

PSEUDO HDR VIDEO USING INVERSE TONE MAPPING

Yu-Chen Lin (林育辰), Chiou-Shann Fuh (傅楸善)

Dept. of Computer Science and Information Engineering,
National Taiwan University, Taiwan
E-mail: r03922091@ntu.edu.tw

ABSTRACT

In the Pseudo-Multiple-Exposure-Based Tone Fusion with Local Region Adjustment [1], it introduces a framework for inverse tone mapping which uses only one Low Dynamic Range (LDR) image to evaluate High Dynamic Range (HDR) image. First, map LDR image to HDR images using S curve. By changing the parameters of S curve, we can get different EV scenes. Then, segment the image into four different luminance regions. Finally, use Gaussian weight sum with respect to regions for all five HDR images to generate the best HDR image. In this paper, we propose some methods to generate Pseudo HDR Video using the inverse tone mapping.

Keywords HDR Video; Inverse Tone Mapping; HDR

1. INTRODUCTION

High-Dynamic-Range Imaging (HDRI or HDR) is a set of techniques used in imaging and photography to reproduce a greater dynamic range of luminosity than is possible with standard digital imaging or photographic techniques.

Defects of traditional HDR:

1. Only suitable for static scenes.
2. Need to capture multiple images of the same scene using different exposures.
3. Not real-time for capturing.

Because of these defects, Inverse Tone Mapping Operators (iTMOs) that works on reproducing real-world appearance images through LDR images become more and more important recently.

In Pseudo-Multiple-Exposure-Based Tone Fusion with Local Region Adjustment [1], it produces an inverse tone mapping method that uses only one LDR image to produce an HDR image.

Because of the broader internet, video became more and more important. In this time, HDR image is not too difficult to produce. On the other hand, HDR video is more difficult than HDR image. It is important

to get HDR video using only one normal video through the development of internet and requirement of high quality video.

Due to this, we propose some methods that produce a Pseudo HDR Video by only one normal video using iTMOs. In next section, we will introduce the iTMOs.

2. INVERSE TONE MAPPING OPERATORS (ITMOS)

Figure 1 shows the overall flowchart of iTMOs. First, we get Pseudo-Multiple-Exposure tone fusion (PMET) by applying Exposure-Dependent Inverse Tone Mapping Function to each pixel of image.

$$L_{wk} = \begin{cases} \frac{10^{-P_k} * L_{ad,k} * L_d(i,j)}{1 - L_d(i,j)} & \text{if } L_d(i,j) < 1 \\ L_{max,k} = 10^{-P_k} * L_{ad,k} * L_{smax} & \text{otherwise} \end{cases}$$

L_{wk} is the world luminance of the k -th pseudo-exposure HDR image.

P_k is a scaling factor of the image and used to control the luminance difference between the adjacent pseudo-exposures.

$L_{ad,k}$ controls the average luminance of the pseudo-exposure image.

$L_{ad,k} = 1 + \exp(\mu EV_k)$ where $\mu = 0.85$.

L_{smax} is a constant and set to be 382.5.

Using $EV = (-1, -0.5, 0, 0.5, 1)$ with respect to $P = (1.6, 1.3, 1, 0.85, 0.75)$ as parameters to produce the five Pseudo-Multiple-Exposures images. Figure 2 shows the inverse function.

After we have five Pseudo-Multiple-Exposure images, we can segment image by luminance. The median luminance L_{middle} of the image histogram is used to separate the image into two regions. For the two regions separated by L_{middle} , the medians L_{low} and L_{high} of these two regions are used again to produce the final four local regions, as show in Figure 3.

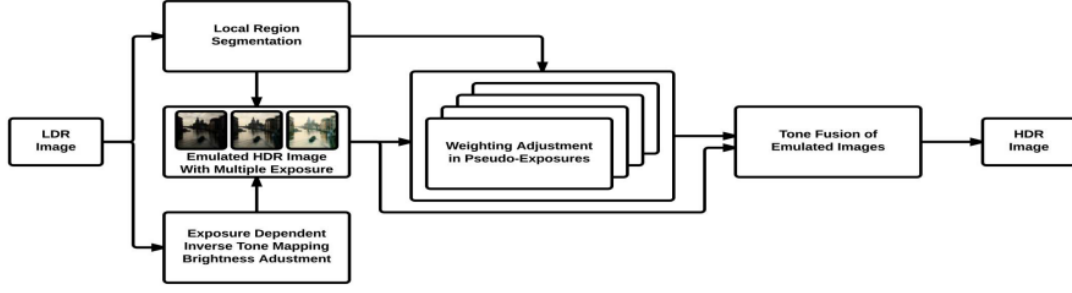


Figure 1: Flowchart

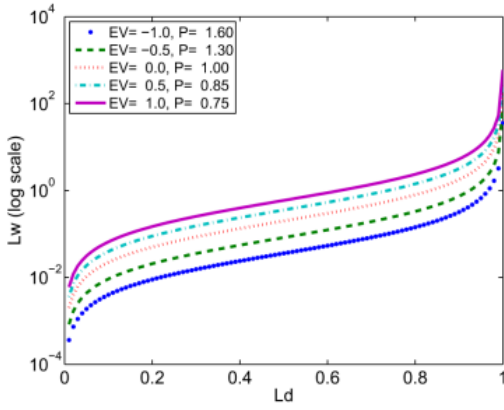


Figure 2: Exposure-Dependent Inverse Tone Mapping Function.

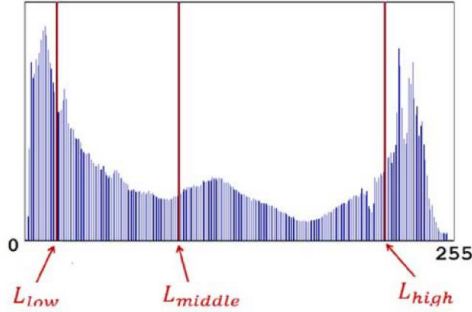


Figure 3: Segmentation by luminance.

After we segment the image into five regions, we can define w_1, w_2, w_3, w_4 and w_5 with respect to five level of HDR luminance. Finally, we can apply Weighting Adjustment in Pseudo-Exposures Tone Fusion of Emulated Images with Multiple Exposures.

$$w_k(i, j) = \exp\left(-\alpha_k * \frac{(L_{wk}(i, j) - L_{med,k})^2}{L_{med,k}^2}\right)$$

$w_k(i, j)$ is the weight value of (i, j) in image k . $L_{wk}(i, j)$ is the pixel value of (i, j) in image k . α_k is an adjustable constant.

$L_{med,k}$ is the average luminance of the region of k -th

estimated image with respect to luminance. For example, in w_1 the $L_{med,k}$ is calculated by average of the lightest region. In w_2 the $L_{med,k}$ is calculated by average of second light region, and so on.

Next, we use a weighted sum to get finally HDR image.

$$L_{w_comb}(i, j) = \frac{\sum_{k=1}^L w_k(i, j) \cdot L_{wk}(i, j)}{\sum_{k=1}^L w_k(i, j)}$$

$L_{w_comb}(i, j)$ Represents the luminance of the reconstructed HDR image.

L is the number of images. (In our case, $L = 5$.)

In the end, apply tone mapping in order to show the result on common device. The result is shown in Figure 4.



Figure 4: Result of iTMOs, the upper image is original and the lower one is the result of iTMOs.

3. PROPOSE METHOD

We have propose three methods for generating a smooth Pseudo HDR video.

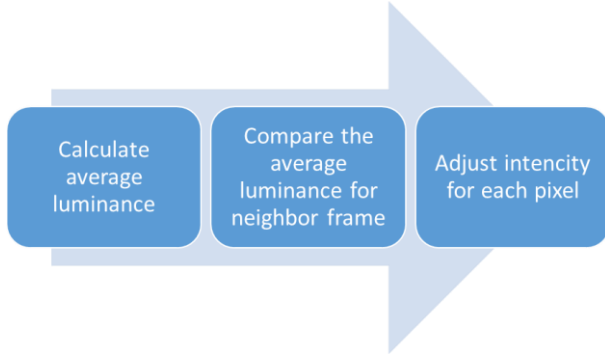


Figure 5: The flow of method 1.

1. Add a luminance constraint for each frame, the average difference of each frame cannot larger than the given constant. The flow of method 1 is shown in Figure 5.

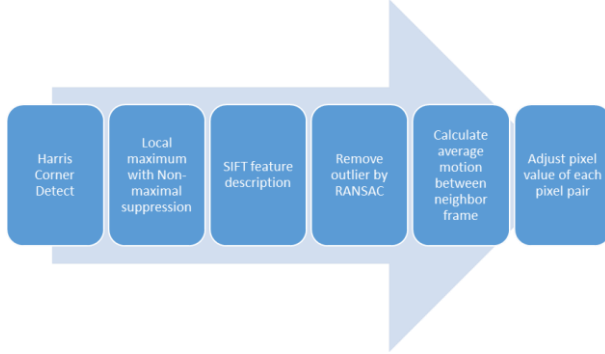


Figure 6: The flow of method 2.

2. In addition, we try to use optical flow to reduce difference of luminance. However, the calculation of optical flow is too large, we use Harris corner feature with SIFT description to get the global motion. Moreover, to assure the feature is robust enough, we find local maximum with Non-maximal suppression. We first sort the R-value (the feature strength of Harris corner) from strongest to weakest. Then choose points from this sorted list with constraint that feature points should have enough distance, r , between each other. $r = \sqrt{30}$ in initial and if we cannot get enough corner points we will reduce r . When we get enough feature points, we will find feature pair by directly calculate Euclidean distance between each feature pair. Then, to assure robustness we use RANSAC to remove outliers. Finally, calculate the average motion of all feature pairs. Use this average motion to predict the location of each pixel in next frame, and let the difference of each pixel pair in luminance is small enough. The flow of method 2 is shown in Figure 6.

3. Moreover, we also propose method 3. In method 3, we collect some neighbor frames as a block. Then, accumulate the total variety of luminance and distribute it to every frame. In this method, the block size we have tried is 5 and 30. See Figure 7.

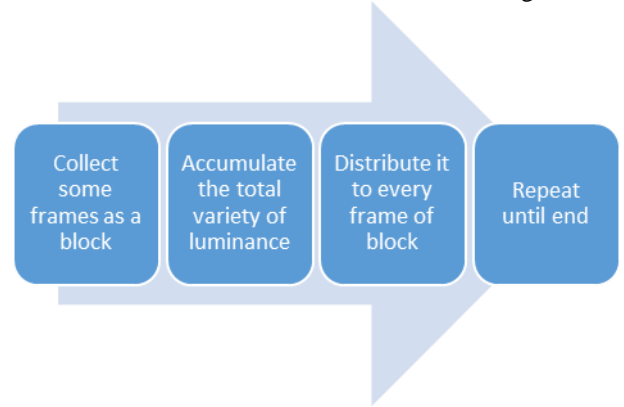


Figure 7: The flow of method 3.

4. RESULT

We apply our propose method to two videos, one has a big motion and the other one is almost stable. Not surprisingly, the result of three methods of video with big motion is bad. In addition, the result of method 1 of stable video is not bad.

Figure 8, 9 and 10 show the result of motion video using method 1, 2 and 3 respectively. The upper images are the 100-*th* frame and the lower images are the 131-*th* frame. We can see that the difference of luminance is large. Thus, the videos will shining when playing.

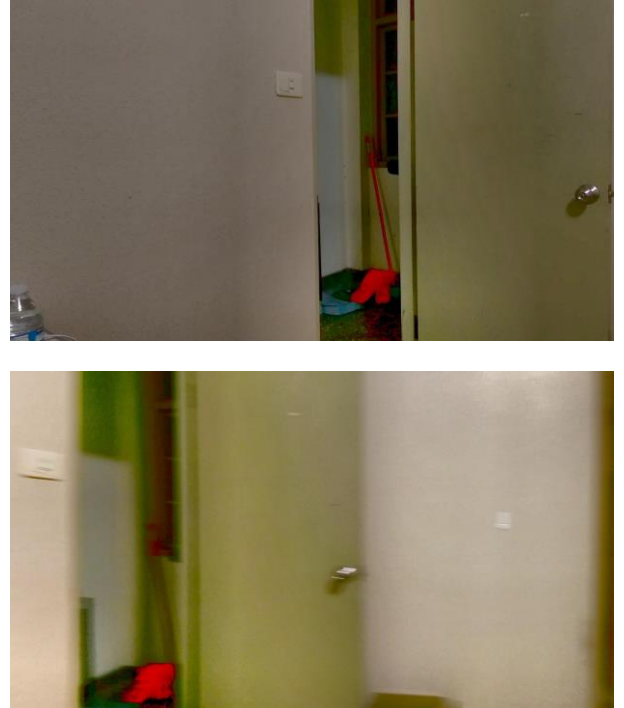


Figure 8: The result of method 1 in video with large motion.



Figure 9: The result of method 2 in video with large motion.

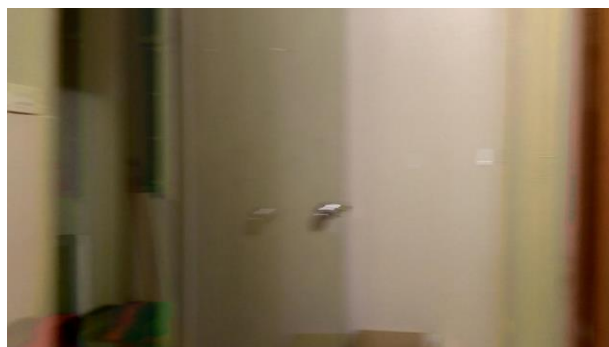


Figure 10: The result of method 3 in video with large motion, because it averages the luminance, we can see the ghost outside the door.

Figure 11 will show some results of method 1 in video without big motion. In stable video, the luminance will change smoother.



Figure 11: The result of method 1 in stable video.



Figure 12: The result of method 2 in stable video.

In Figure 12, we show the results of method 2 for stable movie. Because the motion of foreground and background are not the same, it has some ghosts.

In the last, Figure 13 shows the results of method 3. The same as results of large motion video. It has some ghosts. In addition, we show the origin frame of stable movie in Figure 14.



Figure 13: The result of method 2 in stable video.



Figure 14: The original frame of stable video.

7. REFERENCES

- [1] T. H. Wang, C. W. Chiu, W. C. Wu, J. W. Wang, C. Y. Lin, C. T. Chiu, and J. J. Liou "Pseudo-Multiple-Exposure-Based Tone Fusion with Local Region Adjustment," *IEEE Transactions on Multimedia*, Vol. 17, No. 4, pp. 470 – 484, 2015.

5. FUTURE WORK

Now we propose three methods to generate pseudo HDR video using inverse tone mapping. When the video is stable enough, the result of our method is not bad; but when the video has big motion, the result will be bad. In the future, we will try to research a more robust method to generate pseudo HDR video.

6. CONCLUSION

Due to the development of the internet, video became more and more important. We propose three methods to generate pseudo HDR video, which apply iTMOs for each frame and add some constraints for neighbor frames. In the stable video, it will work; but in normal use, it is not enough. This insufficient is remaining to be future work.