International Journal of Innovative Technologies
WWW.IJITECH.ORG
ISSN 2321-8665
Volume.06, Issue.01,
January-June, 2018,
Pages:0236-0238

A Novel Multi-Exposure Image Fusion: A Structural Patch Decomposition Approach

SAI SANDEEP YANDAMURI[1], AKULA PRAVIN[2]
[1]PG Scholar, Dept of ECE, Bonam Venkata Chalamayya Engineering College, Odalarevu, AP, India.
[2]Professor, Dept of ECE, Bonam Venkata Chalamayya Engineering College, Odalarevu, AP, India.

Abstract: We present a novel deep learning architecture for fusing static multi-exposure images. Current multi-exposure fusion (MEF) approaches use hand-crafted features to fuse input sequence. However, the weak hand-crafted representations are not robust to varying input conditions. Moreover, they perform poorly for extreme exposure image pairs. Thus, it is highly desirable to have a method that is robust to varying input conditions and capable of handling extreme exposure without artifacts. Deep representations have known to be robust to input conditions and have shown phenomenal performance in a supervised setting. However, the stumbling block in using deep learning for MEF was the lack of sufficient training data and an oracle to provide the ground-truth for supervision. To address the above issues, we have gathered a large dataset of multi-exposure image stacks for training and to circumvent the need for ground truth images, we propose an unsupervised deep learning framework for MEF utilizing a no-reference quality metric as loss function. The proposed approach uses a novel CNN architecture trained to learn the fusion operation without reference ground truth image. The model fuses a set of common low level features extracted from each image to generate artifact-free perceptually pleasing results. We perform extensive quantitative and qualitative evaluation and show that the proposed technique outperforms existing state-ofthe-art approaches for a variety of natural images.

Keywords: Multi-Exposure Image Fusion, Highdynamic-Range Imaging, Weighted Mean, Sparse Representation.

I. INTRODUCTION

When we take photos of natural scenes that include very dark and very bright regions, their digital images often lose details of these regions. In general, commonly used digital cameras have narrower ranges of luminance than natural scenes [1, 2]. We cannot obtain details of regions whose luminance is outside camera ranges. These regions are commonly called saturation regions. To clearly represent scenes without saturation regions, multi-exposure image fusion has been proposed [3–22]. It fuses some images into one desired image. The input images are obtained by taking photos of the same scene with different exposure times, and the locations of their saturation regions are different. Hence, the fused image fully represents the scene without saturation regions. Methods for multi-exposure image fusion are mainly classified into two types: weighted mean and gradient cascade. The former type has been actively studied and includes the greatest number of methods [3–19]. These methods fuse images by pixel-wise weighted mean. Various procedures forthe weight calculation have been proposed. Recently, some methods have aimed to prevent visual artifacts, such as motion blurs and ghosts [12–19]. The artifacts are caused by movement of objects, and recent methods try to align objects in the same locations. Consequently, these methods reduce artifacts and produce natural images. In the gradient cascade type, few methods have been proposed [20, 21]. These methods choose the maximum gradients of input images at each pixel, and the resultant gradient field is defined as gradients of the fused image. Finally, the gradients are transformed to the spatial domain, and the result is the fused image. They produce fine edges and textures in fused images. Each of the two types has problems with fused images.

Due to the mean procedure, weighted mean methods produce blurred images. In particular, edges and textures of their resultant fused images are blurred. With the other type, errors caused by noise and saturation are spread all over the image via the transformation from the gradient domain to the spatial domain, and the spreading amplifies the errors. Consequently, unnatural regions occur in fused images. Recently, sparse representation is widely used as fundamental technique in image processing [23–25], because it can approximate images to have sharp edges and textures without slight variations such as noises. Several image applications based on sparse representation have been proposed, and achieve excellent results [23–25]. In the multi-exposure image fusion, a method based on sparse representation has also been proposed [22]. The method divides mean values and residual components of each input image by patch unit, averages the values, and fuses the components based on sparse representation to produce sharp fused components. Unfortunately, since the averaging procedure is poor and the fusionis affected by saturation regions, fused images are visually blurred and artifacts often occur. To overcome the problems of previous methods, we propose a hybrid method for multi-exposure image fusion based on weighted mean and sparse representation.

The proposed method produces averages and details of fused images by using weighted mean and sparse representation, respectively. The details mean edges, local contrasts, and textures. Due to the weighted mean method, the resultant average imagesare visually natural. For fusion of detail components, we use the proposed selection method (which includes sparse representation) to avoid blurs and effects of saturation regions. Due to the proposed fusion, the resultant details have sharp edges and textures. Consequently, the proposed method produces fine fused images without artifacts, and we show that the proposed method outperforms previous methods through simulations objectively and perceptually. We assume that the object alignment is already finished by previous methods in this paper, because the several alignment methods have been proposed, and show their efficacy [2, 18].

II. STRUCTURAL PATCH DECOMPOSITION FOR MEF

In this section, we detail the proposed structural patch decomposition (SPD) approach for MEF. We first describe abaseline version that works for static scenes, and then extend it to dynamic scenes by adding a structural consistency check, resulting in the robust SPD-MEF algorithm.

A. Robust SPD-MEF

We extend the baseline SPD-MEF to account for dynamic scenes in the presence of camera and object motion. We assume that the input source sequence is aligned, for example by setting a tripod or some image registration algorithms [15]–[17]. This assumption is mild because the camera motion is usually small and relatively uniform. In this paper, we implement image registration by first performing SIFT [17] matching and then computing an affine transformation matrix from matched points with an l21-norm loss. It works well on all test sequences that need to be aligned. The use of l21-norm loss is because it is robust to mismatched points and can be efficiently solved using iteratively reweighted least squares. we also pick one exposure as the reference to determine the motion appeared in the fused image and reject inconsistent motions in the rest images w.r.t. it. Throughout the paper, we select the one with normal exposure if the source image sequence contains three input images. Otherwise, we choosethe one that has the least number of under- or over-exposed patches, as suggested in [19], [20]. Within the framework of the proposed SPD, it is very convenient to detect inconsistent motions across exposures by making use of the structure vector **s**k. To be specific, we compute the inner product between the reference signal structure **s**r and the signal structure **s**k of another exposure ρk lies in [−1, 1] with a larger value indicating higher consistency between **s**k and **s**r. Since **s**k is constructed by mean removal and strength. The corresponding binary map generated for each exposure (including the reference which is uniformly one) is referred to as the structural consistency map, as shown in Fig3. From the figure, we observe that the inconsistent motions across exposures are reliably identified with minimal false positive detection, and the structure vectors of over-exposed areas in the reference image (e.g., the clouds in the left part of the4-th image) are consistent with the same regions in other exposures, which verifies our claim of properly handling under- or over-exposed regions.

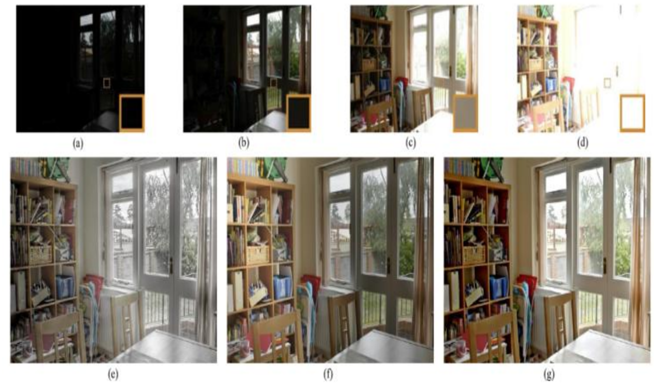

Fig. 2. Making use of color contrast. ck is the average signal strength of the k-th inset patch computed from RGB channels separately. ck is the corresponding signal strength by treating RGB channels jointly. Source image sequence by courtesy of Tom Mertens [9]. (a) c1 = 0.1, c1 = 0.2. (b) c2 = 0.3, c2 = 0.3. (c) c3 = 0.3, c3 = 7.5. (d) c4 = 0.0, c4 = 0.0. (e) Gu12 [8]. (f) Shutao12 [13]. (g) SPD-MEF.

Although we leave open the possibility of filling in the under- or over-exposed regions of the reference image with structures from other exposures, we add another constraint to check whether those structures are proper for fusion inorder to minimize ghosting artifacts by invoking IMF, whichis capable of mapping between intensity values of any two exposures. For example, we can easily create a latent image that contains the same motion as the 4-th image of Fig. 3(a) but has an exposure level like the 2-nd image of Fig. (a) by mapping the intensity values of the former to the latter using IMF. We first create K −1 latent images by mapping the intensity values of the reference image to the rest K −1 exposures and compute the absolute mean intensity difference of co-located patches in the k-th exposure and its corresponding latent image. We again threshold the difference

$$\bar{B}_k = \begin{cases} 1 & \text{if } |l_k - l'_k| < T_m \\ 0 & \text{if } |l_k - l'_k| \geq T_m \end{cases} , \quad \textbf{(1)}$$

Moreover, we are able to adjust the mean intensity of the moving object in the reference image to adapt it to the neighborhood environment, which avoids abrupt intensity changes in a much cheaper way.

III. CONCLUSION AND FUTURE SCOPE

In this paper, we proposed a novel structural patch decomposition (SPD) approach for MEF. Different from most pixel wise MEF methods, SPD-MEF works on color image patches directly by decomposing them into three conceptually independent components and by processing each component separately. As a result, SPD-MEF generates little noise in the weighing map and makes better use of color information during fusion. Furthermore, reliable deghosting performance is achieved by using the direction information of the structure vector. Comprehensive experimental results demonstrated that SPD-MEF produces MEF images with sharp details, vivid color appearance and little ghosting artifacts while maintaining a manageable computational cost. The proposed SPD approach is essentially dynamic range independent. Therefore, it would be interesting to explore its potential use in HDR reconstruction to generate high quality HDR images with little ghosting artifacts. Moreover, the application of SPD is not limited to MEF. As a generic signal processing approach, SPD has been found to be useful in image quality assessment of contrast-changed and stereoscopic images. It is worth considering whether SPD offers any insights that can be transferred to other image processing applications. In addition, although objective quality models for MEF algorithms begin to emerge, the models for objectively comparing MEF algorithms for dynamic scenes are largely lacking. Therefore, it is demanding to switch the focus from developing MEF algorithms for dynamic scenes to developing such objective quality models in order to conduct a fair comparison.

IV. REFERENCES

[1] B. Hoefflinger, High-Dynamic-Range (HDR) Vision, New York: Springer-Verlag, 2007.

[2] E. Reinhard, W. Heidrich, P. Debevec, S. Pattanaik, G. Ward, and K. Myszkowski, High Dynamic Range Imaging, Second Edition: Acquisition, Display, and Image-Based Lighting, CA: Morgan Kaufmann, 2010.

[3] S. Mann and R. W. Picard, "On being 'undigital' with digital cameras: Extending dynamic range by combining differently exposed pictures," in Proc. IS&T Conf., 1995, pp. 422–428.

[4] P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in Proc. ACM SIGGRAPH, 1997, pp. 369–378.

[5] T. Mitsunaga and S. K. Nayar, "Radiometric self calibration," in Proc. IEEE Comput. Society Conf. Comput. Vision Patt. Recognition, 1999, vol. 1, pp. 374–380.

[6] F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," in Proc. ACM SIGGRAPH, 2002, pp. 257–266.

[7] A. A. Goshtasby, "Fusion of multi-exposure images," Image Vision Comput., vol. 23, no. 6, pp. 611–618, 2005.

[8] X. Li, K. M. Lam, and L. Shen, "An adaptive algorithm for the display of high-dynamic range images," J. Vis. Commun. Image Represent., vol. 18, no. 5, pp. 397– 405, 2007.

[9] S. Zheng, W.-Z. Shi, J. Liu, G.-X. Zhu, and J.-W. Tian, "Multisource image fusion method using support value transform," IEEE Trans. Image Process., vol. 16, no. 7, pp. 1831–1839, 2007.

[10] T. Mertens, J. Kautz, and F. Van Reeth, "Exposure fusion: A simple and practical alternative to high dynamic range photography," Comput. Graph. Forum, vol. 28, no. 1, pp. 161–171, 2009.

[11] R. Shen, I. Cheng, J. Shi, and A. Basu, "Generalized random walks for fusion of multi-exposure images," IEEE Trans. Image Process., vol. 20, no. 12, pp. 3634– 3646, 2011.

[12] S. B. Kang, M. Uyttendaele, S. Winder, and R. Szeliski, "High dynamic range video," ACM Trans. Graph., vol. 22, no. 3, pp. 319–325, 2003.

[13] E. A. Khan, A. O. Akyiiz, and E. Reinhard, "Ghost removal in high dynamic range images," in Proc. IEEE Int. Conf. Image Process., 2006, pp. 2005–2008.

[14] W.-C. Kao, "High dynamic range imaging by fusing multiple raw images and tone reproduction," IEEE Trans. Consumer Electronics, vol. 54, no. 1, pp. 10–15, 2008.

[15] P. Sen, N. K. Kalantari, M. Yaesoubi, S. Darabi, D. B. Goldman, and E. Shechtman, "Robust patch-based HDR reconstruction of dynamic scenes," ACM Trans. Graph., vol. 31, no. 6, pp. 203:1–203:11, 2012.

[16] T. Jinno and M. Okuda, "Multiple exposure fusion for high dynamic range image acquisition," IEEE Trans. Image Process., vol. 21, no. 1, pp. 358–365, 2012.

[17] M. Song, D. Tao, C. Chen, J. Bu, J. Luo, and C. Zhang, "Probabilistic exposure fusion," IEEE Trans. Image Process., vol. 21, no. 1, pp. 341–357, 2012.

[18] N. K. Kalantari, E. Shechtman, C. Barnes, S. Darabi, D. B. Goldman, and P. Sen, "Patch-based high dynamic range video," ACM Trans. Graph., vol. 32, no. 6, 2013.

[19] Z. Li, J. Zheng, Z. Zhu, and S. Wu, "Selectively detailenhanced fusion of differently exposed images with moving objects," IEEE Trans. Image Process., vol. 23, no. 10, pp. 4372–4382, 2014.

[20] A. R. V.-Koczy, A. R ´ ovidz, and T. Hashimoto, ¨ "Gradient-based synthesized multiple exposure time color HDR image," IEEE Trans. Instrum. Meas., vol. 57, no. 8, pp. 1779–1785, 2008.

[21] B. Gu, W. Li, J. Wong, M. Zhu, and M. Wang, "Gradient field multi-exposure images fusion for high dynamic range image visualization," J. Visual Commun. Image Represent., vol. 23, no. 4, pp. 604–610, 2012.

[22] J. Wang, H. Liu, and N. He, "Exposure fusion based on sparse representation using approximate K-SVD," Neurocomputing, vol. 135, pp. 145–154, 2014.

[23] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," IEEE Trans. Signal Process.,vol. 54, no. 11, pp. 4311–4322, 2006.

[24] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, "Non-local sparse models for image restoration," in Proc. IEEE Int. Conf. Comput. Vision, 2009, vol. 12, pp. 2272–2279.

[25] K. Zhang, X. Gao, D. Tao, and X. Li, "Single image super-resolution with multiscale similarity learning," IEEE Trans. Neural Netw. Learning Syst., vol. 24, no. 10, pp. 1648–1659, 2013.