

Low Light Image Enhancement via Sparse Representations

Konstantina Fotiadou *, Grigorios Tsagkatakis, and Panagiotis Tsakalides

Institute of Computer Science, Foundation for Research and Technology - Hellas
(FORTH-ICS)

Department of Computer Science, University of Crete,
Heraklion, Crete, Greece

{kfot,greg,tsakalid}@ics.forth.gr

Abstract. Enhancing the quality of low light images is a critical processing function both from an aesthetics and an information extraction point of view. This work proposes a novel approach for enhancing images captured under low illumination conditions based on the mathematical framework of Sparse Representations. In our model, we utilize the sparse representation of low light image patches in an appropriate dictionary to approximate the corresponding day-time images. We consider two dictionaries; a *night dictionary* for low light conditions and a *day dictionary* for well illuminated conditions. To approximate the generation of low and high illumination image pairs, we generated the day dictionary from patches taken from well exposed images, while the night dictionary is created by extracting appropriate features from under exposed image patches. Experimental results suggest that the proposed scheme is able to accurately estimate a well illuminated image given a low-illumination version. The effectiveness of our system is evaluated by comparisons against ground truth images while compared to other methods for image night context enhancement, our system achieves better results both quantitatively as well as qualitatively.

Keywords: De-nighting, Contrast Enhancement, Sparse Representations

1 Introduction

Recently, the demand for the enhancement of low light images has grown tremendously. Images captured during day-time exhibit higher dynamic range, better quality and can be useful for extracting contextual information. Night-time images on the other hand, are characterized by low intensities and usually suffer from the existence of severe noise due to the very small signal power. The aim of *image de-nighting* techniques is to increase the contrast and the sharpness of an image, in order to improve its visual appearance and support the extraction

* This work was funded by the IAPP CS-ORION (PIAP-GA-2009- 251605) grant within 7th Framework Program of the European Community.

of valuable information, that can be used for analysis, detection and segmentation purposes. As a result, image de-nighting operators have attracted broad interests recently, partially due to the numerous applications that require night image enhancement including surveillance, astronomical and medical imaging.

Early approaches in the problem relied on Histogram Equalization (HE) techniques [2],[3]. HE is a relatively simplistic strategy which usually introduces multiple artifacts and leads to significant loss of image details. Authors in [4] demonstrated a Colour Estimation Model (CEM) based on a parameter that controls the range of the image colors on the RGB-scale. Dong *et al.* [5] proposed an algorithm for the enhancement of low-light videos, that inverts the dark input frames and then performs a de-hazing algorithm to improve the quality of the low light images. Another night content enhancement technique, was proposed in [6], where an image is separated into the reflectance and the illuminant component, according to the *retinex theory*, and the enhanced image is generated from the reconstruction of the reflectance component. Due to the challenges associated with the separation of the two components, errors that occur during this process lead to the introduction of artifacts in the reconstructed image. Recently, in [8] a gradient domain technique was proposed, where the authors estimate the mixed gradient field of the intensities from multiple night-time, along with their correspondent day-time images.

Unlike other approaches, our method relies on a single image for the estimation of the enhanced image, using a sparse signal representation technique. Particularly, sparsity suggests that a high-frequency signal can be accurately recovered from its corresponding low-frequency representation. Many ill-posed and inverse problems such as image super-resolution and image denoising [1, 14] introduce the sparsity assumption in the form of a prior, that is able to distinguish the ground truth image content from the degradation effects. Motivated by these approaches, this paper focuses on the problem of recovering the illuminated and enhanced version of a given low-light image, using the sparsity constraint as prior knowledge. We work on image patch pairs sampled from images captured under different exposure times, which are used as proxies to day and night time images. Furthermore, we consider two dictionaries (night and day) for our representation. The dark, under-exposed image patches are used for the creation of the night dictionary \mathbf{D}_n , while the corresponding better exposed image patches (day patches) are used for the creation of the day dictionary, \mathbf{D}_d . The main purpose of this work, is to extract the sparse representation of a night image patch subject to \mathbf{D}_n and directly use it for the reconstruction of corresponding day image patch in terms of \mathbf{D}_d . Our method can be applied for the enhancement of generic images that were captured under poor-illumination conditions. The evaluation of our results are accomplished using both quantitative metrics and visual quality.

The rest of the paper is organized as follows. Section 2, provides an overview of the Sparse Representations (SR) framework and then extensively presents our proposed method. Section 3, presents the results on performance of our and state-of-the-art methods, while our conclusions are given in Section 4.

2 Image De-nighting Using Sparsity

2.1 Sparse Representations Framework

This paper focuses on the problem of recovering the well illuminated version of a given low-intensity image. Our approach is patch-based and we assume that image patches can be represented as linear combinations of elements from an over-complete dictionary. Formally, given a signal $\mathbf{y} \in \mathbb{R}^M$ extracted from an image patch and a dictionary \mathbf{D} , the challenge is to find a vector $\mathbf{x} \in \mathbb{R}^N$ satisfying the relationship: $\mathbf{y} = \mathbf{D}\mathbf{x}$. When the dictionary is overcomplete, *i.e.* $M \ll N$, this problem admits infinite solutions. Motivated by the SR framework, we search for a vector \mathbf{x} that optimizes a certain sparsity measure. An initial approach to this problem is the minimization of the general l_p -norm. In this case the minimization problem becomes:

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{x}\|_p \quad \text{subject to} \quad \|\mathbf{D}\mathbf{x} - \mathbf{y}\|_2 \leq \epsilon \quad (1)$$

where $\|\mathbf{x}\|_p = (\sum_i |\mathbf{x}_i|^p)^{1/p}$ and ϵ is a threshold on the approximation error. When $p = 0$, the sparsity of the coefficient vector \mathbf{x} , is measured by the non-zero counting pseudo-norm, l_0 . Although, the l_0 -norm is the optimal choice for the recovery of the sparse vector \mathbf{x} , it leads to an NP-hard problem. Fortunately, the theory of SR suggests that one can replace the l_0 -norm with the l_1 -norm, leading to following optimization problem:

$$\min \|\mathbf{x}\|_1 \quad \text{subject to} \quad \|\mathbf{D}\mathbf{x} - \mathbf{y}\|_2 \leq \epsilon, \quad (2)$$

where $\|x\|_1 = \sum_i |x_i|$ is the l_1 -norm. This linear regression of the l_1 -norm is denoted as the LASSO problem [12]. Furthermore, one can add the non-negative constraint $x \geq 0$ in order to account for the additive nature of the features we utilize in our representations.

2.2 Dictionary Model

A crucial aspect of the above reconstruction is the proper selection of the dictionary \mathbf{D} , where we seek a dictionary which can sparsify the input data. In general, dictionaries can be analytic or trained [13]. Analytic dictionaries arise from an existing family of transforms, such as histograms, DCT, curvelets, contourlets, wavelets and so on. On the other hand, trained dictionaries are learned from a collection of training data. Although this approach can offer various benefits due to the selection of the most representative examples, it is very computationally expensive. A variant of the later relies on the creation of the dictionary by randomly selecting patches extracted from images that exhibit the same statistical nature (training images). In our setting, we follow the latter approach, working with two coupled dictionaries, \mathbf{D}_n corresponding to the night-time image patches and \mathbf{D}_d , for the day-time image patches. At the training phase, we utilized multiple registered pairs of day-time and their corresponding night-time images for patch extraction. As a consequence, corresponding elements of the

two dictionaries encode the same part of the scheme but at different illumination conditions and can thus serve for the estimation of the high intensity image from the low intensity one.

Feature Selection The features computed from the night-time image patches, have to ensure that the sparse coefficients will have an accurate representation in the appropriate dictionary. State-of-the-art methods use different features for the representation of the degraded components. For instance, Raskar *et al.* [8] reconstructed the final image by integrating first order gradients of the input images, by creating a mixed gradient field. Yang *et al.* [1] used first and second order gradients to represent the low-resolution images. In our representation, for each night-time image patch, we extract the Cumulative Histograms(CH), as features, due to their representation capabilities of varying illuminations conditions. The intensity histogram of a low-illumination patch y_n expresses a discrete representation of the probability density function for the pixel intensities, and may be expressed as:

$$P_{y_n(x,y)}(v) = P(y_n(x,y) = v) = \frac{n_v}{N} \quad (3)$$

where, where $0 \leq v \leq 255$ and N is the total number of pixels in the image patch. Then, the CH of each night-time patch, is measured by summing up the histogram values from gray level 0 to V :

$$CH_{y_n(x,y)} = \sum_{i=0}^V P_{y_n(x,y)}(i) \quad (4)$$

For the day-time dictionary \mathbf{D}_d , we utilized normalized pixel intensities for the reconstruction since they offer a natural approach in modeling the day-time images.

2.3 Image Reconstruction

Given a low-light version of the scene \mathbf{Y} , our task is to generate the corresponding illuminated image \mathbf{X} . According to theory of SR the enhanced image patch \mathbf{x} can be sparsely represented in an over-complete dictionary \mathbf{D}_d created from day-time patches according to:

$$\mathbf{x} = \mathbf{D}_d \mathbf{w}, \quad \mathbf{w} \in \mathbb{R}^K \quad (5)$$

The key insight of our methods is that the sparse representations vector \mathbf{w} can be accurately estimated from low-illumination patches extracted from the input image \mathbf{Y} , by utilizing the night-time dictionary \mathbf{D}_n . By jointly constructing the two dictionaries \mathbf{D}_d and \mathbf{D}_n , we expect to observe the same sparse representations for each patch pair of night-time and day-time images. For each low-illumination image patch, we find a sparse representation with respect to \mathbf{D}_n . Then, in order to extract the illuminated image patch \mathbf{x} we utilize the sparse coefficients with the day-time dictionary \mathbf{D}_d . Since the dictionary \mathbf{D}_n is over-complete, the

equation is under-determined for the unknown coefficients \mathbf{w} . The equivalent l_1 -minimization problem is thus given by:

$$\min \|\mathbf{w}\|_1 \quad \text{s.t.} \quad \|\mathbf{D}_n \mathbf{w} - \mathbf{y}\|_2 \leq \epsilon \quad \text{and} \quad \mathbf{w} \geq 0 \quad (6)$$

Given the optimal sparse coefficients \mathbf{w}^* , the reconstructed illuminated image patch becomes:

$$\mathbf{x} = \mathbf{D}_d \mathbf{w}^* \quad (7)$$

In order to enforce the compatibility between adjacent patches, we process the input image’s patches starting from the upper-left corner with a small overlapping factor in each direction. Due to this fact, the reconstructed image appears with a slight blurring effect. In order to overcome this issue, we perform a back-projection technique, motivated by Yang *et al.* [1]. The main idea is to project the day-time image \mathbf{X}_0 at the solution space $\mathbf{Y} = a\mathbf{H}\mathbf{X}$, where \mathbf{Y} is a dark and blurred version of the day-time image \mathbf{X} , \mathbf{H} is an operator matrix that represents the blurring effect, a low pass Gaussian filter in our case, and a is a small parameter that uniformly changes the illumination of the target day-time image \mathbf{X} . The value of a was set manually after cross validation in order to achieve the best possible result. This operation can be formulated as:

$$\mathbf{X}^* = \arg \min_{\mathbf{X}} \|a\mathbf{H}\mathbf{X} - \mathbf{Y}\|_2 + c\|\mathbf{X} - \mathbf{X}_0\|_2, \quad (8)$$

which can be solved using by gradient descent technique.

3 Experimental Results

3.1 Algorithmic Details

To validate the proposed approach, a series of experiments was conducted using data from a High Dynamic Range (HDR) image dataset [9]. We selected HDR registered images in our experiments motivated by the approximation of a low intensity night image by an image captured with a very short exposure time and a daytime image by a well exposed one. The number of atoms for both the day-time, \mathbf{D}_d and night-time, \mathbf{D}_n dictionaries was set to 550. For the creation of the day-time dictionary, the best performance is achieved using 3x3 patch size, with 1 overlapping factor between adjacent patches.

Our method can be applied on both gray-scale and color images. When color images are processed, instead of processing each color channel separately, we transform the image into the HSV color space and apply our night context enhancement technique to the V channel only. As a pre-processing step, we normalize the dark image by dividing with the mean value of the three components (HSV). In order to evaluate the results of the proposed scheme, we use both quality measures and visual perception. The reconstruction quality is measured via the Structural Similarity Image Quality Index (SSIM)[15], between the illuminated ground-truth image and the reconstructed image.

3.2 Evaluation of the results

Fig. 1 serves as a motivation of our work, depicting a low light (left) and the reconstructed well illuminated (right) image pair along with the corresponding histograms where we observe that the histogram of the reconstructed image is a shifted and distributed version of the original image’s histogram.

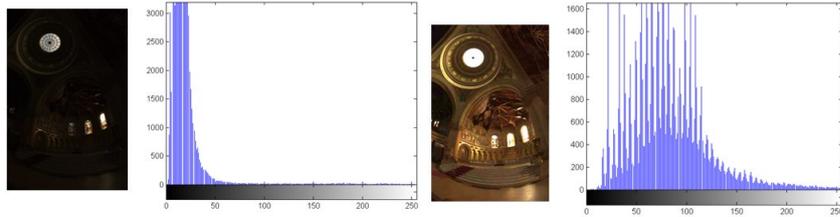


Fig. 1. Memorial dataset:(left) Input dark image along with its histogram, (right) Reconstructed illuminated image with its histogram

We proceed by showing results from experimental scenes to validate our approach. The proposed algorithm is compared against state-of-the-art image denoising methods including the Histogram Equalization method (HE) [2], serving as a baseline processing module,[3], the method proposed by Dong *et al.* [5] for enhancement of low-light video frames, and Color Estimation Model (CEM) [4]. Unlike our method, the other methods under consideration, do not include a dictionary training phase. In order to achieve a fair comparison, we perform the same pre-processing step to all methods, even it is not part of their algorithms. Figs. 2- 7, depict the reconstruction achieved by our algorithm when applied to natural scenes and compared to the previously described methods.

In Fig. 2 one can visualize the performance of the above described methods on the Memorial dataset. We observe that all methods enhance and illuminate the dark test image, but our algorithm produces results closer to the ground truth image. In terms of SSIM [15], our reconstruction achieves better similarity index compared to the other approaches. Figs. 3 and 4, demonstrate the ability of our algorithm to enhance the office and the office very dark datasets, without adding noise or introducing saturation effects, compared to the other methods, which introduce significant noise to the reconstructed image. Furthermore, in Fig. 5, one can notice the visible distortions caused by all three comparable state-of-the-art methods while the proposed method enhances the UWMech image and reveals sufficient information on the dark parts, without amplifying the noise.

Another example of our reconstruction is presented in Fig. 6. Dong’s enhancement approach leads to significant loss of details (especially on the image’s background and on the baby’s face), due to the saturation of the image pixels. HE enhancement, also leads to the introduction of artifacts, artificial colours (especially on the image’s background) and to the amplification of noise. CEM



Fig. 2. Memorial dataset:(Top left) Original Image, (Top middle) Reference Image, (Top right) Dong's method, (Bottom left) HE, (Bottom middle) CEM, (Bottom right) Proposed method



Fig. 3. Office dataset: (Top left) Original Image, (Top middle) Reference Image, (Top right) Dong's method, (Bottom left) HE, (Bottom middle) CEM, (Bottom right) Proposed method

model has provided a noticeable improvement at the image's contrast, but the resulted image is not close to the ground truth ones. We observe that the proposed scheme, reveals sufficient information in the dark parts, without illuminating the already sufficiently illuminated parts.

Finally, Fig. 7 depicts the results of the proposed reconstruction against the compared methods, to the Piano-man dataset. We observe that both HE and Dong's method introduce significant saturation effects on the resulting image. CEM's reconstruction in this case is good and the method does not enhance any artifacts or noise. Our method is able to reconstruct the details present in the



Fig. 4. Office very dark dataset: (Top left) Original Image, (Top middle) Reference Image, (Top right) Dong's method, (Bottom left) HE, (Bottom middle) CEM, (Bottom right) Proposed method

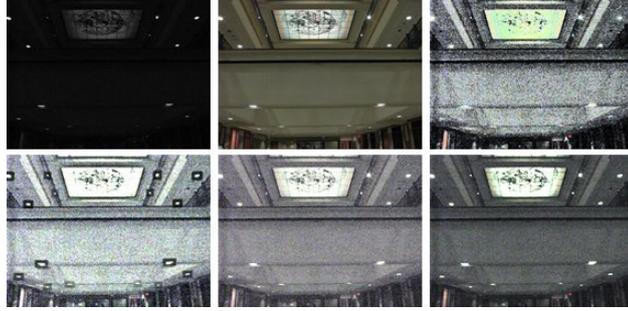


Fig. 5. UWMechDept dataset: (Top left) Original Image, (Top middle) Reference Image, (Top right) Dong's method, (Bottom left) HE, (Bottom middle) CEM, (Bottom right) Proposed method

dark regions of the image and produces an artifact-free result, very close to the ground truth image.

The quantitative results are presented in Table 1. We are able to confirm that our method achieves better results in the terms of the SSIM [15] and visual perception compared to other three state-of-the-art methods.

Table 1. Quality (SSIM [15]) measurements of the enhancement methods

Test image	Memorial	Office	Office very dark	UWMech	Baby	Piano
Dong's [5]	0.5354	0.4116	0.6518	0.2922	0.2973	0.3821
HE [2]	0.5830	0.5654	0.4140	0.2182	0.3014	0.5451
CEM [4]	0.7497	0.7676	0.7314	0.4156	0.5461	0.8011
Proposed	0.8579	0.9498	0.8666	0.6212	0.6504	0.8520



Fig. 6. Baby on grass dataset: (Top left) Original Image, (Top middle) Reference Image, (Top right) Dong's method, (Bottom left) HE, (Bottom middle) CEM, (Bottom right) Proposed method



Fig. 7. Piano man dataset: (Top left) Original Image, (Top middle) Reference Image, (Top right) Dong's method, (Bottom left) HE, (Bottom middle) CEM, (Bottom right) Proposed method

4 Discussion and Future Work

This paper introduced a novel approach for the enhancement of low illumination images. We considered a sparse signal representation approach based on the joint creation of the low and high illumination dictionaries, sampled for appropriate image pairs. The proposed scheme, successfully extracts the necessary information from an image, by illuminating the dark regions, without causing artifacts or saturation effects. Future work includes the investigation of alternative feature operators that could further increase the reconstruction quality.

References

1. J. Yang, J. Wright, T. Huang, Y. Ma: Image Super-Resolution Via Sparse Representations. In: Image Processing, IEEE Trans. on (Vol.19, Issue: 11), pp. 2861–2873.(Nov.2010).
2. Y.T.Kim: Contrast Enhancement Using Brightness Preserving Bi- Histogram Equalization, IEEE Trans., Consumer Electronics, vol. 43,no. 1, pp. 1-8, 1997.
3. Manpreet Kaur, Jasdeep Kaur, Jappreet Kaur: Survey of Contrast Enhancement Techniques based on Histogram Equalization:(IJACSA) Int. Journal of Advanced Computer Science and Applications,Vol. 2, No. 7, 2011.
4. H. Fu, H. Ma, S. Wu: Night Removal by Color Estimation and Sparse Representations. In: 21st Int. Conference on Pattern Recognition, pp. 3656–3659. Tsukuba, Japan (ICPR 2012).
5. X. Dong, Y. Pang, J. Wen, G. Wang, W. Li, Y. Gao, S. Yang: A fast efficient algorithm for enhancement of low lighting video. Journal of Information and Computational Science 7: 10 (2010), pp. 2021-2030.
6. A.Yamasaki, H.Takauji, S.Kaneko, T. Kanade, H. Ohki: Denighting: Enhancement of Nighttime Images for a Surveillance Camera. In: 19th Int. Conference on Pattern Recognition (ICPR 2008).
7. A.Loza, D. Bull, P. Hill, A. Achim: Automatic contrast enhancement of low-light images based on local statistics of wavelet coefficients. Digital Signal Processing, vol. 23, issue 6, pp. 1856–1866. (December 2013).
8. R. Raskar, A. Ilie, and J. Yu: Image fusion for context enhancement and video surrealism. In: 3rd Int. symposium on Non-photorealistic animation and rendering, pp. 85–152, New York (2004).
9. P. Sen, N. Khademi Kalantari, M. Yaesoubi, S. Darabi, D. Goldman, E. Shechtman: Robust Patch-Based HDR Reconstruction of Dynamic Scenes. In: ACM Trans. on Graphics(TOG),SIGGRAPH Asian. Vol.31, No. 6,Art. 203. Technical paper, (November 2012).
10. Q. Shan, J. Jia, M. Brown: Globally Optimized Linear Windowed Tone-Mapping.In: IEEE Trans. on visualization and computer graphics, vol.16 (2010).
11. M. Elad: Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing. Springer (2010).
12. R. Tibshirani: Regression Shrinkage and Selection via the Lasso. Journal of the Royal Statistical Society. ,vol.58, issue 1, pp. 267–288. (1996).
13. M.Aharon, M.Elad, A.Bruckstein, Y.Katz: K-SVD: An Algorithm for Designing of Overcomplete Dictionaries for Sparse Representations. In: IEEE Trans. on Signal Processing, (Volume:54, Issue: 11), pp. 4311–4322, (Nov.2006).
14. M. Elad and M. Aharon: Image denoising via sparse and redundant representations over learned dictionaries. In: Image Processing, IEEE Trans., Vol. 15, No. 12, pp. 3736 –3745 (Dec.2006)
15. Z. Wang, A. Bovik, H. Sheikh and E. Simoncelli: Image Quality Assessment: From Error Visibility to Structural Similarity. In: IEEE Trans. on Image Processing, vol. 13, no. 4, Apr. 2004, pp.600–612, (April 2004).