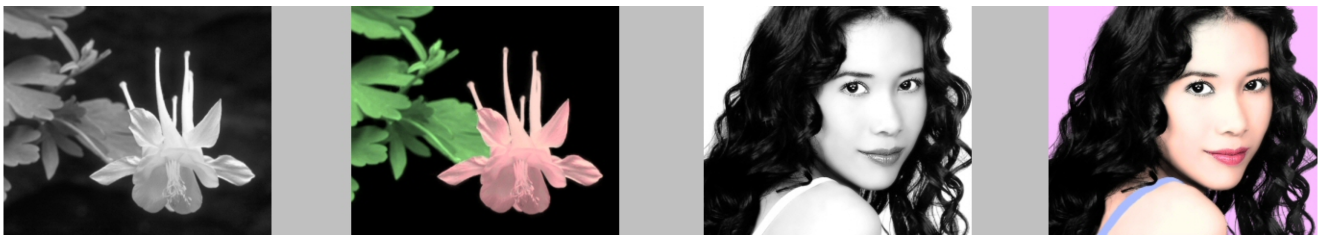
published as: T. Chen, Y. Wang, V. Schillings, Christoph Meinel: Grayscale Image Matting and Colorization;
Proceedings of Asian Conference on Computer Vision (ACCV 2004);
Jeju Island (South Korea), 2004, pp. 1164-1169.

GRAYSCALE IMAGE MATTING AND COLORIZATION *

Tongbo Chen *Yan Wang* *Volker Schillings* *Christoph Meinel*

FB IV-Informatik, University of Trier, Trier 54296, Germany
{chen, schillings, meinel}@ti.uni-trier.de



A A' B B'

Fig. 1. Grayscale image matting and colorization results. A and B are the input grayscale images to our algorithm, while A' and B' are the output color images.

ABSTRACT

This paper presents a novel approach to grayscale image matting and colorization. The first part of this approach is an efficient grayscale image matting algorithm in Bayesian framework. The foreground and background color distributions, and the alpha's distribution are modelled with spatially varying sets of Gaussians. The major novelties of this matting algorithm are the introduction of alpha's distribution and gradient into the Bayesian framework and an efficient optimization scheme. This grayscale image matting algorithm can effectively handle objects with intricate and vision sensitive boundaries, such as hair strands or facial organs. In the second part, by combining the grayscale image matting algorithm with color transferring techniques, an efficient colorization scheme is proposed, which provides great improvement over existing techniques for some difficult cases, such as human faces or images with confusing luminance distribution.

1. INTRODUCTION

Digital image matting is a critical operation in commercial television, film production, and advertisement design. The basic process of matting techniques is to extract embedded foreground objects from a background image by estimating a color and opacity for the foreground element at each pixel. The opacity value at each pixel is typically called its α, and the opacity image, taken as a whole is referred to as the α matte. Usually, the extracted foreground objects and their corresponding mattes are used to composite new images or video clips, which enable designers or directors a great freedom to achieve imaginative and impressive visual effects.

Nevertheless, extracting matte is particularly difficult for some notoriously intricate cases such as thin wisps of fur or hair. Recently have seen great advances in digital matting techniques that achieved impressive results for difficult cases [1, 2, 11]. To our knowledge, Chuang et al.'s Bayesian approach [2] achieved the best results for difficult cases. However, this method only works very efficiently for color image. For grayscale image, the matting problem is less constrained and the direct adaptation of Chuang et al.'s method will lead to failure. The method in this paper follows Chuang et al.'s Bayesian framework and improves it by modelling alpha's distribution and introducing the image gradient into the model. After such improvement, the Bayesian method works very well for grayscale image.

One of the important applications of grayscale image matting algorithm is to combine with color transferring techniques to achieve object-based colorization, where objects in the same image are colorized independently.

* IN PROCEEDINGS OF ACCV2004, JAN 27-30, 2004, JEJU ISLAND, KOREA, PP. 1164-1169

Grayscale image colorization can find its applications in black and white photo editing, classic movies colorization, and scientific illustrations. Colorization can increase dramatically the visual appeal of grayscale images and perceptually enhance scientific illustrations.

Welsh et al. [3] proposed a grayscale image colorization method that works very impressively for natural images and scientific illustration images. In general, Welsh et al.'s method works well on scenes where the image is divided into distinct luminance clusters or where each region has distinct textures. However, their current technique does not work very well with human faces. In this paper, we combine the efficient grayscale image matting algorithm and color transferring techniques to improve the colorization dramatically. Fig. 2 shows the overview of our algorithm.

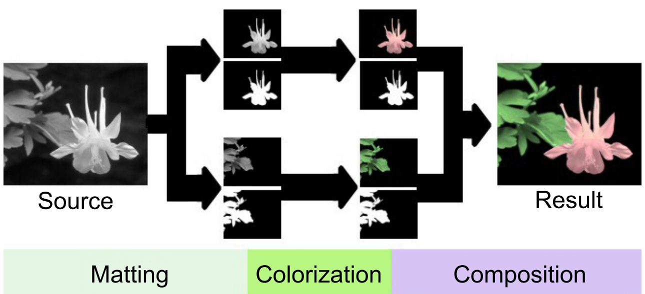


Fig. 2. Algorithm overview. First, the source grayscale image is separated into different objects using the grayscale image matting algorithm. Then, the objects are colorized using color transferring technique. Finally, the colorized objects are composited using alpha blending to reach the ultimate colorization.

2. PREVIOUS WORK

In this section, we describe the main components of previous work: digital matting and color transferring techniques. For each case, we discuss existing work, the particular methods we have selected and the adaptation and improvement.

2.1. Digital matting

In 1984, Porter and Duff [13] introduced the digital analog of the matte — the alpha channel — and showed how synthetic images with alpha could be useful in creating complex digital images. The most common compositing equation is as follows:

$$C = \alpha F + (1-\alpha)B \tag{1}$$

where C, F, and B are the pixel's composite, foreground, and background colors respectively, and alpha is the pixel's opacity component.

Blue screen matting [9] was among the first techniques used for live action matting. One limitation of blue screen is the reliance on a controlled environment. The more general approach to digital matting is to extract foreground objects and their mattes from natural images, known as natural image matting.

Corel's Knockout is the most successful commercial package for natural image matting. The key technique behind this tool is described in patents by Berman et al. [14]. In [1], Ruzon and Tomasi proposed a comparable technique based on statistical observation of natural images. More recently, Hillman et al. [11] used principle component analysis to estimate the optimal foreground, background, and alpha simultaneously. Chuang et al. [2] introduced a Bayesian approach and achieved impressive results for color natural images. For difficult cases, such as fur and hair, the Bayesian approach achieved the best result.

Here we describe concisely the Bayesian approach [2]. Given a know pixel C, the algorithm tries to find the most likely values for F, B, and α in the composition equation (1). Using Bayesian rule, the problem is taken as the maximization over a sum of log-likelihoods:

$$\begin{aligned}
&\arg\max_{F,B,\alpha} P(F,B,\alpha|C) \qquad (2)\\
&\quad = \arg\max_{F,B,\alpha} P(C|F,B,\alpha)P(F)P(B)P(\alpha)/P(C)\\
&\quad = \arg\max_{F,B,\alpha} (L(C|F,B,\alpha)+L(F)+L(B)+L(\alpha))
\end{aligned}$$

where $L(\cdot)$ is the log-likelihood function, i.e. the log of probability $L(\cdot)$=$\log(P(\cdot))$, and the $L(C)$ term is dropped, because it is constant with respect to the optimization parameters (α, F, B). The algorithm proceeds by growing, contour by contour, into the unknown region, heading inward both from the foreground and the background borders. At each unknown pixel, a circular region encompasses a set of trimap foreground and background pixels, as well as any foreground and background values previously computed nearby in the unknown region. The foreground and background samples are then separated into clusters, and weighted mean and covariance matrices are used to derive Gaussian distributions. Given these distributions, the Bayesian matting approach solves for the maximum-likelihood foreground, background, and alpha at the unknown pixel.

In this algorithm, they assume that the log likelihood for opacity $L(\alpha)$ is constant and thus omitted from the maximization in equation (2). To solve the equation efficiently, they break the problem into two quadratic sub-problems. For the first sub-problem, they assume that α is a constant and get the optimized F and B by taking partial derivatives. The value of the assumed α is estimated from the neighborhood of the current pixel. Then, for the second sub-problem, they assume that F and B are constant and yield a quadratic equation in alpha and get the solution by projecting the observed color C onto the line segment FB in color space.

2.2. Color transferring

One of the most common tasks in image processing and editing is to alter an image's color. Recently, Ruderman et al. [5] developed a color space, called $\boldsymbol{l\alpha\beta}$ color space, which minimizes correlation between channels for many natural scenes. Reinhard et al. [4] used this color space to transfer color from one color image to another and achieved impressive visual effect. In [3], Welsh et al. introduced color transfer technique to colorize grayscale images. The basic idea of that paper is to combine the color transferring technique in [4] with texture synthesis techniques. This technique works very well on scenes where the image is divided into distinct luminance clusters or where each of the regions has distinct textures. However, the technique does not work very well with faces. It fails to classifying the difference between skin and lip. More generally, this problem exists in colorizing grayscale images with regions that with similar or confusing luminance distribution. Directly employing this technique will fail to colorize these different regions with user specified colors, even though with the help of swatches.

In our approach, we first extract each object from the grayscale image by employing the grayscale image matting algorithm proposed in this paper. Then each object is colorized with specific colors following Welsh et al.'s algorithm. Finally, the colorized objects are composited to form the colorized version of the original grayscale image. By using this simple scheme, we achieve visually pleasant results for difficult cases.

3. GRAYSCALE IMAGE MATTING

In this section, we describe our grayscale image matting algorithm in detail. This algorithm follows the Bayesian framework and sliding window scheme proposed by Chuang et al. [2]. However, our method differs from theirs in three key aspects. Namely, (1) it models α as a Gaussian distribution and introduces image gradient to weight the standard deviation of $\alpha's$ distribution; (2) it optimizes the objective function in F, B, and α simultaneously; (3) it uses a simple and efficient color clustering algorithm. By introducing these improvements. We can effectively handle the underconstrained grayscale image matting problem.

3.1. Modelling likelihoods

The matting pipeline of our algorithm includes user interaction and solving the Maximum A Posteriori (MAP) problem for each unknown pixel. Given a grayscale image, user segments conservatively the image into three regions: "background," "foreground," and "unknown". For each pixel C in the unknown region, we try to find the most likely estimates of F, B, and α. We take the same formula (2) and define the log-likelihoods, $L(C|F,B,\alpha)$, $L(F)$, $L(B)$, and $L(\alpha)$ in a new way.

The first term in (2) is modelled by measuring the difference between the observed brightness C and the brightness that would be predicted by the estimated F, B, and α:

$$L(C|F,B,\alpha) = -\|C - \alpha F - (1-\alpha)B\|^2/2\sigma_C^2 \qquad (3)$$

This log-likelihood models error in the measurement of C and corresponds to a Gaussian probability distribution with mean $\alpha F + (1-\alpha)B$ and standard deviation σ_C. Here σ_C is a constant and models the noise in imaging process.

The second term $L(F)$ is modelled as the error term in a Gaussian distribution with mean $\overline{F}$ and standard deviation σ_F. Formally, $L(F)$ is expressed as:

$$L(F) = -\|F - \overline{F}\|^2/(2\sigma_F^2 + 2\sigma_C^2) \qquad (4)$$

$\overline{F}$ and σ_F are computed in the neighborhood of pixel C to exploit the spatial coherence of the source image. To more robustly model the foreground brightness distribution, we use the same weighting scheme in [2] to stress the contribution of nearby pixels and pixels with large opacity value. Since the estimated foreground F is also subject to the influence of imaging noise, the image noise term is added to the standard deviation of the Gaussian probability distribution. Such noise is critical to regularize the optimization process and avoid most of the degenerate cases. Similarly, we define $L(B)$ as:

$$L(B) = -\|B - \overline{B}\|^2/(2\sigma_B^2 + 2\sigma_C^2) \qquad (5)$$

For the likelihood of α, instead of take it as constant [2], we model it as the error term in a Gaussian distribution.

$$L(\alpha) = -\|\alpha - \overline{\alpha}\|^2/2\sigma_\alpha^2 \qquad (6)$$

where $\overline{\alpha}$ and σ_α are mean and standard deviation in Gaussian distribution. The computation of $\overline{\alpha}$ and σ_α is weighted by using a Gaussian filter to stress the contribution of nearby pixels. Instead of computing σ_α from neighborhood, we set it constant to model the noise of α.

The introduction of $\alpha's$ distribution constrains the MAP problem and gets better result than only modelling foreground, background and the error between observed C and predicted brightness. But it is a difficult task to set the appropriate standard deviation of $\alpha's$ distribution. The larger the σ_α, the smaller influence of α on the MAP problem and the formula (2) degenerates to a model without $\alpha's$ constraint. On the other hand, the smaller the σ_α, the stronger influence of α on the MAP problem and the edge, where alpha changes rapidly, will be blurred.

To avoid blurring, while keep the constraint of α, we introduce image gradient into the $\alpha's$ distribution based on such observation that when the gradient is large, the α has more chance to change greatly. The adoption of gradient keeps the spatial coherence of alpha in smooth area and looses the constraint for area with large gradient. Formally, we revise formula (6) as:

$$L(\alpha) = -\|\alpha - \overline{\alpha}\|^2 / 2\sigma_\alpha^2 w_g g \tag{7}$$

where g is the normalized gradient of current pixel, w_g is the weight of the influence of gradient on the alpha's distribution. In our experiments, we set w_g around $0.5 \sim 2$, and get satisfying results.

3.2. Optimization

Here, we propose an efficient optimization scheme based on Variable Metric Method (VMM) [12] and a simple and fast color clustering algorithm.

In formula (2), the optimization is in F,B, and α. In [2], the authors optimized the problem in a two-step scheme. For grayscale image matting with the definitions of likelihoods in this paper, we find the two-step optimization scheme doesn't work efficiently. Alternatively, we employ a Variable Metric Method to optimize F, B, and α simultaneously. Furthermore, we also include the constraints: $0 \leq F \leq 255$, $0 \leq B \leq 255$, and $0 \leq \alpha \leq 1$.

In most of the color image matting algorithms [1, 2, 11], the authors employed previous powerful color quantization methods to settle the problem when there are multiple distinct sets of color in the neighborhood. These color quantization algorithms are necessary for color image matting with complex background or foreground color distribution. For grayscale images, the direct simplification of these color quantization algorithms can achieve very good clustering results. However, in experiments, we found these techniques are not necessary respect to the tradeoff between computational cost and performance improvement. We propose a very simple clustering algorithm for grayscale image matting. Given a set of colors S, the objective of clustering is to separate them to n subsets. In our clustering algorithm, we first find the largest and the smallest brightness, I_{min} and I_{max} in S. Then we cluster each color I in this way: $Index(I) = \lfloor (I - I_{min} - \varepsilon)n/(I_{max} - I_{min}) \rfloor$. ε is a small positive number to avoid $Index(I) = n$. This algorithm is very fast and independent of the number of subsets. In experiments, we don't find visual improvement after using very efficient color quantization algorithm [7].

4. COLORIZATION AND COMPOSITION

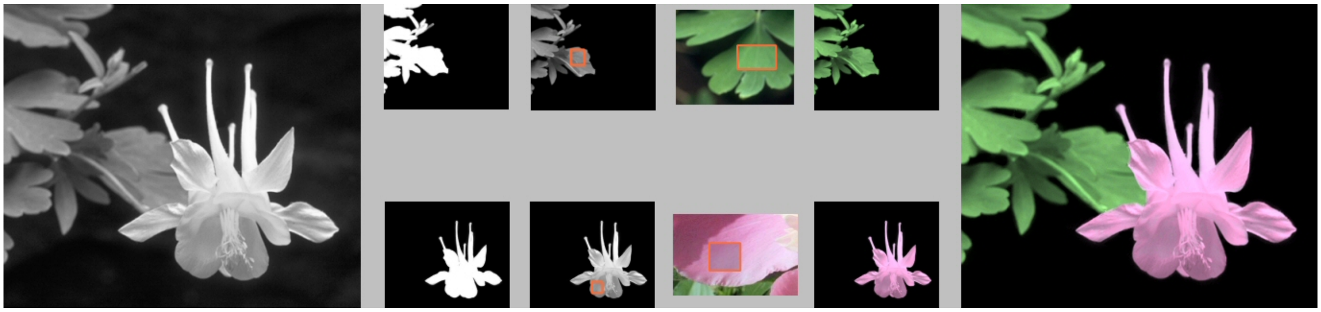
In this section, we combine grayscale image matting algorithm with color transferring techniques to enable region-based colorization. Before colorization process, objects that will be colorized with different color mood are extracted from the grayscale image. Then, each object is colorized using color transferring. Finally, these colorized objects are seamlessly composited to reach the ultimate colorization result.

Our basic color transferring algorithm is based on the method of Welsh et al. [3]. There are two methods described in [3]. One is global colorization which is apt to the cases where the source image and target image have globally similar texture or luminance distributions. When the user wants to colorize the specific regions of the grayscale image with specific color moods in color images, a multi-swatch color transferring method is proposed.

We follow the multi-swatch method and extend it to a more general and accurate method. Here, we assume each extracted object has a uniform color mood. For example, the skin on human face has a uniform color mood while the lip has another uniform color mood. For each extracted objects, we first find color images with aimed color mood. Then we select a pair of swatches from the source image (color image) and the target image (grayscale image). In order to account for global differences in luminance between the two swatches, we perform luminance remapping [6] to linearly shift and scale the luminance histogram of the source swatch to the target swatch. Since the color transferring of a pair of swatches is very fast, user can interactively select an adequate pair of swatches. After getting the right pair of swatches, the rest region of the object is colorized by transferring colors from the colorized swatch. This divide-and-conquer method is more efficient than multi-swatch scheme in [3], because it is not necessary to decide transferring colors from which swatch and the object is only colorized in masked region.

Object-based colorization simplifies the colorization process when the image has no distinct texture or luminance distribution. However, it poses great difficulty for composition. Using traditional segmentation or masking tools to extract objects from image will cause serious ghost effect along boundaries.

In our solution, the objects are efficiently extracted from the background grayscale image. In our experiments, we find the colorized objects can seamlessly composited using standard alpha blending. Even for vision sensitive objects, such as lip and skin, we also get seamless results.



Grayscale image masks objects examples colorized objects colorized image

Fig. 3. Colorizing a flower. The input grayscale image is first separated into flower object and leaves object. The objects are colorized by transferring colors from example color images. Then the colorized objects are composited to reach the colorization result. Grayscale image courtesy ©S. Ballard - California Academy of Science

Grayscale image objects colorized objects colorized image

Fig. 4. Colorizing a human face. The input grayscale image is first separated into seven objects. Four objects are colorized by transferring colors from example color images, while three objects keep grayscale. Then the colorized and uncolorized objects are composited to reach the colorization result. Grayscale image courtesy ©http://www.imagegarden.net

5. EXPERIMENTS AND RESULTS

In this section, we describe how to extract, colorize and composite objects from grayscale image to get the seamless color image. Results for two difficult cases are demonstrated, a flower and a human face.

5.1. Colorizing a flower

This flower image (Fig. 3) is difficult in two points. First, the leaves and the flower have regions with very similar luminance distributions. This similarity will cause the colorization method in [3] failed, even using multi-swatch method. Second, the boundaries between the flower and leaves are very delicate. Typically, the flower and the leaves will be colorized with distinct colors which are in strong contrast each other. Human vision is very sensitive to slight inadequate colorization in such area. Any clean-cut segmentation or insufficient matting algorithm will cause visual defect.

In our solution, we first extract the flower and leaves as two distinct objects. Then we transfer colors from a petal of a peony to the flower. The colors of the leaves are transferred from a color leaf. The results demonstrate that our solution can solve the difficulties effectively.

5.2. Colorizing a human face

Human vision is very sensitive to any small defects on synthetic human face, which poses great challenge for facial image editing. The most difficult problem in facial colorization is how to smoothly colorize the skin and how to keep the seamless connections among hair, skin, lip, eye, eyebrow, and background.

In our experiment (Fig. 4), the facial image is first separated into seven objects: background, hair (including eyebrow), skin (face and body skin), lip, eye white, eyelash (including pupil and eye black), and shoulder strip. Then we colorize the skin and lip by transferring colors from another color facial image. The background and the shoulder strip are only assigned constant chromatic values. The hair, eye white, eyelash are not colorized. The results demonstrate the objects in this facial images are well separated and seamlessly composited. The delicate details near the object boundaries are kept. The whole colorization result is more lively than the grayscale image. It is natural to assume, there are many other possibilities to colorize the girl in realistic or artistically unrealistic way.

6. CONCLUSION AND FUTURE WORK

The problem solved in this paper is how to extract objects from grayscale image, colorize the objects and composite them seamlessly, especially for difficult cases, such as human face and natural images with confusing luminance distribution or delicate boundaries. There is still no efficient technique to solve this problem. We deal with this problem in a three-step, divide-and-conquer way. First, we present an efficient grayscale image matting algorithm in Bayesian framework. The major novelty and improvement over previous algorithms are the introduction of alpha's distribution, where the image gradient is used to weight the standard deviation. Then we transfer colors from example color images to the extracted grayscale objects. Finally, the colorized objects are composited back to reach the ultimate colorization of the grayscale input image. The results demonstrate this solution can handle difficult cases effectively.

There are still a number of avenues for future work. Currently, we just use multi-sized brushes to mask the trimap in matting process. We plan to incorporate object selection tools [10] into our algorithm to facilitate the user interaction. Another possible extension is to colorize facial video clips. In [8], Chuang et al. extended the Bayesian color image matting technique [2] to video matting by combining with robust optical flow technique. Finally, we would like to find new applications for our grayscale image matting algorithm.

7. ACKNOWLEDGEMENTS

The authors would like to thank Yung-Yu Chuang for his explanation and advices about his impressive digital matting technique [2].

8. REFERENCES

[1] M. A. Ruzon and C. Tomasi, "Alpha estimation in natural images," in *Proceedings of Computer Vision and Pattern Recognition*, 2000, pp. 18–25.

[2] Y. Y. Chuang, B. Curless, D. Salesin, and R. Szeliski, "A bayesian approach to digital matting," in *Proceedings of Computer Vision and Pattern Recognition*, 2001, pp. 264 – 271.

[3] T. Welsh, M. Ashikhmin, and K. Mueller, "Transferring color to grayscale images," *ACM TOG*, vol. 20, no. 3, pp. 277–280, 2002.

[4] E. Reinhard, M. Ashikhmin, B. Gooch, and P. Shirley, "Color transfer between images," *IEEE Computer grahics and applications*, vol. 21, no. 5, pp. 34–41, 2001.

[5] D. L. Ruderman, T. W. Cronin, and C. C. Chiao, "Statistics of coneresponses to natural images: Implications for visual coding," *J. Optical Soc. Of America*, vol. 15, no. 8, pp. 2036–2045, 1998.

[6] A. Hertzmann, C. Jacobs, N. Oliver, B. Curless, and D. Salesin, "Image analogies," in *Proceedings of ACM SIGGRAPH*, 2002, pp. 341–346.

[7] X. Wu, "Color quantization by dynamic programming and principal analysis," *ACM Transaction on Graphics*, vol. 11, no. 4, pp. 348–372, 1992.

[8] Y. Y. Chuang, A. Agarwala, B. Curless, D. Salesin, and R. Szeliski, "Video matting of complex scenes," *ACM Transaction on Graphics*, vol. 11, no. 4, pp. 348–372, 2002.

[9] A. R. Smith and J. F. Blinn, "Blue screen matting," in *Proceedings of ACM SIGGRAPH*, 1996, pp. 259–268.

[10] E. N. Mortensen and W. A. Barrett, "Intelligent scissors for image composition," in *Proceedings of ACM SIGGRAPH*, 1995, pp. 191–198.

[11] P. Hillman, J. Hannah, and D. Renshaw, "Alpha channel estimation in high resolution images and image sequences," in *Proceedings of Computer Vision and Pattern Recognition*, 2001, pp. 1063–1068.

[12] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical recipes in C: The art of scientific computing(second edition),* Cambridge University press, Cambridge, 1992.

[13] T. Porter and T. Duff, "Compositing digital images," in *Proceedings of SIGGRAPH*, 1984, pp. 253–259.

[14] A. Berman, A. Dadourian, and P. Vlahos, *Method for removing from an image the background surrounding a selected object*, U.S. Patent 6,134,346, 2000.