Color correction for image stitching by monotone cubic spline interpolation

Fabio Bellavia and Carlo Colombo

Computational Vision Group, University of Florence, Italy,
{fabio.bellavia, carlo.colombo}@unifi.it

Abstract. This paper proposes a novel color correction scheme for image stitching where the color map transfer is modelled by a monotone Hermite cubic spline and smoothly propagated into the target image. A three-segments monotone cubic spline minimizing color distribution statistics and gradient differences with respect to both the source and target images is used. While the spline model can handle non-linear color maps, the minimization over the gradient differences limits strong alterations on the image structure. Adaptive heuristics are introduced to reduce the minimization search space and thus computational time. Experimental comparisons with respect to the state-of-the-art linear mapping models show the validity of the proposed method.

Keywords: color transfer, image stitching, photometric blending

1 Introduction

Color correction is an essential step nowadays in image and video stitching pipelines [8]. After the spatial image registration, colors between corresponding pixels can have strong inconsistencies, due to different acquisition light conditions, such as varying exposure levels and viewpoints, which cannot be eliminated by image blending techniques alone [1]. Different color correction methods have been proposed across the years. In particular, worth to be mentioned are the gain compensation [1] and the Reinhard's method [5], both modelling linear color map functions.
The gain or exposure compensation was introduced to address color balancing in panoramic mosaicing by a least-square minimization approach, in order to get a symmetric blending across the overlapping area across multiple images. On the other hand, the Reinhard's method proposes a linear transformation to make the mean and the standard deviation of color distribution of the target image corresponding to those of the source image. Extensions of both methods have been proposed [7, 8].
Beyond model-based parametric approaches as those described above, modeless non-parametric approaches exist [4], but according to a recent evaluation work [10] the two cited methods are preferable in the general case of image stitching due to their output quality, stability and speed.
This paper introduces a novel color correction algorithm using a monotone spline model as mapping function to better handle non-linear maps. The proposed model, described in Sect. 2, is obtained by exhaustively

searching the knots defining the spline which minimizes an error function based on the intensity values and the gradients of the source and target images. Due to the complexity of the minimization function, not analytically manageable, to speed-up the computation and provide the results in reasonable time, adaptive heuristics have been introduced to reduce the minimization search space. Details are presented in Sect. 2.1.
In order to better handle local image properties, overlapping areas between the source and target images are divided into blocks and the corresponding local color maps are computed. Local color maps are then propagated smoothly with morphological operators and Gaussian blur to provide a smooth color change across non-overlapped areas. This last step is described in Sect. 2.2.
Finally, in Sect. 3 a quantitative experimental evaluation has been carried out on real images. The new spline method is compared against the Reinhard's method and an asymmetric version of the gain compensation, according to quantitative measures, in a similar way to the approach of [10], in Sect. 3.1. As shown in Sect. 3.2, according to the evaluation, the proposed spline method achieves better and robust results in terms of image quality, while maintaining reasonable computational times.
Conclusions and future work are discussed in the end of the paper (Sect. 4).

2 Method Description

2.1 Spline Color Mapping

Given a source image I_1 and a target image I_2, the proposed method aims to obtain a color corrected source image $\tilde{I}_1$ which looks similar to image I_2 by a transformation f between color intensity values, i.e. for a pixel $\mathbf{p}$ in the overlapping areas between the two images it holds

$$\tilde{I}_1(\mathbf{p}) = f(I_1(\mathbf{p})) \approx I_2(\mathbf{p}) \tag{1}$$

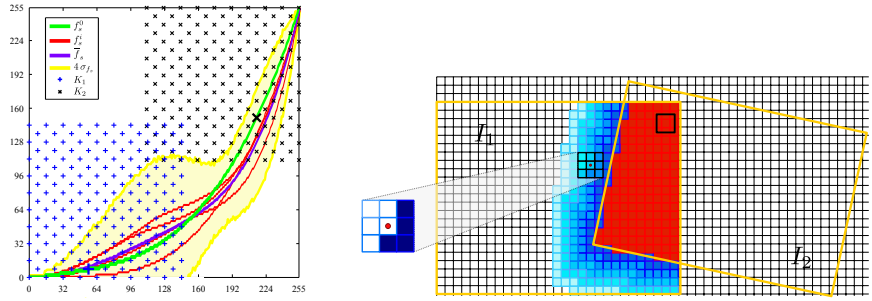
In addition, a smooth transition between overlapping and non-overlapping areas of the target image is required to not alter the corrected source image. The Reinhard's method [5] defines a linear color map function f_R so that the mean and standard deviation of the intensity values between the corrected source image $\tilde{I}_1$ and target image I_2 are equal by imposing

$$f_R(x) = \frac{\sigma_2}{\sigma_1}(x - \mu_1) + \mu_2 \tag{2}$$

where μ_k, σ_k are respectively the mean and standard deviation for a generic image I_k, here assumed to be computed only in the common area $I_1 \cap I_2$. On the other hand, with the gain compensation method, the gains g_k defined as

$$\tilde{I}_k(\mathbf{p}) = g_k \, I_k(\mathbf{p}) \tag{3}$$

are chosen to minimize a quadratic error between the intensity values of the source and target images (see [1] for more details).
In this paper, we propose a more general color mapping function f_s using a monotone Hermite cubic spline [3] with 4 knots. Two of the



(a) (b)

Fig. 1: (a) An example of the search space of the two spline knots K_1 and K_2 for the minimization of the error E_s, bigger marks underline the selected knots, see text for more details. (b) An example of the color map propagation, see text for more details (best viewed in color).

knots are fixed to the extremal mapping values to preserve the bijective property. Taking into account the monotonic constraint, this implies that $f_s(0) = 0$ and $f_s(255) = 255$, assuming a single channel 8-bit integer color range. The choice of the other two non-fixed knots $K_m = (x_m, y_m) = (x_m, f_s(x_m))$, with $m \in \{1, 2\}$ and $x_1 < x_2$ and $y_1 < y_2$, gives rise to different color mapping functions, from linear to sigmoid-like and exponential models (see Fig. 1(a), green and red lines). In particular, we look for the model that minimizes the following weighted error function E_s:

$$E_s = w_\mu \, e_\mu + w_\sigma \, e_\sigma + \sum_{d \in \{x,y\}} \sum_{n=1,2} w_{d,I_n} \, e_{d,I_n} \tag{4}$$

where the w_k are given weight values (according to our experiments, we set $w_\mu = 0.5$ and all the others to $w_k = 0.1$). The errors $e_\mu = |\mu_s - \mu_2|$ and $e_\sigma = |\sigma_s - \sigma_2|$ are respectively the absolute differences between mean values and the standard deviations of the color intensities of the corrected source image and the target image, in a way similar to the Reinhard's method. Here, μ_s, σ_s are the corresponding statistics for the corrected source image. To improve structure similarity with the target image but also to constrain the image structure to the source image, the average derivative absolute differences e_{d,I_n} in the direction d are taken into account in the minimization thus defining

$$e_{d,I_n} = \frac{1}{|I_1 \cap I_2|} \sum_{\mathbf{p} \in (I_1 \cap I_2)} \left| \frac{\partial \tilde{I_1}}{\partial d}(\mathbf{p}) - \frac{\partial I_n}{\partial d}(\mathbf{p}) \right| \tag{5}$$

Since an analytical solution for minimizing the proposed error e_s is not trivial, an exhaustive search for the two knots can be done instead. In the case of 8-bit integer, i.e. $n = 256$, imposing only the monotonic condition, this implies that knot abscissas and ordinates count both for $n\,(n-1)/2$ so that we get $t = (n\,(n-1)/2)^2 = n^2(n-1)^2/4 \approx 10^9$ different error

values to check, which is infeasible in practice. Nevertheless, the search space can be dramatically reduced by observing that, for a given knot, perturbations of its position only slight change the error e_s and some knot positions are redundant, see Fig. 1(a) for a visual explanation. In our case, still referring to Fig. 1(a), defining uniform squared grids in the range of $[0, 144]$ and $[111, 255]$ for each knot respectively, with a step of 8 (i.e., a 19×19 grid size) and a chess-like alternate grid sampling, drastically reduces the search space to approximately 3×10^4 possible knot pairs, still maintaining an almost optimal solution.
The process can still be optimized further by pre-computing the possible splines associated to each knot pair and it can take advantage of parallel CPU processing, reducing to about 4000 error evaluations per thread on a 8 core CPU. Furthermore, according to our experiments, the error on the mean intensity color e_μ is predominant, while other error measures just refine the solution. We then define a double step check: if the error e_μ for the current solution is greater than $e_\mu + 15$ of the best solution so far, we discard the current solution avoiding to compute the full error E_s, thus saving computation.
Finally, if more local color maps have to be computed, as described in the next section, one can take advantage of the already computed neighbourhood color maps, since close color maps change smoothly. Assuming 8-connectivity image blocks, given the color map f_s^0 (green line in Fig. 1(a)) and at least b already optimized of its adjacent block color maps f_s^i (red lines) with $i = 1, \ldots, b$ and $3 \leq b \leq 8$, an "average" spline can be computed on which to sample the knots (yellow area). In particular for a given $0 \leq x_m \leq 255$ we constrain y_m to

$$\overline{f}_s(x_m) - 4\,\sigma_{f_s(x_m)} \leq y_m = f_s^0(x_m) \leq \overline{f}_s(x_m) + 4\,\sigma_{f_s(x_m)} \qquad (6)$$

where $\overline{f}_s(x_m)$ and $\sigma_{f_s(x_m)}$ are respectively the mean and standard deviation of the neighbourhood spline mapping values f_s^i already computed (purple and yellow lines respectively).

2.2 Color Map Propagation

Due to the general inability of a single color map to cover correctly all the color transformations in the overlapped image area, the source and target images are aligned and the area divided in blocks of size 32×32 px as shown in Fig. 1(b), to reflect the color locality property. Color maps are computed for each block with at least 50% of overlap between the two images (red blocks). A wider area of size 64×64 px centered in the block (black square on the red block) is used for the color map computation to avoid abrupt color changes.
In order to achieve a smooth change towards non-overlapping areas of the corrected source image $\overline{I}_1$, color maps of the boundary blocks of the overlapping area are propagated as follows. Considering the block-like versions of the source and target images I_1 and I_2, boundaries of the overlapping region are expanded over non-overlapping blocks of I_1 using the dilation morphology operator [6] with a square 3×3 block kernel.

An increasing label $l = 0, 1, \ldots, u$ is assigned accordingly to the iteration, stopping to grow when a number of blocks equal twice the number of the overlapping blocks has been incorporated or no more blocks can be included (see Fig. 1(b), blue gradient blocks). Since boundary blocks are labelled $l = 0$, starting from the blocks with $l = 1$, a smoothed color map is obtained for these blocks by applying a convolution with a normalized 3×3 Laplacian pyramid kernel [2] on the adjacent blocks with a lower label value. Kernel weights corresponding to equal or higher label values are set to 0 and the kernel is re-normalized (see zoomed grid image); convolution is performed between corresponding map values, i.e. between the $f_s(x)$ for a same x of the adjacent blocks. Finally, we linearly blend the propagated color values with the effective values in the non-overlapping area of the source image. Explicitly

$$\tilde{I}_1(\mathbf{p}) = \frac{l^\star}{u} I_1(\mathbf{p}) + \left(1 - \frac{l^\star}{u}\right) f_s^\star(I_1(\mathbf{p})) \tag{7}$$

where $l^\star$ and $f_s^\star$ are the interpolated values of l and f_s in the color map corresponding to the block containing $\mathbf{p}$ after a Gaussian expansion [2] is used to avoid block-like effects, see Figs. 2(e) for an output example.

3 Experimental Results

3.1 Evaluation Setup

The proposed method was evaluated on a novel dataset of 52 registered color image pairs belonging to 6 real planar scenes, obtained by varying the image exposition. Images and code for the evaluation are freely available[1]. We compare the proposed spline method against Reinhard's method and an asymmetric version of the gain compensation (see the additional material[2] for more details). The color map propagation described in Sect. 2.2 was applied to all methods in order to achieve a fair comparison, thus only varying the color map function computation.
For each image pair we run two different tests to evaluate the methods. In the first test T_1, we just considered the error with respect to the target image as ground-truth in the overlapping area $R = I_1 \cap I_2$ between the two aligned images. To evaluate the color propagation, in test T_2, a random chosen connected subset of 40% of the overlap area $R^\star \subset R$ was considered as the effective overlapping area to obtain the color map, while color propagation was made in the remaining overlapping area $R^c = R \setminus R^\star$ (see Fig. 2).
In both setups we considered the results obtained on grayscale images (luminance channel) and color images. In the latter case the methods were applied to each of the RGB channels independently. Using the ground-truth target image I_2, we evaluated the color corrected source image $\tilde{I}_1$ according to *Stuctural SIMilarity* ($SSIM$) index [9] and error measures defined on the intensity values and gradient, denoted respectively as E_c

[1] http://cvg.dsi.unifi.it/download/spline/spline.zip
[2] http://cvg.dsi.unifi.it/download/spline/spline_additional_material.zip

(a) (b) (c) (d) (e)

Fig. 2: Source image (a), target image (b), overlapping area $R = I_1 \cap I_2$ (c), $R^\star$ (gray) and R^c (black) regions (d) and spline method output (e) for an example T_2 test setup (best viewed in color).

and E_d. The $SSIM$ index is an image quality index that measures the structure coherence of one image with another. It was already used in color transfer evaluations [10] and works on grayscale images only. Assuming the RGB colorspace, the error E_c on the image intensities is defined instead as the average *root mean square* (RMS) error for each considered pixel $\mathbf{p}$, which reduces to the average absolute error in the case of grayscale images

$$E_c = \frac{1}{|A|} \sum_{\mathbf{p} \in A} RMS \left(G_{\sigma_e} * \tilde{I}_1(\mathbf{p}) - G_{\sigma_e} * I_2(\mathbf{p}) \right) \tag{8}$$

where A is the image region among $\{R, R^\star, R^c\}$ on which to evaluate the error. Convolution with a Gaussian kernel G_{σ_e} with standard deviation $\sigma_e = 8$ has been applied to both the corrected source and target images in order to cope with small image misalignment errors. The gradient error E_d is defined in similar way.
In order to rank the methods, since the dataset contains image pairs going from small color variations to challenging ones, we promote strong error differences between the methods with respect to small ones. For a given image pair and evaluation criterion, we define the *normalized smooth rank* $r(m)$ for the value m associate to a method as

$$r(m) = \frac{|m - m_b| + \delta}{\sum_{n \in \{m\}} (|n - m_b| + \delta)} \tag{9}$$

where m_b is the best value among those of the compared methods (the maximum for the $SSIM$ index, the minimum for the E_c and E_d errors), and δ is a tolerance factor to avoid strong rank variations in the case of small measure differences, set to $\delta = 1$ in the case of the $SSIM$ index and E_c and to $\delta = 0.1$ for E_d.

3.2 Results

Table 1 shows the average smooth ranks for each method in the case of color and grayscale images for the different tests and evaluation criteria, while in Fig. 3 some challenging examples are shown. Corrected source images are over-imposed on the target image, without applying

any blending to better appreciate the outputs. For detailed results on each image pair, see the additional material[2]. According to the results, the proposed spline method provides in general the best outputs, followed by Reinhard's method, while the asymmetric gain compensation gives the worst results. This holds for the $SSIM$ index and the E_c and E_d errors in both the color and grayscale cases. As an additional observation, it can be noted that color images give higher errors, since some information is lost by handling the color channels independently.
About the computing time, the average time per block are 550 ms, 2 ms and 2 ms respectively for the spline, Reinhard's and asymmetric gain compensation methods. Furthermore, the average time for block in the case of the spline method double from 550 ms to 1200 ms when the adaptive search space according to adjacent color maps is not implemented (see 6). The computing time of the spline method is reasonable (about 5 minutes for a color image pair) and feasible for off-line tasks, also considering the better visual output quality of the method with respect to the second ranked Reinhard's method in the case of challenging image pairs. In this cases, the Reinhard's method look more unrealistically contrasted or with wrong colors due to strong exposure variations (see Fig.3, top and middle rows) or misalignment issues (bottom row).

Table 1: Average smooth rank (%) for the difference evaluation measures, best values are in bold.

	T_1			T_2								
	R			$R^\star$			R^c			R		
	Spline	Reinhard's	A. Gain	Spline	Reinhard's	A. Gain	Spline	Reinhard's	A. Gain	Spline	Reinhard's	A. Gain
$SSIM$ index	**12.9**	32.0	55.0	**13.1**	30.8	56.0	**18.3**	22.1	59.4	**14.8**	27.4	57.7
E_c (RGB)	**05.2**	12.9	81.7	**06.0**	12.4	81.4	14.0	**10.8**	75.0	**08.5**	10.8	80.5
E_c (gray)	**07.8**	12.6	79.5	**08.4**	12.3	79.1	16.8	**12.6**	70.4	**11.5**	11.8	76.6
E_d (RGB)	**14.7**	44.8	40.4	**15.3**	42.7	41.9	29.3	**27.9**	42.6	**29.4**	39.8	30.7
E_d (gray)	**21.0**	44.9	33.9	**21.2**	43.6	35.0	**33.2**	33.3	33.3	32.0	38.1	**29.7**

4 Conclusions

This paper describes a novel color correction method for image stitching, which adopts a spline model and can handle non-linear color map functions. Different heuristics have been introduced to reduce the model search space and, consequently, the computation time. The comparison with state-of-the-art color correction algorithms shows the robustness and the validity of the approach. Future work will include to test adaptive block shapes according to the image segmentation as in [7] and further computation speed-up improvements.

Acknowledgement

This work has been carried out during the ARROWS project, supported by the European Commission under the Environment Theme of the "7th Framework Programme for Research and Technological Development".

(a) (b) (c) (d) (e)

Fig. 3: Source images (a), target images (b), spline (c), Reinhard's (d) and asymmetric gain compensation (e) outputs for some image pairs used in the evaluation (best viewed in color).

References

1. Brown, M., Lowe, D.G.: Automatic panoramic image stitching using invariant features. International Journal of Computer Vision 74(1), 59–73 (2007)
2. Burt, P.J., Adelson, E.H.: A multiresolution spline with application to image mosaics. ACM Transactions on Graphics 2(4), 217–236 (1983)
3. Fritsch, F.N., Carlson, R.E.: Monotone piecewise cubic interpolation. SIAM Journal on Numerical Analysis 17(2), 238–246 (1980)
4. Jia, J., Tang, C.: Tensor voting for image correction by global and local intensity alignment. IEEE Transaction on Pattern Analysis and Machine Intelligence 27(1), 36–50 (2005)
5. Reinhard, E., Ashikhmin, M., Gooch, B., Shirley, P.: Color transfer between images. IEEE Computer Graphics and Applications 21(5), 34–41 (2001)
6. Soille, P.: Morphological Image Analysis: Principles and Applications. Springer-Verlag, 2 edn. (2003)
7. Taiand, Y., Jia, J., Tang, C.: Local color transfer via probabilistic segmentation by expectation-maximization. In: Computer Vision and Pattern Recognition. pp. 747–754 (2005)
8. Uyttendaele, M., Eden, A., Szeliski, R.: Eliminating ghosting and exposure artifacts in image mosaics. In: Computer Vision and Pattern Recognition. pp. 509–516 (2001)
9. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing 13(4), 600–612 (2004)
10. Xu, W., Mulligan, J.: Performance evaluation of color correction approaches for automatic multi-view image and video stitching. In: Computer Vision and Pattern Recognition. pp. 263–270 (2010)