IMAGE COLORIZATION USING HYBRID DOMAIN TRANSFORM

Hongbo Ao, Yongbing Zhang and Qionghai Dai

Shenzhen Key Lab of Broadband Network and Multimedia,
Graduate School at Shenzhen, Tsinghua University, Shenzhen 518055, China

ABSTRACT

Image colorization is the process of spreading the user specified colors to all the expected regions in image. It is a great challenge to distinguish between texture edge and object boundary. To improve the performance of colorization, we incorporate depth and texture to accurately extract boundary information in an implicit way. Inspired by the low complexity and high efficiency properties of recently proposed domain transform, we proposed a hybrid domain transform, taking corresponding depth image of the processed texture image into account, to perform colorization and recoloring. Various experimental results demonstrate that hybrid domain transform is able to achieve better edge-aware property while maintaining the property of low complexity.

Index Terms— colorization, recoloring, propagation, hybrid domain transform, depth

1. INTRODUCTION

Image colorization, which propagates user scribbles to the whole image, is a common topic in image and video processing field. Many efficient theories and methods have been brought out to achieve better performance of colorization. Previous algorithms mainly focus on the similarity of image data itself or structure of image feature vectors and then use global optimization or local manipulation to colorize images and videos. Levin et al. adopt an interactive method to refine the color propagation [1]. But the propagation over far regions is mainly controlled by smooth constraint , which would result in poor performance. Structural similarity method needs to estimate distance between pixels and find the relations between neighbour pixels [2], but this will consume a great many times to solve the linear system. Furthermore, it will fail to work due to the lack of semantic constraints in some cases. A latest method, sparse function approximation [3], assumes that the difference of color values between neighbour pixels is given as a monotonically increasing function of the difference of grayscale values between neighbour pixels.Such a rigorous hypothesis will not always meet the practical condition and perform color propagation as well as expected although its efficiency is high. Sparse control model for color propagation [4] utilizes an adaptive strategy to automatically determine the influence of edit samples across the whole image jointly considering spatial distance, sample location, and appearance of pixels. Domain transform algorithm [5] derives a simple transform based algorithm on mathematical theory. It reduces the computational complexity to a large extent while performing good color propagation at the same time,but it fails to propagate edit on a fast-changing region in images.

Colorization result is closely related to image structures such as color, texture, as well as edge information. We have tried several ways to improve the result of colorization and found that depth information of a scene can be used to perform better propagation, since depth image has plain regions within internal objects while sharp edge can be observed within external objects.

In this paper,we devise an algorithm, termed as hybrid domain transform, to jointly incorporate texture and depth information into colorization to determine where should the user scribbles actually be propagated. Various experimental results validate the superiority of the proposed method.

The remainder of this paper is organized as follows. We will give a brief introduction of domain transform in section 2 and the detailed description of hybrid domain transform is introduced in section 3. The experimental results will be provided in section 4 to show the advantages of our algorithm. Finally, the paper is concluded in section 5.

2. DOMAIN TRANSFORM FILTER

Domain transform [5] is a simple transform to preserve the geodesic distance between points on the processed image curves. Applying domain transform, we can filter one image in the transformed domain via performing the filtering only in row order or column order rather than filtering it in a two dimensional manner, which is the main cause of heavy computational burden. Consequently, domain transform is extremely efficient to perform edge-aware filtering for its low computational complexity and distance preserving property.

For a multichannel signal, assume it can be expressed as $I(x) = (I_1(x), I_2(x), \cdots, I_c(x))$,its domain transform has the form as shown below

$$ct(u) = \int_0^u 1 + \sum_{k=1}^{c} |I_k^{'}(x)| dx, \quad (1)$$

where $I_k(x)$ is the $k-th$ channel of the signal.

Taken the smooth property of the signal and the ability to control the signal's space and range into consideration, domain transform can be scaled as

$$ct(u) = \int_0^u \frac{\sigma_H}{\sigma_s} + \sum_{k=1}^{c} \frac{\sigma_H}{\sigma_{r_k}} |I_k^{'}(x)| dx, \quad (2)$$

where $\sigma_H, \sigma_s, \sigma_{r_k} (k = 1, \cdots, c)$ are parameters to control the smooth effect of the result. After performing domain transform, we should design a filter in the transformed domain to carry out image filtering. In image colorization, the filter can be expressed as:

$$J(p) = \frac{1}{K_p} \sum_{q \in D(\Omega)} I(q) H(t(\hat{p}), t(\hat{q})) \quad (3)$$

where $t(\hat{p}) = ct(p)$ and H is a box filter expressed as:

$$H(t(\hat{p}), t(\hat{q})) = \delta\{|t(\hat{p}) - t(\hat{q})| \leq r\}, \quad (4)$$

with r being the radius of the filter.

3. HYBRID DOMAIN TRANSFORM FILTER

Domain transform have exploited the gray or color edge information to perform colorization and it works well in many cases. But sometimes it may fail to propagate user specific scribbles to the whole expected region in the image. We have studied these cases and found that many different adjacent regions in an image often have the similar gray level or color components which unable to make proper propagation according to the unclear edge information. Furthermore, objects in videos and images often have different depth in the real scene even if they have a similar or indeed the same color or gray level.

The main feature of depth image is that it contains abundant edge information of its related color image or gray image. Consequently, it is a perfect feature to perform better colorization in non-smooth regions in images and videos. Non-smooth region in images often lead to a wrong judgment of region's edge thus make colorization fail to perform an expected propagation. In contrary, depth image can help to get a correct judgment of edge in a scene with considerable depth variation.

To incorporate the depth image into color propagation, we view a pixel of an image in a high dimensional space, which is formed by pixel's color, coordinate and depth(i.e. R^6 space). After we extend original domain transform into this space, we can express our hybrid domain transform algorithm as

$$ct(u) = \int_0^u \frac{\sigma_H}{\sigma_s} + \sum_{k=1}^{c} \frac{\sigma_H}{\sigma_{r_k}} |I_k^{'}(x)| + \sigma_d |D^{'}(x)| dx, \quad (5)$$

where $\sigma_H, \sigma_s, \sigma_{r_k} (k = 1, \cdots, c), \sigma_d$ are the parameters controlling the signal's space, range and propagation performance, $D(x)$ is the depth image.

The algorithm can be actually expressed in the form

$$ct(u) = ct_I(u) + ct_d(u), \quad (6)$$

$$ct_d(u) = \int_0^u \sigma_d |D^{'}(x)| dx, \quad (7)$$

In the general form of hybrid domain transform, $ct_I(u)$ is the color term, which is actually the original domain transform, $ct_d(u)$ is the depth term, which utilizes the depth image. Here, σ_d is an important parameter which will influence the propagation result greatly by introducing more edge information of the regions in the images. Larger σ_d will enhance the influence of edge information to colorization, while smaller σ_d will eliminate the influence of edge information largely. In image colorization, we awalys set the parameter as $\sigma_H = \sigma_s$, $\sigma_s > 0$, $\sigma_{r_k} > 0 (k = 1, \cdots, c)$ and $\sigma_d > 0$.

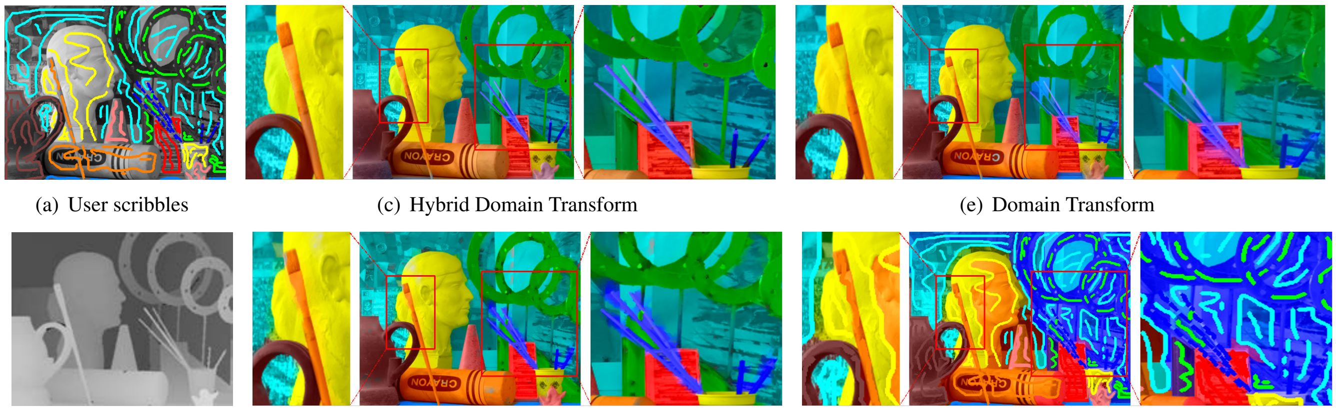
We can use the proposed hybrid domain transform to perform a better colorization in many cases where poor performance would be obtained by domain transform.

4. EXPERIMENTAL RESULTS

We apply the proposed hybrid domain transform algorithm for image colorization and recoloring in this section. Various experiments are carried out to validate the advantages of the proposed hybrid domain transform over all the competing reference methods.

4.1. Colorization

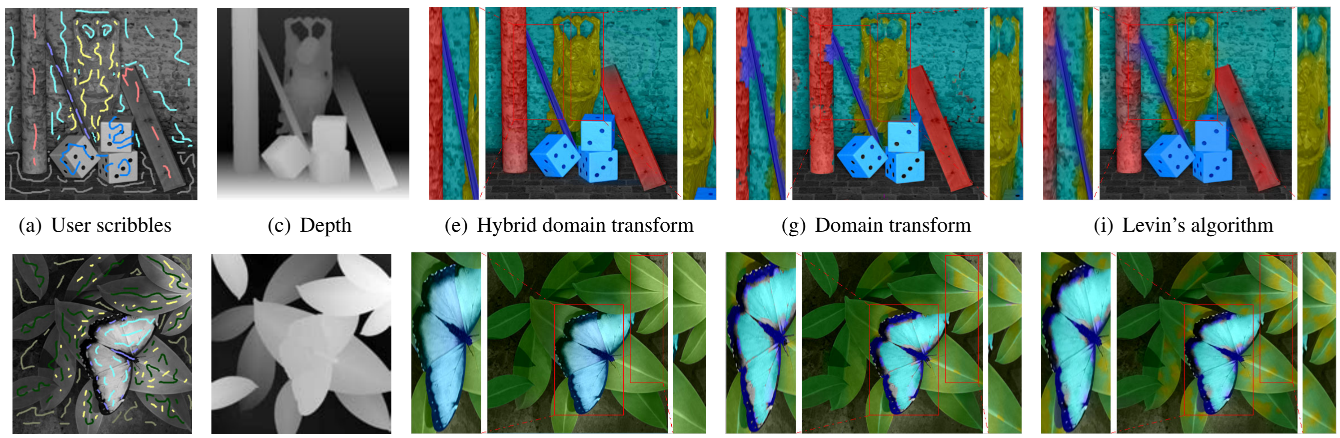
For each input gray image, user puts some color scribbles, which are then propagated to the whole image by colorization. For the convenience of expression, user scribbles are defined to be S, and a normalization function N is also defined to keep track of how many colors should be propagated to each pixel in the result. It should be noted that in function N, one is assigned to positions where scribbles are provided and zero otherwise. We filter image S and image N in a way similar to Eq. (5), and the filtered versions are expressed as $\tilde{S}$ and $\tilde{N}$. We can get the propagated color as $\tilde{S}/\tilde{N}$ and then add the original luminance of the input gray scale image to construct a new color image R in YUV color space. Finally, we convert the image R to RGB color space to get the colorization result $\tilde{R}$.



(a) User scribbles (c) Hybrid Domain Transform (e) Domain Transform

(b) Depth (d) Levin's Algorithm (f) Uruma's Algorithm

Fig. 1. Colorization result of Art.(Parameters of our algorithm: $\sigma_H = \sigma_s = 180, \sigma_{r_k} = 1.0 (k = 1, 2, 3), \sigma_d = 5.0$)



(a) User scribbles (c) Depth (e) Hybrid domain transform (g) Domain transform (i) Levin's algorithm

(b) User scribbles (d) Depth (f) Hybrid domain transform (h) Domain transform (j) Levin's algorithm

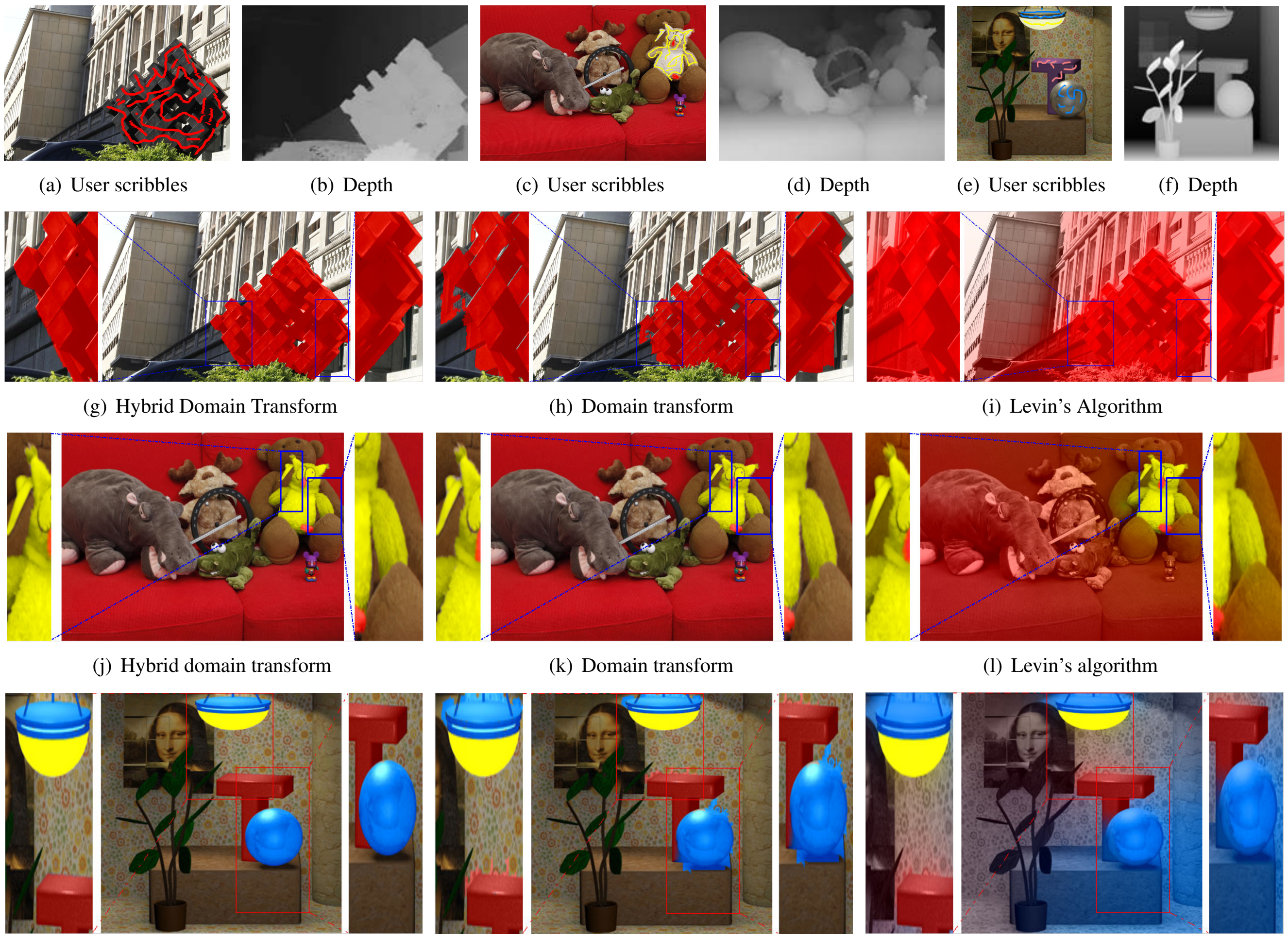
Fig. 2. Colorization result.(Parameters of our algorithm: $\sigma_H = \sigma_s = 200, \sigma_{r_k} = 0.8 (k = 1, 2, 3), \sigma_d = 8.0$)

We select Levin's method [1], the most recent method by Uruma [3], and the traditional domain transform [5] to perform and compare colorizations with the proposed hybrid domain transform under the same input scribbles, as indicated in Fig. 1 (a). Here, the depth image we utilize is the ground truth downloaded from [6]. From the results in Fig. 1 (c)-Fig. 1 (f), we can conclude that the proposed hybrid domain transform achieves a better colorization performance over all the competing algorithms. Especially, Uruma's method [3] can not propagate the colors quite well under the given scribbles. From the left parts of Fig. 1 (c)-Fig. 1 (f), we can see clearly that the edge of the brush is more sharp and correct in the result generated by our method while all the other algorithms make an incorrect distinguishing between the brush and the plaster statue. Furthermore, clear edge can be seen from the right part of Fig. 1 (c) which is colorized by the proposed hybrid domain transform while the boundary is obscure on right parts of Fig. 1 (d)-(f). More colorization results can be found in Fig. 2. All the detail results validate the superior performance of the hybrid domain transform in terms of edge preserving in colorization.

4.2. Recoloring

In recoloring, a set of regions of interest are defined by color scribbles, e.g. a color c_i defines a region R_i . The contribution of each p is defined as $\tilde{N}_{R_i} / \sum_j \tilde{N}_{R_j}(p)$, where$\tilde{N}_{R_i}$ can be obtained in a similar way as the processing steps in colorization.

In Fig. 3, we recolor three images in Fig. 3 (a), Fig. 3 (c), Fig. 3 (e). Clear boundary can be observed in the depth images Fig. 3 (b), Fig. 3 (d), Fig. 3 (f) and this will enable the hybrid domain transform to achieve superior recoloring result. Here, the utilized depth images are downloaded from [7]. Fig. 3 (g), Fig. 3 (h) and Fig. 3 (i) are three recolored results generated by hybrid domain transform, domain transform and Levin's method. We can conclude that Fig. 3 (g) is much better than Fig. 3 (h) and Fig. 3 (i), since Fig. 3



(a) User scribbles (b) Depth (c) User scribbles (d) Depth (e) User scribbles (f) Depth

(g) Hybrid Domain Transform (h) Domain transform (i) Levin's Algorithm

(j) Hybrid domain transform (k) Domain transform (l) Levin's algorithm

(m) Hybrid Domain Transform (n) Domain transform (o) Levin's Algorithm

Fig. 3. Recoloring results.(Parameters of our algorithm: $\sigma_H = \sigma_s = 150, \sigma_{r_k} = 1.2(k = 1, 2, 3), \sigma_d = 10.0$)

(h) has a worse result around the vicinity of statues boundary and the scribbles are propagated to the whole image badly in Fig. 3 (i). To verify such observation, amplified figures of local regions are provided in Fig. 3, where we can see obvious incorrect recolored regions on the left parts of Fig. 3 (h) and Fig. 3 (i) while Fig. 3 (g) seems to be perfect. The same conclusion can be drawn from the amplified local regions in right parts of images Fig. 3 (g), Fig. 3 (h) and Fig. 3 (i). Fig. 3 (j)-Fig. 3 (l) and Fig. 3 (m)-Fig. 3 (o) are another two group results generated by hybrid domain transform, domain transform and Levin's method, respectively. Amplified portions of recolored results are depicted on the left and right parts of Fig. 3 (j)-Fig. 3 (l)and Fig. 3 (m)-Fig. 3 (o). It can be observed that Levin's algorithm achieves a bad result, as indicated in image Fig. 3 (l) and Fig. 3 (o). Domain transform performs a better recoloring in image Fig. 3 (k) and Fig. 3 (n) while hybrid domain transform generates the best result among all of them.

5. CONCLUSION

A hybrid domain transform is proposed in the paper to improve the performance of colorization and recoloring. Depth image is taken into consideration in the geodesic distance preserving algorithm as an extra information. Since depth image is much smooth within internal object while exhibits distinct variations across the object boundaries, it is of great help to propagate scribbles within plain regions while stop the propagation around the edge regions. Besides, the hybrid domain transform can be finished by iterations of 1D filtering along the horizontal and vertical directions,which leading to a fast implementation of the proposed algorithm.

6. ACKNOWLEDGEMENT

This work was partially supported by National Natural Science Foundation of China(61170195) and the Joint Funds of National Science Foundation of China(U1201255,U1301257).

7. REFERENCES

[1] A. Levin, D. Lischinski, and Y. Weiss, "Colorization using optimization," in *ACM Transactions on Graphics (TOG)*, vol. 23, no. 3. ACM, 2004, pp. 689–694.

[2] X. Chen, D. Zou, Q. Zhao, and P. Tan, "Manifold preserving edit propagation," *ACM Transactions on Graphics (TOG)*, vol. 31, no. 6, p. 132, 2012.

[3] U. Kazunori, K. Katsumi, T. Tomohiro, and F. Toshihiro, "Image colorization algorithm using series approximate sparse function," *Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1215–1219, 2014.

[4] L. Xu, Q. Yan, and J. Jia, "A sparse control model for image and video editing," *ACM Transactions on Graphics (TOG)*, vol. 32, no. 6, p. 197, 2013.

[5] E. S. Gastal and M. M. Oliveira, "Domain transform for edge-aware image and video processing," in *ACM Transactions on Graphics (TOG)*, vol. 30, no. 4. ACM, 2011, p. 69.

[6] The Disney Research Datasets. [Online]. Available: http://vision.middlebury.edu/stereo/data/

[7] The Middlebury Stereo Datasets. [Online]. Available: http://www.disneyresearch.com/project/lightfields/