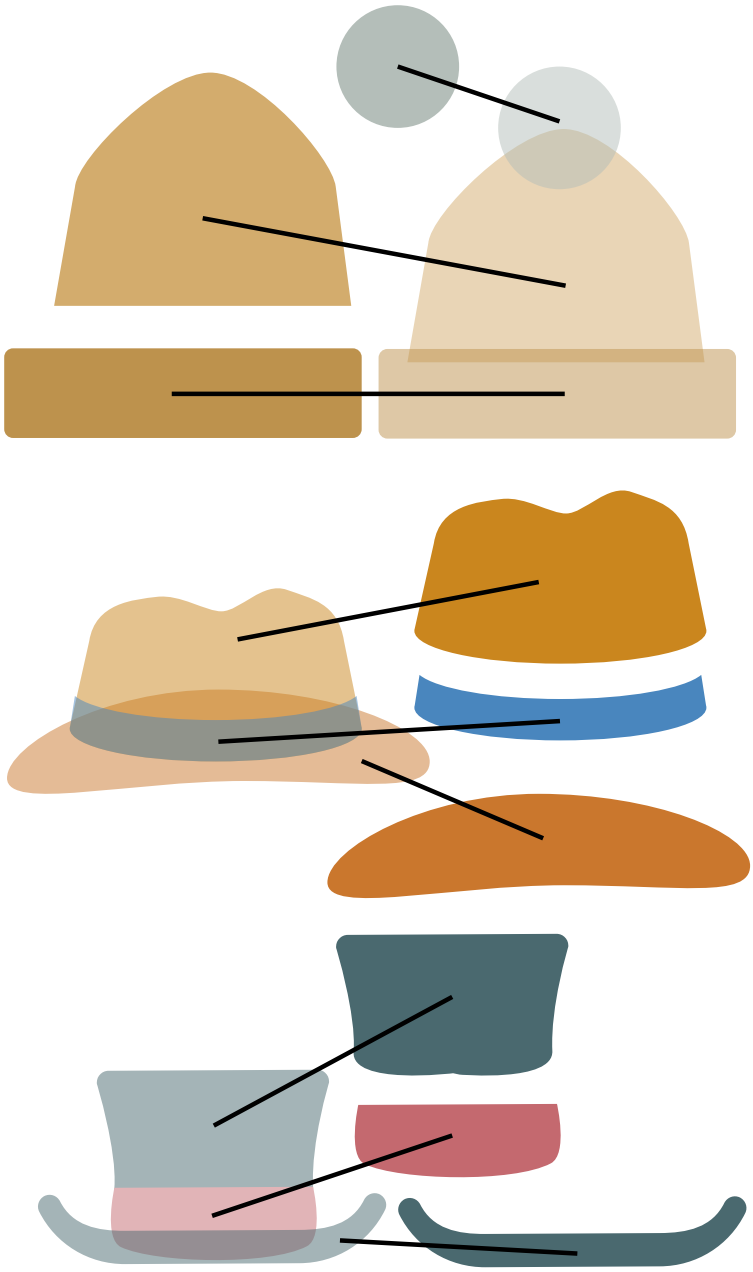

Figure 37. Layering of the synthesized hat clipart.

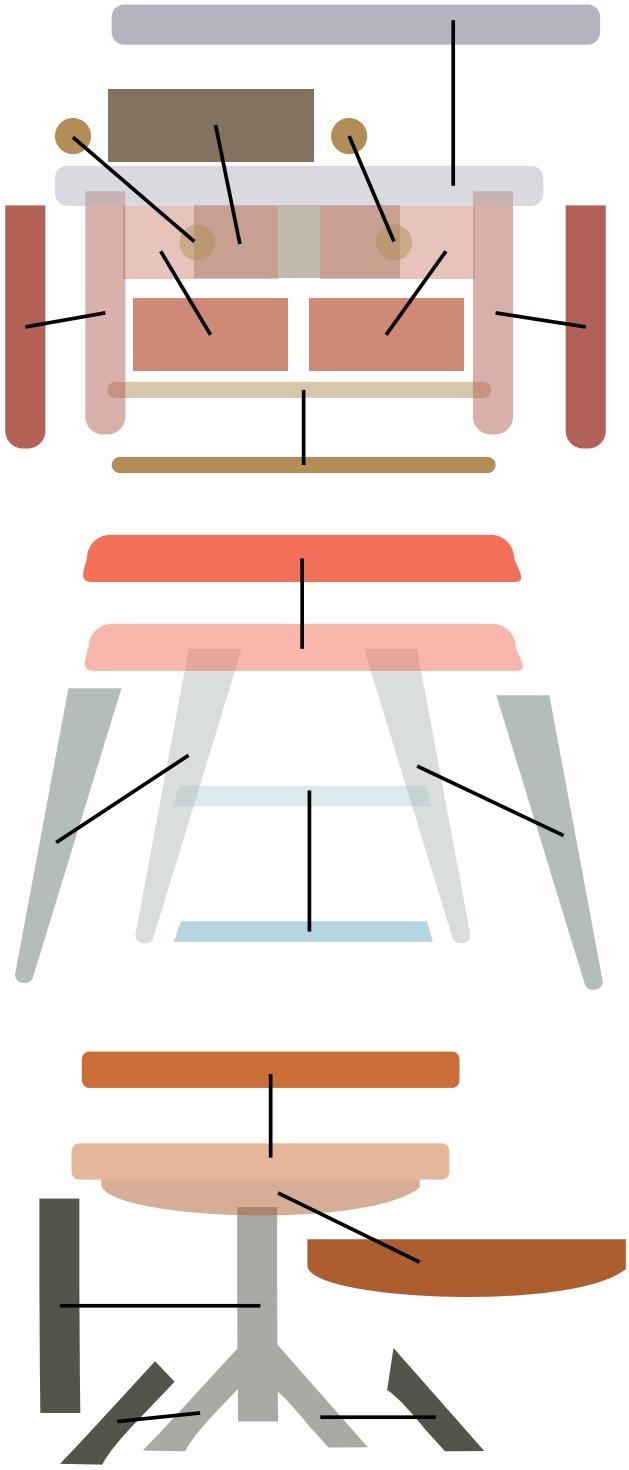

Figure 38. Layering of the synthesized table clipart.

References

FAVREAU, J.-D., LAFARGE, F., AND BOUSSEAU, A. 2017. Photo2clipart: Image abstraction and vectorization using layered linear gradients. *ACM Transactions on Graphics (SIGGRAPH Asia Conference Proceedings) 36*, 6 (November). URL: http://www-sop.inria.fr/reves/Basilic/2017/FLB17. 6

HOSHYARI, S., DOMINICI, E. A., SHEFFER, A., CARR, N., WANG, Z., CEYLAN, D., SHEN, I., ET AL. 2018. Perception-driven semi-structured boundary vectorization. *ACM Transactions on Graphics (TOG) 37*, 4, 118. 6, 7