

# ARF: Artistic Radiance Fields

## Supplemental Document

Kai Zhang<sup>1</sup> Nick Kolkin<sup>2</sup> Sai Bi<sup>2</sup> Fujun Luan<sup>2</sup>  
Zexiang Xu<sup>2</sup> Eli Shechtman<sup>2</sup> Noah Snavely<sup>1</sup>

<sup>1</sup>Cornell University <sup>2</sup>Adobe Research

**More implementation details.** To discard view-dependence of the photorealistic radiance fields, we adopt different strategies for different radiance fields representations. For Plenoxels, we simply remove all spherical harmonics components except the first one; for TensorRF, we zero out the view directions when inputting them to the MLP; for NeRF, because the network capacity for the appearance part is small once we fix the density part, we use a new bigger MLP that maps each 3D point to a color vector for representing stylized appearance. At each stylization iteration, we render an image for computing losses from a viewpoint randomly selected out of all the training viewpoints used for reconstructing photo-realistic radiance fields. We perform the stylization optimization for 10 epochs. Learning rate settings follow the photorealistic radiance fields reconstruction stage for each radiance field representation.

**Ablation of network layers.** The pretrained VGG-16 network consists of 5 layer blocks: conv1, conv2, conv3, conv4, conv5. Each layer block begins with a max-pooling layer that downsamples the feature map by 2. Inside each layer block, feature maps are of the same spatial resolution and hence can be concatenated to form a single feature map for this block. We concatenate the output feature maps by all relu layers in the conv3 block to form our feature maps used in our NNFM loss and content preservation loss, because we find that the conv3 block features better capture style details than the other blocks in Fig. 1.

**Derivation of the affine color transformation  $\mathbf{A}$  in Section 4.3.** Let  $\mathbf{c}$  denote the color vectors of the image set to be re-colored in order to match the style image’s colors  $\mathbf{s}$ . We transform  $\mathbf{c}$  as below:

$$\mathbf{Ac} = \mathbf{U}_s \mathbf{A}_s^{\frac{1}{2}} \mathbf{U}_s^T \mathbf{U}_c \mathbf{A}_c^{-\frac{1}{2}} \mathbf{U}_c^T (\mathbf{c} - \mathbb{E}[\mathbf{c}]) + \mathbb{E}[\mathbf{s}], \quad (1)$$

where  $\mathbf{U}_s, \mathbf{A}_s, \mathbf{U}_c, \mathbf{A}_c$  are obtained via the following eigen-decompositions of covariance matrices  $\text{Cov}[\mathbf{c}], \text{Cov}[\mathbf{s}]$ :

$$\text{Cov}[\mathbf{c}] = \mathbf{U}_c \mathbf{A}_c \mathbf{U}_c^T \quad (2)$$

$$\text{Cov}[\mathbf{s}] = \mathbf{U}_s \mathbf{A}_s \mathbf{U}_s^T. \quad (3)$$



**Fig. 1.** Ablation studies of extracting features for our NNFM loss using different layer blocks of the pretrained VGG-16 network. We find that conv3 layer block generates more visually pleasing results than other layer blocks.

*Proof* of  $E[\mathbf{Ac}] = E[\mathbf{s}]$ :

$$E[\mathbf{Ac}] = \mathbf{U}_s \mathbf{\Lambda}_s^{\frac{1}{2}} \mathbf{U}_s^T \mathbf{U}_c \mathbf{\Lambda}_c^{-\frac{1}{2}} \mathbf{U}_c^T (E[\mathbf{c}] - E[\mathbf{c}]) + E[\mathbf{s}] \quad (4)$$

$$= \mathbf{0} + E[\mathbf{s}] = E[\mathbf{s}]. \quad (5)$$

*Proof* of  $\text{Cov}[\mathbf{Ac}] = \text{Cov}[\mathbf{s}]$ :

$$\text{Cov}[\mathbf{Ac}] = \text{Cov}[\mathbf{U}_s \mathbf{\Lambda}_s^{\frac{1}{2}} \mathbf{U}_s^T \mathbf{U}_c \mathbf{\Lambda}_c^{-\frac{1}{2}} \mathbf{U}_c^T \mathbf{c}] \quad (6)$$

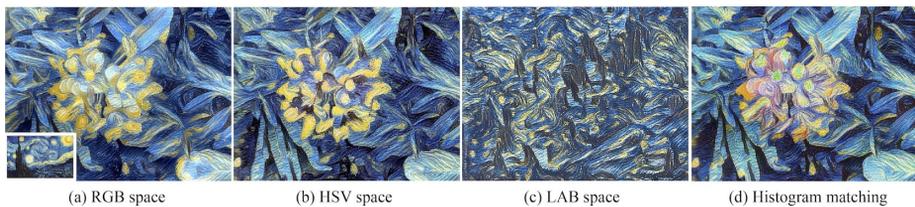
$$= \mathbf{U}_s \mathbf{\Lambda}_s^{\frac{1}{2}} \mathbf{U}_s^T \cdot \text{Cov}[\mathbf{U}_c \mathbf{\Lambda}_c^{-\frac{1}{2}} \mathbf{U}_c^T \mathbf{c}] \cdot (\mathbf{U}_s \mathbf{\Lambda}_s^{\frac{1}{2}} \mathbf{U}_s^T)^T \quad (7)$$

$$= \mathbf{U}_s \mathbf{\Lambda}_s^{\frac{1}{2}} \mathbf{U}_s^T \cdot \mathbf{U}_c \mathbf{\Lambda}_c^{-\frac{1}{2}} \mathbf{U}_c^T \cdot \text{Cov}[\mathbf{c}] \cdot (\mathbf{U}_c \mathbf{\Lambda}_c^{-\frac{1}{2}} \mathbf{U}_c^T)^T \cdot (\mathbf{U}_s \mathbf{\Lambda}_s^{\frac{1}{2}} \mathbf{U}_s^T)^T \quad (8)$$

$$= \mathbf{U}_s \mathbf{\Lambda}_s^{\frac{1}{2}} \mathbf{U}_s^T \cdot \mathbf{I} \cdot (\mathbf{U}_s \mathbf{\Lambda}_s^{\frac{1}{2}} \mathbf{U}_s^T)^T \quad (9)$$

$$= \text{Cov}[\mathbf{s}]. \quad (10)$$

**Ablation of color spaces.** We show an ablation study of using different color spaces for the color transfer algorithm described in Sec. 4.3. We find RGB color space tends to yield more visually pleasing results than other color spaces.



**Fig. 2.** Ablation study of color spaces. We execute the color transfer algorithm described in Sec. 4.3 in RGB color space (a); this tends to generate better results than running it in HSV (b) and LAB (c) color spaces, and replacing it with histogram matching (d).