

Constructing Ghost Free High Dynamic Range Images Using Convolutional Neural Network and Structural Similarity Index

¹Shahid Khan, ²Husnain Mansoor Ali

^{1,2}*Shaheed Zulfiqar Ali Bhutto Institute of Science and Technology, Karachi, Pakistan*

shahid.khan@szabist.edu.pk
husnain.mansoor@szabist.edu.pk

Abstract—A foreign object, commonly called as a ghost artifact, is integrated in the HDR output image when there is a moving object in the photography scene. The problem is persisting even after numerous models proposed by researchers. The most advanced techniques for capturing HDR photographs are still struggling with the produced output till date. The majority of the existing techniques are unable to tackle the situation where luminance varies in the input images. This ultimately causes the algorithms to compromise on quality by reducing the input images. In this research, two techniques are presented that help in getting rid of ghost artifacts from HDR output. The first method that is used in this research is a vintage method that uses structural similarity index measurement. The second method is using Convolutional Neural Network that is proved to be the best configuration of neural network for image recognition purpose. The first method using structural similarity is based on the vintage method of matching two images. This method works by comparing the objective image with the reference image and the algorithm estimates the degree of similarity in context of luminance, contrast and structural information. The second method is the use of convolutional neural network. Since convolutional neural network is a specialized tool for image processing, a convolutional neural network is designed and trained to estimate the index of similarity of two or more images.

Keywords—SSIM, Convolutional Neural Network, HDR, SSIM, LDR, Ghost Artifact

I. INTRODUCTION

High Dynamic Range images are increasingly becoming important in the field of photography and videography sciences [1]. More than ever HDR content is delivered over a variety of medium these days. In recent era, many software

application packages are needing digital photographs to mesh with the digital images seamlessly [1]. What human perceives with their eyes cannot be captured on a digital image as of yet with a single click of a regular camera [2]. In a natural scene, whatever detail in present cannot be obtained due to the way the luminance work.

In a bright scene, the color that is present in the brightest area cannot be obtained with full detail. Same goes with the dark areas of the scene. The color information that is present in the darkest area of the scene cannot be perceived with full detail. High Dynamic Range Imaging is the technique that solves this problem. It works by taking multiple images with different exposure times first and subsequently add them up with respect to the luminance each image has.

This way high dynamic range image construction works is very prone to have objects in the result image that should not be there[2]. Such foreign objects are due to moving objects in the scene.

Though there are many techniques present in the literature and several of them can be accepted as an suitable approach, there is still lot more to be done and there is a huge room for improvement. Every effort to remove the ghost artifact contributed to a certain extent but majority of them could not handle the different exposure values of input images.

This is yet another effort to help to deal with the anti-ghosting in high dynamic range images. The proposed methods using structural similarity index method and convolutional neural network methods not only helpful in removing ghost artifacts, but it also deals effectively with the problem of different exposure time values.

The research in hand is presenting two methods for ghost artifact removal.

1. Ghost Artifact removal using Structural Similarity Index SSIM estimation[3]
2. Ghost Artifact removal using Convolutional Neural Network Image Comparison[4]

The first method of Structural Similarity Index SSIM estimation is originally used for estimating the quality of video that is aired on TV [3]. The original film is taken as the reference and the video that is coming from the TV are compared. The algorithm estimates the index of similarity with respect to the luminance, contrast and structure.

The second method that is used in this research for removing any potential foreign objects is the use of convolutional neural network. A convolutional neural network is considered as the specialized type of neural network that helps best in comparing images for similarity [4]. In this research, a convolutional neural network is designed and trained. Once we have a trained neural network, patches of the input images are passed into it and compared with the counterpart patches in the rest of the images. After analysis, the convolutional neural network comes with a degree and index of similarity of the input patches. On this point, the system takes a decision on that if the compared patches are similar or else. In cases where the similarity index is high enough, the patches are considered as same and no further action is taken at that point. But when the similarity index is a very small fractional value, it is learnt that both the patches are dissimilar. The dissimilar patches are then replaced with the closest matching patch from any of the counterpart patch.

II. RELATED WORK

The recent work [2] to removed ghost artifacts works by splitting the input LDR images in 64 x 64 pixel patches. This technique first orders the input images by exposure time values and then it makes pairs. It then detects the mismatching pairs and replace when needed. Another approach [5] is based on the technique where is tries to find inconsistent pixels across all input images and then correct each one using the reference images. Researchers have also proposed a method of fast moving object detection in HDR imaging using a sub-sampling based technique [6] which helps in detecting the ghost artifact using thresholding technique.

III. METHODOLOGY

A. Structural Similarity Index Measurement (SSIM):

Structural Similarity Index (SSIM) is a measure that tells the level of similarity between two images. It works by accepting two images as parameters[3]. The first image is the

reference image. The reference image is the template image and the second parameter are the target image which will be compared with the reference image. The structural similarity index algorithm tells the percent of similarity of the target image with the reference image.

SSIM estimation is based on three properties [3].

1. Luminance
2. Contrast
3. Structural information

The mathematical model for calculating the similarity index between images X and Y is calculated as below. This this model, the image X is used as template or a reference image and image Y is the target image whose similarity is calculated with the image Y.

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (1)$$

The tricky part that most of the earlier proposed [5] algorithms were not able to handle was the difference of the exposure time value. Though SSIM with the default settings can handle the different exposure to an acceptable extent, as it is indicated in the experiments section ahead. It is worthwhile to test SSIM with customization of removing the luminance and contrast part and only working with the structural similarity.

a. Experiments:

The first step is to input he array of raw images into the system. Second, we break the input images into equal size of patches. The size that we choose for the patches is a variable size which is adjusted according to the size of potential ghost artifact as described in [2]. Once the patches are created, we start matching each patch with the counterpart patches. Each comparison gives an index values of similarity. An array of the differences is created. For each patch in the test. Each index value of similarity is a non-negative number between 0 and 1. The patches with odd values are considered as a potential ghost artifact carrier and thus they are replaced with the patch with the largest value. Once this process is done, we assemble the images and start constructing the high dynamic range images. This technique is found to be effective in avoiding the ghost artifact to be incorporated in the result high dynamic range image.

b. Results:

Following are the results of applying SSIM in figure 1-3.



Fig. (1). Similarity Index 7.889

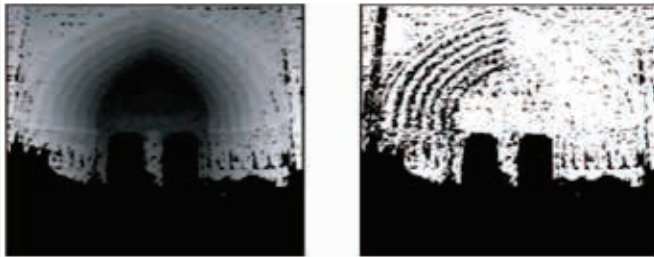


Fig. (2). Similarity Index 8.101



Fig. (3). Liberty Images with different light conditions and similarity index more than 6.541 for each pair

B. Patch Matching Using Convolutional Neural Network (CNN):

Convolutional Neural Network is getting famous for image comparison[4]. The recent research has proved that the convolutional neural network is able to detect the image similarity. The good thing about convolutional neural network is that one doesn't need to specify the parameters and features or ground truths manually during the training. This research uses the convolutional neural network to that can learn directly from the raw images.

The convolutional neural network in this research is trained on a dataset of more than 120,000 images. Each training example is labeled at either as a matching pair or mismatching pair. Half of the training examples are matching pairs and half of the pairs are mismatching pairs. The accuracy of this method is found to be at the extent where it is outperforming the state-of-the-art [4].

It is proposed that this convolutional network can be used to estimate the patch similarity during the problem in hand. A patch having a ghost artifact when matched with a patch that does not have the ghost artifact can be effectively detected using this method of convolutional neural network.

In the proposed research methodology, it is proposed that the trained convolutional neural network [4] can be used to detect the similarity and dissimilarity to identify the ghost artifact. It will also be helpful in replacing the removed infected patch with the most relevant patch present in the rest of the images as elaborated in figure 4.

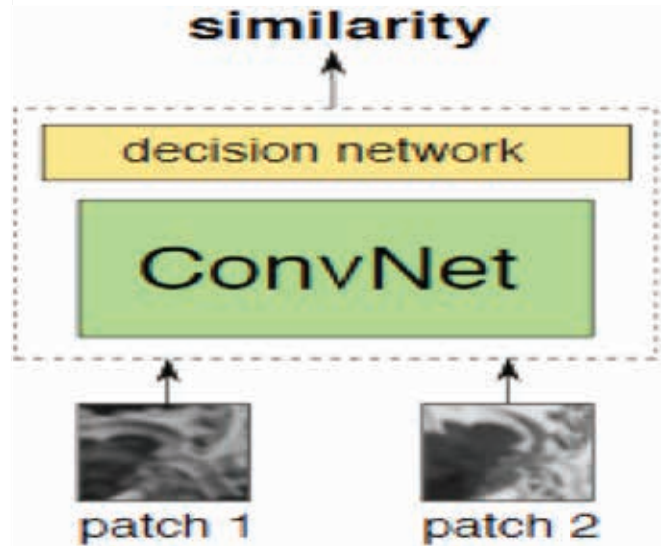


Fig. (4). ConvNet Similarity Model

a. Experiments:

To test the patch matching in the research [4], the pre-trained convolutional neural network is integrated and tested. The input images are preprocessed as per the needs of the requirements of the pre-trained convolutional neural network. The basic requirement of this pre-trained convolutional neural network is that all input images must be converted into greyscale.

The second requirement is to break the image patches into 64 x 64 dimensionally and the orientation was normalized. When testing on this pre-trained convolutional neural network is carried on, the following results have been generated.

b. Results:

Following are the results that are generated using ConvNet in figure 5-9.

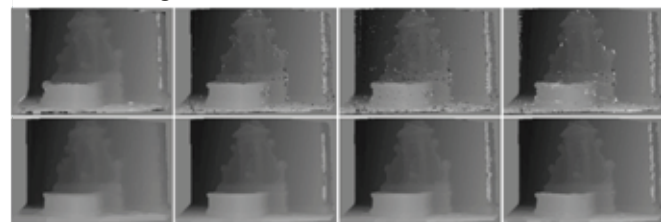


Fig. (5). Patches with average similarity of 6.395

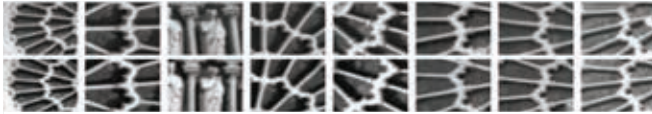


Fig. (6). True Positives



Fig. (7). False Negatives



Fig. (8). True Negatives

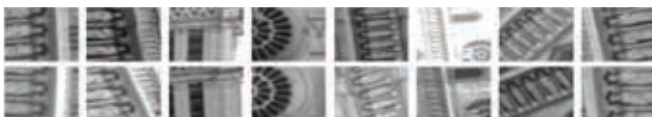


Fig. (9). False Positives

IV. CONCLUSION

Two methods have been proposed to prevent the ghost artifact from integrating in HDR images. SSIM when used for detecting differences in pair of patches, was able to draw a line between matching and mismatching patches. The ConvNet on the other hand was able to clearly separate the matching and mismatching patches. ConvNet is even favorable to detect the matching and mismatching patches in cases where both patches had different exposure time value. ConvNet comparatively is more favorable than SSIM to detect moving objects in LDR images.

V. FUTURE WORK

As there are several shortcomings in the both proposed methods, it is worthwhile to do execute more research in the following areas. The first proposed method of using structural similarity index function SSIM, there is a huge potential room

for improvement. The result can be improved by using the spatial information from image only and to ignore the other two features of luminance and contrast. While it is expected that it will improve the patch matching accuracy, it is also suspected that the due to lack of luminance and contrast will lead towards increasing the error and the generation of false positives.

REFERENCES

- [1] Y. Moon, "A Survey on Applications of High Dynamic Range Technologies in Consumer Electronic Devices," In *Proceedings of 16th International Conference on Control, Automation and Systems (ICCAS)*, 2016. DOI: 10.1109/ICCAS.2016.7832356
- [2] H. Mahmoudabadi, M. J. Olsen, and S. Todorovic, "Detecting Sudden Moving Objects in a Series of Digital Images with Different Exposure Times," *Computer Vision and Image Understanding*, vol. 158, pp: 1-14, 2017.
- [3] DOI: 10.1016/j.cviu.2017.01.004
Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Transactions on Image Processing.*, vol. 13, no. 4, pp. 600-612, 2004.
- [4] DOI: 10.1109/TIP.2003.819861
S. Zagoruyko and N. Komodakis, "Learning to Compare Image Patches via Convolutional Neural Networks," In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
- [5] DOI: 10.1109/CVPR.2015.7299064
Z. Li , S. Wu , Z. Zhu , S. Xie , and S. Rahardja, "Anti-Ghost Of Differently Exposed Images With Moving Objects," In *Proceedings of 19th IEEE International Conference on Image Processing*, 2012.
- [6] DOI: 10.1109/ICIP.2012.6467467
Z. Li , Z. Zhu and S. Rahardja, "Fast Movement Detection For High Dynamic Range Imaging," In *Proceedings of 18th IEEE International Conference on Image Processing*, 2011. DOI: 10.1109/ICIP.2011.6116512

© Author(s) 2017. CC Attribution 4.0 License. (<http://creativecommons.org/licenses/by-nc/4.0/>)

This article is licensed under the terms of the Creative Commons Attribution Non-Commercial License which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.