



# 2021「中技社科技獎學金」

## 2021 CTCI Foundation Science and Technology Scholarship

### 境外生研究獎學金

#### Research Scholarship for International Graduate Students



## Neural Style Palette: a multimodal style transfer from a single style image

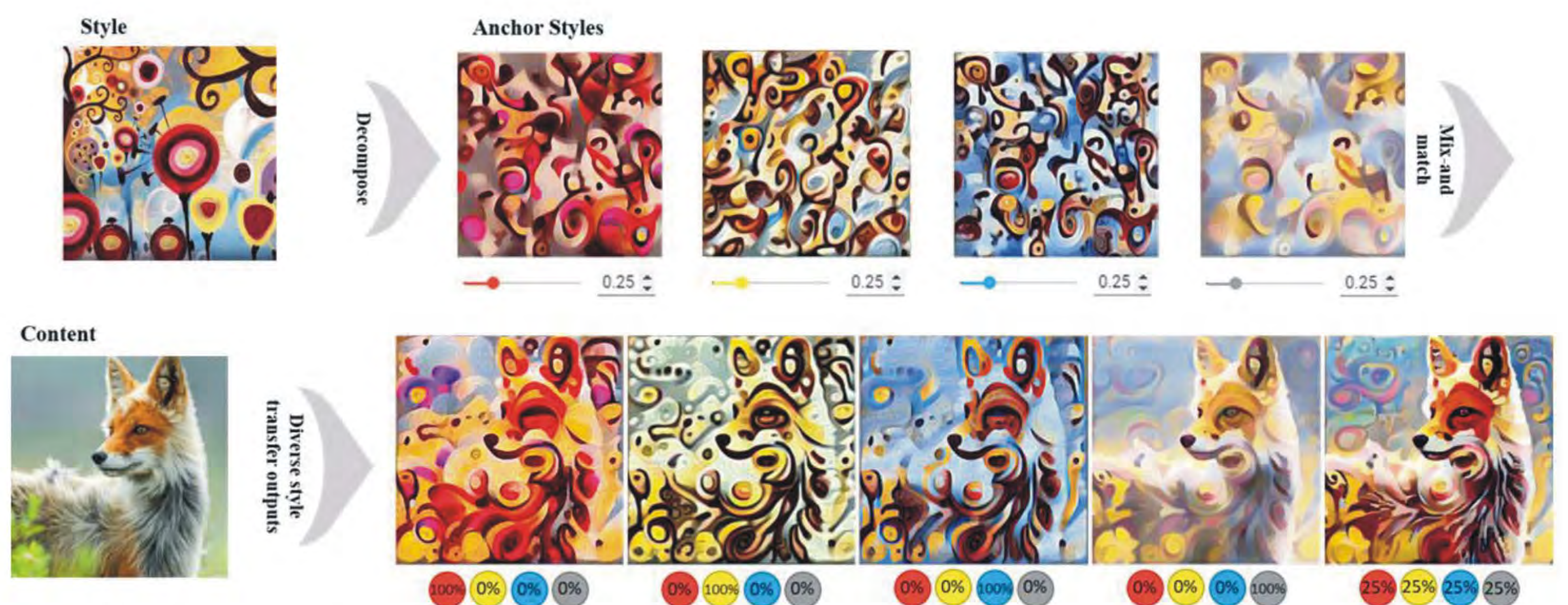
John Jethro Virtusio, Kai-Lung Hua

Department of Computer Science and Information Engineering  
National Taiwan University of Science and Technology

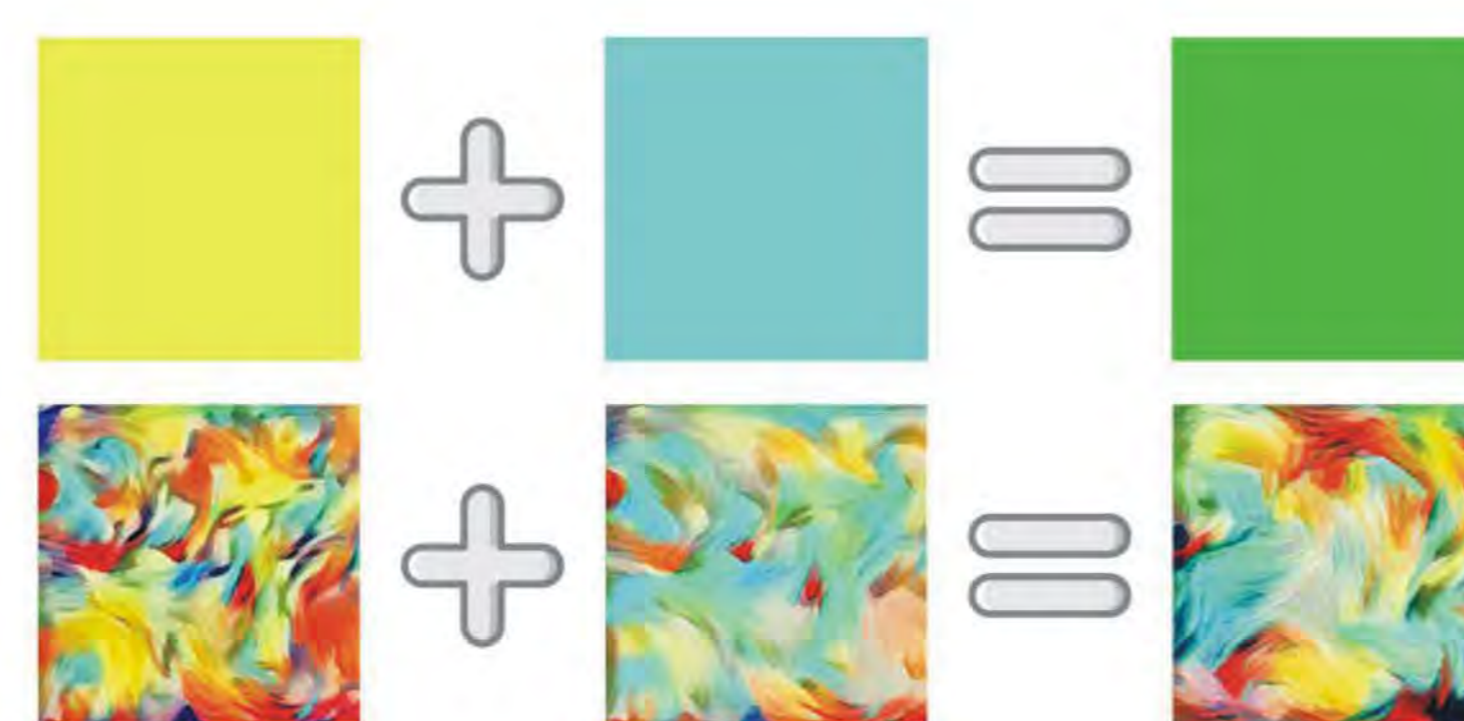
**Abstract:** Despite the myriad of elements found in a single style image, existing neural style transfer methods produce outputs with limited variety—typically only a single realization of the styled image. They also do not provide an easy way to control the stylization process, limiting the creative freedom of users. In this paper, we propose Neural Style Palette (NSP), a method for interactively generating a variety of stylized images from only a single style input.

**Introduction:** Neural Style Transfer has enabled everyday people to generate spectacular works of art. Despite its obvious potential in creativity apps, little work has been dedicated to making it more user interactive and controllable. Different users have unique preferences, which makes user interaction a critical feature. We observe that a single style image is composed of multiple unique attributes like colors, textures, and patterns. In this paper, we propose an approach that extracts these different style attributes of a single image, which the users can mix-and-match to produce diverse and guided outputs. In Figure 1 (top row), we show an example of our approach extracting sub-textures, which we also refer to as anchor styles. It can be observed that each anchor style emphasizes different style attributes, including unique colors and strokes. To the best of our knowledge, we are the first to explicitly extract and visually represent various attributes of a single style image—all without training from a large dataset. Because we can visually represent the anchor styles, the users can use it to achieve their desired stylizations. In Figure 1 (bottom row), we show how users can interactively mix-and-match different anchor styles to produce diverse outputs.

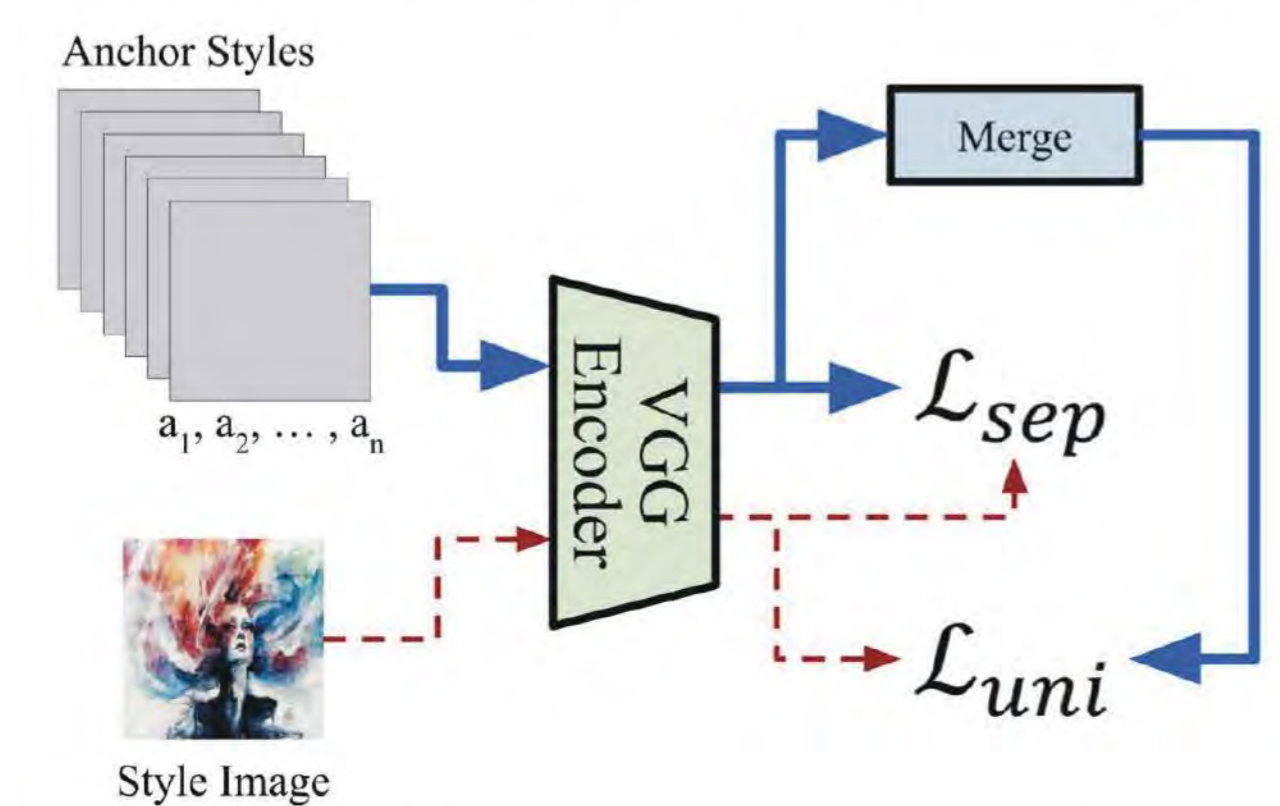
**Motivation:** A style image is a combination of multiple unique perceptual elements like textures and colors. We motivate our neural style pastiche from this interesting property of style and provide a way to extract these perceptual elements through a specialized texture synthesis. By explicitly generating the style compositions of a single style image, we enable a user interactive mix-and-match style transfer. In Figure 2, we show a toy example motivating our approach. For example, the color green can be represented as a combination of different colors. Similarly, a single style image, such as the one in the figure, can be represented as a combination of different style images (referred to as anchor styles in our work). Explicitly decomposing the style image into multiple anchor styles is a critical feature in interactive style transfer because it presents the user's opportunity to mix-and-match components according to their preference.



**Figure 1:** Neural Style Palette (NSP) allows the user to control and produce diversified stylized images from a single style image. (Top row) several anchor styles capturing unique colors and texture of the input style image. (Bottom row) diverse and user-guided style transfer results via blending of anchor styles.



**Figure 2:** (top row) a toy example shows that color green can be represented as a combination of two other colors. (bottom row) example of our NSP representing the  $I_s$  as a combination of two anchor styles.

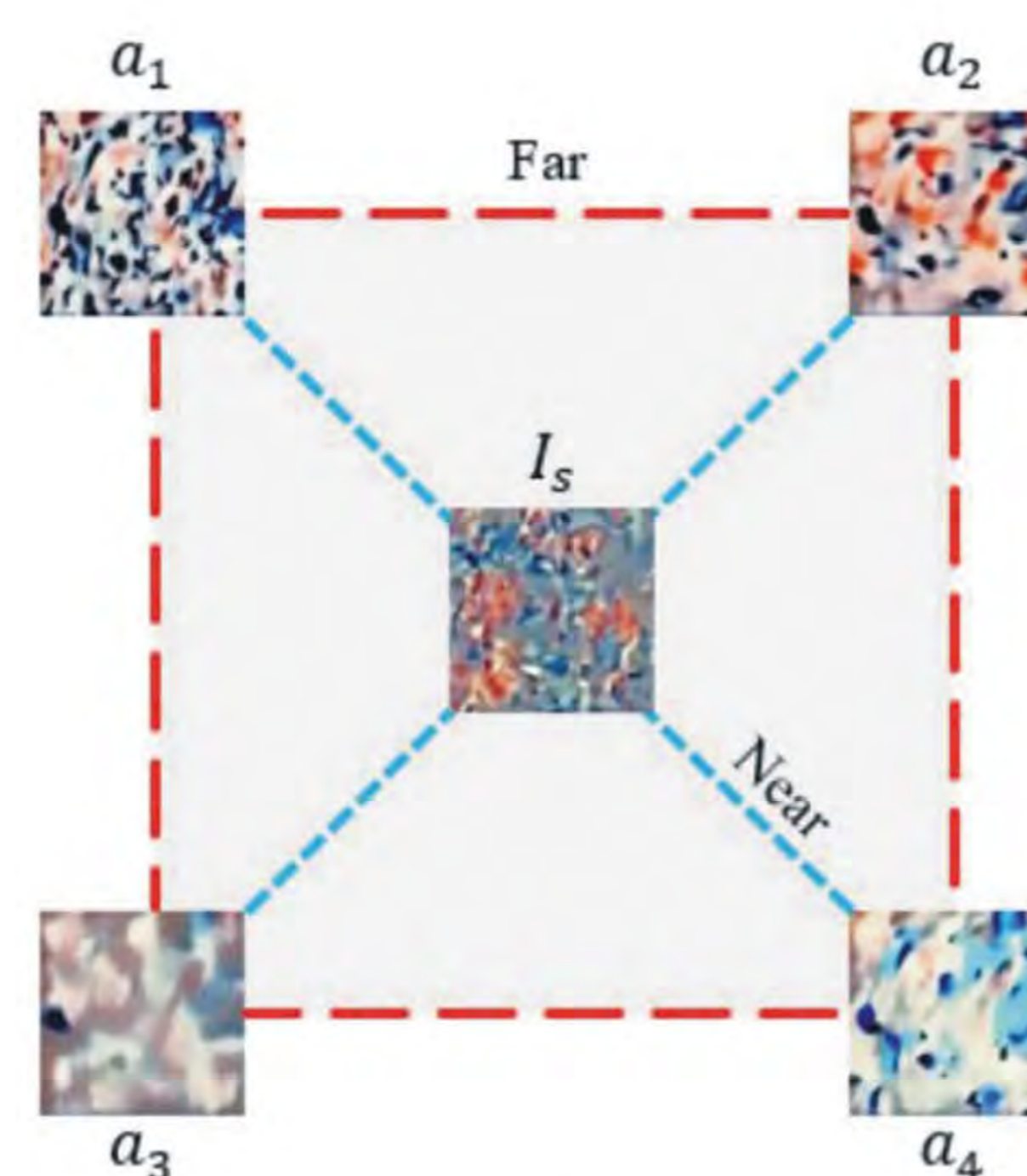


**Figure 3:** Overview of our proposed method for generating multiple anchor styles, forming the Neural Style Palette (NSP). We optimize the pixel values of the anchor styles  $a_1, \dots, a_n$ , in order to jointly minimize the style separation loss  $L_{sep}$  and style unification loss  $L_{uni}$ .

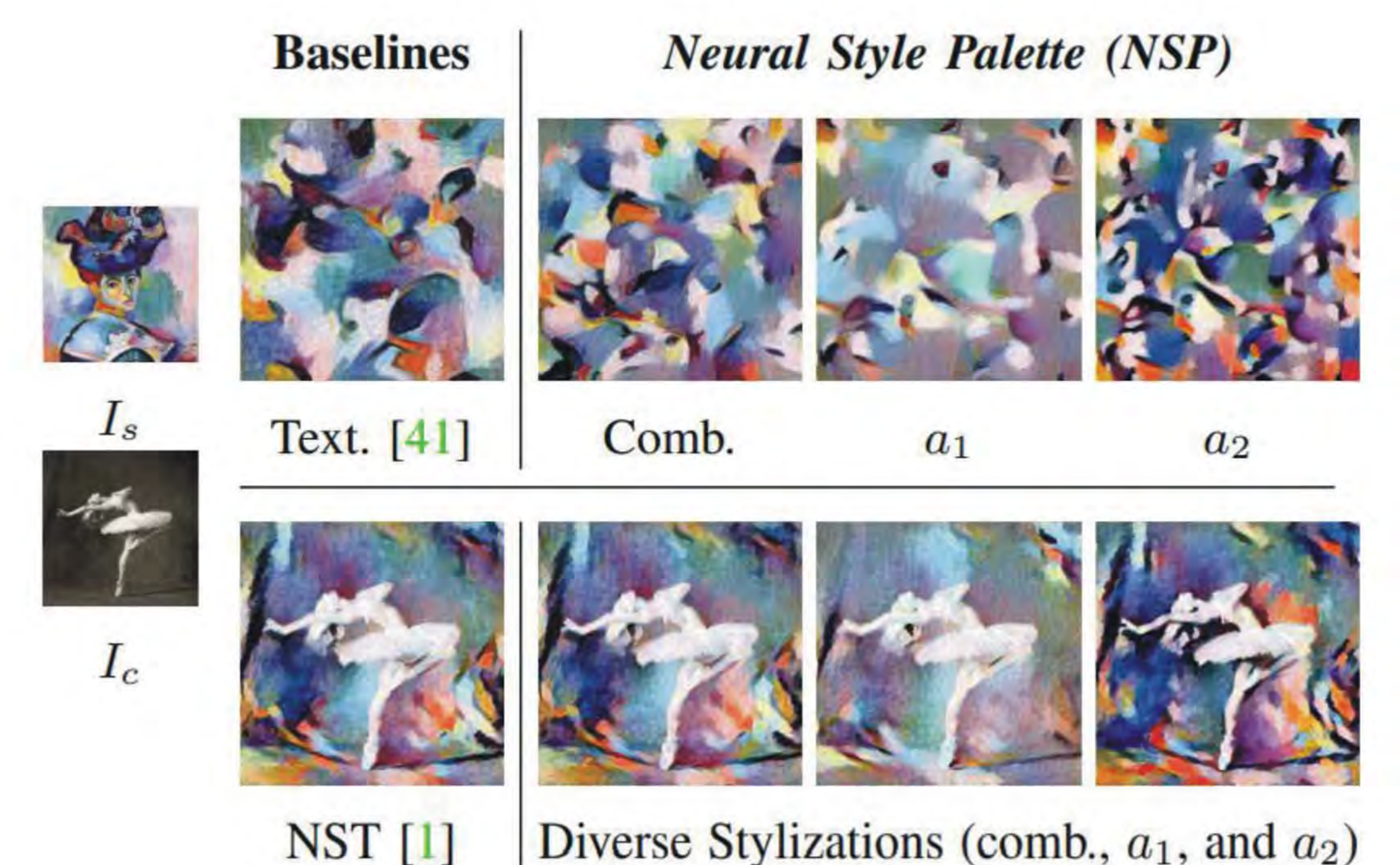
**Methodology:** Generating the anchor style can be interpreted as a specialized texture synthesis where we represent the anchor styles as a set of  $N$  images  $a_i \in \mathbb{R}^{3 \times H \times W}$ . It is also possible to represent anchor styles as Gram matrices but representing them as images have two main advantages. First, humans easily interpret images, allowing users to imagine how the stylized output would look like. Second, it will enable us to incorporate our method with other style transfer approaches easily. We generate the anchor styles by directly optimizing the pixel values according to two constraints: (1) each must be distant from one another, (2) merging the set of anchor styles via average must result to the style of input image  $I_s$ . For instance, the texture synthesis of the input style and the merged anchor styles should be similar. We enforce these constraints using style separation loss and style unification loss, respectively.

#### Summary of Contributions:

- We propose Neural Style Palette (NSP), a visual guide for the users that enables a diverse, multimodal style transfer— all from a single style input.
- We propose a novel style transfer interaction through our anchor styles that allow the users to magnify, minimize, or remove certain style features.
- We developed a sample user interface for our interactive style transfer tool, enabling novice artists to create inspired works of art.



**Figure 4:** (top row) a toy example shows that color green can be represented as a combination of two other colors. (bottom row) example of our NSP representing the  $I_s$  as a combination of two anchor styles.



**Figure 5:** Baselines Neural Style Transfer (NST), and Neural Texture Synthesis (Text.) with and without our NSP. Combined (Comb.) refers to the equal blending of anchor styles  $a_1$  and  $a_2$ .

