Importance-Driven Composition of Multiple Rendering Styles

Jiazhou Chen, Yujun Chen, Xavier Granier, Jingling Wang, Qunsheng Peng

Importance-Driven Composition of Multiple Rendering Styles

Jiazhou Chen [1,2] Yujun Chen [1] Xavier Granier [2] Jingling Wang [1] Qunsheng Peng [1]

[1] *State Key Lab of CAD&CG - Zhejiang University* [2] *IPARLA - INRIA Bordeaux University*

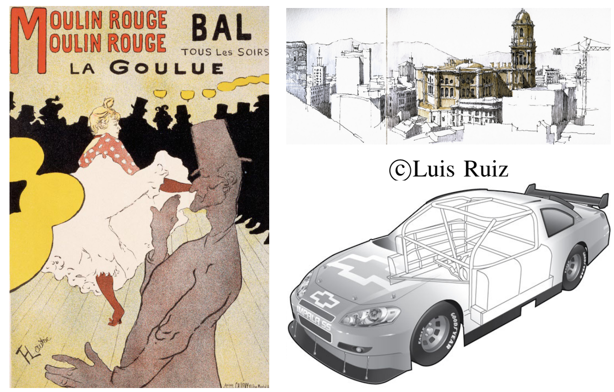
***Abstract*—We introduce a non-uniform composition that integrates multiple rendering styles in a picture driven by an importance map. This map, either issued from saliency estimation or designed by a user, is introduced both in the creation of the multiple styles and in the final composition. Our approach accommodates a variety of stylization techniques, such as color desaturation, line drawing, blurring, edge-preserving smoothing and enhancement. We illustrate the versatility of the proposed approach and the variety of rendering styles on different applications such as images, videos, 3D scenes and even mixed reality. We also demonstrate that such an approach may help in directing user attention.**

***Keywords*-stylization; nonphotorealistic rendering; importance map; composition**

I. Introduction

The use of multiple styles has been adopted in many artworks, mostly modern ones, to highlight the important meanings. However, it can only be conducted by skillful artists. Some examples are shown in Figure 1. DeCarlo de al. [1] cites Toulouse-Lautrec's poster as an example to show that different regions can be abstracted in different levels. We also cite this poster in the left of Figure 1 to show that artists also employ significantly different styles for different regions. In the other two pieces of artwork in the right, different styles, such as watercolor painting, line drawing, luminance, are combined in different but reasonable manners to exhibit the creators' special design intents for different contents, which also presents the maximal of visual information.

In Computer Graphics, many rendering papers focusing on a single style have been published, while few works consider multiple styles. Image abstraction techniques [1], [2], [3], [4], [5], [6] adopt edge-preserving smoothing that gives the freedom to omit or remove visual information under the guide of user interaction, eye tracking data or automatic saliency estimation. Level-of-Detail rendering of 3D scenes [7], [8] is achieved by either pre-building the hierarchical structure of the mesh or adjusting the size of strokes adaptively. Importance-driven volume rendering [9], [10] shares the same idea with X-Ray vision [11], [12], [13] in Augmented Reality that focus and context are rendered differently to display the interior or hidden structures. Unfortunately, they severely restrict the choices of styles. A pioneer work to extend this compatibility is called stylized focus [14], which directs the attention of viewers to areas of



©Henri de Toulouse-Lautrec ©Vic Kulihim Illustration

Figure 1: Three examples of artistic illustration with multiple styles. The poster in the left, "Mouline Rouge-La Goulue", adopts detailed painting to exhibit the pose of the woman dancer and the material of her dress, uniform color plus line drawing for her partner, silhouette filled with black color for all the spectators in the background. In the right-top painting, the cathedral painted with watercolor stands out from surrounding building drawn by lines, and shows a clear overall arrangement of this city. In the right-bottom illustration, the line drawing in the middle of the car exhibits the interior structures while the luminance in the head and the tail shows the global appearance of the car.

interests in architectural rendering of 3D models. User can select several elaborate focus models and rendering effects, but gaze-based importance definition can not satisfy complex compositing tasks.

In this paper, we introduce a non-uniform compositing approach that combines multiple rendering styles. The compositing is driven by an importance map. It independently renders the images, videos or 3D scenes with different styles selected by users conforming to this importance map. Different images are then composited in screen space. Our approach shares a lot of similarities with the classical compositing tools, but brings three key advantages:

- It composites multiple rendering styles non-uniformly guided by an importance map instead of one or two styles;
- It can process different input data, such as images, videos, 3D scenes, or even mixed reality data;

- It is compatible with different importance maps, even user-created or user-edited ones.

The goal of this paper is not to introduce a new rendering style, but an importance-driven approach using multiple existed styles. We introduce our general approach on an image example in Section III, after reviewing the previous work in Section II. We then describe all the details of each step in our compositing pipeline in Section IV and V. We also demonstrate the compatibility for various input data, styles, and applications in Section VI. The final Section VII presents a conclusion and potential future extensions.

II. Previous work

Most of the previous work only combines two styles: on one side the lines and strokes and on the other side, color abstractions. DeCarlo and Santella [1], [2] first proposed the idea of "Image Abstraction" that offers such a style with a spatially variant abstraction omits or removes redundant visual information, as mentioned in Section I. Orzan et al. [15] uses the Gaussian scale-space theory to deal with the discontinuity problem for computing a perceptually meaningful hierarchy representation.

Inspired by these pioneer works, many extensions on image or video abstraction have been published [3], [4], [5], [6], [16]. In general, they reduce the contrast in low-contrast regions while increase the contrast in high-contrast regions. Different from the hierarchical structure analysis used in [1], [2], they employ various image filters, such as edge-preserving smoothing filter, Flow-based Difference-of-Gaussian (FDoG) filter, shock filter and Kuwahara filter, to achieve the goal of reducing or increasing the contrast. Recent techniques take the advantages of faster performance with parallel implementations and improved temporal coherence, but still severally restrict in a particular rendering effect.

All these techniques are based on an underlying importance map either extracted from Eye-tracker data [2], or from information extracted from images. One more and more popular way to extract visual importance information is to compute a saliency map [17], [18], [5], [19]. These techniques detect low level features (such as intensity variance, orientation and color contrast, texture difference) and higher level features (such as horizon lines, human faces, cars) using bottom-up computation, but they don't match actual eye movements [20]. Thus, Judd et al. [20] introduces the eye tracking data to learn a saliency model to predict where humans look. Zeng et al. [21] analyzes top-down semantics of the image by image parsing, but requires much more user interactions to segment and label the image. Kang et al. [22] defines an importance map based on outline, and controls the abstraction level by adaptively adjusting the stroke attributes.

Using saliency or any other perception-based importance map only allows controlling the level of stylization according to what a common user will perceive in an image. A wider range of applications are possible if we want to use such a map to control the gaze. Focus+Context rendering is a well-known expressive rendering mechanism in Augmented Reality (AR) and Volume Rendering domain. Kalkofen et al. [11], Chen et al. [12] and Sandor et al. [13] render the focus (the hidden scenes in the X-Ray vision of AR) and the context (other visible scenes) with different methods, and composite them in the screen to reveal the correct occluding order. Viola et al. [9] and Krüger et al. [10] composite interior structures and exterior surfaces by dynamically adjusting the transparency in volume rendering.

The gaze simulation in [14] is the closest related work of our approach. The authors introduce stylized focus to draw viewer's gaze to the emphasized area through local variations in shading effects and line qualities. Though four focus models are proposed, some easy-use interaction tools are expected but missed unfortunately. And the stylized focus supports single-point based emphasized area only, which loses the compatibility with either eye tracking data or saliency map, and also the ability to fulfill complex stylization tasks.

In this paper, we present an improved approach that allows the combination of a larger set of styles while allowing direct user controls of the spatial variation of both styles and combination.

III. General approach

Our approach is a generalization of image abstraction systems employed by many previous works [3], [5], [20] and first described in DeCarlo and Santella's work [1]. This generalization is in two-folds: 1) it permits more than two rendering styles to be integrated into an image; 2) it allows users to modify easily the important map. The former brings richer visual information and more stylization effects and the latter provides a controllable interface to alter the visual importance.

To demonstrate the proposed approach, we adopt three rendering styles, including original color (original color is considered as a stylization for simplification), color desaturation and line drawing (see the overview of our approach in Figure 2). The importance for each style is illustrated by different colors, such as red, green and blue color in Figure 2. We may employ an image saliency estimation technique [20] to initialize the importance, to produce results similar to [1]. Or, we can let the user define it with interaction tools.

Based on the importance map $\mathcal{I}$, we are now ready to compose the stylized result. For each position x in the image, a non-uniform composition of N stylized images $S_i(x,\mathcal{I})$, $i \in [1,...,N]$, using weights $W_i(x,\mathcal{I})$ is:

$$C(x) = \frac{1}{K}\sum_{i=1}^{N} W_i(x,\mathcal{I}) S_i(x,\mathcal{I}), \qquad (1)$$

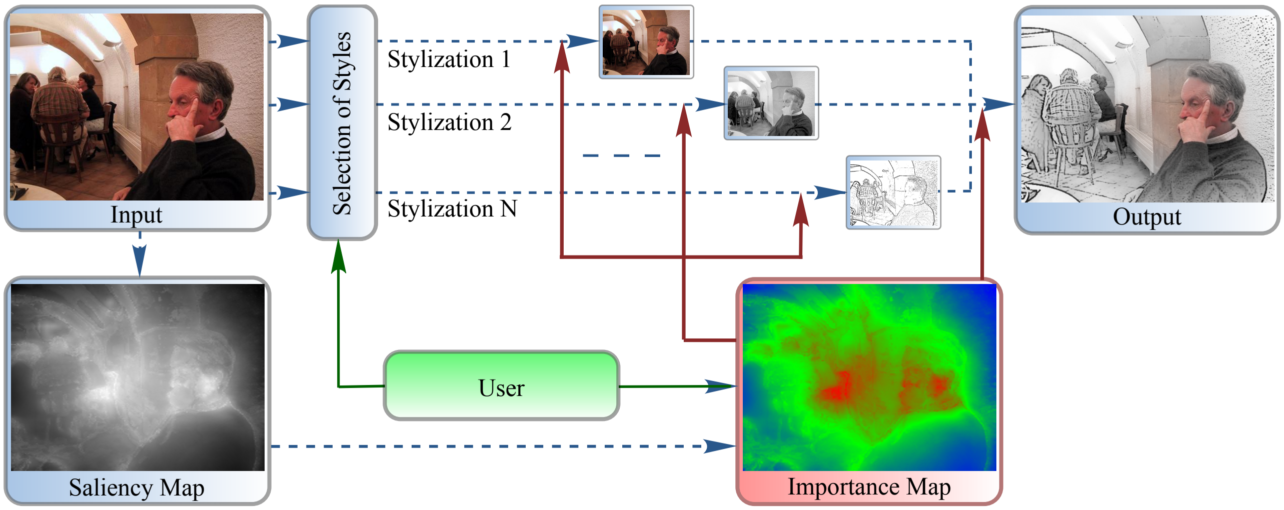


Figure 2: Overview of our system. Three example styles (original color, desaturated color, line drawing) are presented by red, green and blue color respectively in the importance map. One of the main contributions of our paper is that we use the importance map to guide not only stylization, but also the composition, illustrated by thick red arrows. The user, presented by a green box, has to select the styles, and modify the importance map optionally.

where $K = \sum_{i=1}^{N} W_i(x, \mathcal{I})$ is a normalizing term. We take the importance map as the weighting functions to control the effect of stylization.

IV. Importance Map

Our importance map is presented as a gray-scale map with values between 0 and 1. In order to composite more than 2 styles, we subdivide the entire gray-scale interval into several sub-intervals at E_i, with $1 \geq E_0 > E_1 > ... > E_{N-1} \geq 0$. Each partitioning point corresponds to a rendering style. Then the importance value of pixel x must fall into a specific sub-interval and determines $W_i(x, \mathcal{I})$ in equation 1 according to some rules.

There are many ways to create the importance map. They can be roughly divided into two categories: saliency estimation or feature extraction, and user interaction. Our approach does not try to improve any of these techniques, but provides the compatibility to utilize any of them.

A. Construction from original image

We construct the saliency map $\mathcal{I}$ according to Judd et al [20]. This map can be further improved using different image-processing tools. In our approach, we have experimented the following two. One tool involves cross bilateral filtering (CBF) the importance map taking the color difference in neighborhood pixels of the original image into account:

$$\mathcal{I}^c(x) = \frac{1}{\mathcal{K}^c(x)} \sum_{y \in \mathcal{S}} G_{\sigma_s}(|x-y|) G_{\sigma_r}(|C(x) - C(y)|) \mathcal{I}(y) \tag{2}$$

where $\mathcal{K}^c(x) = \sum_{y \in \mathcal{S}} G_{\sigma_s}(|x-y|) G_{\sigma_r}(|C(x) - C(y)|)$ is a normalizing term, $G_\sigma(x) = exp(-x^2/(2\sigma^2))/\sqrt{2\pi\sigma^2}$ is Gaussian smoothing, $\mathcal{S}$ is a set of neighboring pixels and $C(x)$ presents the color of pixel x in original image. A CBF-based improved importance map $\mathcal{I}^c$ and its relative result are shown in the third row of Figure 3. Compared with the original result (zoomed-in image in blue frame), $\mathcal{I}^c$ (zoomed-in image in red frame) removes small spots while preserves the apparent edges.

The second one concerns averaging importance values in the same segmented region, which is produced by Mean-Shift image segmentation [23]:

$$\mathcal{I}^s(x) = \frac{1}{\mathcal{K}^s(x)} \sum_{L(y)=L(x)} \mathcal{I}(y) \tag{3}$$

where L is the segmentation label for each pixel, and $\mathcal{K}^s(x)$ is the amount of pixels whose label is the same as $L(x)$. A segmentation-based importance map $\mathcal{I}^s$ and its relative result are shown in the forth row of Figure 3. Compared with the original result and CBF-based improved one, $\mathcal{I}^s$ (zoomed-in image in green frame) ensures the consistency in each segmented regions.

B. User interaction

Users may decide themselves the expected importance value E_i for each rendering style. To facilitate the users to assign a selected style to a region on the image with specified importance value distribution, we provide them with a classical gradient slider to adjust the location of E_i at the gray-scale interval (see the top image in Figure 3). Initially, the interval is partitioned uniformly. Then users adjust the sub-intervals by dragging the bars.

Thanks to this gradient slider, a full gradient-based interaction tool is naturally developed. Users can directly

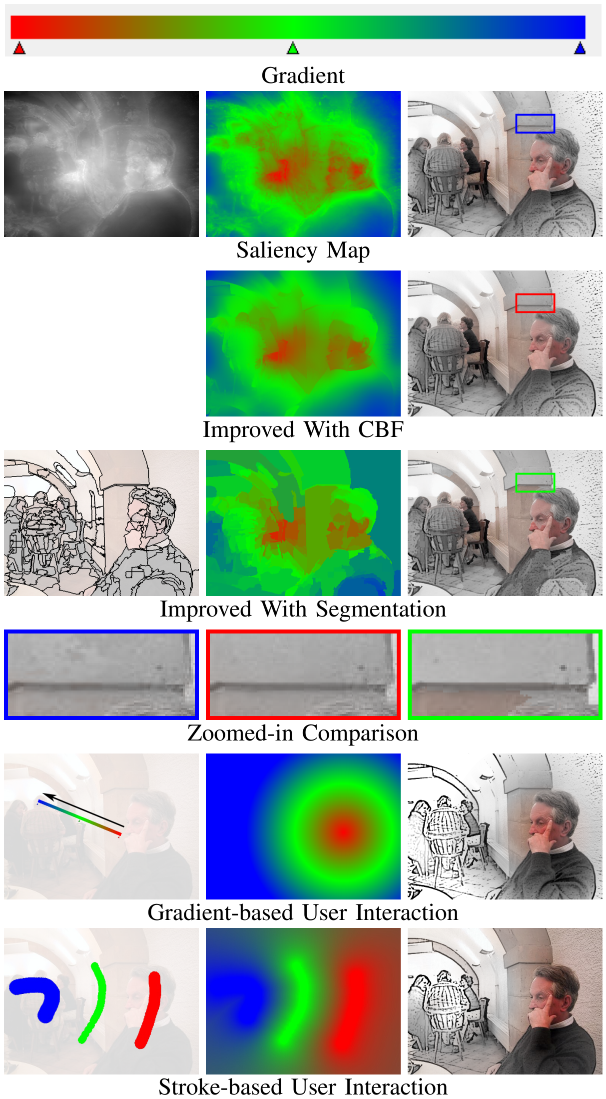


Figure 3: Five importance map examples computed by different user inputs. We take three styles for example, and use red, green and blue color to present these styles, whose importance is 1, 0.5 and 0 respectively. Different user inputs are shown in the left column. In the middle column, we visualize the importance maps using one color for each style and composite them together using the weights described in equation 8, the composition results are shown in the right column.

draw a line segment on the image to modify or create a new importance map. Similar to many painting softwares, we provide three gradient modes: radial gradient, linear gradient and symmetrical gradient, using the following three equations:

$$\mathcal{I}^r(x) = 1 - c(|x - x_B|/|x_E - x_B|) \tag{4}$$

$$\mathcal{I}^l(x) = 1 - c((x - x_B) \cdot (x_E - x_B)/|x_E - x_B|^2) \tag{5}$$

$$\mathcal{I}^{sy}(x) = |1 - 2c((x - x_B) \cdot (x_E - x_B)/|x_E - x_B|^2)| \tag{6}$$

where $c(\cdot)$ is a function clamping value to $[0, 1]$ and x_B, x_E are the beginning and end points of the line segment. A radial gradient example is the sixth row in Figure 3.

To provide even more freedom to satisfy complex tasks, we also design a stroke-based interaction tool. Users can directly draw strokes on the image, composed by a series of circles with different radii, to indicate different styles in different regions. We propose the following interpolation scheme to generate a whole importance map: 1) it ensures a smooth importance value transition between neighboring pixels, 2) it favors pure selected style for user drawn regions. Although, radial basis function is a well-known smooth interpolation function, its computation costs a lot, since it requires to solve a linear equation whose dimension is the amount of all the drawn pixels.

We propose a fast approximate solution based on the distance function:

$$\mathcal{I}^{st}(x) = \frac{1}{K'} \sum_{i=0}^{N} E_i \prod_{j=1, j \neq i}^{N} D(x, j) \tag{7}$$

where $K' = \sum_{i=0}^{N} \prod_{j=1,j\neq i}^{N} D(x, j)$ is a normalizing term and $D(x, j)$ is the distance between the current pixel x and the nearest circle related to the jth style. This distance can be easily computed with GPU, by the distance to the center of circle subtracted by its radius, thus the proposed solution achieves fast performance. Our GPU implementation includes two steps: 1) recording and storing the centers and radii of every circle of each stroke in a 1D texture, 2) computing the minimal distance from each pixel on the screen to these circles in parallel. The acceleration is owing to the fact that the amount of centers is much less than the pixels covered by these circles. This fast computation meets all the required properties, see the bottom row in Figure 3.

V. Non-uniform composition

A. Multiple weights

During the multi-style composition, each style is associated with a specific weight map, thus N weight maps are needed in total. But, our importance map initialized by saliency map is only a scalar image whose values are within $[0, 1]$. Therefore, we need to find a mapping to set up N weight maps based on the importance map and E_i regarding the N rendering styles. We create the weight maps using a piecewise linear interpolation:

$$W_i(x, \mathcal{I}) = \begin{cases} \frac{|\mathcal{I}(x) - E_{i-1}|}{|E_i - E_{i-1}|}, & \mathcal{I}(x) \in [E_i, E_{i-1}] \\ \frac{|\mathcal{I}(x) - E_{i+1}|}{|E_i - E_{i+1}|} & if\ \mathcal{I}(x) \in [E_{i+1}, E_i) \\ 0, & others \end{cases} \tag{8}$$

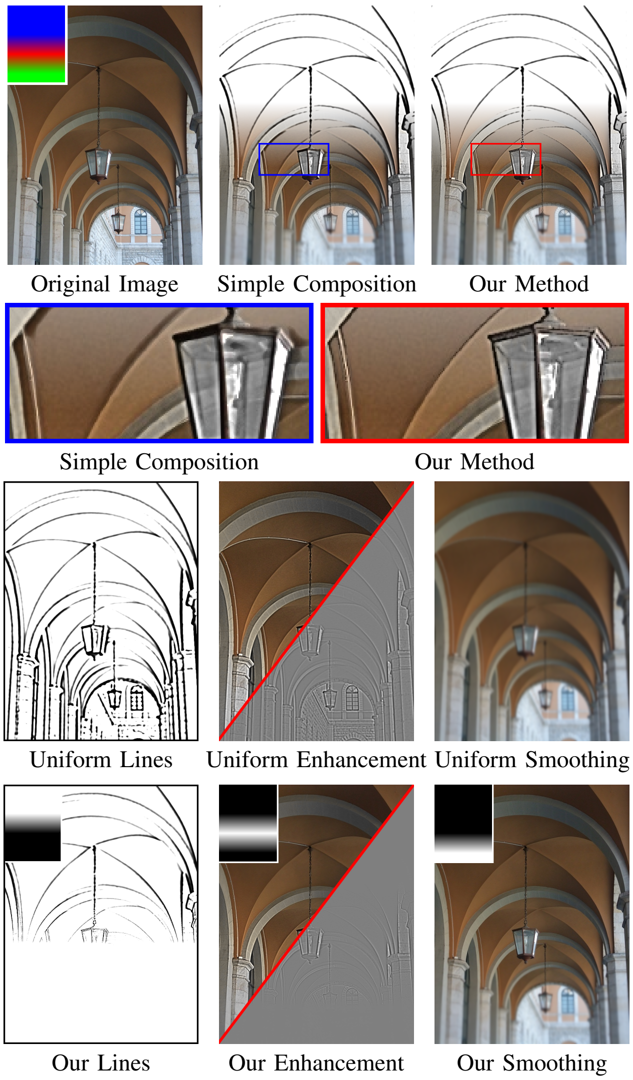


Figure 4: Compare with uniform stylization. In this experiment, we apply line extraction, image enhancement and blurring driven by an importance map. The first row contains the original image (with an importance map), a simple composition result and a composition produced by our method. A zoomed-in comparison is shown below. To compare each component of these two compositions, we illustrate all the stylized images one-by-one in the last two rows.

where $|\mathcal{I}(x) - E_i|$ is a distance function between the importance value of current pixel and the expected importance level of the ith style. To simplify the formula, we mark E_{-1} as 1 and E_N as 0. In the gradient tool, the user can see the interpolation result illustrated by different colors, as long as all E_i are given. The importance map is illustrated in the same way before the stylization, as the middle column of Figure 3. Note that highest importance is presented by red, while least importance is presented by blue in this paper, but the user can select any color for them.

B. Spatial varying stylization

Besides the composition, the importance map is also used to guide the stylization step. A uniform stylization without importance will introduces ghosting artifacts in the final compositing results, see Figure 4 for example. The goal of the composition in this example is to draw lines in near-by scene (the top part of the image), blur the distant scene (the bottom part) and enhance the color in the middle as focus. The original image and an gradient importance map is given in the left top image. A simple composition can be obtained by a composition of uniform stylized images. But, the color ghosting artifacts occurs in the regions where more than two styles exist, see the blue zoomed-in image.

Our method adapts the kernel size of filter according to the weights, which are computed by equation 8. In the regions with zero weight, the kernel size become zero, thus filter is invalid in these regions. Compared with simple color composition using uniform stylization, our method achieves higher quality. In Figure 4, the lines extracted by our method become thinner when the weights are getting smaller, while the lines in simple composition have uniform thickness which leads to ghosting artifacts. The same phenomenon can be observed for image enhancement and blurring.

We have implemented various stylization techniques to demonstrate the effectiveness of our approach:

- **Color desaturation** is an easy method to reduce the color cues. We reduce the two chromatic channels in the CIE Lab color space [24]. Because it's a pixel-wise conversion, its conversion does not need to be adapted according to the importance as the previous ones.
- **Enhancement** technique increases the sharpness in the region with high contrast. We employ Unsharp-masking [25] to enhance both images and 3D objects in high important regions.
- **Gaussian smoothing** is an ideal effect for distant scenes.
- **Edge-preserving smoothing** has the ability to smooth the low contrast regions while preserve the sharp edges. Bilateral filter [26] is a spatial variable technique which meets the requirement.
- **Line drawing** is the very common but effective method to outline the shape of the objects. We use Difference-of-Gaussian operator [3] to detect lines for images and 3D objects.

All listed rendering techniques, except the color desaturation, are based on local filters. We thus adapt the kernel size of their filters, which use smaller filter kernel where the importance is low, to ensure the composition quality. The original color is sometimes needed to be preserved in some case, such as the important region in Figure 2, we consider it as a special style. We apply these styles directly on the

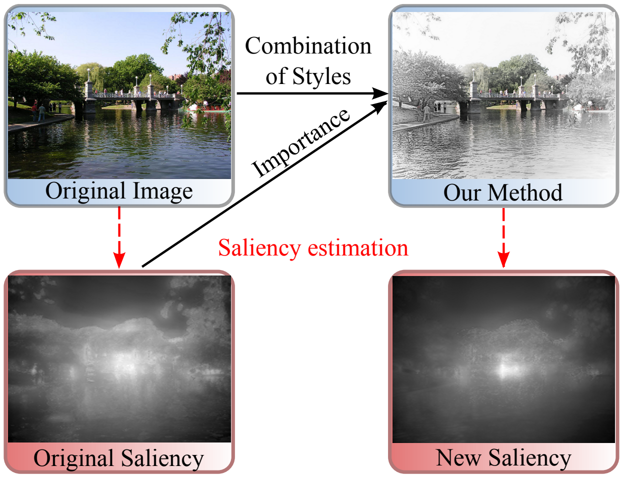


Figure 5: Overview of our evaluation on estimating the visual attention before and after our stylization by the way of saliency Estimation.

images or each frame of videos. For 3D scenes, we use them either on luminance of their shading or normal map of the 3D meshes. Combination with three styles are employed in Figure 2, 3, 4, more results combined with more styles will be shown in next section.

VI. Results

The parameters of our approach are E_i (expected importance value for each style) for the composition and the ones from each stylization, such as the maximal enhancement strength, maximal Gaussian smoothing kernel, spatial kernel and range kernel for bilateral filter, maximal spatial scale, sensitivity and sharpness of DoG edge extraction.

Figure 6 contains various stylized results generated by our system. In the first row, the left image show an example of architectural 3D models, which are rendered by photo-realistic manner and then stylized with lines, desaturated colors and enhancement from left to right. The image in the middle mimics depth-of-field phenomenon approximately, by applying a linear gradient Gaussian blurring. For 3D scenes, we first render them using classic photorealistic manner, and then stylize the rendered image. A X-ray vision in AR example is given on the right, the virtual underground pipelines are rendered using X-toon shading [27], the moving car is segmented out [28] and presented in the front of the pipelines as the original color while lines extracted to represent shape features are preserved in both background and composited region. The second row displays three frames from a video experiment, the dolphin is highlighted by increasing the luminance, the color on the boat behind the dolphin is desaturated and smoothed with edge-preservation while lines are used for sea in the background. A example on large scale image, part of the Chinese ancient painting "Qingming Festival by the Riverside" painted by Song painter Zhang Zeduan, the bottom row of Figure 6, shows the ability of our approach for complex stylization tasks.

One interesting potential of our approach is to direct viewer's attention by combining multiple rendering styles. The saliency in regions where a high-preserving style is applied will be increased, while the one in regions stylized by a low-preserving style will be reduced, as demonstrated in Figure 5. According to the saliency map of the original image, full color information (as a high-preserving style) is preserved in high-saliency regions, while desaturated (as a low-preserving style) in low-saliency regions. We then estimate a new saliency map of the processed result. Compare with the original saliency map, the new one gathers the saliency to the high-saliency regions, which shows the potential of our method to direct viewer's attention.

The performance of our approach mainly depends on the stylization approaches we employed. Fortunately, our system prefers spatial variable stylization techniques, which are naturally parallel. Therefore, we have implemented our system in GPU architecture, achieve a real time performance.

VII. Conclusion

In this paper, we have introduced a non-uniform compositing approach that combines multiple rendering styles driven by an importance map. This map is employed to guide both stylization and composition, which improves the quality and compatibilities with different input data and user interactions.

We have shown that various styles can be adopted in our system, but they are still limited in spatially varying stylization techniques. More sophisticated use of importance map may solve this limitation. Moreover, temporal coherence has not been discussed yet, though it is not a big issue in our video experiments.

Another interesting future extension is to exploit the potential ability of our approach in controlling the attention. We have discussed this possibility in Section VI, but some further development has to be done. To achieve this goal, there are at least two main issues to deal with: 1) the importance map depends mostly on saliency map, while saliency estimation techniques in the state-of-art are far from detecting accurate visual attention; 2) our system requires the user to select styles and adjust the expected importance values for each style. Though a default mode and user-friendly gradient tools are provided, arbitrarily selection by fresh users may produce unexpected results or even reverse the visual attention. A reasonable strategy to help the users to do a good selection will probably improve our system for attention control. We believe this direction of research deserves future endeavor.

Acknowledgements

This research is supported by National Basic Research Program of China (973 Program, Grant No.2009CB320802), National Natural Science Foundation of China (Grant No.60970020), and the Animare Project of French (ANR-08-JCJC-0078-01).

References

[1] D. DeCarlo and A. Santella, “Stylization and abstraction of photographs,” in *Proc. SIGGRAPH*. ACM, 2002, pp. 769–776.

[2] A. Santella and D. DeCarlo, “Abstracted painterly renderings using eye-tracking data,” in *Proc. international symposium on Non-photorealistic animation and rendering*. ACM, 2002, pp. 75–83.

[3] H. Winnemöller, S. C. Olsen, and B. Gooch, “Real-time video abstraction,” *ACM Trans. Graph. (Proc. SIGGRAPH)*, vol. 25, no. 3, pp. 1221–1226, 2006.

[4] H. Kang, S. Lee, and C. K. Chui, “Flow-based image abstraction,” *IEEE Trans. Visualization and Comput. Graph.*, vol. 15, no. 1, pp. 62–76, 2009.

[5] H. Zhao, X. Mao, X. Jin, J. Shen, F. Wei, and J. Feng, “Real-time saliency-aware video abstraction,” *Vis. Comput.*, vol. 25, pp. 973–984, 2009.

[6] J. Kyprianidis, H. Kang, and J. Döllner, “Image and video abstraction by anisotropic kuwahara filtering,” in *Comput. Graph. Forum*, vol. 28, no. 7. Eurographics/John Wiley and Sons Ltd., 2009, pp. 1955–1963.

[7] D. Luebke, M. Reddy, J. Cohen, A. Varshney, B. Watson, and R. Huebner, *Level of Detail for 3D Graphics*. Morgan Kaufmann, 2003.

[8] A. Ni, K. Jeong, S. Lee, and L. Markosian, “Multi-scale line drawings from 3D meshes,” in *Proc. symposium on Interactive 3D graphics and games*. ACM, 2006, pp. 14–17.

[9] I. Viola, M. Feixas, M. Sbert, and M. E. Gröller, “Importance-driven focus of attention,” *IEEE Trans. Visualization and Comput. Graph.*, vol. 12, pp. 933–940, 2006.

[10] J. Krüger, J. Schneider, and R. Westermann, “Clearview: An interactive context preserving hotspot visualization technique,” *IEEE Trans. Visualization and Comput. Graph.*, vol. 12, no. 5, pp. 941–948, 2006.

[11] D. Kalkofen, E. Mendez, and D. Schmalstieg, “Interactive focus and context visualization for augmented reality,” in *Proc. IEEE and ACM International Symposium on Mixed and Augmented Reality*. IEEE Computer Society, 2007, pp. 1–10.

[12] J. Chen, X. Granier, N. Lin, and Q. Peng, “On-line visualization of underground structures using context features,” in *Proc. ACM Symposium on Virtual Reality Software and Technology (VRST)*. ACM, 2010, pp. 167–170.

[13] C. Sandor, A. Cunningham, A. Dey, and V. Mattila, “An Augmented Reality X-Ray system based on visual saliency,” in *Proc. IEEE International Symposium on Mixed and Augmented Reality*. IEEE Computer Society, 2010, pp. 27–36.

[14] F. Cole, D. DeCarlo, A. Finkelstein, K. Kin, K. Morley, and A. Santella, “Directing gaze in 3d models with stylized focus,” *Proc. Eurographics Symposium on Rendering*, pp. 377–387, 2006.

[15] A. Orzan, A. Bousseau, P. Barla, and J. Thollot, “Structure-preserving manipulation of photographs,” in *Proc. international symposium on Non-photorealistic animation and rendering (NPAR)*. ACM, 2007, pp. 103–110.

[16] H. Kang and S. Lee, “Shape-simplifying Image Abstraction,” in *Comput. Graph. Forum*, vol. 27, no. 7. Eurographics/John Wiley and Sons Ltd., 2008, pp. 1773–1780.

[17] L. Itti, C. Koch, and E. Niebur, “A model of saliency-based visual attention for rapid scene analysis,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, pp. 1254–1259, 1998.

[18] T. Liu, J. Sun, N. ning Zheng, X. Tang, and H. yeung Shum, “Learning to detect a salient object,” in *Proc. IEEE conf. Comput. Vision and Pattern Recognition*. IEEE Computer Society, 2007, pp. 1–8.

[19] P. Bhat, C. Zitnick, M. Cohen, and B. Curless, “GradientShop: A gradient-domain optimization framework for image and video filtering,” *ACM Trans. Graph.*, vol. 29, pp. 1–14, 2010.

[20] T. Judd, K. Ehinger, F. Durand, and A. Torralba, “Learning to predict where humans look,” in *IEEE International Conf. on Comput. Vision (ICCV)*, 2009.

[21] K. Zeng, M. Zhao, C. Xiong, and S.-C. Zhu, “From image parsing to painterly rendering,” *ACM Trans. Graph.*, vol. 29, no. 1, pp. 2:1–2:11, 2009.

[22] H. Kang, C. Chui, and U. Chakraborty, “A unified scheme for adaptive stroke-based rendering,” *The Visual Computer*, vol. 22, no. 9, pp. 814–824, 2006.

[23] D. Comaniciu and P. Meer, “Mean shift: A robust approach toward feature space analysis,” *IEEE Trans. pattern analysis and machine intelligence*, vol. 24, no. 5, pp. 603–619, 2002.

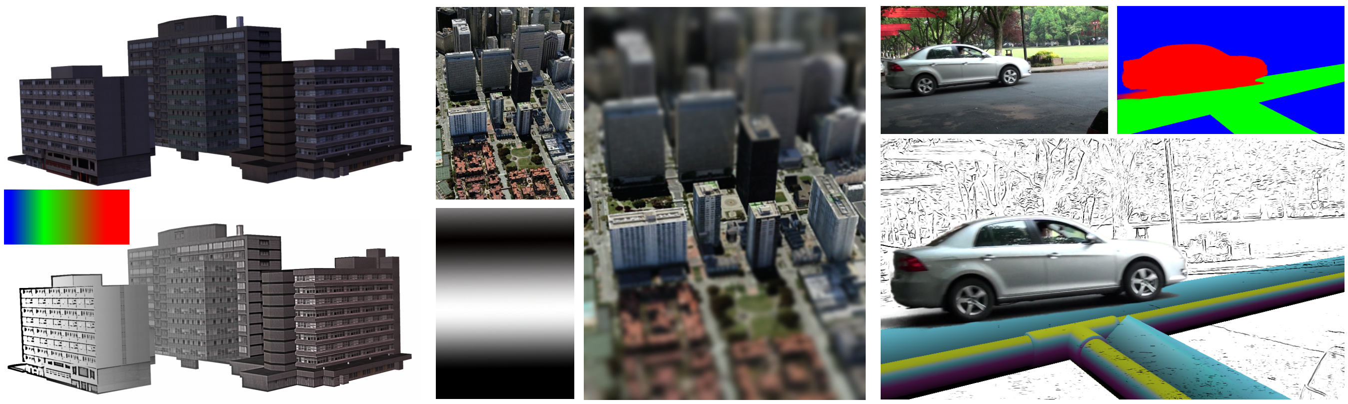
[24] G. S. W. Wyszecki, *Color Science: Concepts and Methods, Quantitative Data and Formulae*. John Wiley and Sons Ltd., 1982.

[25] T. Ritschel, K. Smith, M. Ihrke, T. Grosch, K. Myszkowski, and H.-P. Seidel, “3D Unsharp Masking for Scene Coherent Enhancement,” *ACM Trans. Graph. (Proc. of SIGGRAPH 2008)*, vol. 27, no. 3, 2008.

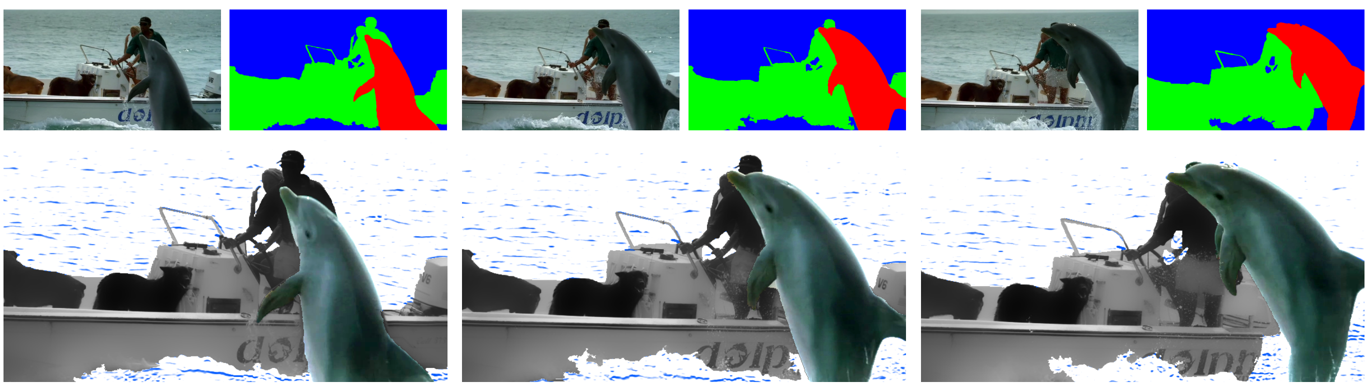
[26] C. Tomasi and R. Manduchi, “Bilateral filtering for gray and color images,” in *Pro. IEE International Conf. Comput. Vision (ICCV)*. IEEE Computer Society, 1998.

[27] P. Barla, J. Thollot, and L. Markosian, “X-toon: An extended toon shader,” in *International Symposium on Non-Photorealistic Animation and Rendering*. ACM, 2006.

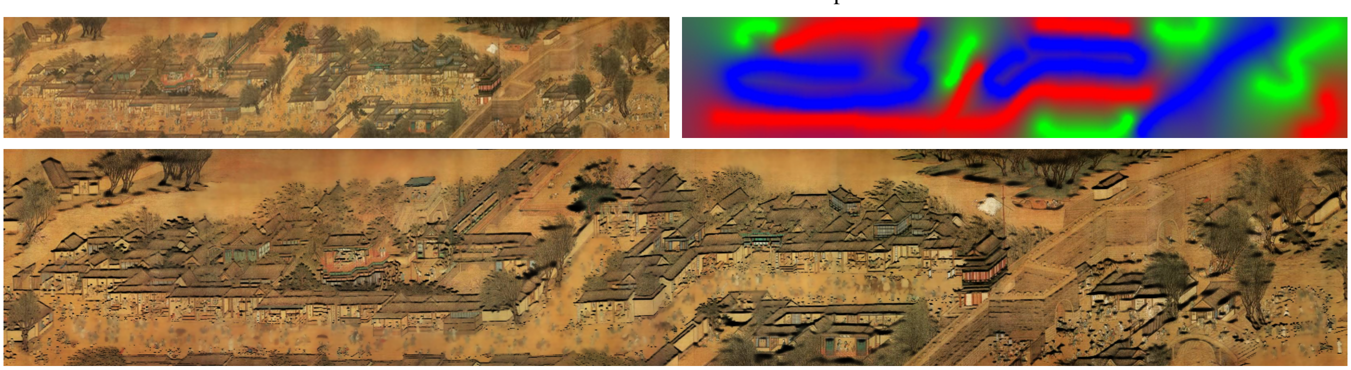
[28] F. Zhong, X. Qin, J. Chen, W. Hua, and Q. Peng, “Confidence-Based Color Modeling for Online Video Segmentation,” *Computer Vision–ACCV 2009*, pp. 697–706, 2010.



Architectural 3D models Appr. depth-of-field X-Ray vision in AR



Three frames from our video example



Complex stroke-based interaction

Figure 6: Various results. Each result is shown with its scaled-down original image and importance map. In the top row, an example of architectural 3D models, which includes a photorealistic rendered image, gradual and final result from top to bottom, is given on the left, a depth-of-field effect is approximated using symmetrical gradient importance in the middle, and the right image exhibits an example for X-Ray vision in AR. The middle row shows three result frames in our video experiments, their importance maps produced by video segmentation are also given. We also experiment complex stroke-based user interaction on large scale image in the bottom: trees are all outlined with thick lines, and buildings are silhouetted by thin lines, and the edge-preserving smoothing is applied on the pedestrians.