Image and video decomposition and editing.

JIANCHAO TAN, George Mason University

My research has centered around understanding the colorful appearance of physical and digital paintings and images. My work focuses on decomposing images or videos into more editable data structures called layers, to enable efficient image or video re-editing.

Given a time-lapse painting video, we can recover translucent layer strokes from every frame pairs by maximizing translucency of layers for its maximum re-usability, under either digital color compositing model or a physically inspired nonlinear color layering model, after which, we apply a spatial-temporal clustering on strokes to obtain semantic layers for further editing, such as global recoloring and local recoloring, spatial-temporal gradient recoloring and so on.

With a single image input, we use the convex shape geometry intuition of color points distribution in RGB space, to help extract a small size palette from a image and then solve an optimization to extract translucent RGBA layers, under digital alpha compositing model. The translucent layers are suitable for global and local image recoloring and new object insertion as layers efficiently.

Alternatively, we can apply an alternating least square optimization to extract multi-spectral physical pigment parameters from a single digitized physical painting image, under a physically inspired nonlinear color mixing model, with help of some multi-spectral pigment parameters priors. With these multi-spectral pigment parameters and their mixing layers, we demonstrate tonal adjustments, selection masking, recoloring, physical pigment understanding, palette summarization and edge enhancement.

Our recent ongoing work introduces an extremely scalable and efficient yet simple palette-based image decomposition algorithm to extract additive mixing layers from single image. Our approach is based on the geometry of images in RGBXY-space. This new geometric approach is orders of magnitude more efficient than previous work and requires no numerical optimization. We demonstrate a real-time layer updating GUI. We also present a palette-based framework for color composition for visual applications, such as image and video harmonization, color transfer and so on.

CCS Concepts: • **Computing methodologies → Image manipulation**; **Image processing**;

Additional Key Words and Phrases: images, layers, time-lapse, video, painting, palette, generalized barycentric coordinates, harmonization, contrast, convex hull, RGB, color space, recoloring, compositing, mixing

ACM Reference Format:
Jianchao Tan. 2018. Image and video decomposition and editing.. 1, 1 (June 2018), 4 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

My research investigates the colorful appearance of images and videos for re-editing purposes. I target digital paintings, digitized physical paintings, time-lapse painting history, and natural images and videos.

The colorful appearance of digital paintings are modeled by Porter and Duff "over" compositing model (PD), as described in Porter and Duff [1984], which is also called alpha compositing model. This is a standard color compositing model in industry. It models current canvas color as a linear interpolation between previous canvas color and new coming color. The interpolation value is called alpha values or opacity values. Artists usually use a data structure called layers to organize paintings in painting software. These layers are RGBA image format, where the fourth channel is per pixel alpha values, as used in Tan et al. [2016]. There is also an additive linear mixing model, which model each pixel color as a convex interpolation of small set of pigment colors, as used in Aksoy et al. [2017]. Artists use palette to create either digital paintings with painting software or real world paintings with brushes. The palette colors are representative colors of one painting. Many recent works focused on extracting palette from paintings or photo images, such as Abed [2014]; Aharoni-Mack et al. [2017]; Chang et al. [2015]; Lin and Hanrahan [2013]. Some works use extracted palette to help recolor images, as described in Aharoni-Mack et al. [2017]; Chang et al. [2015]; Mellado et al. [2017].

The colorful appearance of a physical painting is determined by the distribution of paint pigments across the canvas, which can be modeled either as a per-pixel linear mixture of a small number of palette pigments with absorption and scattering coefficients (mixing model), or a per pixel layering of a small number of palette pigments with reflectance and transmittance coefficients (layering model). The model was proposed by Kubelka and Munk (KM), as described in Barbarić-Mikočević and Itrić [2011]; Kubelka and Munk [1931], which is a nonlinear color compositing model. This model is also widely used in recent painting analysis and non-photo realistic rendering works, such as Abed [2014]; Aharoni-Mack et al. [2017]; Baxter et al. [2004]; Curtis et al. [1997]; Tan et al. [2018a, 2015].

Decomposing layers from timelapse paintings. The creation of a painting, in the physical world or digitally, is a process that occurs over time. Later strokes cover earlier strokes, and strokes painted at a similar time are likely to be part of the same object. In the final painting, this temporal history is lost, and a static arrangement of color is all that remains. The rich literature for interacting with image editing history cannot be used. To enable these interactions, we presented a set of techniques to decompose a time lapse video of a painting (defined generally to include pencils, markers, etc.) into a sequence of translucent "stroke" images. We presented translucency-maximizing solutions for recovering physical (Kubelka and Munk layering) or digital (Porter and Duff "over" compositing operation) paint parameters from before/after image pairs and enable spatial temporal video recoloring. We also presented a pipeline for processing real-world videos of paintings capable of handling

Author's address: Jianchao Tan, George Mason University, tanjianchaoustc@gmail.com.

XXXX-XXXX/2018/6-ART
https://doi.org/10.1145/nnnnnnn.nnnnnnn

long-term occlusions, such as the painter's hand and its shadow, color shifts, and noise. This work is described in Tan et al. [2015].

Decomposing layers from images. We presented a decomposition technique to decompose an image into layers via RGB-space geometry. In our decomposition, each layer represents a single-color coat of paint applied with varying opacity. In RGB-space, the linear nature of the standard Porter-Duff "over" pixel compositing operation implies a convex hull structure. The vertices of the convex hull of image pixels in RGB-space correspond to a palette of paint colors. These colors may be "hidden" and inaccessible to algorithms based on clustering visible colors (Chang et al. [2015]). For our layer decomposition, users choose the palette size (degree of simplification to perform on the convex hull), as well as a layer order for the paint colors (vertices). We applied a modified progressive hull method in Sander et al. [2000] to simplify convex hull to get target size of palette. We then solve a constrained optimization problem to find translucent, spatially coherent opacity for each layer, such that the composition of the layers reproduces the original image. We demonstrate the utility of the resulting decompositions for recoloring (global and local) and object insertion. Our layers can be interpreted as generalized barycentric coordinates, as described in Floater [2015]; we compare to these and other recoloring approaches. This work is described in Tan et al. [2016].

Pigment-based image analysis and editing. We also presented an alternating nonlinear least square optimization method based on Kubelka-Munk mixing color model to efficiently recover multispectral pigment parameters (absorption and scattering) from a single RGB image, yielding a plausible set of pigments and their mixture maps with a low RGB reconstruction error. Using our decomposition, we repose standard digital image editing operations as operations in pigment space rather than RGB space, with interestingly novel results. We demonstrate tonal adjustments, selection masking, cut-copy-paste, recoloring, palette summarization, and edge enhancement. This work is described in Tan et al. [2018a].

Ongoing: Image decomposition and harmonization. A detailed description of this work is in our first version draft on arXiv (Tan et al. [2018b]).

At the core of our and other recent approaches Aksoy et al. [2017]; Chang et al. [2015]; Tan et al. [2016]; Zhang et al. [2017] to image editing, images are decomposed into a palette and associated per-pixel compositing or mixing parameters. We propose a new, extremely efficient yet simple and robust algorithm to do so. Our approach is inspired by the geometric palette extraction technique of Tan et al. [2016]. We consider the geometry of 5D RGBXY-space, which captures color as well as spatial relationships and eliminates numerical optimization. After an initial palette is extracted automatically (given an RMSE reconstruction threshold), the user can edit the palette in our GUI and obtain new decompositions instantaneously. Our algorithm's performance is extremely efficient even for very high resolution images ($\geq$ 100 megapixels)—20x faster than the state-of-the-art Aksoy et al. [2017]. We demonstrate applications like color harmonization, color transfer, are greatly simplified by our framework.

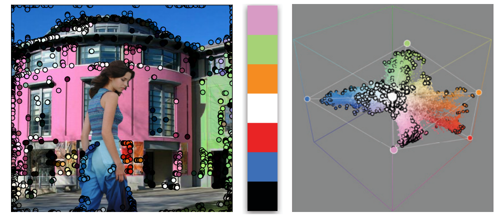


Fig. 1. Visualization of the two convex hulls. The simplified RGB convex hull is the basis for the methods in Tan et al. [2016], capturing the colors of an image but not their spatial relationships. Our 5D RGBXY convex hull captures color and spatial relationship at the same time. We visualize its vertices as small circles; its 5D simplices are difficult to visualize. Our approach splits image decomposition into a two-level geometric problem. The first level are the RGBXY convex hull vertices that mix to produce any pixel in the image. The second level are the simplified RGB convex hull vertices, which serve as the palette RGB colors. Since the RGBXY convex hull vertices lie inside the RGB convex hull, we find mixing weights that control the color of the RGBXY vertices. The two levels combined allow instant recoloring of the whole image. The left figure shows the locations of the RGBXY vertices in image space. The right figure shows the geometric relationships between the 3D and 5D convex hull vertices, and how the simplified RGB convex hull captures the same color palette for both algorithms.

2 IMAGE DECOMPOSITION

2.1 Image decomposition via RGBXY convex hull

In this work, we extract additive linear mixing layers from input image. We provide a fast and simple, yet spatially coherent, geometric construction.

Spatial Coherence To provide spatial coherence, our key insight is to manipulate color mixing weights in 5D RGBXY-space, where XY are the coordinates of a pixel in image space, so that spatial relationship are considered along with color in a unified way (Figure 1). We first compute convex hull of the image in RGBXY-space. We then compute Delaunay generalized barycentric coordinates (weights) for every pixel in the image in terms of the 5D convex hull. Pixels that have similar colors *or* are spatially adjacent will end up with similar weights, meaning that our layers will be smooth both in RGB and XY-space. These mixing weights form an $Q \times N$ matrix W_{RGBXY}, where N is the number of image pixels and Q is the number of RGBXY convex hull vertices. We also compute W_{RGB}, generalized barycentric coordinates (weights) for the RGBXY convex hull vertices in the 3D simplified convex hull. W_{RGB} is a $P \times Q$ matrix, where P is the number of vertices of the simplified RGB convex hull (the palette colors). The final weights for the image are obtained via matrix multiplication: $W = W_{\text{RGB}} W_{\text{RGBXY}}$, which is a $P \times N$ matrix that assigns each pixel weights solely in terms of the simplified RGB convex hull. These weights are smooth both in color and image space. To decompose an image with a different RGB-palette, one only needs to recompute W_{RGB} and then perform matrix multiplication. Computing W_{RGB} is extremely efficient, since it depends only on the palette size and the number of RGBXY convex hull vertices. It is independent of the image size and allows users to experiment with image decompositions based on interactive palette editing (Figure 3b).

2.2 Evaluation

We generate state-of-the-art decompositions in terms of quality Figure 2b compares recolorings created with our layers to those from Aksoy et al. [2017], Tan et al. [2016], and Chang et al. [2015]. Figure 2a shows a direct comparison between our additive mixing layers and those of Aksoy et al. [2017] for direct inspection. In Figure 2c, we compare the running time of additive mixing layer decomposition techniques. We also ran an additional 6 extremely large images containing 100 megapixels (not shown in the plot). The 100 megapixel images took on average 12.6 minutes to compute. Peak memory usage was 15 GB. For further improvement, our approach could be parallelized by dividing the image into tiles, since the convex hull of a set of convex hulls is the same as the convex hull of the underlying data. A working implementation (48 lines of code) of the RGBXY decomposition method can be found in **this code link**. The "Layer Updating" performance is nearly instantaneous, taking a few milliseconds to, for 10 MP images, a few tens of milliseconds to re-compute the layer decomposition given a new palette.

Interactive Layer Decompositions To take advantage of our extremely fast layer decomposition, we implemented a web GUI for viewing and interacting with layer decompositions (Figure 3a). An editing session begins when a user loads an image and precomputes RGBXY weights. Our GUI allows users edit palettes and see the resulting layer decomposition in real-time. See Figure 3b for a result created with our GUI. In this example, the automatic palette (right) corresponding mixing weights become sparser as a result of interactive editing.

3 APPLICATIONS

3.1 Layer-based editing

My research can enable various editing applications on images and videos with the help of various extracted layers. For translucent layers extracted from time-lapse painting video, we can enable temporal-spatial gradient recoloring, as shown in Figure 4a. We can also do spatial-temporal clustering to obtain semantic structure for further editing, as shown in Figure 4b For alpha compositing layers from single image, we can recoloring by changing layer palette colors in real time, as shown in Figure 2b. We can also insert new object as new layer to merge naturally into original image, as shown in Figure 5a. For additive mixing layers extracted from digitized physical paintings, we can enable more editing beyond RGB, as shown in Figure 5b.

3.2 Harmonization and color transfer

We can simplify color harmonization fitting procedure of Cohen-Or et al. [2006] by replacing their whole pixel histograms with just our palette colors, and we project palette colors onto template axis in LCh-space to enable color harmonization. Because we use a spatially coherent image decomposition, no additional work is needed to prevent discontinuous recoloring as in Cohen-Or et al. [2006]. Figure 5c shows different harmonic templates enforced over the same input image. Our pipeline can naturally extend to video input, by simply applying our image decomposition and harmonization on each frame independently. Surprisingly, we can obtain a pretty temporal coherent layer decomposition without additional processing beyond the proposed framework. Additionally, our palette based harmonization template fitting can enable color transfer between input image and reference image. More details can be found in Tan et al. [2018b]

4 CONCLUSION

My work focus on manipulating color of images and videos, exploring several methods to decompose images and videos into more editable layers, followed by several editing applications, such as recoloring and insertion. Our ongoing work presented a very efficient, intuitive and capable framework for color composition by exploring RGBXY space geometry. It allows us to formulate previous and novel approaches to color harmonization and color transfer with very robust results. Our palette manipulations can be plugged into any palette-based system. Our image decomposition can be used generally by artists for manual editing or in other algorithms.

4.1 Future work

Palette for vector graphics We are also exploring the color palettes editing on vector graphics input, for example, palette based suggestive colorization on gray scale vector graphics, which may involve the usage of several deep learning techniques.

Spatial-temporal palette of videos We also also plan to extend our framework to video, exploring the spatial-temporal coherence of palettes, to potentially provide more robust color grading methods. Additionally, combining with semantic segmentation informations of video frames, we can provide semantic spatial-temporal palettes model to enable better semantic video color grading.

REFERENCES

Farhad Moghareh Abed. 2014. Pigment identification of paintings based on Kubelka-Munk theory and spectral images. (2014).

Elad Aharoni-Mack, Yakov Shambik, and Dani Lischinski. 2017. Pigment-Based Recoloring of Watercolor Paintings. (2017).

Yağiz Aksoy, Tunç Ozan Aydin, Aljoša Smolić, and Marc Pollefeys. 2017. Unmixing-based soft color segmentation for image manipulation. *ACM Transactions on Graphics (TOG)* 36, 2 (2017), 19.

Vesna Džimbeg-Malčić Željka Barbarić-Mikočević and Katarina Itrić. 2011. Kubelka-Munk theory in describing optical properties of paper (I). *Technical Gazette* 18, 1 (2011), 117–124.

William Baxter, Jeremy Wendt, and Ming C Lin. 2004. IMPaSTo: a realistic, interactive model for paint. In *Proceedings of the 3rd international symposium on Non-photorealistic animation and rendering*. ACM, 45–148.

Huiwen Chang, Ohad Fried, Yiming Liu, Stephen DiVerdi, and Adam Finkelstein. 2015. Palette-based Photo Recoloring. *ACM Trans. Graph.* 34, 4 (July 2015).

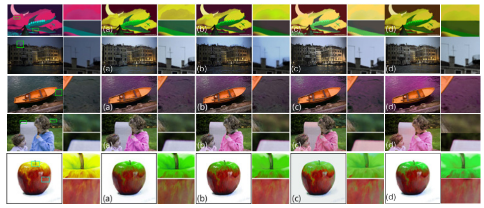
Daniel Cohen-Or, Olga Sorkine, Ran Gal, Tommer Leyvand, and Ying-Qing Xu. 2006. Color Harmonization. In *ACM SIGGRAPH 2006 Papers (SIGGRAPH '06)*. ACM, New York, NY, USA, 624–630. https://doi.org/10.1145/1179352.1141933

Cassidy J Curtis, Sean E Anderson, Joshua E Seims, Kurt W Fleischer, and David H Salesin. 1997. Computer-generated watercolor. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*. ACM Press/Addison-Wesley Publishing Co., 421–430.

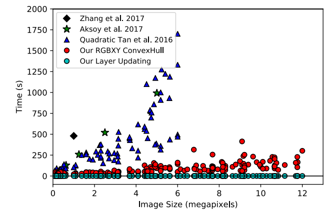
Michael S Floater. 2015. Generalized barycentric coordinates and applications. *Acta Numerica* 24 (2015), 161–214.

Paul Kubelka and Franz Munk. 1931. An article on optics of paint layers. *Z. Tech. Phys* 12, 593-601 (1931).

(a) Layer comparison: Aksoy et al. [2017] creates an overabundance of layers (two red layers above) and does not extract the blueish tint, which appears primarily in mixture. Our RGBXY technique identifies mixed colors is able to separate the translucent purple haze in front of the girl's face. Additionally, our GUI allows editing the palette to modify layers in real time, results are shown in Figure 3b

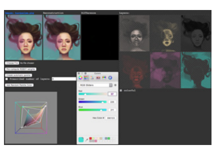


(b) To evaluate our RGBXY decomposition algorithm, we compare our layers with previous approaches in a recoloring application. From left to right: (a) Aksoy et al. [2017], (b) Tan et al. [2016], (c) Chang et al. [2015] and (d) our approach. Our recoloring quality is similar to the state of the art, but our method is orders of magnitude faster and allows interactive layer decomposition while editing palettes.

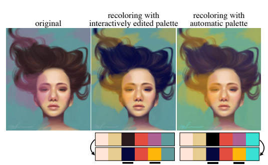


(c) Running time comparison. We evaluated our RGBXY algorithm on 170 images up to 12 megapixels and an additional six 100 megapixel images (not shown; average running time 12.6 minutes). Our algorithm's performance is orders of magnitude faster and scales extremely well with image size. Our layer updating is real time.

Fig. 2. Layer comparison, recoloring comparison and running time comparison.

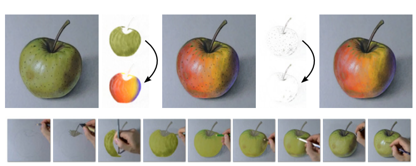


(a) Our GUI for interactively editing palettes. As users either drag the palette colors or click layers to edit color wheel, the layer decomposition updates lively.

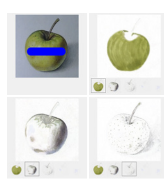


(b) The user edits the automatically generated palette in GUI to ensure that the background and hair colors are directly represented. As a result, editing the purple haze and hair no longer affects the background color.

Fig. 3. Live updating GUI and recoloring comparison.

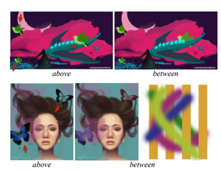


(a) From a time lapse recording of the painting process (bottom) we extract translucent paint layers into a temporal creation history. This allows artists to produce complex effects such as color gradients controlled by time (top left) or temporal eraser (top right).

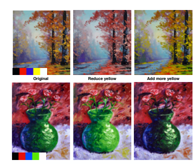


(b) The user draws a selection scribble, the system then suggests several temporal clusters.

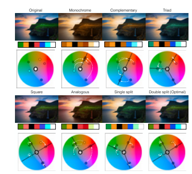
Fig. 4. Timelapse painting decomposition and editing application



(a) Inserting graphics as new layers between our decomposed layers produces a more natural result than pasting graphics above.



(b) Adjusting the mixing weight of yellow pigment and the scattering coefficients of the green pigment.



(c) Results of our harmonic templates fit to the an input image.

Fig. 5. Different editing applications based on different layers

Sharon Lin and Pat Hanrahan. 2013. Modeling how people extract color themes from images. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 3101–3110.

Nicolas Mellado, David Vanderhaeghe, Charlotte Hoarau, Sidonie Christophe, Mathieu Brédif, and Loic Barthe. 2017. Constrained palette-space exploration. *ACM Transactions on Graphics (TOG)* 36, 4 (2017), 60.

Thomas Porter and Tom Duff. 1984. Compositing digital images. In *ACM Siggraph Computer Graphics*, Vol. 18. ACM, 253–259.

Pedro V Sander, Xianfeng Gu, Steven J Gortler, Hugues Hoppe, and John Snyder. 2000. Silhouette clipping. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*. ACM Press/Addison-Wesley Publishing Co., 327–334.

Jianchao Tan, Stephen DiVerdi, Jingwan Lu, and Yotam Gingold. 2018a. Pigmento: Pigment-Based Image Analysis and Editing. *IEEE Transactions on Visualization and Computer Graphics, to appear* (2018).

Jianchao Tan, Marek Dvorožňák, Daniel Sýkora, and Yotam Gingold. 2015. Decomposing Time-lapse Paintings into Layers. *ACM Trans. Graph.* 34, 4, Article 61 (July 2015), 10 pages. https://doi.org/10.1145/2766960

Jianchao Tan, Jose Echevarria, and Yotam Gingold. 2018b. Pigmento: Pigment-Based Image Analysis and Editing. *arXiv preprint arXiv:1804.01225* (2018).

Jianchao Tan, Jyh-Ming Lien, and Yotam Gingold. 2016. Decomposing Images into Layers via RGB-space Geometry. *ACM Trans. Graph.* 36, 1, Article 7 (Nov. 2016), 14 pages. https://doi.org/10.1145/2988229

Qing Zhang, Chunxia Xiao, Hanqiu Sun, and Feng Tang. 2017. Palette-Based Image Recoloring Using Color Decomposition Optimization. *IEEE Transactions on Image Processing* 26, 4 (2017), 1952–1964.