Digital Signal Processing 93 (2019) 1–12

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

Directed color transfer for low-light image enhancement

Laura Florea *, Corneliu Florea

Image Processing and Analysis Laboratory (LAPI), University Politehnica of Bucharest, Romania

ARTICLE INFO

Article history:
Available online 5 July 2019

Keywords:
Low-light
Color transfer
Clustering
Extreme channels
Deconvolution

ABSTRACT

Underexposed, low-light, images are acquired when scene illumination is insufficient for a given camera. Camera limitation originates in the high chance of producing motion blurred images due to shaky hands. In this paper we suggest to actively use underexposing as a measure to prevent motion blurred images to appear and propose a novel color transfer as a method for low light image amplification. The proposed solution envisages a dual acquisition, containing a normally exposed, possibly blurred image and an underexposed/low-light, but sharp one. Good colors are learned from the normal exposed image and transferred to the low light one using a framework matching solution. To ensure that the transfer is spatially consistent, the images are divided into luminance perceptual consistent patches called frameworks and the optimal mapping is piece-wise approximated. The two image may differ by colors and subject to improve the robustness of the spatial matching, we added supplementary extreme channels. The proposed method shows robust results from both an objective and a subjective point of view.

1. Introduction

The development of mobile phones and the wide spread use of integrated camera devices require miniaturization of the camera module. This further leads to design changes such as diminishing the optic size or shrinking the photo-sensible area of the image sensor. In the image sensor, the decrease of photo-sensible area, indirectly, reduces the correlation between the incident light and the reported image intensity, thus forcing increased exposure time.

In many situations the device containing a camera module is held directly in hand by the photographer. One characteristic of the human beings is the existence of the hand tremor. The reduced photo-sensible area of the image sensor decreases the picture angle, while the human hand jitter is always present. The combination of these two factors together with the resulting large exposure time gives rise to increased chances that the relative hand tremor induces motion blur. This phenomenon is, probably, the most annoying degradation of the photographs visual quality, so that camera manufactures and photographers are frequently searching for methods to contain its effects.

Various solutions have been proposed over the years for the problem of image degradation due to motion blur. In-camera solutions are based on Optical Image Stabilization which can be delivered in many embodiments; usually a prismatic block of glass or the sensor itself is moved in opposition with camera movement which is recorded by motion sensors. Such a solution acts simultaneous with the image acquisition and its main disadvantage is related to cost and size of panning block, which, in many cases, does not fit into small camera modules.

Alternatives are related to post-processing: the blurry acquisition is allowed and afterwards measures to compensate and restore are envisaged. Such an alternative is to estimate the degradation kernel (known as Point Spread Function – PSF) and compensate it. If the estimation is done directly from the degraded image, the whole solution is called "blind deconvolution" (and we refer to the works of Levin et al. [1] and Ruiz et al. [2] for reviews on the topic). The dominant idea in this case is to use the image edginess to re-construct the PSF [3] and further use the found PSF in a non-blind deconvolution.

In another class of solutions, one may estimate the movement by motion sensors and follow by non-blind deconvolution, as in the work of Joshi et al. [4].

Methods that build a PSF have the main disadvantage that, usually, they assume the motion kernel to be spatially invariant (or uniform). Yet, this assumption, according to the measurements reported by Singhy and Riviere [5], is not realistic. Human tremor contains significant components on the Z axis and translational components (that lead to different trajectories for pixels corresponding to different depths [4]) resulting to heavily non-stationary PSFs. In the later years, proposals to address non-uniform blur appeared. Such examples that use non-stationary PSF models are, for instance, in the works of Whyte et al. [6], Pan et

* Corresponding author.
E-mail address: laura.florea@upb.ro (L. Florea).

https://doi.org/10.1016/j.dsp.2019.06.014
1051-2004/

al. [7] or Sun et al. [8]; however, in these cases the non-uniformity assumed fails to realistically model the natural variation of the human tremor.

An alternative to PSF estimation and deconvolution is to avoid the circumstances that generate the unwanted motion blur by reducing the exposure time below the "motion limit". The "motion limit" is found in the world of photographers by the empirical "1 over f35" rule: a hand held 35 mm camera should have an exposure in seconds that is not longer than the inverse of the focal length in millimeters; an arbitrary camera optics and image sensors may be referenced to a 35 mm camera. The "motion limit" was more thoroughly determined by Xiao et al. [9] as "q over f35" (with $q > 1$) and depending on the camera weight and photographer experience. In our work we set the exposure time based on the rules proposed in later mentioned work.

Our proposal draws inspiration by the previous works [10,11], as it assumes the acquisition of two input images: one is normally exposed, but potentially blurred and one is underexposed but sharp and still. The issue of enhancing the underexposed image is treated as a problem of transferring color from the normally exposed image (that in the remainder of the paper will be named the reference image) to the underexposed one (also named subject image).

This paper is a continuation over our previous work on low light enhancement via color transfer [12]. Our previous solution suffered, in some specific cases, from poor matching between the two input images. As they have different exposure values, the images should be different with respect to pixel intensity and in this work we introduce *extreme channels* as a measure to alleviate such effects. The use of the extreme channels is also theoretically motivated.

Overall, the main contribution of this paper is the introduction of a perceptually inspired color transfer method adapted to the following dual-image input scenario: (1) the underexposed image has sharp edges but it lacks good colors; (2) the normally exposed image is potentially blurred due to hand tremor, but has good colors; (3) the two images contain *almost* the same scene. This claim will be validated by intensive experimentation. A secondary contribution of the current work is a thorough discussion of the reasons why underexposing and amplification is more practical than deconvolution.

The remainder of the paper is organized as follows: prior works related to color transfer and low-light enhancement is reviewed in section 2. The discussion of practical impact in contrast to deconvolution based methods follows in section 3. The proposed algorithm is described in section 4 at both intuitive and theoretical levels, with an emphasis on the effectiveness of using extreme channels. Implementation details and achieved results are detailed in the next sections. The paper ends with conclusions.

2. Related work

The main contribution of this paper is a color transfer method designed for enhancing low-light images. Thus the current section will briefly survey color transfer and low light enhancement methods.

Color transfer (or color mapping) algorithms aim to recolor a subject image by computing a transfer function (mapping) between that image and another one (called the reference image). Following the recent reviews on the topic by Faridul et al. [13] and, respectively, Finalyson et al. [14], color transfer methods may be divided on point-based or region based. In the point based category, one should note the very influential work of Reinhard et al. [15] which matches the first two statistical moments of the two images in the $l\alpha\beta$ uncorrelated color space introduced earlier, [16]. Yet the work while being simple and intuitive is general and many further enhancement have been proposed to address various scenarios. Also in this category fall the methods proposed by Pitie et al. [17] which maps the N-dimensional color distribution of the reference image onto subject image, or the one introduced by Pouli and Reinhard [18] which performs the mapping pixel-wise but do stage it sequentially by considering pyramidal resolutions. These solutions lead to good quality results, but we consider that they are general transformations which do not adapt well to certain situations such as the one described here. Mechrez et al. [19] proposed a more complex framework that is able to address style transfer; in the color transfer part the method relies on deep-net based semantic segmentation followed by a mapping computed based on solving a set of Poisson equations.

In the category of region based methods falls the region consistent method [20]; yet its application is restricted by the assumption that region pairs preserve their monotonicity in the two images, which may not be necessarily fulfilled in our scenario due to different acquisition time. Also Olivera et al. coarsely register two images, segment images into regions (by Expectation-Maximization [21] or mean-shift [22]) and perform transfer from one to the other based on region impairment. Conceptually we differ by the fact that our solutions does not assume any registration step, thus it does not encode rigid spatial correspondences between the two images, but only color intensities correspondences. Furthermore, we introduce a general mathematical model out of which, given a probabilistic approach and specific choices, these previously proposed methods may be retrieved.

Another similar sets of methods are based on optimal transport [23,24]. In either case, for color transfer, the image is first split into segments. In the first work, the split is coarse and based on K-means [23], while the second is fine-grained with superpixel algorithm. Next, the mapping is computed via optimal transport with some adaptations to the problem: the optimization is regularized and relaxed (approximate), as bijective matching may overfit and cause artifacts. The segments used are further refined [24] by smoothing based on the spatial distance. The here-proposed method differs by the transfer mapping (which in our case is piecewise linear), the fuzzy segments and the use of the extreme channels as better references for spatial matching.

Low-light image enhancement is another area that captured a lot of interest. In the later years, Fotiadou et al. [25] proposed to enhance low-light image by constructing day and respectively night dictionaries based on sparse representations. Lore et al. [26] showed that low-light enhancement is achievable by the same auto-encoder based deep-net topology that was previously shown to perform denoising. Ko et al. [27] derived their method on variational framework where $\mathcal{L}_2$ norm is minimized for pixel smoothness and $\mathcal{L}_1$ for noise control. Zhang et al. [28] also use a variational model, this time in conjunction with Retinex theory to enhance low-light images. Ren et al. [29] used the camera response function to select an adapted amplification factor for each pixel within a single image framework. However all methods are based on single image enhancement and it is reasonable to assume that a reference image should improve the resulting image quality.

Multiple images were used by Fu et al. [30]; yet the images used originates in the single low-light image and are derived on different paths for consistency with human intuition and for better control over denoising, illumination and contrast. The same idea of dual path development for low light enhancement is found in the work of Jung et al. [31], which relied on wavelet decomposition for separation of data into luminance and contrast. We differ from these works by the fact that in our case two images are acquired, while they extract the pair from a single acquired one, and the nature of the images: in our case one is color reference, while the other is content (object) reference.

HDR imaging is an area that bears similarities with the proposed method. To acquire HDR scenes, consecutive frames with different exposures are typically acquired and combined into a HDR image that is viewable on regular displays. Mainly two approaches are identifiable in the prior art: irradiance fusion [32] which acknowledges that the camera recorded frames are non-linearly related to (1) the scene reflectance, thus, it does fusion in the radiance space and follows with a tone mapping to return in to display space; (2) exposure fusion [33] which directly combines the acquired frames into the final image. While many solutions have been proposed, our approach differs by the fact that the fusion is directed with respect to a reference (normally exposed image for colors and underexposed image for texture/content), while in HDR the fusion is uniform, without reference. Furthermore, while most HDR methods assume identical scenes in the acquired images, in our case the scene is never identical due to hand motion between and during the acquisition process.

3. Hand tremor and deconvolution

In still image acquisition, the motion blur may appear due to the involuntary hand tremor. A formal definition of the tremor is provided by Crawford and Zimmerman: "Tremor is a common disturbance of movement, and it is defined as a rhythmic and oscillatory movement of a body part, caused by involuntary repetitive muscle contractions" [34]. From a statistical point of view, the tremor is a random signal. No matter the exposure time, there is a non-zero probability to have either motion blur degraded image (i.e. the tremor cumulative motion is larger than a pixel), or a perfectly sharp image (i.e. cumulative motion is smaller than the size of a pixel). With the increase of the exposure time, the probability of blur increases, while the probability of sharp decreases.

An intuition about the probable size of a motion blur kernel (PSF) can be found by briefing studies of human tremor. The hand tremor was substantially studied in the bioengineering domain as it interferes with microsurgeons ability to keep hands still. Veluvolu and Ang [35] performed a comparative study of microsurgeons and normal people and found that amplitude of movement for microsurgeons is at half of the normal people. More recently Papini et al. [36] performed extensive studies of the hand tremor while aiming to build haptic interface; they have found the majority of the spectrum between 3.1 and 6.1 Hz.

Using inertial sensors, Singhy and Riviere [5] measured the absolute deviation of the human tremor in microsurgeons and found comparable amplitudes for all three axes. In terms of image processing, this finding means that the amplitude of rotational components generating a certain size of motion blur during an acquisition also exists on the Z axis and thus it contributes to the deep non-stationarity nature of the PSF. Assuming that the spatially variable PSF is completely retrieved (which was not yet achieved in works related to deconvolution), typically the non-stationary deconvolution is highly computationally intensive [8]. For instance Gupta et al. [37] report one hour on CPU to solve 1 Mpixel image, while Hirsch et al. [38] report 440 secs with GPU acceleration for the same image size. More recent methods reports small time for PSF estimation such as below 2 seconds in [39]. However the preferred method for deconvolution with spatially variant kernel is by minimization of the expected log likelihood (EPLL) [40], which being computationally intensive takes around 100 sec for 1 Mpixel image

In parallel, taking into account that Singhy and Riviere [5] report an average displacement for the hand tremor of 22 µm, while the dominant frequency of 4–5 Hz [35], [36], and noting that a high end smartphone has a camera with the pixel size of 1.12 µm for an exposure of 1/4 sec = 4 Hz, a PSF size of 19 pixels may be produced. Also for the same exposure the PSF may be completely non-uniform across the image: the PSF in top left corner may get to be near-perpendicular from the one in bottom right corner.

This paper argues that it is computationally more efficient and the results are more robust if, instead of deconvolution, an underexposing followed by a color transfer oriented method for compensation of the low light is used. The results further presented, show that up to 2 exposure stops may be compensated by such a solution.

4. Framework oriented color transfer

In summary, the proposed method assumes two images as inputs: one normally exposed (with $EV = 0$[1]) and an underexposed one (having $EV < 0$). For each of the two images, the two extreme, dark and bright, channels are computed. Each 5 (R, G, B + dark + bright) dimensional image is decomposed in intensity consistent frameworks using a clustering algorithm. Next, the frameworks are matched in pairs: given a framework of the low-light image, its pair is found as most similar framework, in terms of color range, from the normally exposed one. The colors are transferred from the normally exposed image to the underexposed one; the transfer is performed per each pair framework. The proposed schematic is presented in Fig. 1. In the following subsections, we discuss insights of the method.

4.1. Color transfer model

The first strong solution to the color transfer problem was proposed by Reinhard et al. [15]. In this method, the tones, u, from the source image, I_s, are adjusted based on the ones from the reference image, I_r, on each chosen color plane independently, according to:

$$g(u) = au + b \tag{1}$$

The specificity of this mapping is given by the choice of the color planes (taken as uncorrelated planes) and of the constants a and b. These are computed as the ratio of the standard deviations $a = \frac{\sigma_r}{\sigma_S}$ and respectively as difference of statistical means of the two images $b = a \cdot \mu_r - \mu_s$. This approach, while being simple thus general, was amended by various consecutive improvements [13].

One may observe that for the proposed solution, the content of the two images is highly similar as being consecutive acquisitions. Thus we may choose to add more adaptability by implementing the transfer as:

$$g(u) = \sum_{i=1}^{N} c_i \nu_i(u)(a_i u + b_i) \tag{2}$$

where c_i are algorithm depending weights, while $\nu_i(u)$ are pixel dependent weights that in or solution will determine the branch of the transfer.

While vectorial transfer is possible, due to the too large space, it is custom to restrict u from eq. (2) to be scalars (i.e. intensities from a color plane). Intuitively the transfer is implemented as linear on pieces. From an intensity perspective the function is piecewise linear, where "pieces", i are compact ranges of values. From a spatial point of view, the mapping is a convex combination, where the weights c_i are membership degrees of a location

[1] EV refers to the relative exposure value. The exposure value results as a combination of the terms from the APEX systems where shutter speed, diaphragm opening and amplification aims to balance the scene illumination. Normal exposure is reached when the APEX equation is balanced. For more detail kindly see [41].

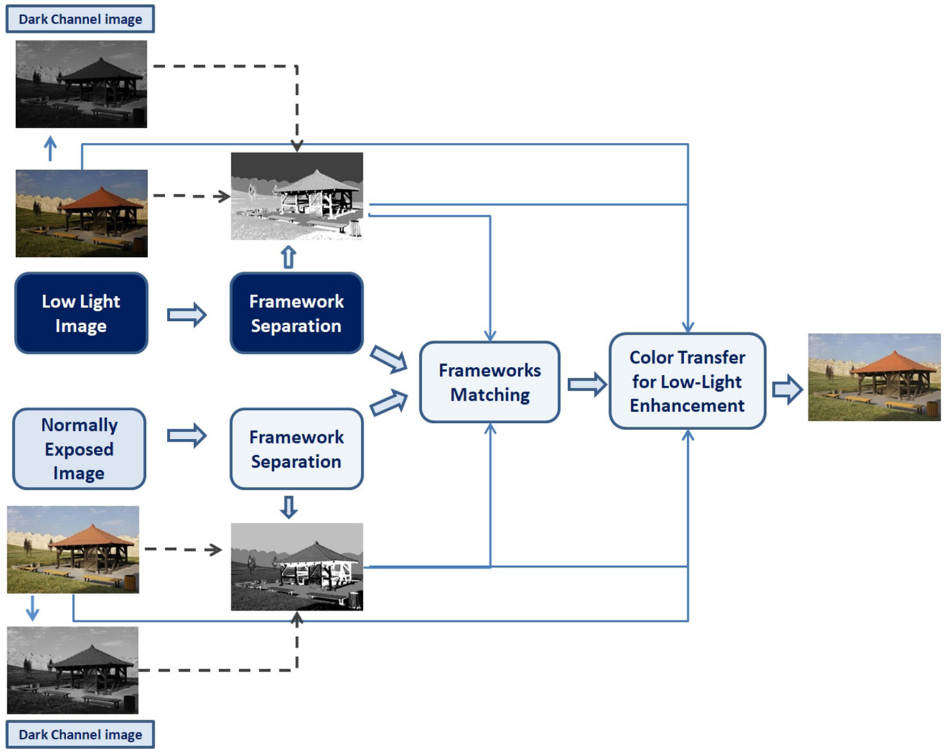


Fig. 1. The schematic of the proposed algorithm. Two input images containing differently exposed views of the same scene are given. They first are segmented into frameworks; following a framework matching, the low-light image receives the colors of the normally exposed one, thus being enhanced. The consistency of the frameworks between the two input images is based on the use of the extreme channel images, obtained from the two inputs. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

to a specific range and act to prevent the appearance of artifacts in the middle of objects.

By selecting $\nu_i(u) = \begin{cases} u, & u \in [u_i^{(m)}, u_i^{(M)}] \\ 0, & \text{otherwise} \end{cases}$ and $u_i^{(m)} = u_{i-1}^{(M)} + 1$, the eq. (2) retrieves the piecewise-consistent color mappings method [20]. Instead of the boxcar function with very steep transition, in this work we opted for smoother transition typical of fuzzy logic; this aspect will be detailed later in this section.

Given the histogram of the reference image, $h(I_r)$ and the histogram of the reconstructed image, $h(g(I_s))$ the solution of mapping depicted in eq. (2) is retrievable by solving the minimization problem:

$$argmin_{\mathbf{a},\mathbf{b},\mathbf{c},\nu(u)} (h(I_r) - h(g(I_s))^2 \;\; s.t. \sum_{i=1}^{N} c_i = 1 \tag{3}$$

where $\mathbf{a} = [a_1, \dots a_N]$, $\mathbf{b} = [b_1, \dots b_N]$, $\mathbf{c} = [c_1, \dots c_N]$, $\nu(u) = [\nu_1(u), \dots \nu_N(u)]$. First, one has to choose a parametric form for the function ν to have the minimization possible. Opting for a boxcar function and solving directly eq. (3), the solution retrieved has $u_i^{(m)} + 1 = u_{i-1}^{(M)}$ and $\mathbf{c}$ as one of the N-dimensional unit vectors; thus it implements the piecewise linear approximation of the eq. (1), that was previously proposed [20]. Modeling with a Gaussian Mixture Model using a maximum posterior probability inference, the mosaicing preprocessing solution [21] is found. Alternatively one may model the histogram as multivariate kernel density and retrieve the mean-shift oriented method [22].

Additional boundary constraints often lead to results that are not necessarily perceptually pleasant. Thus, we consider a different approach inspired from the human perception: the functions ν are taken so to select the color frameworks of the scene, the weights c_i allow even more overlapping between frameworks, while the linear parameters, $\mathbf{a}$, $\mathbf{b}$ are still inspired from the original approach of Reinhard et al. [15].

4.2. Color frameworks

Although many studies attempted to explain the human perception of complex scenes, no definite model exists. Yet, the reformulation by Gilchrist et al. [42] of the anchoring theory for complex scenes proved to pass many perceptual tests and explained many phenomena. This anchoring theory focuses on luminance interpretation and states that when depicting a scene, the relation between the representation luminance and the scene lightness can be correctly perceived only through a mapping between the luminance value and the value on the scale of perceived level, process called *anchoring*.

For increasingly complex scenes, the anchoring theory asserts that scenes are perceived by the humans in terms of consistent areas, named *frameworks*. A framework is defined as a region of common illumination [42]. For image perception, the human brain estimates the lightness within each framework through the anchoring to the luminance perceived as white, followed by the computation of the global lightness. While the framework theory was developed for luminance images, we assume the same strategy for color images. Intuitively color quantization assumes image organization in frameworks and the perception of quantized scene is appropriate. We consider that scene decomposition in frameworks and performing the transfer between matching frameworks to solve eq. (3) could lead to an interesting color transfer method.

The first computational model of the anchoring theory for complex images was provided by Krawczyk et al. [43] for rendering high dynamic range images. This paper follows the same guidelines, with the major difference that for extraction of frameworks instead of the mean-shift, we rely on a thresholded version of Fuzzy C-Means as they allow some image data to be in more than one framework.

We recall that for Fuzzy C-Means (FCM) [44], [45], the following objective function has to be minimized:

$$J_{FCM} = \sum_{k=1}^{P}\sum_{i=1}^{N} \nu_{ik}||\mathbf{x}_k - \mathbf{v}_i||_2, \quad s.t. \sum_{i=1}^{N} \nu_{ik} = 1, \forall k \tag{4}$$

where $\mathbf{x}_k$ are the n-dimensional image pixels (here $n = 5$), P is the total number of pixels, N is the total number of clusters and $\mathbf{v}_i$ are the centroids/means of the clusters. In the interpretation of the anchoring theory [42], $\mathbf{v}_i$ act as anchors. $\|\cdot\|_2$ is the $\mathcal{L}_2$ norm.

The solution $(\nu_{ik}, \mathbf{v}_i)$ is found iteratively once the number of clusters, N is chosen.

Yet, to increase the practical robustness of the standard FCM, two adaptations are used. First FCM has the known drawback of converging into local optima which leads to non-overlapping frameworks and to the visual failure of the color transfer method. In order to solve this aspect we used the previously proposed method based on simulated annealing [46]. Secondly, sometimes the FCM converges (truthfully) in unsatisfactory clusters. More precisely cases with large near-saturated areas (in the normally exposed image) or near-black ones (in the underexposed image) are separated on different clusters, while the rest of the pixels are in pushed in wide range clusters. Such cases are detected and cluster are merged back. An illustrative example of the last situation is presented in Fig. 3.

Intuitively, instead of direct minimization of eq. (2) the optimization is done sequentially, first determining $\nu_i(u) = \nu$ via FCM. In other words, the images are clustered on sets with compact color levels, which may be perceived as a color extension of the frameworks from the anchoring theory.

If only the images (represented in a tri-dimensional color space, such as RGB or CieLab) are considered, the two resulting frameworks from the segmented images are similar, but *not identical*, due to the differences between the initial images. An example can be seen in Fig. 1.

Let us denote the frameworks of the reference image by R_i and those of subject image by S_i. Let us assume that given N frameworks in both images, after the matching, the indexes are in increasing order such that S_i is paired with R_i, $i = 1 \ldots, N$. The mean square error (MSE) is given as:

$$\overline{d_i} = \frac{1}{N_{si}N_{ri}} \sum_{k=1}^{N_{si}}\sum_{p=1}^{N_{ri}} \|\mathbf{s}_k - \mathbf{r}_p\|_2 \tag{5}$$

where N_{si} is the number of locations described by vectorial values $\mathbf{s}_k$ in the i-th framework from the source image, while N_{ri} and $\mathbf{r}_p$ are their counterparts in the reference image. This can be developed using each cluster centroid (mean) as follows:

$$MSE = \overline{d_i} = \frac{1}{N_{si}N_{ri}} \sum_{k=1}^{N_{si}}\sum_{p=1}^{N_{RSi}} \|\mathbf{s}_k - \mathbf{v}_{Si} - \mathbf{r}_p + \mathbf{v}_{Ri} + \mathbf{v}_{Si} - \mathbf{v}_{Ri}\|_2 \tag{6}$$

Based on the observation that $\mathcal{L}_2$ norm follows the triangle inequality, the MSE has as upper bound $\overline{d_i} \le \overline{d_i}^{sup}$:

$$\begin{aligned}
\overline{d_i}^{sup} &= \frac{1}{N_{si}N_{ri}}\left(\sum_{k=1}^{N_{si}}\sum_{p=1}^{N_{ri}}\big(\|\mathbf{s}_k - \mathbf{v}_{Si}\|_2 + \|\mathbf{r}_p - \mathbf{v}_{ri}\|_2 \right.\\
&\quad \left. + \|\mathbf{v}_{Si} - \mathbf{v}_{ri}\|_2\big)\right)\\
&= \frac{1}{N_{si}N_{ri}}\left(\sum_{k=1}^{N_{si}}\sum_{p=1}^{N_{ri}}\|\mathbf{s}_k - \mathbf{v}_{si}\|_2 + \sum_{k=1}^{N_{si}}\sum_{p=1}^{N_{ri}}\|\mathbf{r}_p - \mathbf{v}_{ri}\|_2 \right.\\
&\quad \left. + \sum_{k=1}^{N_{si}}\sum_{p=1}^{N_{ri}}\|\mathbf{v}_{si} - \mathbf{v}_{Ri}\|_2\right)\\
&= \frac{1}{N_{si}N_{ri}}\left(N_{ri}\sum_{k=1}^{N_{si}}\|\mathbf{s}_k - \mathbf{v}_{si}\|_2 + N_{si}\sum_{p=1}^{N_{ri}}\|\mathbf{r}_p - \mathbf{v}_{ri}\|_2 \right.\\
&\quad \left. + N_{si}N_{ri}\|\mathbf{v}_{si} - \mathbf{v}_{Ri}\|_2\right)\\
&= \frac{1}{N_{si}}\sum_{k=1}^{N_{si}}\|\mathbf{s}_k - \mathbf{v}_{si}\|_2 + \frac{1}{N_{ri}}\sum_{p=1}^{N_{ri}}\|\mathbf{r}_p - \mathbf{v}_{ri}\|_2\\
&\quad + \|\mathbf{v}_{si} - \mathbf{v}_{Ri}\|_2
\end{aligned} \tag{7}$$

This can be summarized as:

$$\overline{d_i}^{sup} = \Psi_{si} + \Psi_{ri} + \|\mathbf{v}_{si} - \mathbf{v}_{Ri}\|_2 \tag{8}$$

where $\Psi_{si} = \frac{1}{N_{si}}\sum_{k=1}^{N_{si}}\|\mathbf{s}_k - \mathbf{v}_{si}\|_2$ and $\Psi_{ri} = \frac{1}{N_{ri}}\sum_{p=1}^{N_{ri}}\|\mathbf{r}_p - \mathbf{v}_{ri}\|_2$ are the variances over each axis for values in the i-th cluster/framework from the source image and the reference image.

The first comment with respect to eq. (8) is that the result is also intuitive. Given the linear nature of the transfer between pairing frameworks, as described by eq. (2), the highest stress/error is at the boundaries of each interval. Statistically, the amount of pixels at the boundaries is given by the variance inside the cluster. At the limit, if all the pixels have the same values and are equal with the mean, the transfer is perfect or errorless.

The second point of discussion is with respect to the methods for reducing such error. The two images have differently exposure values, thus the same object should be described by pixels having different values in the two images. The problem is to bring the values closer, in order to reduce errors.

One may assume a scenario where the centroids are close to each other, yet this means that one should have smaller difference between the exposure values. Since the normally exposed image is fixed at $EV = 0$, the change may be only to the low-light one. Reducing the negative amplitude of the EV for the low light image, it does contradicts the main idea of the solution, which is to use the lowest exposure value possible, so to ensure the smallest hand shake and thus the smallest motion blur.

The alternative is to use a scenario where the variances inside clusters are smaller. A choice is to increase the number of clusters, yet this makes the matching harder, as more alternatives can exist for each pair. The proposed solution and the major improvement with respect our previous work [12] is to use channels (description) that have *smaller variance*.

A topological interpretation of eq. (8) may be retrieved stating from the observation of Domingos [47]: "most of the volume of a high-dimensional orange is in the skin, not the pulp"; a rigorous discussion on the topic may be found in the work of Aggarwal et al. [48]. Eq. (8) contains three "oranges": one with fixed/given size, $\|\mathbf{v}_{si} - \mathbf{v}_{Ri}\|_2$ and two adjustable, Ψ_{si} and Ψ_{ri}. The proposed approach is to reduce the volume of the two adjustable "oranges" by bringing their "skin" closer to the center.

4.3. Extreme channels

Beside the theoretical formulation from eq. (8) there is another intuitive approach to the proposed development. In our previous work [12], we noted that one main reason for failure was the occasional imperfect matching between the frameworks of the subject image and the ones of the reference image. This is related to their different nature: the low-light image is degraded by noise, while the normally exposed may be degraded by motion blur. The effects of the degradations appear more evident on object with a slowly varying color, as, for instance, the small motion blur may lead to creating false boundaries. In such cases, the segmentation may, incorrectly, broke an object into multiple segments. In the current proposal, it is important to use channels that offer more stability with respect to the degradations.

One such solution draws inspiration from the concept of dark channel introduces by He et al. [49]. Instead of a single dark channel, two are utilized: a dark and a bright channel. We recall that for an image I, these are defined as follows:

$$\begin{aligned} I_{dark}(\mathbf{x}) &= \min_{c\in\{r,g,b\}}\left(\min_{\mathbf{y}\in\Omega(\mathbf{x})} I^c(\mathbf{y})\right) \\ I_{bright}(\mathbf{x}) &= \max_{c\in\{r,g,b\}}\left(\max_{\mathbf{y}\in\Omega(\mathbf{x})} I^c(\mathbf{y})\right) \end{aligned} \tag{9}$$

where I^c is a color channel of I and $\Omega(\mathbf{x})$ is a local patch centered in current pixel $\mathbf{x}$. In this work the small patch contains 3×3 neighbors.

The expected effect is to have a reduced variance, which is achieved given the use of min or max, and to provide more stable values for pixels inside a framework. Statistically, the extreme channels have a variance reduced with at least $\frac{1}{3}$ over any original color channel given any large enough image patch.

5. Implementation

We consider the two input images (normally exposed and underexposed). The color transfer implementation follows the procedure:

- *Extreme channels*: Given the two input images in the RGB color space compute the additional 2 extreme channels for each image.
- *Color space*: The input images are transformed from RGB space into the CieLab color space.
- *Pixel description*: Given the two extreme channels, the pixel at one location will be described by a 5-dimensional tuple: L, a, b triple followed by the dark and the bright resulting values.
- *Frameworking*: Determine the frameworks on each of the two images by applying FCM, separately, on both of them. Taking into account that the images contain almost the same scene, for speed-up purposes, we compute the clustering on one image and we use its result to initiate the FCM algorithm on the second image (by using the positions of the pixels on the image, not the colors itself). This way we also diminish the probability that the FCM clustering converges in different local minima. Hard threshold the membership weights, ν so that to select only one framework for each location.
- *Matching*: Match the frameworks of the low-light image with the frameworks of the normally exposed one. The reference image framework, R_k matching S_i is found as:

  $$k = \arg\max_j S_i \cap R_j \quad . \tag{10}$$

  The eq. (10) comes from the fact that the two images contain almost the same scene (i.e. mis-alignment is small), thus we search for maximal spatial overlapping. The match is found by comparing all possible combinations.
- *Get statistics*: For each framework, either in subject image, S_i, or in the reference image, R_i, compute the mean (μ_i^s and respectively, μ_i^r) and the standard deviations (σ_i^s, σ_i^r). The computation is only on the three color axes, as they will be transferred.
- *Framework transfer*: For each pair of frameworks, compute a transfer function using eq. (1). If one denotes by $\theta = \{S_i, R_i, \nu_i\}$, $i = 1\dots N$ as the model of the frameworking process, the conditional probabilities $p_{ij}(R_j/\theta)$ of having pixels in the framework R_j that originate in the framework S_i are computed.

  Next, one computes the linear parameters, a_j, b_j, of a subject pixel considered to be in the framework S_i by:

  $$a_j = \frac{\sigma_j^r}{\sigma_i^s}; \;\; b_j = \mu_j^r - a_j \cdot \mu_i^s \tag{11}$$
- *Global transfer*: Compute the image transfer using eq. (2), where the c_i are the framework confusion conditional probabilities: $c_i = p_{ij}(R_j/\theta)$.

We note that while FCM considers 5-dimensional input data, the rest of algorithm is implemented on each color plane (L, a, b) separately. At the end of the transfer procedure, the resulting image is converted back to the original color space (RGB) for storing.

The model of transfer implemented in Eq. (1) assumes that in a matching pair of frameworks, the color gamut can be reliable modeled with a single Gaussian and thus only the mean and the variance are needed. While the results reported in the next section show the capabilities of this rather simple model, one might seek a more elaborated model. Into this direction, the methods based on optimal transport [23,24], are the complete and optimal model. However, in practical color transfer, as noted in the mentioned work [23,24], the optimal transfer needed to be relaxed, and a coarse model is preferred. The full model which uses a bijection is not expected to work due to different color gamuts. Yet, in certain scenarios, a more elaborate model than a Gaussian mode may lead to more pleasant images. Also one might seek improvement with respect to local vicinity, as it was used for instance in the LECARM algorithm [29].

Low and high resolution Noting that small content differences may exist due to camera motion between acquisitions, spatial matching cannot be perfect. To accelerate the overall process and to reduce the impact of mis-alignment, the FCM runs on images with reduced resolution. We have chosen the width of 640 and the original aspect ratio. On the small resolution images, the framework means and variances necessary for eqs. (1), (11) are found, while the weights ν_{ik} are computed on the full resolution image to ensure smooth transitions.

6. Results and discussions

6.1. Database

To test the proposed algorithm we collected a specific database using three cameras: a professional one (digital SLR), a consumer one and a smartphone. We have considered two types of differences between the two images forming a set: while the reference image is normally (well) exposed, the low-light images are underexposed with either $EV = -1$ or $EV = -2$ (i.e. exposure time is half and respectively a quarter from normal). The images were acquired with hand-held camera, thus they are not perfectly aligned and the normal exposed image is often blurred.

The photographed scene ranged from indoor, with and without people, to outdoor images. Outdoor cases contain both landscapes (during sunset, to have a lower light), where the focus is on far objects, and scenes where the focus is closer, also in dimmer light.

In total more than 100 pairs of underexposed images with $EV = -1$ and 100 pairs of underexposed images with $EV = -2$ have been gathered.[2] For each of these pairs we also acquired a normal exposed image without blur by placing the camera on a tripod. We will use this image as a reference for the objective evaluation metrics. We note that this image is not perfectly aligned with the hand-held ones either.

[2] Database is available at http://imag.pub.ro/steadycam/steadycam_download.

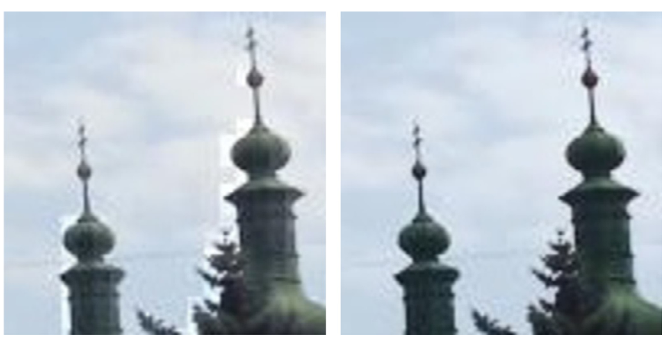
6.2. Evaluation metrics

As mentioned, the proposed color transfer algorithm takes as input two images: a blurred normally exposed one and a sharp underexposed one. These two images are acquired with a hand-held camera and are not perfectly aligned. The result of the algorithm should be a normally exposed, sharp image.

To evaluate the correctness of the color transfer method, we compare its result with the reference, normally exposed image, acquired with a tripod placed camera. We will call this image *the evaluation reference image* to distinguish it from the color reference image, which may be blurred.

For evaluation purposes, all the resulting corrected images are compared with the evaluation reference image and peak signal-to-noise-ratio (PSNR) and structural similarity – SSIM [50] between the two images are computed. These two measures are typically used to assess the accuracy of reproduction for color transfer methods. We, again, note that the evaluation reference image is not perfectly aligned with the ones taken with the hand-held camera, thus is not perfectly aligned with the image resulting from the color transfer method. Additionally as the resulting image appeals to perceptual pleasantness under an exposure modification scenario we also use the "Statistical Naturalness" component from Tone Mapped Image Quality Index (TMQI) metric [51]. The other component, "Structural Fidelity" is a non-linear version of SSIM constructed for comparing radiance maps (HDR images) with final displayed images (LDR); the non-linearity is needed to cope with range compression (which is not present in the current scenario). In contrast "Statistical Naturalness", denoted by SN, is a non-reference metric that evaluates, based on perceptual experiments, how pleasant a final image is.

Compared to our previous work [12], for a more accurate evaluation we manually aligned the evaluation reference image with the one resulted after the color transfer method was applied. Such a step is necessary as both quality measures report inconclusive values in cases of unregistered pairs.



(a) (b)

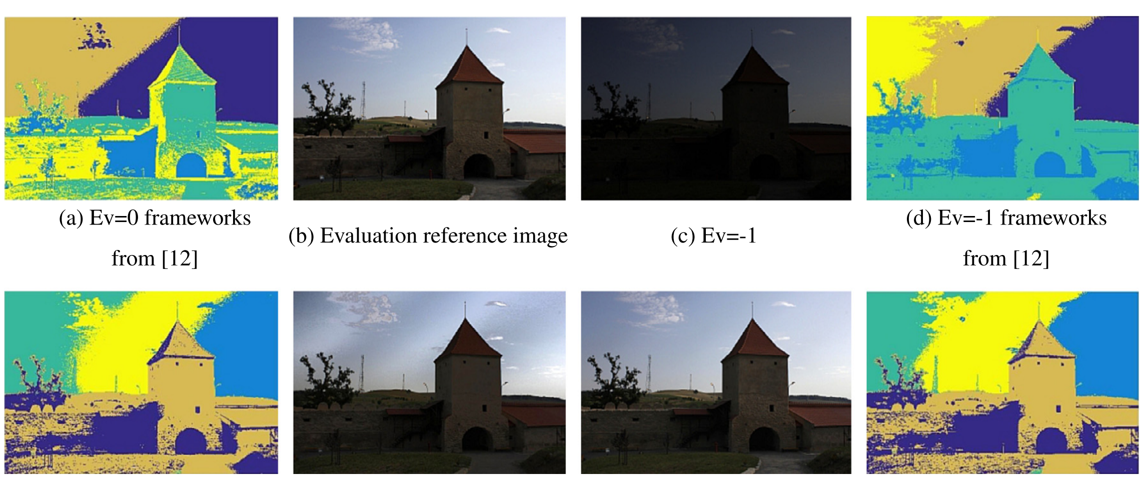
Fig. 2. Artifacts at transition may appear if one computes the weights for the FCM clustering at small resolution (a), compared to computing them at full resolution in figure (b).

6.3. Results

FCM resolution The first encountered problem was due to the time required by the clustering algorithm to run on a high resolution image. In order to make this time acceptable, one reduces the resolution of the images during clustering, which leads to another problem: it introduced visible artifacts at the transition between frameworks. For smaller images, the transitions from one framework to others are more noticeable, thus producing disturbing artifacts. These transition artifacts appear mainly in the regions that are over-segmented by the FCM algorithm. However, computing the pixels weight at full resolution avoided this downside. An example can be seen in Fig. 2.

The usage of lower resolution allowed an average speed-up of 4× on our database. We note that the speed-up depends on the content of the image, on the image resolution and on the choice of the initial centroids for the FCM. The database contains images with resolution varying from 18 MPixels (DSLR camera) to 5 MPixels (smartphone camera) and with very different contents (some very colorful, others with few colors), thus the speed-up for each image can vary.

Framework merging and extreme channels If only the CieLab color channels are used for clustering, the clustering algorithm may produce, at times, an over-segmentation, by artificially splitting near-saturated areas or almost black ones. The initial proposed solution [12] inspects such frameworks and, at necessity (i.e. framework's means are too close) merges them. However this solution is not always working. By using the extreme channels as additional data for the clustering algorithm, these kinds of problems appear less often. In this case, the clustering results are more similar between



(a) Ev=0 frameworks from [12] (b) Evaluation reference image (c) Ev=-1 (d) Ev=-1 frameworks from [12]

(e) Ev=0 frameworks using extreme channels (f) Result from [12] (g) Proposed method (h) Ev=-1 frameworks using extreme channels

Fig. 3. An example where the use of extreme channels is beneficial. (b) Evaluation reference image ($Ev = 0$) and (c) Underexposed image ($Ev = -1$). (a)–(d) The frameworks from [12]. (e)–(h) The frameworks obtained using the additional information given by the extreme channels. (f) Result with the method from [12]. Note the artifacts on the sky and the color/texture of the tower. (f) Proposed method.

Table 1
Achieved performance of the proposed method with respect to camera used. SN denotes "Statistical Naturalness" and larger values are better.

Camera	PSNR			SSIM			SN		
	$EV = -1$	$EV = -2$	All	$EV = -1$	$EV = -2$	All	$EV = -1$	$EV = -2$	All
Smartphone	22.96	22.35	22.65	0.72	0.69	0.70	0.59	0.47	0.53
Consumer	21.73	21.22	21.47	0.75	0.71	0.72	0.57	0.46	0.52
DSLR	22.80	21.68	22.26	0.76	0.73	0.75	0.38	0.34	0.36

Table 2
Numerical comparison between the proposed method and prior related methods.

Method	PSNR			SSIM			SN		
	$EV = -1$	$EV = -2$	All	$EV = -1$	$EV = -2$	All	$EV = -1$	$EV = -2$	All
Proposed	22.5	21.75	22.13	**0.74**	**0.71**	**0.72**	**0.51**	0.42	**0.47**
Florea et al. [12]	20.54	18.05	19.29	0.58	0.56	0.57	0.44	**0.43**	0.44
Reinhard et al. [15]	18.12	16.75	17.44	0.60	0.58	0.59	0.40	0.25	0.33
Pitie et al. [17]	19.23	18.34	18.79	0.62	0.60	0.61	0.50	0.30	0.40
Pouli et al. [18]	18.66	17.52	17.10	0.62	0.59	0.61	0.44	0.26	0.35
Mean Shift	19.65	18.35	19.00	0.61	0.60	0.60	0.44	0.30	0.37
Mechrez et al. [19]	**22.99**	**22.05**	**22.6**	0.69	0.65	0.66	0.46	0.41	0.44
Ren et al. [29]	17.57	15.49	16.53	0.64	0.58	0.61	0.44	0.27	0.36

(a) The evaluation reference image (b) Proposed (c) Florea et al. [12]
(d) Under-exposed image (e) Pitie et al. [17] (f) Pouli et al. [18]

Fig. 4. Comparison between proposed method and classical image transfer methods. One may note that the proposed method is closest to the evaluation reference image.

the underexposed and the normal exposed images. An illustrative example is presented in Fig. 3 (f), where the lack of merging causes visible artifacts in the center of the sky.

Camera related performance In Table 1, the achieved performance with respect to the camera used is reported. As discussed in section 3, the PSF size (thus the amount of blur) and the pixel size are closely related. The quality of acquired images is increasing from the smartphone (which has 1.12 μm pixel size), to the consumer camera (with 1.76 μm pixel size) and to the DSLR (with 4.99 μm pixel size). SSIM numerical values indicate that the image quality retrieved using the proposed color transfer method is in accordance with the input image quality. The Statistical Naturalness says how pleasant an image is overall and it depends on camera color tuning. This metric, yet shows that the larger an amplification the least pleasant an image is.

Comparison with related work We extensively compare the proposed method with related work on color transfer, [15], [18] and [17], as the authors provide code, and to our previous work [12]. We also compare with the recent work on photorealist style transfer [19] and camera dependent low-light enhancement (LECARM) [29] using authors provided code. In the case of the latter, we sought, from the camera models provided, the one which lead to best results.

We have also replaced the FCM with mean-shift clustering since mean-shift is a method that usually performs well on natural images and does not require the user to specify the number of clusters, but the results were not as good as using the FCM.

Numerical results are shown in Table 2, while visual, comparative, results are presented in Figs. 4 and 5. We stress that the proposed method is tested on a significantly larger database than other similar solutions: in many cases, [17], [20], [22], etc. at most 15 images are used; we test on over 200 image sets. Yet, although on particular examples other methods may produce results leading to higher numerical values, overall, and on each category, the proposed method reaches the top performance.

From a subjective point of view there are further observations to be made. Artifacts of the proposed method are rarer and usually less disturbing than those of other solutions. Typical artifacts are related to slightly incorrect colors due to under-segmentation

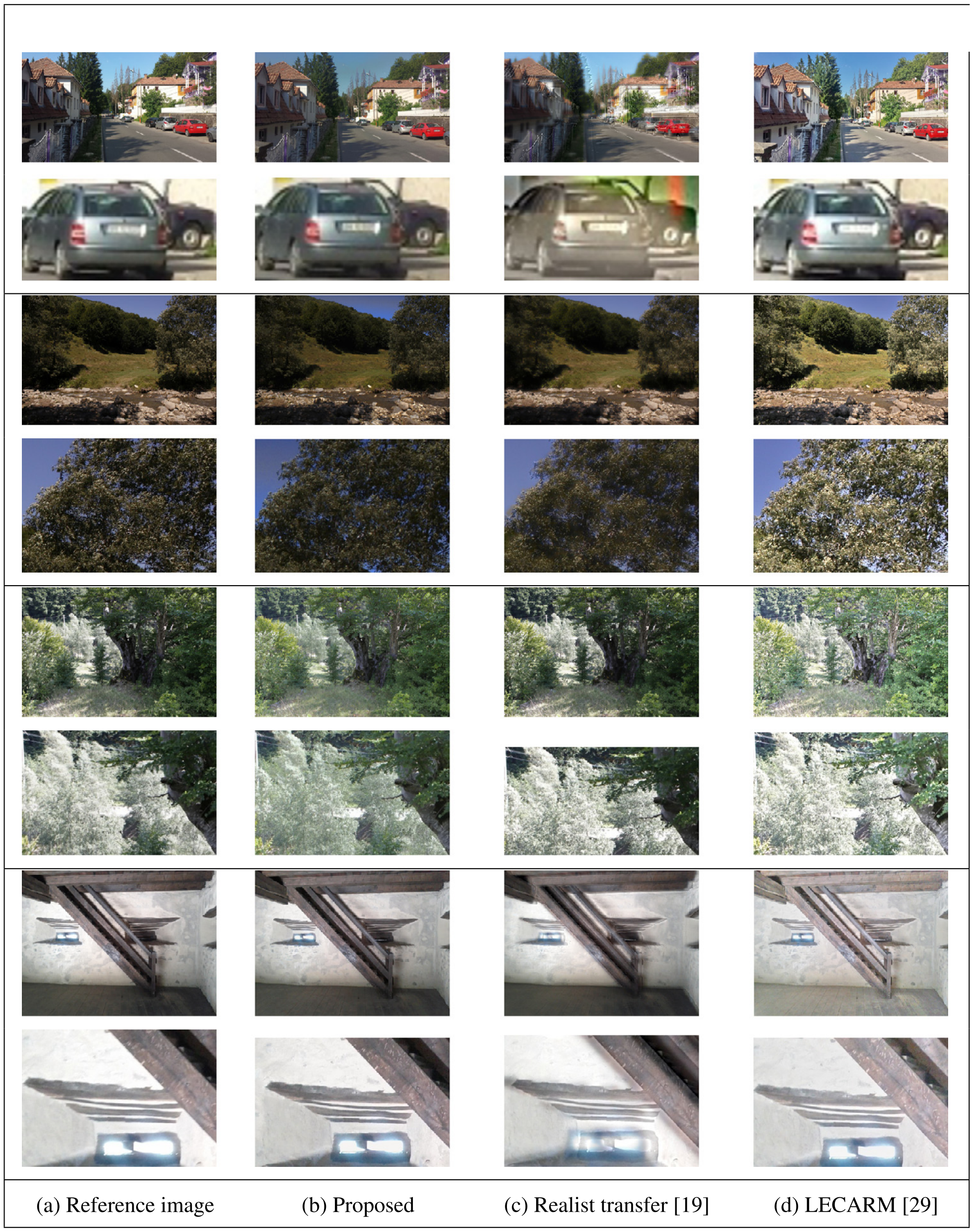


Fig. 5. Comparison between the proposed method and other alternative solutions to the problem. Original images in odd rows and details in even ones. The images are best viewed when zooming in the electronic version of the paper. One may note that the proposed method is closest to the evaluation reference image. The images obtained with realistic style transfer suffer from artifacts due to mis-alignment. The images obtained with LECARM, while independently may look fine, are oversharpened (as showed in the zoom from the rows 2 and 8) and overcontrasted (as showed in rows 4 and 6).

and some visible transition due to large displacements between frames. By contrast, in the initial color transfer algorithm [15], there are not any transition artifacts since the image is considered as a whole. However global transfer leads to much poorer colors in smaller regions, thus explaining the lowest reported results from Table 2.

Fig. 4 contains an outdoor image where all discussed methods performed reasonably well. The current method gives the result which is closest to the evaluation reference image. The sky and the buildings are wrongly colored by the algorithm proposed by Pouli et al. [18], also. The method by Pitie et al. [17] exhibits some artifacts on the closest building. Our previous solution [12] incorrectly merged the clouds with the buildings at the clustering step, which lead to more yellow clouds and whiter buildings. For more subjective comparisons between our previous method and older ones we refer the reader to the results section from [12].

Fig. 5 contains a series of comparisons with more recent prior works [19], [29]. Overall the so-called photorealist style transfer [19] assumes that the two images are perfectly aligned, and if this condition is not met the results are less impressive and occasionally with artifacts.

Since the currently proposed method is developed starting from our previously proposed one [12], in Fig. 6 comparative results can be seen. For a better view we selected a region from each image and we enlarged it in Figs. 6 (d), (e) and (f). One may notice that the current method seems to give sharper results. This is mostly due to the resolution used for the clustering and to the semi-random initialization of the FCM. By ensuring that the weights are computed on full resolution we do not artificially blur the segmented image. The better matching of the clusters due to the use of the extreme channels results in better overall colors for the reconstructed image.

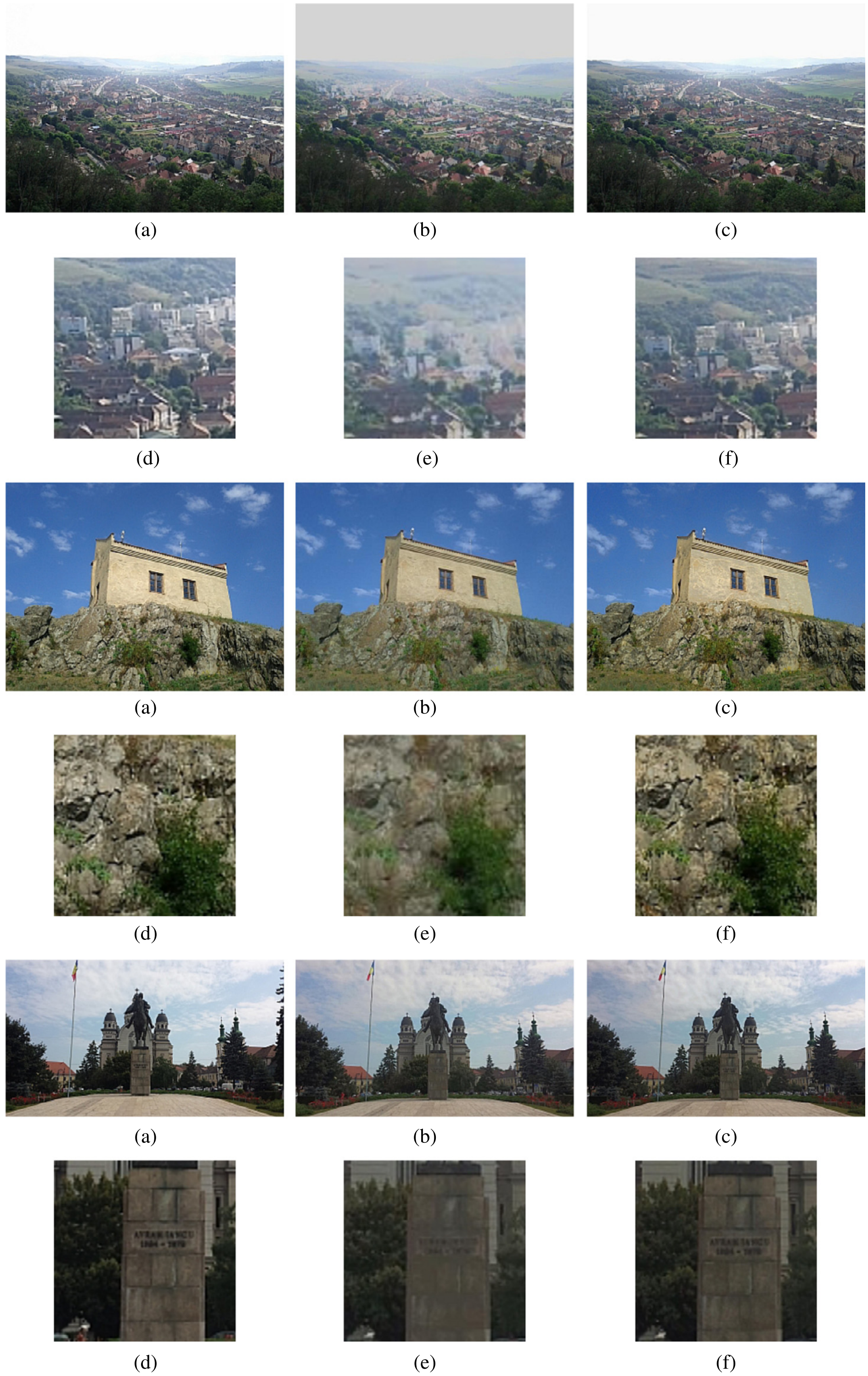
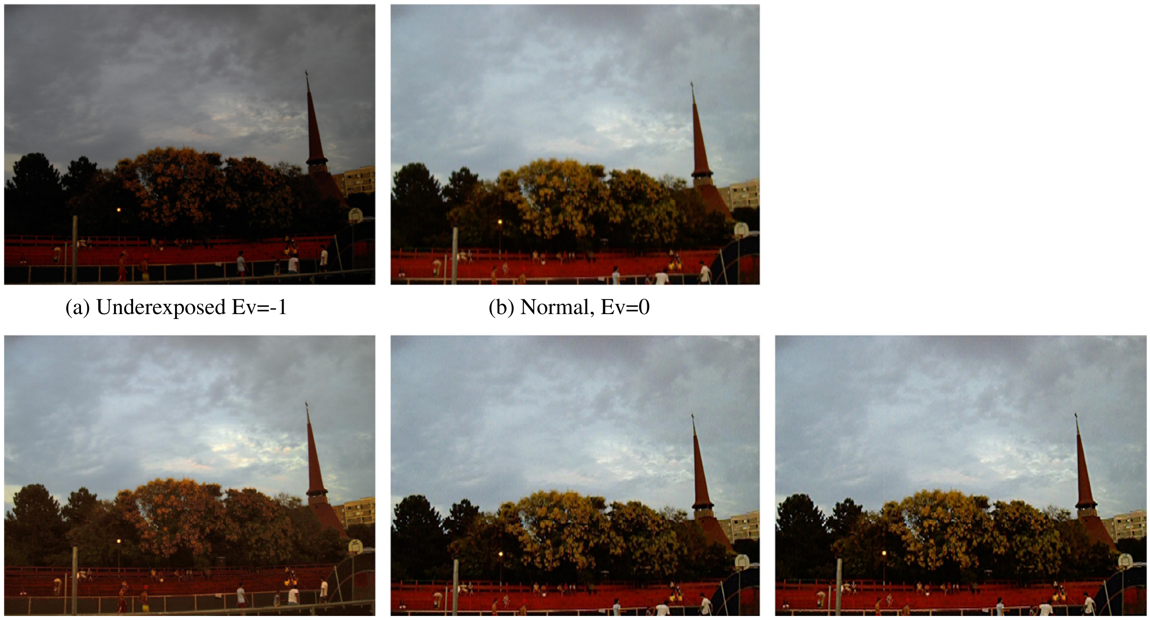


Fig. 6. Comparison between proposed method and our initially proposed one [12]. (a) Evaluation reference image; (b) [12]; (c) Proposed method; (d), (e), (f) Cropped regions from above examples.

Comparison with blind deconvolution In the previous sections, we have claimed that the proposed method is a practical alternative to the blind deconvolution. To support this claim in Fig. 7 the proposed approach is illustrated on a case where the normally exposed image is visibly blurred. For deconvolution we show the results from two state of the art solutions: the blind patch recurrence solution [52] and the sparse blind regularization for blind deconvolution [53].

Blind deconvolution introduces visible artifacts by reducing the actual resolution. The blurring of the image means that some of the image high frequencies become zero (by multiplying them with the zeros of the PSF). In order to restore the image via deconvolution, these frequencies need to be deducted from assumptions on the image. These assumptions usually lead to losing the image details, thus the result of deconvolution solutions looks patchy. In contrast, using color transfer, these details are preserved better as no frequency canceling takes place.



(a) Underexposed Ev=-1 (b) Normal, Ev=0

(c) Proposed (d) Blind deblurred [52] (d) Blind deblurred [53]

Fig. 7. Comparison between proposed method and a blind deconvolution method [52]. The best view for comparison is when zooming in on the electronic version. One may note that the deconvolution introduces noticeable artifacts.

Table 3
Comparative duration (in minutes) of evaluated method. All times have been obtained using author release Matlab code on the same CPU (Intel i7 3.0 GHz), single thread while processing an image with 18 Mpixels.

Method	Proposed	Pitie et al. [17]	Pouli et al. [18]	Mechrez et al. [19]	Ren et al. [29]
Duration [min]	1.5	1.9	8.12	145	0.25

Another argument for using color transfer as opposed to blind deconvolution is its higher computational efficiency. The time required for blind deconvolution algorithms is quite large. For the image in Fig. 7, with a resolution of 0.7 MPixels, the blind patch recurrence solution [52] needed 20 minutes (with 10% in PSF estimation and 90% in the actual deconvolution), while the sparse blind regularization [53] required 25 minutes. By comparison, for the same image, our method takes 6 seconds.

6.4. Duration

The proposed solution is implemented in Matlab. On an Intel i7 3.0 GHz, running on a single core, it requires 1.5 minutes to enhance an image of 18 Mpixels (the resolution of DSLR images). Comparative durations are presented in Table 3. Our method is average as duration given prior art, which has as extremes the recent methods [19] – 2.3 hours and [29] – 15 seconds for the same platform and images. The conclusion is confirmed by the details of the algorithm too: the method of Ren et al. [29] does not perform adaptation with respect to the local scene (thus has reduced complexity), while the solution of Mechrez et al. [19] is concerned with local adaptation and, thus, yields intensive computation. The general transfer methods, including ours, perform only a coarse adaptation and consequently have an average complexity.

7. Conclusions

This paper proposes a method that addresses the potential motion blur arising in images acquired in low light by underexposing and color transfer. The method implements a piece-wise transfer based on decomposing the reference and source image on frameworks. The frameworks consistency is increased by the use of extreme channels as it has been proven that due to having smaller variance reduce the gap between the two images. Also we argue that, while facing potential motion blur is more efficient to underexpose images and perform color transfer for low-light compensation than implement blur deconvolution.

On the theoretical side, we contributed by the introduction of a generative model for color transfer and we show that many previously introduced methods may be retrieved as particular cases of it. At last we have introduced a color transfer method that is shown to outperform related methods on a substantially large image database.

Subjective evaluations show that images without visible quality degradation are computed while underexposing with 2 EV stops (i.e. taking a quarter from the exposure time required by the scene nominal illumination). The algorithm is subject to full optimization and may be implemented inside the camera.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors wish to thank Ciprian Ionascu for his contribution in the initial phases of the research.

References

[1] A. Levin, Y. Weiss, F. Durand, W. Freeman, Understanding blind deconvolution algorithms, IEEE Trans. Pattern Anal. Mach. Intell. 33 (12) (2011) 2354–2367.
[2] P. Ruiz, X. Zhou, J. Mateos, R. Molina, A.K. Katsaggelos, Variational bayesian blind image deconvolution: a review, Digit. Signal Process. 47 (2015) 116–127.

[3] J. Dong, J. Pan, Z. Su, Blur kernel estimation via salient edges and low rank prior for blind image deblurring, Signal Process. Image Commun. 58 (2017) 134–145.
[4] N. Joshi, S.B. Kang, C. Lawrence-Zitnick, R. Szeliski, Image deblurring using inertial measurement sensors, ACM Trans. Graph. 29 (4) (2010) 30.
[5] S.P.N. Singhy, C.N. Riviere, Physiological tremor amplitude during retinal microsurgery, in: Proc. of the IEEE Bioengineering Conference, 2002, pp. 171–172.
[6] O. Whyte, J. Sivic, A. Zisserman, J. Ponce, Non-uniform deblurring for shaken images, Int. J. Comput. Vis. 98 (2) (2012) 168–186.
[7] J. Pan, R. Liu, Z. Su, X. Gu, Kernel estimation from salient structure for robust motion deblurring, Signal Process. Image Commun. 28 (9) (2013) 1156–1170.
[8] J. Sun, W. Cao, Z. Xu, J. Ponce, Learning a convolutional neural network for non-uniform motion blur removal, in: Proc. of IEEE Computer Vision and Pattern Recognition, 2015, pp. 769–777.
[9] F. Xiao, J. Pincenti, J. Farrell, Camera motion and effective spatial resolution, in: Proc. Int. Congress of Imaging Science, Vol. I, 2006, pp. 33–36.
[10] A. Drimbarean, A. Zamfir, F. Albu, V. Poenaru, C. Florea, P. Bigioi, E. Steinberg, P. Corcoran, Low-light video frame enhancement, US Patent 8199222, 2007.
[11] L. Yuan, J. Sun, L. Quan, H. Shum, Blurred/non-blurred image alignment using sparseness prior, in: Proc. International Conference on Computer Vision, 2007, pp. 1–8.
[12] L. Florea, C. Florea, C. Ionascu, Avoiding the deconvolution: framework oriented color transfer for enhancing low-light images, in: Computer Vision and Pattern Recognition Workshops, 2016, pp. 936–944.
[13] H.S. Faridul, T. Pouli, C. Chamaret, J. Stauder, E. Reinhard, D. Kuzovkin, A. Trémeau, Colour mapping: a review of recent methods, extensions and applications, Comput. Graph. Forum 35 (1) (2016) 59–88.
[14] G.D. Finlayson, H. Gong, R. Fisher, Color homography: theory and applications, IEEE Trans. Pattern Anal. Mach. Intell. (2017) 1–14.
[15] E. Reinhard, M. Ashikhmin, B. Gooch, P. Shirley, Color transfer between images, IEEE Comput. Graph. Appl. 21 (5) (2001) 34–41.
[16] D. Ruderman, T. Cronin, C. Chiao, Statistics of cone responses to natural images: implications for visual coding, J. Opt. Soc. Am. A 15 (8) (1998) 2036–2045.
[17] F. Pitie, A. Kokaram, R. Dahyot, Automated colour grading using colour distribution transfer, Comput. Vis. Image Underst. 107 (2007) 123–137.
[18] T. Pouli, E. Reinhard, Progressive color transfer for images of arbitrary dynamic range, Comput. Graph. 35 (4) (2011) 67–80.
[19] R. Mechrez, E. Shechtman, L. Zelnik-Manor, Photorealistic style transfer with screened Poisson equation, in: BMVC, 2017.
[20] S. Kagarlitsky, Y. Moses, Y. Hel-Or, Piecewise-consistent color mappings of images acquired under various conditions, in: Proc. of International Conference on Computer Vision, 2009, pp. 2311–2318.
[21] M. Oliveira, A.D. Sappa, V. Santos, Unsupervised local color correction for coarsely registered images, in: Proc. of Computer Vision and Pattern Recognition, 2011, pp. 201–208.
[22] M. Oliveira, A.D. Sappa, V. Santos, A probabilistic approach for color correction in image mosaicking applications, IEEE Trans. Image Process. 24 (2) (2015) 508–523.
[23] S. Ferradans, N. Papadakis, G. Peyré, J.-F. Aujol, Regularized discrete optimal transport, SIAM J. Imaging Sci. 7 (3) (2014) 1853–1882.
[24] R. Giraud, V.-T. Ta, N. Papadakis, Superpixel-based color transfer, in: 2017 IEEE International Conference on Image Processing, ICIP, IEEE, 2017, pp. 700–704.
[25] K. Fotiadou, G. Tsagkatakis, P. Tsakalides, Low light image enhancement via sparse representations, in: Proc. of Int. Conf. on Image Analysis and Recognition, in: LNCS, vol. 8814, 2013, pp. 84–93.
[26] K.G. Lore, A. Akintayo, S. Sarkar, LLNet: a deep autoencoder approach to natural low-light image enhancement, Pattern Recognit. 61 (2017) 650–662.
[27] S. Ko, S. Yu, S. Park, B. Moon, W. Kang, J. Paik, Variational framework for low-light image enhancement using optimal transmission map and combined $\mathcal{L}_1$ and $\mathcal{L}_2$-minimization, Signal Process. Image Commun. 58 (2017) 99–110.
[28] R. Zhang, X. Feng, L. Yang, L. Chang, C. Xu, Global sparse gradient guided variational retinex model for image enhancement, Signal Process. Image Commun. 58 (2017) 270–281.
[29] Y. Ren, Z. Ying, T.H. Li, G. Li, LECARM: low-light image enhancement using the camera response model, IEEE Trans. Circuits Syst. Video Technol. 29 (4) (2019) 968–981.
[30] X. Fu, D. Zeng, Yue Huang, Y. Liao, X. Ding, J. Paisley, A fusion-based enhancing method for weakly illuminated images, Signal Process. 129 (2016) 82–96.
[31] C. Jung, Q. Yang, T. Sun, Q. Fu, H. Song, Low light image enhancement with dual-tree complex wavelet transform, J. Vis. Commun. Image Represent. 42 (2017) 28–36.
[32] P. Debevec, J. Malik, Recovering high dynamic range radiance maps from photographs, in: ACM SIGGRAPH, 1997, pp. 369–378.
[33] T. Mertens, J. Kautz, F.V. Reeth, Exposure fusion, in: Proceedings of Pacific Graphics, 2007, pp. 382–390.
[34] P. Crawford, E. Zimmerman, Differentiation and diagnosis of tremor, Am. Fam. Phys. 83 (6) (2011) 697.
[35] K. Veluvolu, W.T. Ang, Estimation and filtering of physiological tremor for real-time compensation in surgical robotics applications, Int. J. Med. Robot. Comput. Assist. Surg. 6 (2010) 334–342.
[36] G.P.R. Papini, M. Fontana, M. Bergamasco, Desktop haptic interface for simulation of hand-tremor, IEEE Trans. Haptics 9 (1) (2016) 33–42.
[37] A. Gupta, N. Joshi, L. Zitnick, M. Cohen, B. Curless, Single image deblurring using motion density functions, in: Proc. of European Conference on Computer Vision, 2010, pp. 171–184.
[38] M. Hirsch, C.J. Schuler, S. Harmeling, B. Scholkopf, Fast removal of non-uniform camera shake, in: Proc. of International Conference on Computer Vision, 2011, pp. 463–470.
[39] C. Schuler, M. Hirsch, S. Harmeling, B. Scholkopf, Learning to deblur, IEEE Trans. Pattern Anal. Mach. Intell. 38 (7) (2016) 1439–1451.
[40] D. Zoran, Y. Weiss, From learning models of natural image patches to whole image restoration, in: Proc International Conference on Computer Vision, 2011, pp. 479–486.
[41] E. Bilissi, S. Triantaphillidou, E. Allen, Exposure and image control, in: The Manual of Photography, 10th edition, Focal Press, Oxford, 2011, pp. 227–243.
[42] A. Gilchrist, C. Kossyfidis, F. Bonato, T. Agostini, J. Cataliotti, X. Li, B. Spehar, V. Annan, E. Economou, An anchoring theory of lightness perception, Psychol. Rev. 106 (4) (1999) 795–834.
[43] G. Krawczyk, K. Myszkowski, H.-P. Seidel, Lightness perception in tone reproduction for high dynamic range images, in: Computer Graphics Forum, vol. 24, 2005, pp. 635–645.
[44] J.C. Dunn, A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters, J. Cybern. 3 (1973) 32–57.
[45] J.C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, New York, 1981.
[46] P. Liu, L. Duan, X. Chi, Z. Zhu, An improved fuzzy c-means clustering algorithm based on simulated annealing, in: Proc. of Fuzzy Systems and Knowledge Discovery, 2013, pp. 39–43.
[47] P. Domingos, A few useful things to know about machine learning, Commun. ACM 55 (10) (2012) 78–87.
[48] C.C. Aggarwal, A. Hinneburg, D.A. Keim, On the surprising behavior of distance metrics in high dimensional space, in: International Conference on Database Theory, 2001, pp. 420–434.
[49] K. He, J. Sun, X. Tang, Single image haze removal using dark channel prior, IEEE Trans. Pattern Anal. Mach. Intell. 33 (12) (2011) 2341–2353.
[50] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612.
[51] H. Yeganeh, Z. Wang, Objective quality assessment of tone mapped images, IEEE Trans. Image Process. 22 (2) (2013) 657–667.
[52] T. Michaeli, M. Irani, Blind deblurring using internal patch recurrence, in: Proc. of European Conference on Computer Vision, 2014, pp. 783–798.
[53] H. Zhang, D. Wipf, Y. Zhang, Multi-image blind deblurring using a coupled adaptive sparse prior, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2013, pp. 1051–1058.

Laura Florea received her Ph.D in 2009 and M.Sc in 2004 from University "Politehnica" of Bucharest, Romania. From 2004 she teaches classes in the same university, where she is currently an Associate Professor. Her interests include computational photography, automatic understanding of human behavior by analysis of portrait images, image processing algorithms for digital still cameras, medical image processing and statistic signal processing theory. She is author of more than 50 peer-reviewed journal and conference papers.

Corneliu Florea was born in 1980 in Bucharest. He got his master degree from University "Politehnica" of Bucharest, Romania in 2004 and the Ph.D from the same university in 2009. There he is an Associate Professor and teaches classes on statistical signal and image processing, computer vision and machine learning and has introductory courses in computational photography. His research interests include machine leaning and computer vision methods for portrait understanding and non-linear image processing algorithms for digital still cameras. He is author of more than 50 peer-reviewed papers and of more than 20 US patents and patent applications.