Image editing with spatiogram transfer

Nicolas Papadakis, Aurélie Bugeau, Vicent Caselles

Image editing with spatiograms transfer

Nicolas Papadakis, Aurélie Bugeau and Vicent Caselles [*†‡]

Abstract

Histogram equalization is a well known method for image contrast enhancement. Nevertheless, as histograms do not include any information on the spatial repartition of colors, their application to local image editing problems remains limited. To cope with this lack of spatial information, spatiograms have recently been proposed for tracking purposes. A spatiogram is an image descriptor that combines a histogram with the mean and variance of the position of each color. In this paper we address the problem of local retouching of images by proposing a variational method for spatiogram transfer. More precisely, a reference spatiogram is used to modify the color value of a given region of interest of the processed image. Experiments on shadow removal and inpainting demonstrate the strength of the proposed approach.

1 Introduction

Image editing, also called photo retouching, refers to processes that create a transformed version of a given image. It can be used directly for image enhancement, seamless cloning [19], etc. or as a post-processing step for applications such as inpainting [5]. The referred image modifications can either be global (for the whole image) or local (limited to a selected area).

One of the most popular tools for global image processing is histogram equalization. It has been widely used for contrast enhancement [12], 3D reconstruction [16] (to compare objects or scenes that have been observed by different cameras under different illuminations) or for colorization [18]. Histogram equalization techniques search for a transfer function that maps one image to another with an approximately uniform histogram. To transfer the original histogram to a reference one that is not uniform, more advanced techniques have been proposed [10] for gray-scale images. Even if in some applications each color channel can be treated separately, usually they are correlated and one needs to consider the transfer of 3D histograms [20]. Instead of using the histograms as variables in the transfer system, the image can be directly processed, as demonstrated

[*]N. Papadakis is with the MOISE team (INRIA/CNRS), Laboratoire Jean Kuntzmann (LJK, UMR 5224), Campus de Saint Martin d'Hères, 38041 Grenoble, France. (e-mail: nicolas.papadakis@imag.fr)

[†]A. Bugeau is with the AIV team, Laboratoire Bordelais de Recherche en Informatique (LaBRI, UMR 5800) , 351 cours de la Libération F-33405 Talence, France. (e-mail: aurelie.bugeau@labri.fr)

[‡]V. Caselles is with the Departamento de Tecnologías de la Información y las Comunicaciones, Universitat Pompeu Fabra, Carrer de Roc Boronat 138, 08018 Barcelona, Spain. (e-mail: vicent.caselles@upf.edu)

in [17]. As a consequence, and contrary to the aforementioned methods, two pixels having the same color in the original image can have different color values in the final corrected image. However, to obtain realistic image synthesis, the transfer of color, in histogram space or in image space, is not sufficient and it is necessary to consider an additional spatial processing of the images [22]. Indeed, the change of content between the two images from which the histograms are extracted has to be taken into account. Consider for instance a pair of landscape images where the sky in one picture covers a larger area than in the other [20].

Many tools also exist for local image editing, most of them addressing only one of the following applications: shadow removal [25], highlight removal [13, 29], Poisson editing [19], Poisson cloning [11], texture synthesis [8], image inpainting [1]. Even if many applications were proposed in [19], they all require a particular setting of the reference gradient field used to do the Poisson editing. Namely for highlight or shadow removal, the gradient amplitude is artificially modified to make the scene more dark or bright. For general shadow removal applications, this process is not able to recover the missing colors. Hence, it would be more convenient to have the same methodology to achieve for example shadow removal and inpainting. Histograms are a powerful tool which has not been completely studied in this context of local editing. For inpainting application, the histogram transfer has been applied as a post-processing step in [14] to homogenize the color of the resulting image inside and outside the inpainting mask. Few works on histogram modification that use information on the spatial distribution of colors exist. Such methods are nevertheless only dedicated to contrast enhancement [27] or image cloning [28] and can even require several steps [28], whereas we rather prefer a single energy minimization as proposed for contrast enhancement in [2, 24].

In this paper, we aim at doing local image editing through histogram transfer. In order to consider the spatial distribution of colors, we use the concept of spatiogram and propose a general formulation based on spatiogram transfer that allows dealing with various image editing problems through a single energy minimization. The concept of spatiogram is related to the mean and covariance of the position of each color and was proposed in [3] as a useful descriptor of images for tracking applications. It has also proven to be a powerful tool in video [31] and image retrieval [7, 26] applications.

Let us describe the organization of this paper. In Section 2, we briefly review the notion of spatiogram proposed in [3] and the variational transfer of histograms of [17]. In Section 3 we show how to combine both works in order to realize local color enhancement of images. The visualization of spatiograms is illustrated in Section 4. Finally, in Section 5 we present some experiments, where we use the proposed model for image inpainting, and shadow and highlight removal.

2 Related works

In this section, we introduce the basic ingredients that will be necessary for spatiogram transfer purposes.

2.1 Introduction to spatiograms

Let $I(x)$ be a color image defined on the spatial domain Ω, such that $I : x \in \Omega \mapsto (I_1, I_2, I_3) \in [0; 255]^3$. The RGB color space is considered here. Other color spaces that have less correlated components like L$\alpha\beta$ [23], Lab or YUV could have also been used. However, we rather consider the general problem of color transfer in 3D, since there exists no color space where the channels are completely uncorrelated.

For a color $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3) \in [0; 255]^3$, the histogram h_I of I is defined as:

$$h_I(\boldsymbol{\lambda}) = \frac{1}{|\Omega|} \int_\Omega \delta(I(x) - \boldsymbol{\lambda}) dx, \tag{1}$$

where $\delta(.)$ is the Dirac function whose value is 1 if its argument is 0 and 0 otherwise. In practice, the histogram is discretized using a reduced number of $N = (N_1 \times N_2 \times N_3)$ bins, but we will keep the continuous formulation all along the paper to simplify our notation.

Histograms give a global description of the color distribution of images. In order to add some spatial information to this description, Birchfield and Rangarajan proposed in [3] to consider also the mean position μ_I and the spatial covariance Σ_I of the pixels that contribute to each bin:

$$\begin{aligned} \mu_I(\boldsymbol{\lambda}) &= \frac{1}{|\Omega| h_I(\boldsymbol{\lambda})} \int_\Omega x \delta(I(x) - \boldsymbol{\lambda}) dx, \\ \Sigma_I(\boldsymbol{\lambda}) &= \frac{1}{|\Omega| h_I(\boldsymbol{\lambda})} \int_\Omega (x - \mu_I(\boldsymbol{\lambda}))(x - \mu_I(\boldsymbol{\lambda}))^T \delta(I(x) - \boldsymbol{\lambda}) dx. \end{aligned} \tag{2}$$

The spatiogram of image I is finally defined as

$$S_I = (h_I, \mu_I, \Sigma_I). \tag{3}$$

Obviously, the above means and covariances are only defined for the bins $\boldsymbol{\lambda}$ where $h_I(\boldsymbol{\lambda}) > 0$, i.e. if there is at least one pixel of I that contributes to the corresponding color.

In this paper, we will introduce a model to transfer spatiograms. Our approach is based on the principle of variational histogram transfer that we now shortly review.

2.2 Variational transfer of histograms

In this section, we briefly review the variational histogram transfer as proposed in [17]. Let us consider two images I^0 and I^r, where I^r is a reference image. We want to modify I^0 in such a way that its histogram is as close as possible to the one of I^r. Building an energy that compares the histogram values of the two images (such as $\int (h_I(\boldsymbol{\lambda}) - h_{I^r}(\boldsymbol{\lambda}))^2 d\boldsymbol{\lambda}$) is the intuitive way to perform this transfer. However, due to the Dirac function involved in definition (1), such an energy is not differentiable with respect to I. As a consequence, as in [17], we proposed to consider the cumulated histograms. Using the continuous values of $\boldsymbol{\lambda} \in [0; 255]^3$, the cumulated histogram of I can be expressed as:

$$H_I(\boldsymbol{\lambda}) = \frac{1}{|\Omega|} \int_\Omega g_{\boldsymbol{\lambda}}(I(x)) dx, \tag{4}$$

where the Heaviside function

$$g_{\boldsymbol{\lambda}}(I(x)) = \begin{cases} 1 & \text{if } I_1(x) \leq \lambda_1,\, I_2(x) \leq \lambda_2,\, I_3(x) \leq \lambda_3, \\ 0 & \text{otherwise,} \end{cases}$$

is differentiable with respect to I for a small perturbation $dI = (dI_1, dI_2, dI_3)$ and gives a dirac operator:

$$\nabla_{I_i} g_{\boldsymbol{\lambda}}(I) dI_i = -\delta(\lambda_i - I_i) dI_i. \tag{5}$$

Cumulative histograms are indeed the right tool to compute the mean of two histograms [9]. As the cumulative histogram H_I corresponds to an integration of h_I, it is equivalent to compare histograms or cumulative histograms: $H_I = H_J \Leftrightarrow h_I = h_J$.

An energy comparing the L^2 distance between cumulative histograms can be defined as:

$$E_T(I) = \int_D \left(H_I(\boldsymbol{\lambda}) - H_{I^r}(\boldsymbol{\lambda})\right)^2 d\boldsymbol{\lambda}, \tag{6}$$

where $D = [0, 255]^3$ is the range of the color channels.

A data term is added in order to penalize the separation of I and the initial color values I^0:

$$E_0(I) = \int_\Omega f(x) ||I(x) - I^0(x)||^2 dx, \tag{7}$$

where f is taken as:

$$f(x) = e^{\left(-\frac{\min_{z \in \partial\Omega} |x-z|^2}{\sigma^2}\right)}. \tag{8}$$

In this expression, $|.|$ denotes the euclidean distance and $\sigma > 0$ determines the domain of influence. This gives a simple data weight inversely related to the distance of a point to the mask boundary $\partial\Omega$. It will be of particular interest for post-processing inpainting results, where the information in the inner of the mask boundary should be kept. This will not be the case for other applications, such as shadow removal, where we do not want to preserve the colors inside the mask.

The variational transfer of histograms can finally be done by minimizing the energy

$$E_T(I) + \alpha E_0(I)$$

with respect to I. The weight $\alpha \geq 0$ balances the influence of the initial data I^0 with respect to the distance between the cumulative histograms of the variable I and the reference image I^r.

The minimization can be performed by means of a gradient descent approach, as shown in [17]. Even if the energy is not convex (due to the presence of the term $E_T(I)$), the local minima computed gives good results. Histogram transfer could be directly performed in the histogram space as it is done in [20] by looking for the best mapping (in the sense of the optimal transportation cost) between the histograms h_{I^0} and h_{I^r}. However, a main advantage of the formulation described here is that it allows preserving some local attributes of the original image I^0. For instance, as in [17], one could consider an additional term preserving the gradient orientations of the original image I^0. The function f is also a way to incorporate some specific spatial processes. Hence, two pixels having originally the same color can end up with different final colors after

the variational transfer. Such local processing is obviously not possible when dealing with mappings in the histogram space. In next section we describe how this approach can be extended to spatiogram transfer.

3 A model to transfer spatiograms

Let us describe how to combine the previous tools in order to implement the transfer of spatiograms.

3.1 Transferring the spatial information

The direct extension of the variational method to spatiograms could be done by considering the L^2 distance between the means and covariances defined in (2). However, as these quantities involve Dirac functions, and are not defined for empty bins, they are not differentiable with respect to I. Moreover, as the mean and the covariance of colors are not densities, considering their cumulation has no sense, and the previous framework cannot be adapted.

It is therefore necessary to model the transfer of mean and covariance in different ways. The purpose of transferring spatiograms is to encourage a pixel x with color $I^0(x)$ to have a final color $I(x)$ that is consistent with the spatiogram of the reference image S_{I^r}. In other words, x should not end with the color $I(x)$ if the mean position $\mu_{I^r}(I(x))$ is far from x. This condition could be added in the variational model with an additional term that aims at minimizing the distance:

$$\int_\Omega ||x - \mu_{I^r}(I(x))||^2_{\Sigma_{I^r}(I(x))} dx, \tag{9}$$

where $||u||^2_\Sigma$ is the norm $u^T \Sigma^{-1} u$. In practice, minimizing such a term with a gradient descent approach is not efficient, since the means and covariances are not defined for all the bins. A solution could be to reformulate the problem as a multi-label one where a color has to be assigned to each pixel. Techniques such as graph cuts [4] or convexification [21] could then be used. Nevertheless, even if we consider a small number of bins in the spatiograms, the computational cost involved in such approaches becomes prohibitive (e.g. 8 bins by color channel corresponds to 512 labels). The solution we propose is to add to the model a penalization term:

$$E_S(I) = \int_\Omega \int_{\tilde{D}} e^{\left(-||x-\mu_{I^r}(\boldsymbol{\lambda})||^2_{\Sigma_{I^r}(\boldsymbol{\lambda})}\right)} ||I(x) - \boldsymbol{\lambda}||^2 d\boldsymbol{\lambda} dx. \tag{10}$$

By minimizing this term, we encourage a pixel x to end up with the color $\boldsymbol{\lambda}$ if its position is close enough to the mean position $\mu_{I^r}(\boldsymbol{\lambda})$. With this new energy term containing the spatial information of the spatiogram, there is no need of differentiating the mean and covariance at each bin. As a consequence, the problem is now well posed, since it only considers the integral on the domain of non empty bins: $\tilde{D} = \{\boldsymbol{\lambda} \mid h_{I_r}(\boldsymbol{\lambda}) \geq \epsilon\}$. More precisely, we consider the bins that contain at least three pixels (i.e. $\epsilon = 3/|\Omega|$) in order to be able to compute a covariance. Note also that if all pixels of a bin $\boldsymbol{\lambda}$ belong to the same line $L_{\boldsymbol{\lambda}}$, then the determinant of the associated covariance matrix $\Sigma_{I^r}(\boldsymbol{\lambda})$ will be zero and the matrix will not be invertible. In this case, the argument of the exponential in (10) is simply given by the opposite of the distance between the pixel x and its projection onto the line $L_{\boldsymbol{\lambda}}$.

3.2 Dealing with mask boundaries

By doing a transfer of histogram or spatiogram inside a mask, some artifacts can appear at the mask boundary as there is no specific processing ensuring the spatial continuity of colors. Hence, we have to consider the color values at the boundary of the mask in order to diffuse the information across the boundary, as it is done in Poisson editing methods [19]. This is realized by considering the energy term:

$$E_D(I) = \sum_{i=1}^{3} \int_{B(\Omega)} ||\nabla I_i||^2, \tag{11}$$

together with the constraint $I|_{\partial\Omega} = I^0|_{\partial\Omega}$, where $\partial\Omega$ denotes the boundary of the mask Ω. In this equation $B(\Omega)$ is a band defined as $\{x \in \Omega : \text{dist}(x, \partial\Omega) \leq \eta\}$, $\eta > 0$. This additional term permits to diffuse the known data on the boundary into $B(\Omega)$ to ensure the spatial continuity of colors.

3.3 Final energy and minimization

The transfer of spatiograms into a region of interest is finally represented by the energy:

$$E(I) = E_T(I) + \alpha E_0(I) + \beta E_S(I) + \gamma E_D(I), \tag{12}$$

where $\alpha \geq 0$, $\beta \geq 0$ and $\gamma \geq 0$ weight the influence of the data, spatial and diffusion terms. By minimizing the whole energy (12), the image I will be modified so that its spatiogram S_I is closer to the spatiogram of the reference image S_{I^r} than to the one of the original image S_{I^0}.

We simply use a gradient descent strategy to minimize (12). Hence, we have the following numerical scheme for the variable $I = (I_1, I_2, I_3)$:

$$I_i^{k+1} = I_i^k - \tau \nabla_{I_i^k} E,$$

where $\tau > 0$ and the initial condition is given by the original image I^0. Following [17], the derivative of the comparison term between the two cumulative histograms (6) is

$$\begin{aligned} &\nabla_{I_i^k} E_T(I(x)) \\ =&\nabla_{I_i^k} \int_D (H_{I^k}(\boldsymbol{\lambda}) - H_{I^r}(\boldsymbol{\lambda}))^2 \, d\boldsymbol{\lambda} \\ =&-\frac{1}{|\Omega|} \int_{I_j(x)}^{255} \int_{I_k(x)}^{255} [H_{I^k}(I_i(x), \lambda_j, \lambda_k) - H_{I^r}(I_i(x), \lambda_j, \lambda_k)] \, d\lambda_k d\lambda_j, \end{aligned} \tag{13}$$

since, from relation (5), we have that $\nabla_{I_i^k(x)} H_{I^k}(\boldsymbol{\lambda}) = -\delta(\lambda_i - I_i^k(x))$. In (13), we used the notation $(I_i(x), \lambda_j, \lambda_k)$ to denote $(I_1(x), \lambda_2, \lambda_3)$ for $i = 1$, $(\lambda_1, I_2(x), \lambda_3)$ for $i = 2$ and $(\lambda_1, \lambda_2, I_3(x))$ for $i = 3$.

The data attachment term (7) simply gives

$$\nabla_{I_i^k} E_0 = f(I_i^k - I_i^0), \tag{14}$$

and the derivative of the term (10) is:

$$\nabla_{I_i^k} E_S = \int_{\tilde{D}} e^{\left(-||x - \mu_{I^r}(\boldsymbol{\lambda})||^2_{\Sigma_{I^r}(\boldsymbol{\lambda})}\right)} (I_i^k - \boldsymbol{\lambda}) d\boldsymbol{\lambda}. \tag{15}$$

Finally, the derivative of the diffusion energy term (11) is given by

$$\nabla_{I_i^k} E_D = -\Delta I_i^k. \tag{16}$$

Note that the primitive involved in equation (13) can be precomputed at each step from the current values of h_I and h_{I^r}. Similarly, the weights represented by the exponential in equation (15) can be computed and stored just once, for all $x \in \Omega$ and $\boldsymbol{\lambda} \in \tilde{D}$. As the minimization relies on an explicit scheme, the gradient descent can then be parallelized in the image domain.

4 Spatiogram visualization

In this section, we propose to apply the proposed energy minimization process in order to recreate a full image. The topic of image reconstruction from its descriptors has recently been studied in [32] in case of local descriptors. In our case, this application is equivalent to visualizing the spatiogram. As we process the whole image, it is not necessary to consider the mask boundary and we set $\gamma = 0 E_D(I) = 0$. Starting with the reference image in Fig. 1(a) and computing its spatiogram, we then only use the model formed by the sum of the spatiogram energy terms (6) and (10), with $\beta = 1$ to recreate the image, see Fig. 1(d). In this example, the unknown image I has been initialized with a constant gray value. The data term is thus ignored with $\alpha = 0$. In Fig. 1(c) we show the result obtained by minimizing only the spatial energy E_S of spatiograms. For comparison, we display in Fig. 1(b) the result obtained with the visualization process proposed in [3], where the color assigned to each pixel minimizes the cost (9):

$$I(x) = arg \min_{\boldsymbol{\lambda} \in \tilde{D}} ||x - \mu_{I^r}(\boldsymbol{\lambda})||^2_{\Sigma_{I^r}(\boldsymbol{\lambda})}. \tag{17}$$

This experiment shows how the spatiogram information can be used to reconstruct an image whose topology and colors are relatively close to the original image. This gives us a tool to visualize a spatiogram. We can naturally observe that the method does not allow to reconstruct texture. However, the proposed model can be successfully applied to specific image processing problems as shown in the next section.

5 Applications

In this section, we present some applications of the proposed model to image editing problems. First, we show how our process can realize (limited) image inpainting. Then, some pertinent examples for inpainting enhancement and shadow removal are presented. Some experiments on highlight removal are finally discussed as a perspective for future work.

5.1 Discussion on parameters

Before going any further, we would like to add some comments about the chosen parameters. In all the subsequent experiments, we used a discretization of 32 bins for each color channel, which corresponds to 2 minutes for treating a region containing 10^4 pixels with our current implementation that is not parallelized.

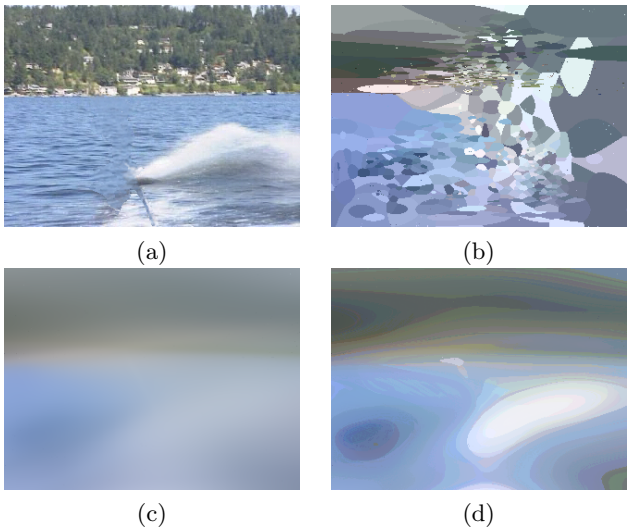


Figure 1: Example of image creation. (a) Original reference image. (b) Result obtained by assigning the best color to each pixel according to (17). (c) Result obtained by minimizing the energy term (10). (d) Result obtained by minimizing the sum of the spatiogram energy terms (6) and (10).

When considering 8 bins by color channel, the process only takes a few seconds to reach convergence. Note that the computational cost also depends on the number of pixels in the mask. In the experiments in sections 5.2.2, 5.3, 5.4 and 5.5, the reference spatiograms have been computed on a band of 20 pixels around the input mask. The results shown in the paper have been obtained with a time step $\tau = 10^{-3}$, the value $\sigma = 3$ for the domain of influence, a data coefficient $\alpha = 0.1$, a spatiogram coefficient $\beta = 1$ and a diffusion coefficient $\gamma = 10$. This large value of the diffusion coefficient is only used in a band of 5 pixels to reduce the visual artifacts that occur at the mask boundary. These parameters have been chosen empirically and a small variation of their values, almost preserving their order of magnitude, does not perturb the results. Because of the cumulative histogram, the number of bins has an influence on the weight of the histogram transfer energy term. For that reason, this term has been weighted with a factor $1/N^3$.

5.2 Image completion

The proposed model can be used for image completion in order to reconstruct the unknown part of an image using the spatiogram properties of the known region. For this application, there is no original image inside of the area to completed. Therefore, in the following completion experiments, we do not consider the data term by simply setting $\alpha = 0$.

5.2.1 Influence of the energy terms

As a first example, let us study a simple toy image (Fig. 2) and show the influence of the different terms involved in our model (12). The domain Ω is the unknown central area of the image. The reference spatiogram is computed on the the whole image outside the mask. The result in Fig. 2(b) has been obtained by only considering the energy term E_T (6) acting on histograms. As the mask has been initialized with a uniform gray value, the process applies the same perturbation to all the pixels. The obtained result consists of a constant color, which is the barycenter of the input colors, weighted by the frequency of each color.

Fig. 2(c) presents the result obtained by minimizing only the term E_S (10), *i.e.* without histogram transfer. New blurred colors that do not respect the histogram corresponding to the known part of the image are created. Notice however that, thanks to the use of spatial information, the reconstruction respects the spatial repartition of the colors at the mask boundaries. In Fig. 2(d) we show the result obtained by minimizing the spatiogram energy terms E_T and E_S, by setting α and γ in (12) to 0. This simple example shows the key role of the histogram transfer energy term $E_T(I)$, which allows imposing a global criteria (the histogram) to the unknown area. Such property is not ensured with classical inpainting approaches.

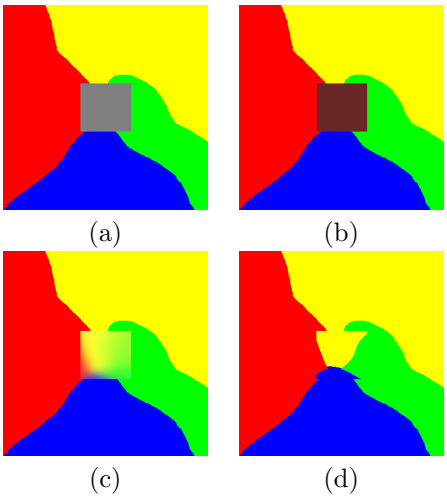


Figure 2: Example of inpainting and influence of the spatiogram energy term E_S. (a) Original image. (b) Result obtained by minimizing the single histogram energy term (6): $E = E_T$. (c) Result obtained by minimizing the single energy term (10): $E = E_S$. (d) Result obtained by minimizing the spatiogram energy terms (6) and (10): $E = E_T + E_S$.

With respect to state of the art diffusion-based inpainting approaches [30], such results are too sharp to be visually plausible. Therefore, we incorporate E_D, the diffusion term (11). We note that the goal of this subsection is not to propose a new inpainting method, but only to show what can be expected from spatiograms transfer. More sophisticated terms with higher order derivatives

[1, 6] could have been used instead of the diffusion (11). This would improve the results obtained by minimizing this single diffusion energy term E_D (11) (see Fig. 3(b)) but would still contain blurred edges. For instance, the diffusion method of [30] gives the improved result[1] in Fig. 3(c), which approximately preserves the known colors (even if some new ones are created) but not their histogram (see Fig. 4(a) for details).

For completeness, we also applied an exemplar-based inpainting method [8], as illustrated in Fig. 3(d). The result obtained with this copy/paste patch-based approach shows strong visual artifacts as the known part of the image does not contain enough pertinent information (see Fig. 4(b)).

As illustrated in Fig. 3(e), the minimization of the histogram E_T (6) and diffusion E_D (11) terms leads to a smooth result that respects exactly the histogram information of colors. However, the lack of spatial information implies the presence of a blue area inside the mask, on the upper part. As shown in Fig. 3(f), minimizing the full energy (12) solves this problem. Note that the influence of the diffusion term (11) is important when we compare Fig. 3(f) with the result in Fig. 2(c), as there is now a continuity of colors at the mask boundary (see Fig. 4(c)). The most interesting property of this approach is that it allows realizing a variational diffusion of the colors inside the mask while limiting the creation of new colors, which is often an important limitation of diffusion-based approaches. With respect to classical methods [30], very few new colors are created at the boundaries of two different color areas with the model (12).

5.2.2 Inpainting

Some results of inpainting on more complex images are shown in Fig. 5. We compared our results with the diffusion-based [30] and exemplar-based approaches [8]. In these examples, as we have no prior color information on the mask to inpaint, α is set to 0.

On the first example (see the first row of Fig. 5), the diffusion method of [30] and the proposed approach give both good global results by correctly rebuilding the line between sea and grass. The texture is obviously better with the exemplar-based approach [8]. On the second example which contains natural tree and grass textures, the diffusion method and the proposed process create, as expected, a smooth image which is not visually plausible. The patch-based method is in this case more realistic. By zooming on this last result (see Fig. 6 (a)), it can nevertheless be observed that the texture that has been inpainted is not the one we would think of. The histogram of inpainted colors is indeed not so close to the histogram of the exterior band around the mask. Most of the colors have indeed been copied from a particular region.

We will now show how the spatiogram transfer can be used to enhance the results obtained with exemplar-based inpainting methods.

[1] As suggested in [30], we set the contour preservation parameter $p1 = 10^{-3}$, the structure anisotropy $p2 = 10^2$ and the time step $dt = 150$.

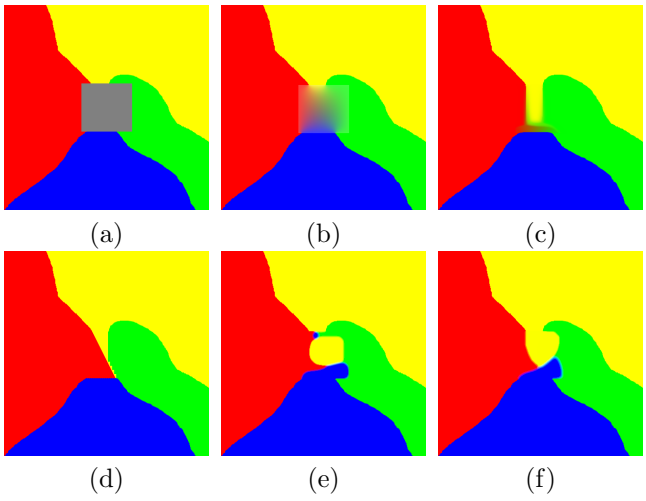


Figure 3: Example of inpainting. (a) Original image. (b) Result obtained by minimizing the single diffusion energy term (11): $E = E_D$. (c) Result obtained using the diffusion method in [30]. (d) Result obtained with [8] using patches of size 9×9. (e) Result obtained by minimizing the histogram energy (6) and the diffusion term (11): $E = E_T + 10E_D$. (f) Result obtained by minimizing the whole energy (12): $E = E_T + 10E_D + E_S$.

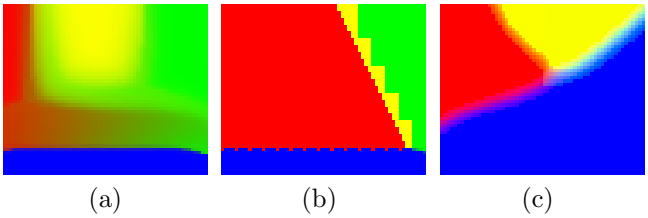
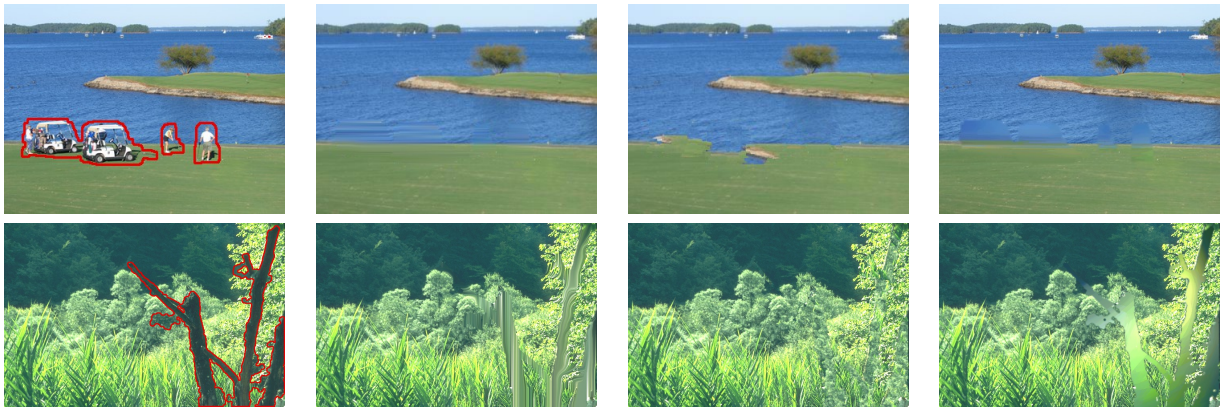


Figure 4: Zoom of the inpainted images (c), (d) and (f) of Fig. 3. (a) Result obtained with the diffusion method in [30]. (b) Result obtained with [8] using patches of size 9×9. (c) Result obtained by minimizing the whole energy (12):$E = E_T + 10E_D + E_S$.

5.3 Correcting exemplar-based inpainting

The patch-based inpainting methods allow reconstructing natural textures. As previously mentioned, the patch effectively used during the process can come from undesired regions of the image. Post-processing the exemplar-based result I^0 with the spatiogram transfer allows correcting this default. In order to preserve the exemplar-based texture estimation at the boundary of the mask, we now make use of the data term (7) which involves the function f that gives a data weight inversely related to the distance of a point to the mask boundary $\partial\Omega$. Hence, pixels inside the mask and close to the boundary are encouraged to keep their original color values, while pixels deeply inside the mask will be able to change their color values.



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

Figure 5: Exemplar-based inpainting. (a) Original image and the inpainting mask in red. (b) Results obtained with [30] . (c) Results obtained with [8] using patches of size 9×9. (d) Results obtained by minimizing (12).

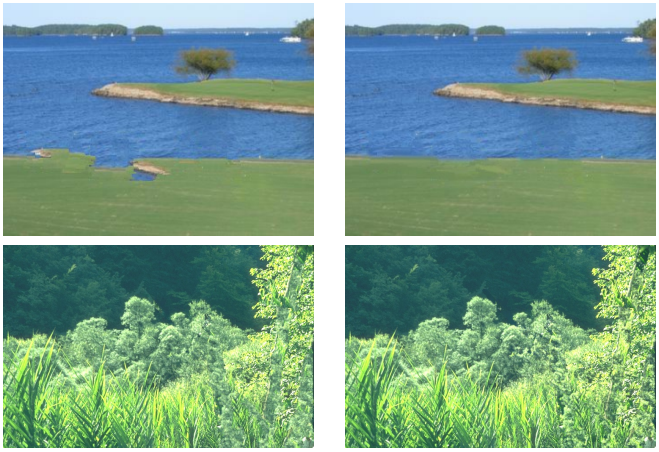
(a) (b)

Figure 6: (a) Zoom on the exemplar-based inpainting result of image (c) in the second row of Fig. 5. (b) Zoom on the corresponding spatiogram transfer correction.

From Fig. 7, we can see that the spatiogram transfer allows correcting the default of exemplar-based inpainting results while preserving the textures. This combination produces really plausible reconstructions (see Fig. 6(b) for details on the tree example).

5.4 Shadow removal

We now focus on the shadow removal problem [25] which is a pertinent field of application for spatiogram transfer. As in the case of exemplar-based inpainting enhancement, the mask already contains some structures/textures that should be preserved through the transfer process. Let I^0 be an image containing areas with shadows due to sun exposure that we would like to remove. Such areas can be automatically estimated with state-of-the-arts method [33]. Once the mask Ω containing shadow areas has been estimated, the proposed process can



(a) (b)

Figure 7: Exemplar-based inpainting correction. (a) Result obtained with [8] using patches of size 9×9. (b) Post-processing result obtained by minimizing the energy (12) and using the data term function (8).

be applied to transfer the histogram of the exterior region to the area inside the mask. The use of spatiograms allows to preserve the distribution of outside colors inside the mask.

An example of shadow removal taken from [25] is illustrated in Fig. 8(b). The result obtained with our approach is shown in image 8(d), we can see that the known information (the texture inside the mask and the colors outside) is combined to estimate an image that seem real. In Fig. 8(c) we show the result obtained with the transfer of colors realized in histogram space, using an adaptation of the color grading code of [20][2]. Excepting the mask boundary, the result obtained seems very similar to the one produced by our spatiogram transfer, but if we look carefully at the details (see Fig. 9), we observe that the transfer in histogram space creates some un-natural spots as it contains no particular spatial processing.

A further example shown in Fig.10(a), taken from [15], displays the original image and the mask of shadowed pixels. Fig. 10(b) shows the result of [15] which requires an accurate user selection of the shadow transition boundary (i.e. the pixels that are not totally in the shadow nor in the light) to perform a specific process. Note that we do not need to have such an accurate band estimation, as the diffusion and data terms are able to deal with these shadows transition areas. Fig. 10(c) and 10(d) show the results obtained with color transfer in histogram space (using the code of [20]) and with our method, respectively. We note that using spatiogram encourages the spatial repartition of final colors.

The results obtained with [20] are good inside the mask, but not satisfactory at its boundary: the color of the pixels that belong to the mask and are not in the shadow are too bright on the final result. This is due to the fact that his-

[2]To the best of our knowledge, there is no existing work that consider histogram transfer to solve the shadow removal problem. The use of [20] is thus also an original way of addressing the problem

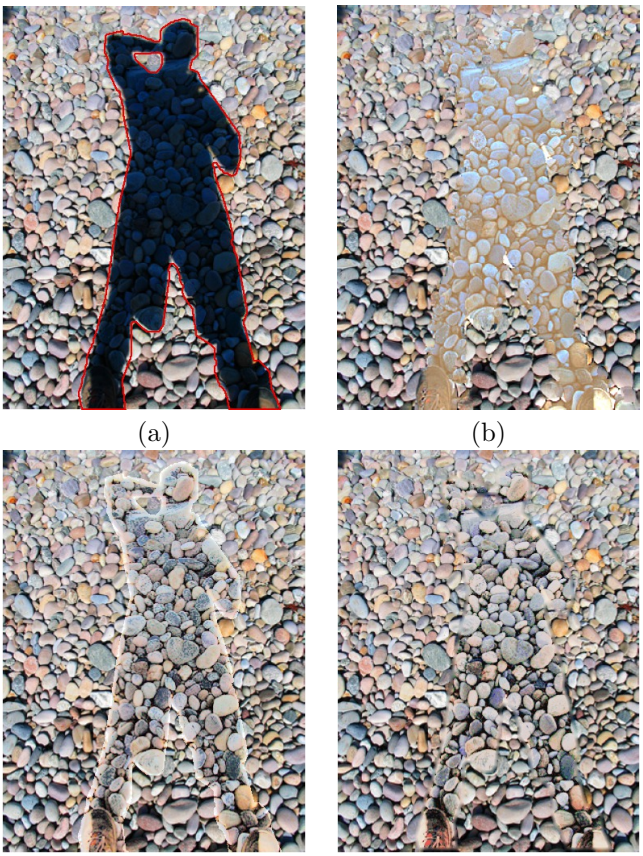


Figure 8: Shadow removal. (a) Original image. The mask of shadowed pixels to be processed is delimited with a red curve. (b) Result taken from [25]. (c) Result obtained with a transfer of colors in histogram space [20]. (d) Solution with the model (12).

Figure 9: Zoom on shadow removal. (a) Zoom of the result in Fig. 8(c) obtained with a transfer of colors in histogram space [20]. (b) Zoom of the result in Fig. 8(d) obtained with the model (12).

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

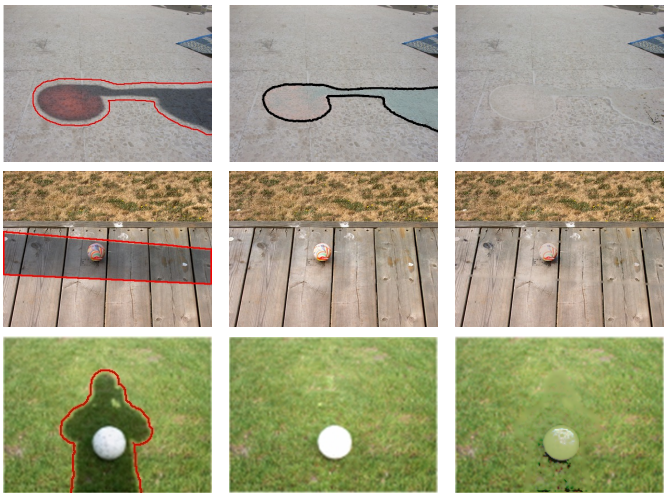
Figure 10: Shadow removal. (a) Original image with the mask of shadowed pixels and the shadow boundary. The image is taken from [15]. (b) Result in [15]. (c) Result obtained by histogram transfer using [20]. (d) Result obtained using the full model (12).

togram transfer methods try to brighten all pixels inside the mask, while pixels close to the boundary are not completely covered by the shadow. There is indeed a smooth transition between shadow and non-shadow that must be treated with detail. This is an important problem that has been carefully addressed in [15] by underlining the problem of mask definition: to be efficient, shadow removal methods need: (i) a mask that only contains pixels with shadow, and (ii) the knowledge of the band where shadow transition occurs. In this paper, in order to circumvent this limitation, we introduced the diffusion energy (11) in a band inside the mask. This allows removing the bad boundary effect previously observed. Note that this addition does not only smooth the image, as the spatiogram transfer is simultaneously taken into account in the band. However, as these two terms do not allow creating textures, the band should be taken small enough in practice (5 pixels). It is important to mention that such a process running in the image domain cannot be used with the method of [20] that considers the transfer of colors in histogram space. A posterior diffusion should then be applied in this case.

Fig. 11 shows other examples of shadow removal. The original images are taken from [25]. We compare our results with the ones in [25]. In the the first row of Fig. 11(b), we see that [25] is not able to recover all the colors in the shadowed area[3], whereas the proposed method removes all red and blue colors inside the mask while preserving the texture of the original image. In the last two examples, our method gives results comparable to the ones in [25], except for the ball color. Indeed, in this case, our method cannot recover the ball colors: as there is no other ball available outside the shadow mask, the histogram have empty bins for these colors.

These shadow removal experiments have been run with an image completion method such as [8] or [30]. However, such approaches would have created structures without preserving the original textures (see for instance the rocks in Fig. 8 or the balls in Fig. 11).

[3]The black boundary just denotes the mask boundary



(a) (b) (c)

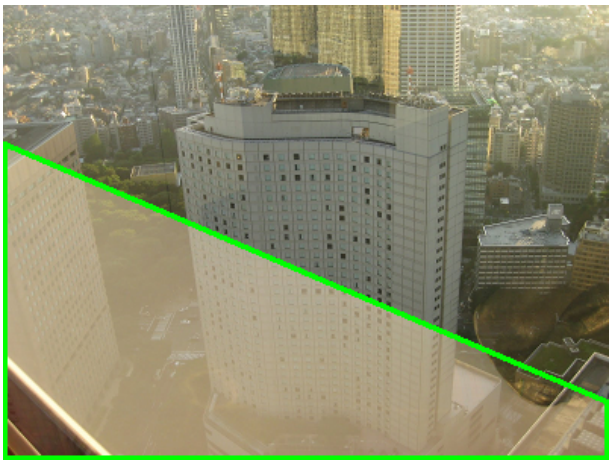
Figure 11: Shadow removal. (a) Original image with the mask of shadowed pixels. (b) Result taken from [25]. (c) Result obtained using the full model (12).

5.5 Highlight removal

The same process can be used to remove visual artifacts due to camera flash in a glass or sun exposure. In the general case of highlight removal, we show in Fig. 12 that the proposed process is imperfect as the mask boundary is not well restored. Indeed, the textures at the mask boundary cannot be handled correctly with the diffusion term. The colors inside the mask are nevertheless visually plausible.

We compared our approach with the Poisson editing method [19] in Fig. 13. As mentioned in the original paper, the implementation we used to realize these experiments includes an artificial reduction of the reference gradient amplitude in order to remove the highlight effects[4]. With this modification of the data, Poisson editing tends to smooth the image and is not able to remove the highlights. Nevertheless, neither Poisson editing [19] nor the proposed method work correctly for highlight removal. In the case of Poisson editing, it can be impossible to find the correct artificial reduction of gradient amplitude that produces plausible colors. With the proposed model, the failure is mainly due to the color saturation. Contrary to shadows, there is no preserved texture to correct. We have studied the influence of the color space representation for computing histograms in this particular case, as we thought that working in an appropriate uncorrelated color space could lead to a solution. With YCbCr or HSV color spaces, the illumination and color information can be represented by different channels that are less correlated. We then divided the channels into two groups, depending on whether or not the highlight is perturbing the channel information. The process has then been applied separately on each group,

[4] http://www.howardzzh.com/research/poissonImageEditing/index.htm



(a)

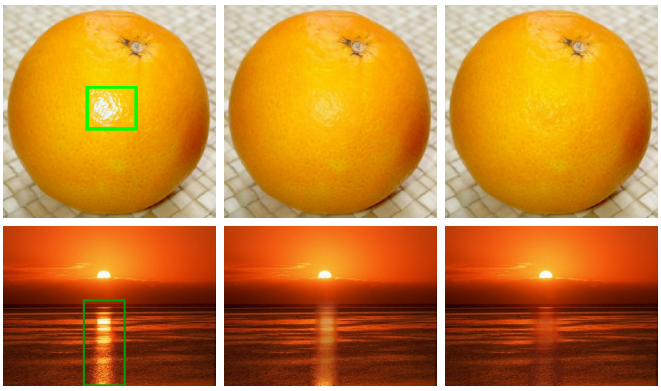
(b)

Figure 12: Flash removal. (a) Original image with the mask of highlight pixels to be removed delimited with a green curve. (b) Estimated solution.

while considering correlations inside the channel of a group. Unfortunately, the obtained results were not better than the ones obtained with RGB color space. Some other direction should be studied in the future.

6 Conclusion and perspectives

In this paper, we have proposed a variational formalism to realize spatiogram transfer, and addressed different image processing problems such as image completion or shadow and flash removal. We showed with various examples and comparisons that the methodology allows improving inpainting results obtained with diffusion-based or patch-based methods. We also applied the process to shadow removal and obtained some significant results. The extension to high-



(a) (b) (c)

Figure 13: Highlight removal. (a) Original image with the mask of highlight pixels to be removed delimited with a green curve. (b) Result with Poisson Editing [19]. (c) Estimated solution.

light removal is more complex, and will be studied in the near future.

The enhancement of the proposed model deserves also further attention. Namely, spatiogram representation could benefit from the use of multi-modal Gaussian densities, as suggested in [3], in order to capture more information on the spatial distribution of colors. Finally, some advanced model (such as [15]) dealing with mask boundary effects will also be studied.

Acknowledgments

Nicolas Papadakis acknowledges the support of the French Agence Nationale de la Recherche (ANR) under reference ANR-11-BS01-014-01. Vicent Caselles acknowledges partial support by MICINN project, reference MTM2009-08171, by GRC reference 2009 SGR 773 and by "ICREA Acadèmia" prize for excellence in research, the last two funded by the Generalitat de Catalunya.

References

[1] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," in *SIGGRAPH: ACM Transactions on Graphics (SIGGRAPH'00)*, 2000, pp. 417–424.

[2] M. Bertalmío, V. Caselles, E. Provenzi, and A. Rizzi, "Perceptual color correction through variational techniques," *IEEE Transactions on Image Processing*, vol. 16, no. 4, pp. 1058–1072, 2007.

[3] S. T. Birchfield and S. Rangarajan, "Spatiograms versus histograms for region-based tracking," in *IEEE Conference on Computer Vision and Pattern Recognition(CVPR'05)*, 2005, pp. 1158–1163.

[4] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 11, pp. 1222–1239, 2001.

[5] A. Bugeau, M. Bertalmio, V. Caselles, and G. Sapiro, "A Comprehensive Framework for Image Inpainting," *IEEE Transactions on Image Processing*, pp. 2634 – 2645, Oct. 2010.

[6] T. F. Chan and J. Shen, "Variational image inpainting," *Communications on Pure and Applied Mathematics*, vol. 58, pp. 579–619, 2005.

[7] K. Choudhary, M. Pundlik, and D. Choukse, "An integrated approach for image retrieval based on content," *Journal of Computer Science*, vol. 7, no. 3, pp. 42–47, 2010.

[8] A. Criminisi, P. Perez, and K. Toyama, "Object removal by exemplar-based inpainting," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'03)*, vol. 2, 2003, pp. 721–728.

[9] J. Delon, "Midway image equalization," *Journal of Mathematical Imaging and Vision*, vol. 21, pp. 119–134, 2004.

[10] J. Delon, "Movie and video scale-time equalization application to flicker reduction," *IEEE Transactions on Image Processing*, vol. 15, no. 1, pp. 241–248, 2006.

[11] T. Georgiev, "Covariant derivatives and vision," in *European Conference on Computer Vision (ECCV'06)*, 2006.

[12] Y. Jin, L. Fayad, and A. Laine, "Contrast enhancement by multi-scale adaptive histogram equalization," in *Wavelet Applications in Signal and Image Processing IX*, 2001.

[13] P. Koirala, M. Hauta-Kasari, and J. Parkkinen, "Highlight removal from single image," in *Advanced Concepts for Intelligent Vision Systems (ACIVS'09)*, vol. 5807, 2009, pp. 176–187.

[14] H. Li, S. Wang, and W. Z. M. Wu, "Image inpainting based on scene transform and color transfer," *Pattern Recognition Letters*, vol. 31, pp. 582–592, May 2010.

[15] F. Liu and M. Gleicher, "Texture-consistent shadow removal," in *European Conference on Computer Vision (ECCV'08)*, 2008, pp. 437–450.

[16] S. Malassiotis and M. G. Strintzis, "Stereo vision system for precision dimensional inspection of 3d holes," *Machine Vision and Applications*, vol. 15, pp. 101–113, December 2003.

[17] N. Papadakis, E. Provenzi, and V. Caselles, "A variational model for histogram transfer of color images," *IEEE Transactions on Image Processing*, vol. 20, no. 6, pp. 1682 – 1695, June 2011.

[18] M. Pappas and I. Pitas, "Digital color restoration of old paintings," *IEEE Transactions on Image Processing*, vol. 9, pp. 291–294, 2000.

[19] P. Pérez, M. Gangnet, and A. Blake, "Poisson image editing," in *SIGGRAPH: ACM Transactions on Graphics (SIGGRAPH'03)*, 2003, pp. 313–318.

[20] F. Pitié, A. C. Kokaram, and R. Dahyot, "Automated colour grading using colour distribution transfer," *Computer Vision and Image Understanding*, vol. 107, pp. 123–137, July 2007.

[21] T. Pock, A. Chambolle, D. Cremers, and H. Bischof, "A convex relaxation approach for computing minimal partitions," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'09)*, 2009, pp. 810 –817.

[22] J. Rabin, J. Delon, and Y. Gousseau, "Regularization of transportation maps for color and contrast transfer," in *International Conference on Image Processing (ICIP'10)*, 2010, pp. 1933 –1936.

[23] D. L. Ruderman, T. W. Cronin, and C.-C. Chiao, "Statistics of cone responses to natural images: implications for visual coding," *Journal of the Optical Society of America*, vol. 15, no. 8, pp. 2036–2045, Aug 1998.

[24] G. Sapiro and V. Caselles, "Histogram modification via differential equations," *Journal of Differential Equations*, vol. 135, no. 2, pp. 238–268, 1997.

[25] Y. Shor and D. Lischinski, "The shadow meets the mask: pyramid-based shadow removal," *Computer Graphics Forum*, vol. 27, no. 2, pp. 577–586, 2008.

[26] B. Singh, G. Sinha, and I. Khan, "Comparison of histogram and spatiograms for content based retrieval of remote sensing images," *Communications in Computer and Information Science*, vol. 70, pp. 152–156, 2010.

[27] J. Stark, "Adaptive image contrast enhancement using generalizations of histogram equalization," *IEEE Transactions on Image Processing*, vol. 9, no. 5, pp. 889 –896, 2000.

[28] K. Sunkavalli, M. Johnson, W. Matusik, and H. Pfister, "Multi-scale image harmonization," in *ACM SIGGRAPH 2010 papers*, 2010, pp. 125:1–125:10.

[29] P. Tan, S. Lin, L. Quan, and H. yeung Shum, "Highlight removal by illumination-constrained inpainting," in *International Conference on Computer Vision (ICCV'03)*, 2003, pp. 164–169.

[30] D. Tschumperlé, "Fast anisotropic smoothing of multi-valued images using curvature-preserving pde's," *International Journal of Computer Vision*, vol. 68, pp. 65–82, 2006.

[31] A. Ulges, C. H. Lampert, and D. Keysers, "Spatiogram-based shot distances for video retrieval," in *Trecvid 2006 Workshop*, 2006.

[32] P. Weinzaepfel, H. Jégou, and P. Pérez, "Reconstructing an image from its local descriptors," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'11)*, Jun. 2011.

[33] T.-P. Wu and C.-K. Tang, "A bayesian approach for shadow extraction from a single image," in *International Conference on Computer Vision (ICCV'05)*, vol. 1, 2005, pp. 480–487.