HiEI: A Universal Framework for Generating High-quality Emerging Images from Natural Images

Jingmeng Li[1], Lukang Fu[1], Surun Yang[1], and Hui Wei*[1,2]

[1] Laboratory of Algorithms for Cognitive Models, Fudan University, Shanghai, China
{jmli21,lkfu23,sryang22}@m.fudan.edu.cn, and weihui@fudan.edu.cn
[2] Innovation Center of Calligraphy and Painting Creation Technology, MCT, China

1 Effect of different N_l on color quantization

We set the parameters N_l to 5 and 16, respectively, and Fig. 1.1 displays some results of color quantization with 1-bit color space. As can be seen from the results presented, the use of shallow content features leads to more detail in the results, and the use of deeper features leads to the loss of detailed content.

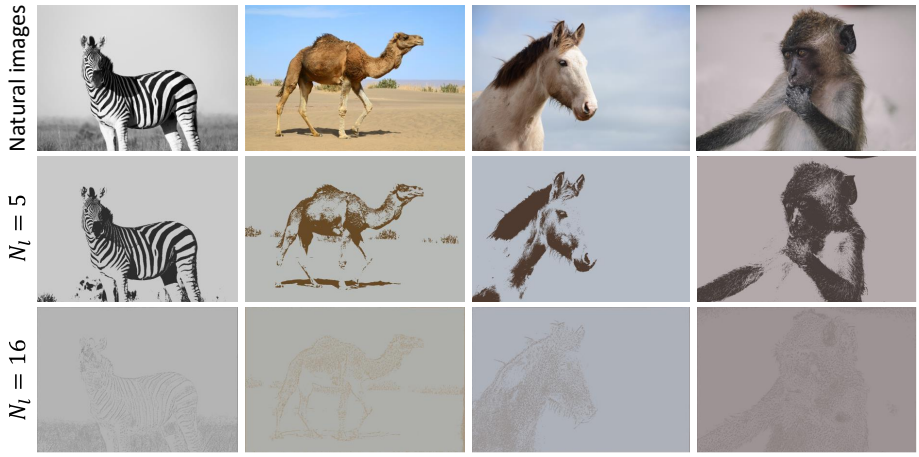

Fig. 1.1: Color quantization results with different settings of parameter N_l. Using shallower network layer features results in more local detail, and using deeper network layers results in more local detail being lost.

* Corresponding author.

2 Color quantization with different size of color spaces

We compared the color quantization results of TTNet and MedianCut, OCTree and ColorCNN in different sizes of color spaces based on the BSDS500 dataset [1]. BSDS500 is a classical dataset in the field of edge detection and contains 500 natural images that are divided into three parts: 200 images in the training set, 200 images in the test set, and 100 images in the validation set. We set the color space to 2-bit, 4-bit and 8-bit in order. Some results are displayed in Fig. 2.1. Visually, the color quantization results of TTNet and ColorCNN in 2-bit and 4-bit spaces outperform those of two traditional methods MedianCut and OCTree.

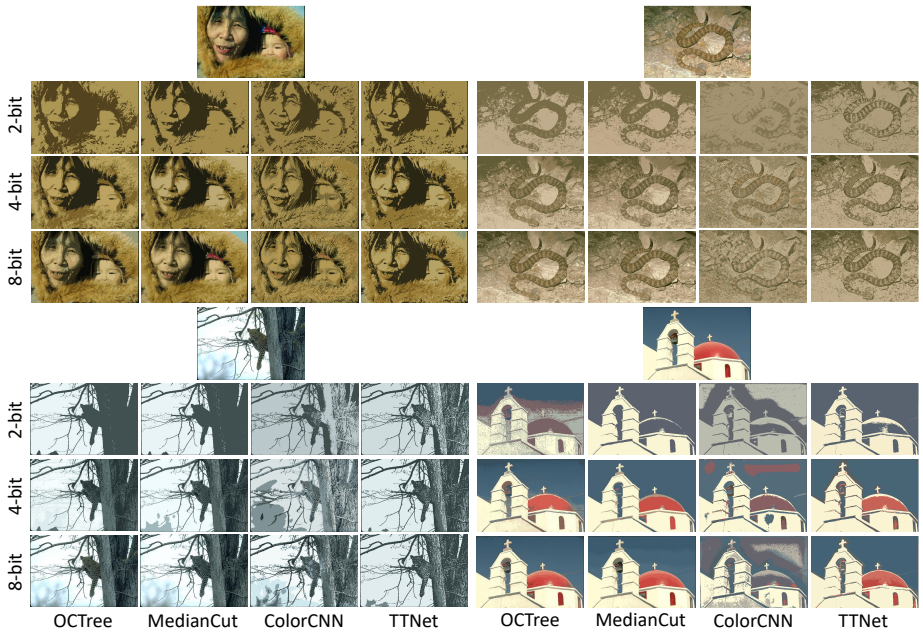

Fig. 2.1: Visual comparison of color quantization results with different color spaces (*e.g.*, 2-bit, 4-bit, 8-bit).

Both shape and texture information of objects in an image can be presented by edges, and the BSDS500 dataset provides a groungtruth for image edge detection.Therefore, we use the number of edges in the color quantization results to measure to the richness of visual cues. If the edge recall is higher, the richer the visual cues reserved in the color quantization result. Fig. 2.2 presents the edge recall the four methods in four different sizes of color spaces: 1-bit, 2-bit, 4-bit, and 8-bit. Overall, our model TTNet performs best, especially in smaller color spaces (*e.g.*, 1-bit, 2-bit, 4-bit). In 8-bit color space, our method slightly underperforms MedianCut and OCTree.

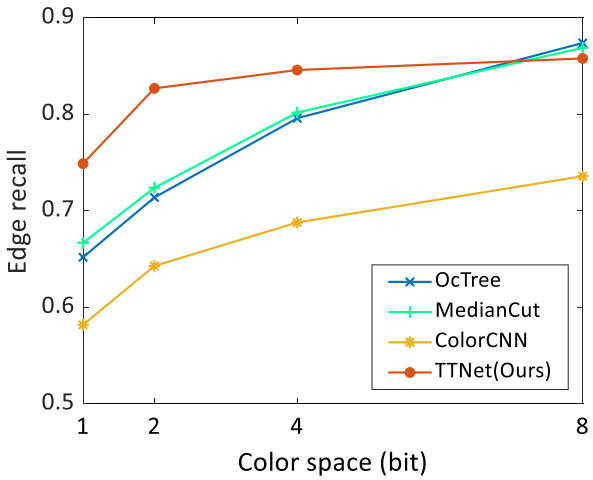

Fig. 2.2: Quantitative evaluation of color quantization results with different sizes of color spaces.

3 Ablation studies for PDC and TV modules

PDC moudle. The PDC module is designed to enable HiEI to generate EIs with different perceived difficulties. Here, we use $\text{HiEI}_{(\{1,0.7,0.5\},\{0,0.2,0.4,0.6,0.8,1\},+)}$ to generate EIs. We calculate the first content loss between these EIs and the natural image, and the second content loss between $\text{EIs}_{(\{0.7,0.5\},\{0.2,0.4,0.6,0.8,1\},+)}$ and $\text{EI}_{(1,0,+)}$. An example is illustrated in Fig. 3.1. Visually, different parameter settings have a significant impact on the process of find and recognize the objects in EIs. Quantitatively, the differences between these EIs and the natural images gradually increase as the parameters are adjusted, and substantial differences exist between these EIs.

TV moudle. The TV module aims to solve the problem of distortion when zooming in the generated results based on low-resolution images. Fig. 3.2 presents the visual comparison between the rendering templates generated based on images with 96×96 resolution and their vectorization results. When we zoom in these samples, the rendering templates are distorted, but the vectorized results are not.

References

1. Arbelaez, P., Maire, M., Fowlkes, C., Malik, J.: Contour detection and hierarchical image segmentation. IEEE transactions on pattern analysis and machine intelligence **33**(5), 898–916 (2010) 2

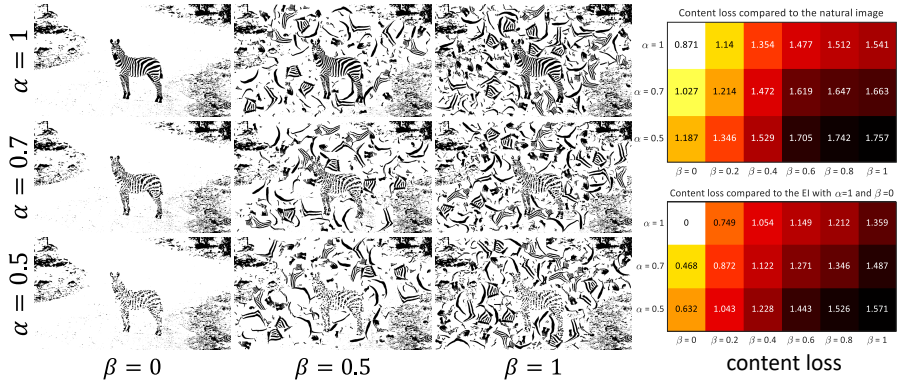

Fig. 3.1: Explaination for the diversity of EIs generated by HiEI. The PDC module can quantitatively adjust the generation results through two parameters α and β, thus affecting the perception of human vision and deep networks.

Fig. 3.2: Examples of EIs generated by HiEI using natural images with 96×96 resolution. Please zoom in on the images for comparison.