Intensity Mapping Functions For HDR Panorama Imaging: Weighted Histogram Averaging

Yilun Xu, Zhengguo Li, Senior Member, IEEE, Weihai Chen*, Member, IEEE, and Changyun Wen, Fellow, IEEE

Abstract—It is challenging to stitch multiple images with different exposures due to possible color distortion and loss of details in the brightest and darkest regions of input images. In this paper, a novel intensity mapping algorithm is first proposed by introducing a new concept of weighted histogram averaging (WHA). The proposed WHA algorithm leverages the correspondence between the histogram bins of two images which are built up by using the non-decreasing property of the intensity mapping functions (IMFs). The WHA algorithm is then adopted to synthesize a set of differently exposed panorama images. The intermediate panorama images are finally fused via a state-of-the-art multi-scale exposure fusion (MEF) algorithm to produce the final panorama image. Extensive experiments indicate that the proposed WHA algorithm significantly surpasses the related state-of-the-art intensity mapping methods. The proposed high dynamic range (HDR) stitching algorithm also preserves details in the brightest and darkest regions of the input images well. The related materials will be publicly accessible at https://github.com/yilun-xu/WHA for reproducible research.

Index Terms—Image stitching, color mapping, intensity mapping functions, weighted histogram averaging, multi-scale exposure fusion.

I. INTRODUCTION

Panorama is a visual representation of the environment viewed from one 3D position. Since the field of view (FoV) of mobile phones or digital cameras is limited, a panorama image is often obtained by stitching a series of still sub-images which cover certain areas of the scene [1]. Exposures of sub-images corresponding to different areas are often different for a scene, especially for a high dynamic range (HDR) scene. If all the sub-images are shot with the same exposure, the limited dynamic range of the camera makes the panorama images lose a lot of information in the brightest or darkest areas [2]. The exposures are thus flexibly adjusted when shooting different areas such that the generated panorama images can record richer scene information [3]. The resultant sub-images are differently exposed low dynamic range (LDR) images. There are two alternative ways to stitch all the LDR sub-images. One is to produce an LDR panorama image as in [4], [5]. The other is to generate an HDR panorama image as in [3], [6], [7]. An example is shown in Fig. 1. Input images in Fig. 1(a) can be captured via the normal panorama mode of the latest Google pixel phone [8]. They can also be captured by other smart phones with inertial measurement units (IMUs) [9]. Details in the brightest and darkest regions are preserved much better in the Fig. 1(c) than LDR panorama images the Fig. 1(b)-(d). Therefore, HDR panorama imaging is highly demanded due to its outstanding performance.

Existing HDR stitching algorithms [3], [6], [7] used camera response functions (CRFs) [11] to map the LDR sub-images to the corresponding HDR irradiance maps for processing, and then tone-mapped [12] the synthesized HDR panorama image into a high-quality LDR panorama image to match the dynamic range of the display. In recent years, multi-exposure image fusion (MEF) algorithms were proposed to improve the computational efficiency and convenience of HDR imaging chain [13]–[17]. The algorithms can directly obtain high-quality LDR images for displays by weighted fusion of the LDR images. It can simplify the pipeline and get rid of the restriction from the lighting condition and devices. It is thus desired to develop a simpler HDR stitching algorithm than the algorithms in [3], [6], [7] by leveraging the existing MEF algorithms.

In this paper, a novel HDR stitching algorithm is proposed by introducing a histogram bins-based intensity mapping algorithm [1]. The proposed intensity mapping algorithm includes estimation and correction of intensity mapping functions (IMFs). The proposed intensity mapping method is inspired by two methods in [2], [18]. The method in [18] first sets up a pixel level correspondence between the two differently exposed images, and then computes the mapped value through averaging over all the corresponding pixels in the other image. The method in [18] is very accurate if there is neither camera movement nor moving objects in the two images. However, the pixel-level correspondence is very sensitive to the camera movement and moving objects. The method in [2] shows that the histogram-bin-level correspondence is robust to the camera movement and moving objects even though the accuracy of the method in [2] is an issue. Inspired by the above two methods, a novel histogram-bin-level correspondence is first built up between the two differently exposed images, i.e., each bin in the histogram of one image corresponds to one unique segment bins in the histogram of the other image, by using the non-decreasing property of the IMFs. The mapped value is defined as the weighted average of intensities in the matched segment rather than the intensity of a single bin in the matched segment as in [2]. The proposed method is thus termed

* Corresponding author: Weihai Chen.

Yilun Xu and Weihai Chen are with the School of Automation Science and Electrical Engineering, Beihang University, Beijing, 100191 (e-mail: yilunxu_buaa@163.com, and whchen@buaa.edu.cn).

Zhengguo Li is with SRO Department, Institute for Infocomm Research, 1 Fusionopolis Way, Singapore (email: egzl@i2r.a-star.edu.sg).

Changyun Wen is with the School of Electrical and Electronics Engineering, Nanyang Technological University, Singapore (email: ecywen@ntu.edu.sg).

1Intensity mapping algorithms can also be referred to as color mapping or color correction algorithms.
weighted histogram averaging (WHA). Clearly, the proposed WHA algorithm is totally different from the methods in [2], [18] even though it is inspired by them. Since the overlapping area of the sub-image does not necessarily contain all the pixel values in the dynamic range, the calculated IMFs will contain some empty values. To make IMFs cover the whole dynamic range, linear interpolation and extrapolation are used to fill in the empty values after the IMFs are estimated. The proposed WHA algorithm is more accurate than the algorithm in [2] and is more robust than the algorithm in [18] with respect to the camera movement and moving objects.

Instead of correcting all the sub-images using the IMFs to directly produce a final LDR panorama image as in [4], a set of differently exposed panorama images are synthesized by using the IMFs and fused into a final HDR panorama image in this paper. The brightness of each sub-image is selected as a benchmark, and the brightness of the remaining sub-images is calibrated using the proposed IMFs. All the corrected sub-images are stitched together to form an LDR panorama image. By continuously changing the benchmark sub-image, the differently exposed LDR panorama images can be obtained. The IMFs can effectively prevent the brightness distortion from appearing in the differently exposed LDR panorama images [19]. The MEF algorithm in [20] is finally adopted to fuse all the differently exposed panorama images because the experimental results in [10] show that the MEF algorithm adopted outperforms other MEF algorithms from the MEF-SSIM point of view [21]. Details in the brightest and darkest regions of the input images are well preserved by the proposed stitching algorithm. Experimental results validate the proposed algorithms. Overall, two main contributions of this paper are: 1) a novel IMF estimation algorithm which outperforms existing IMF estimation algorithms including [2], [18]; and 2) a simpler HDR stitching algorithm than the algorithms in [3], [6], [7].

The rest of this paper is organized as follows. Relevant works are reviewed in Section II. Details of the proposed algorithm are presented in Section III. Extensive experimental results are provided in Section V. Lastly, concluding remarks are listed in Section VI.

II. RELATED WORKS

Due to the limited FoV of mobile phone or digital camera, a set of still sub-images which cover certain areas of the scene is captured and stitched together to produce a panorama image [1]. Exposure of all the sub-images is usually differently because of the different lighting conditions in the covered areas, especially for those sub-images with large camera view movements [22]. Due to the different exposures, there could be large intensity changes among differently exposed sub-images [13], [20]. Intensity mapping algorithms are usually applied to normalize all the sub-images before they are stitched together. Relevant works on intensity mapping are first summarized, and details on the two most relevant intensity mapping algorithms are then highlighted in this section.

A. Existing Intensity Mapping Algorithms

Many color matching algorithms have been proposed in a dozen years [2], [4], [5], [18] because they have broad application in many fields such as panorama stitching, image fusion, video intensity mapping, etc [23], [24]. Intensity mapping algorithms can be divided into model-based parametric algorithms and statistics-based non-parametric algorithms [24]. The model-based parametric algorithms assume that the intensity mapping from the original image to the target image follows a known mapping model. The statistics-based non-parametric algorithms do not build an explicit mapping model.
but estimate the look-up table from the relevant statistics of the original image and target image, and then use the look-up table as IMFs to calibrate the image.

Among the model-based parametric algorithms, the intensity mapping algorithm in [25] rearranges the color distribution of the source image so that the mean and variance of the original image and the target image are consistent. In order to decouple color channels, color space is converted into the log space representation so that the transformation can be applied on the special space. The gain compensation algorithm is introduced in [26] to address symmetric color balancing by a least-square minimization and is widely used in panoramic stitching tasks. The affine mapping algorithm in [27] uses probabilistic moving least squares to solve for a fully nonlinear and nonparametric intensity mapping in the RGB color space. When images come from entirely different domains, model-based parametric algorithms will lack a certain degree of versatility and convenience, and statistics-based non-parametric algorithms may be more appropriate [25]. Accuracy of the model-based parametric algorithms could be an issue even though they are simple.

Many statistics-based non-parametric algorithms have been proposed to estimate the look-up table as IMFs. The intensity mapping algorithm based on the CRFs [11] is widely used in HDR imaging [16], [28]. The CRFs which are calculated by a sequence of multiple exposure images and camera parameters can be used to calibrate the brightness of the images. Algorithms based on truncated Gaussian models [29] and algorithms based on spline models [30], [31] are proposed to estimate the IMFs from overlapping regions. The algorithm in [2] estimates the IMFs by matching the cumulative histogram of two images. They can be applied for image stitching, removing ghosting artifacts [32], etc. A intensity mapping algorithm is proposed in [4] to improve the cumulative histogram-based algorithm by using hybrid histogram matching, and it achieves state-of-the-art (SOTA) performance in the multi-camera intensity mapping. But when the exposure differences among the input images become large, the performance of the algorithm in [4] will drop. An algorithm is proposed in [18] for inpainting of differently exposed images. The algorithm uses geometrical correspondence between images to calculate the IMFs. The accuracy of the algorithm in [18] is higher than the algorithms in [2], [4] for two perfectly aligned images. However, when there is camera movement between two images, the accuracy of the algorithm in [18] will drop significantly. The accuracy is also an issue when there are moving objects. A new statistics-based non-parametric intensity mapping method will be proposed to overcome these problems in this paper. Same as the existing statistics-based non-parametric algorithms, the proposed algorithm has a high accuracy. It is also robust to the camera movement and moving objects.

B. Two Most Relevant Intensity Mapping Algorithms

The IMFs between two images are determined by the camera, the exposures between two images and the scene information [2]. In this subsection, the details of the two most relevant works on the estimation of IMFs are provided in order to facilitate readers to better understand the new algorithm in Section 4. The IMFs are calculated in the RGB color space. Here, an example is shown for a single color channel, but the IMFs can be performed separately for each channel.

1) Cumulative Histogram Based IMFs: Let \( Z_i \) and \( Z_j \) be two differently exposed images. The pixel's position index and intensity of \( Z_i \) are denoted by \( p \) and \( z \), respectively, where \( p = (p_x, p_y) \). Let \( \Lambda_{i\rightarrow j}(z) \) be the IMF from the image \( Z_i \) to the image \( Z_j \), and the mapped image \( Z_i \rightarrow j \) can be calculated as follows

\[
Z_i \rightarrow j(p) = \Lambda_{i\rightarrow j}(Z_i(p))
\]  

(1)

The position index set of pixels with the intensity value \( z \) in the image \( Z_i \), \( \Omega_i(z) \), is defined as

\[
\Omega_i(z) = \{p|Z_i(p) = z\}.
\]  

(2)

The IMF \( \Lambda_{i\rightarrow j}(z) \) is estimated by using the cumulative histograms of the images \( Z_i \) and \( Z_j \) in [2]. Let \( H_i(k) \) be the cardinality of the set \( \Omega_i(z) \), and it represents the value of \( z \)th histogram bin of image \( Z_i \). The width of the histogram bins defaults to 1. The cumulative histogram \( C_i(z) \) of the image \( Z_i \) is then created by

\[
C_i(z) = \sum_{k=0}^{z} H_i(k).
\]  

(3)

The estimation method in [2] is based on two assumptions: 1) the IMFs have non-decreasing property, and 2) no saturation and quantization problems in the image. Suppose that the intensity value \( z_1 \) in the image \( Z_i \) maps to the intensity value \( z_2 \) in the image \( Z_j \). The number of pixels in \( Z_i \) whose intensity values are not greater than \( z_1 \) approaches the number of pixels in \( Z_j \) whose intensity values are not greater than \( z_2 \). Mathematically, the function \( \Lambda_{i\rightarrow j}(z) \) can be estimated as

\[
\Lambda_{i\rightarrow j}(z) = \arg\min_{z'} |C_i(z) - C_j(z')|.
\]  

(4)

The method in [2] exploits the non-decreasing properties of IMFs to build up the correspondence among the histogram bins of the images \( Z_i \) and \( Z_j \). The IMFs between the images \( Z_i \) and \( Z_j \) are estimated by using the histogram-bin-level correspondence. The estimated IMFs are robust to the camera and object motion in \( (Z_i, Z_j) \) since scene motion does not significantly change the distribution of image histograms (or cumulative histograms). However, when \( Z_i \) and \( Z_j \) have a large exposure ratio, the image saturation and quantization problems are obvious. Multiple values will map to one value, or one value will map to multiple values. Clearly, the accuracy of the estimated IMFs by the equation (4) is an issue.

2) Geometrical Correspondence Based IMFs: The estimation method in [18] is based on the assumption: the pixels in images \( Z_i \) and \( Z_j \) are aligned, i.e., there is neither camera nor object motion in the two images. This algorithm estimates the IMFs by 1) finding all pixels with the same intensity in one image, and 2) calculating the average of all the co-located
pixels in the other image. In other words, the function $\Lambda^2_{i\rightarrow j}(z)$ is estimated as

$$\Lambda^2_{i\rightarrow j}(z) = \sum_{p \in \Omega_i(z)} Z_j(p) H_i(z), \ H_i(z) \neq 0. \quad (5)$$

When there exists at least one $z$ such that $H_i(z) = 0$, there will be an 'empty value' problem. The corrected process is divided into two steps. First, a median filter is adopted starting from the middle of the valid values towards left and right separately. The second step extends two ends of the curve by using the neighbourhood slope.

When ignoring the 'empty value' problem, it can be easily verified that $\Lambda^2_{i\rightarrow j}(z)$ is actually obtained by solving the following optimization problem:

$$\arg\min_{z'} \sum_{p \in \Omega_i(z)} (z' - Z_j(p))^2. \quad (6)$$

Subsequently, the PSNR will be maximized by $\Lambda^2_{i\rightarrow j}(z)$.

In the equation (5), the IMFs are estimated by averaging over all the corresponding pixels in the image $Z_j$, this results in IMFs having sufficient accuracy for any exposure ratio between the images $Z_i$ and $Z_j$ if they are aligned. However, the assumption used in [18] is to build up a geometrical correspondence between pixels in the two differently exposed images, the pixel level correspondence is not true if there are camera and object motion in the two differently exposed images. The robustness of this method with respect to camera movement and/or moving objects is thus an issue.

III. The Proposed IMF Estimation Method

In this section, a novel IMF estimation algorithm that inherits the advantages of [2], [13] is introduced. The proposed algorithm outperforms the algorithms in [2], [18].

A. Initial Estimation of IMF via WHA

This new method uses the histogram bin level correspondence rather than the pixel level correspondence of the two differently exposed images as in [18]. The correspondence among the histogram bins of the images $Z_i$ and $Z_j$ can be built up by using the non-decreasing property of the IMFs as follows:

Consider two pixels $Z_i(p_{x_i}, p_{y_i})$ and $Z_j(p'_{x_j}, p'_y_j)$ which satisfy $Z_i(p_{x_i}, p_{y_i}) > Z_j(p'_{x_j}, p'_y_j)$. Suppose that the value of histogram bins in the image $Z_j$ corresponding to $Z_i(p_{x_i}, p_{y_i})$ and $Z_j(p'_{x_j}, p'_y_j)$ are $H_j(z)$ and $H_j(z')$, respectively. $z$ is not smaller than $z'$.

Since scene motion does not significantly change the distribution of image histograms, the above correspondence among the histogram bins is also true for the images $Z_i$ and $Z_j$ with camera movement and/or moving objects. The correspondence will be utilized to find all the sub-bins (or bins) in the image $Z_j$ corresponding to the bin $H_i(z)$. A non-decreasing mapping function $\psi_{i\rightarrow j}(z)$ is defined as

$$C_j(\psi_{i\rightarrow j}(z) - 1) < C_i(z) \leq C_j(\psi_{i\rightarrow j}(z)), \quad (7)$$

and $\psi_{i\rightarrow j}(-1)$ is predefined as 0.

According to the definitions of $C_i(z)$ and $C_j(z)$, $H_j(\psi_{i\rightarrow j}(z - 1))$ and $H_i(\psi_{i\rightarrow j}(z))$ are the first and last bins corresponding to the bin $H_i(z)$, respectively. Let the sizes of sub-bins (or bins) in the image $Z_j$ corresponding to the bin $H_i(z)$ be denoted by $\hat{H}_{i\rightarrow j}(k)$, and it is defined as in the following two cases:

**Case 1:** $\psi_{i\rightarrow j}(z - 1) < \psi_{i\rightarrow j}(z)$. $\hat{H}_{i\rightarrow j}(k)$ is defined as

$$\hat{H}_{i\rightarrow j}(k) = \begin{cases} C_j(k) - C_i(z - 1), & \text{if } k = \psi_{i\rightarrow j}(z - 1) \\ C_i(z) - C_j(k - 1), & \text{if } k = \psi_{i\rightarrow j}(z) \\ H_j(k), & \text{otherwise} \end{cases} \quad (8)$$

**Case 2:** $\psi_{i\rightarrow j}(z - 1) = \psi_{i\rightarrow j}(z)$. $\hat{H}_{i\rightarrow j}(k)$ is defined as

$$\hat{H}_{i\rightarrow j}(k) = H_i(z). \quad (9)$$

It can be easily verified that

$$\sum_{k=\psi_{i\rightarrow j}(z-1)}^{\psi_{i\rightarrow j}(z)} \hat{H}_{i\rightarrow j}(k) = H_i(z). \quad (10)$$

By using the correspondence among the histogram bins, the proposed IMF $\Lambda_{i\rightarrow j}(z)$ is then defined as

$$\Lambda_{i\rightarrow j}(z) = \sum_{k=\psi_{i\rightarrow j}(z-1)}^{\psi_{i\rightarrow j}(z)} \frac{\hat{H}_{i\rightarrow j}(k)}{H_i(z)} k, \ H_i(z) \neq 0. \quad (11)$$

It can be shown from the equation (11) that the proposed IMF $\Lambda_{i\rightarrow j}(z)$ is a weighted average of all the sub-bins (or bins) corresponding to the bins $H_i(z)$. Thus, the proposed IMF estimation algorithm is called the WHA. The Matlab code of the proposed WHA is given in the appendix. The definition of the function $\psi_{i\rightarrow j}(z)$ in the equation (7) plays a crucial role in the proposed WHA. It builds up an elegant histogram-bin-level correspondence between the two differently exposed images. Unlike the pixel-level correspondence in [18], the histogram-bin-level correspondence is robust to both the camera movement and the moving objects.

The difference between the proposed IMF and the IMF in the equation (4) is now presented. Let $z^* = \arg \min_{z' \in \{\psi_{i\rightarrow j}(z) - 1, \psi_{i\rightarrow j}(z)\}} \{ |C_i(z) - C_j(z')| \}$. (12)

It can be derived from the equations (4), (7), (10), and (12) that

$$\Lambda_{i\rightarrow j}(z) = \sum_{k=\psi_{i\rightarrow j}(z-1)}^{\psi_{i\rightarrow j}(z)} \frac{\hat{H}_{i\rightarrow j}(k)}{H_i(z)} z^*, \ H_i(z) \neq 0. \quad (13)$$

Clearly, $k$ in the formula (11) is replaced by $z^*$ in the formula (13). In other words, $\Lambda_{i\rightarrow j}(z)$ is determined by the intensity of a single bin in the matched segment.

Similar to the formula (5), the proposed WHA has a higher accuracy for any exposure ratio than the formula (4). In addition, compared with the intensity $Z_j(p)$ used in the formula (5), the histogram bin indicator $k$ used in the formula (11) is not sensitive to either the camera movement or the moving objects, so the proposed algorithm is more robust than formula (5). Overall, this new method is more accurate than [2] and more robust than [18].
B. Interpolation and Extrapolation of the Proposed IMF

Since the image $Z_i$ does not necessarily include pixels in all the dynamic range, there is often a problem of empty bins. When $H_i(z) = 0$, the function $A_{i \rightarrow j}(z)$ can be calculated by using two neighboring $A_{i \rightarrow j}(z_1)$ and $A_{i \rightarrow j}(z_2)$, and the formula is given as follows:

$$A_{i \rightarrow j}(z) = \frac{(z - z_1)(A_{i \rightarrow j}(z_1) - A_{i \rightarrow j}(z_2))}{z_1 - z_2} + A_{i \rightarrow j}(z_1), \quad (14)$$

where the $z_1$ and $z_2$ can be estimated from

$$\min \{|z - z_1| + |z - z_2|\} \quad \text{s. t.} \quad \lambda(z - z_1)(z - z_2) < 0, \quad (15)$$

where $\lambda$ is set as 1 for the interpolation, and -1 for the extrapolation.

IV. APPLICATION TO HDR PANORAMA IMAGING

Inputs of the proposed algorithm are $N(\geq 2)$ images to be stitched. Without loss of generality, consider the case that $N = 3$. All the three images $Z_1$, $Z_2$, and $Z_3$ have different exposures and partially overlapping areas. All the input images are supposed to be geometrically aligned, for example by the algorithms in [33]–[35]. Three differently exposed LDR images are synthesized as below.

1) Taking the brightness of $Z_1$ as the benchmark, the images obtained by calibrating the images $Z_2$ and $Z_3$ are denoted as $Z_{2 \rightarrow 1}$ and $Z_{3 \rightarrow 1}$. They are calculated as

$$\begin{align*}
Z_{2 \rightarrow 1} &= A_{2 \rightarrow 1}(Z_2) \\
Z_{3 \rightarrow 1} &= A_{3 \rightarrow 1}(Z_3),
\end{align*} \quad (16)$$

where the IMF $A_{3 \rightarrow 1}(z)$ is

$$A_{3 \rightarrow 1}(z) = A_{2 \rightarrow 1}(A_{3 \rightarrow 2}(z)). \quad (17)$$

The first intermediate LDR panorama image, $\tilde{Z}_1$, is obtained by stitching all the images $Z_1$, $Z_{2 \rightarrow 1}$, and $Z_{3 \rightarrow 1}$. The brightness of the image $\tilde{Z}_1$ is same as that of the image $Z_1$.

2) Taking the brightness of $Z_2$ as the benchmark, the images obtained by calibrating the image $Z_1$ and $Z_3$ are denoted as $Z_{1 \rightarrow 2}$ and $Z_{3 \rightarrow 2}$. They are computed as

$$\begin{align*}
Z_{1 \rightarrow 2} &= A_{1 \rightarrow 2}(Z_1) \\
Z_{3 \rightarrow 2} &= A_{3 \rightarrow 2}(Z_3),
\end{align*} \quad (18)$$

The second intermediate LDR panorama image, $\tilde{Z}_2$, is obtained by stitching all the images $Z_{1 \rightarrow 2}$, $Z_2$, and $Z_{3 \rightarrow 2}$. The brightness of the image $\tilde{Z}_2$ is same as that of the image $Z_2$.

3) Taking the brightness of $Z_3$ as the benchmark, the images obtained by calibrating the images $Z_1$ and $Z_2$ are denoted as $Z_{1 \rightarrow 3}$ and $Z_{2 \rightarrow 3}$. They are calculated as

$$\begin{align*}
Z_{1 \rightarrow 3} &= A_{1 \rightarrow 3}(Z_1) \\
Z_{2 \rightarrow 3} &= A_{2 \rightarrow 3}(Z_2),
\end{align*} \quad (19)$$

where the IMF $A_{1 \rightarrow 3}(z)$ is computed as

$$A_{1 \rightarrow 3}(z) = A_{2 \rightarrow 3}(A_{1 \rightarrow 2}(z)). \quad (20)$$

The third intermediate LDR panorama image, $\tilde{Z}_3$, is obtained by stitching all the images $Z_{1 \rightarrow 3}$, $Z_{2 \rightarrow 3}$, and $Z_3$. The brightness of the image $\tilde{Z}_3$ is same as that of the image $Z_3$.

The final HDR panorama image $\tilde{Z}$ can be obtained by fusing all the three differently exposed panorama images $\tilde{Z}_1$, $\tilde{Z}_2$, and $\tilde{Z}_3$ through the MEF algorithm in [20]. It was shown in [10] that the MEF algorithm in [20] outperforms other MEF algorithms from the MEF-SSIM point of view [21]. Details in the brightest and darkest regions of the three images $Z_1$, $Z_2$ and $Z_3$ are preserved better in the image $\tilde{Z}$ than $\tilde{Z}_1$, $\tilde{Z}_2$, and $\tilde{Z}_3$.

V. EXPERIMENT RESULTS

Extensive experimental results are provided to verified the proposal algorithms. Readers are invited to view to electronic version of figures and zoom in them so as to better appreciate differences among all images.

A. Implementation Details

1) Dataset Description: Experiments are conducted on two datasets, VETHDR-Nikon dataset and HDR-SICE dataset. Between them, the VETHDR-Nikon dataset is used to evaluate the performance of IMFs, and the HDR-SICE dataset is used to show the results of the MEF-based HDR stitching algorithm.

The VETHDR-Nikon dataset is from [36], collected by a Nikon 7200 camera. It consists of 495 pairs of images in 8-bit JPEG files format. Each pair consists of two full-resolution images, with long and short exposure times, respectively. The exposure time ratio and ISO are fixed as 8 and 800, during shooting respectively. All the images are resized to 1000 × 1600. Camera shaking, object movement are strictly controlled to ensure that only the illumination is changed.

The HDR-SICE dataset is from [10], collected by 7 different models of cameras. It includes 589 sequences from indoor and outdoor scenes, containing a total number of 4, 413 multiexposure images. All the images are resized to 520 × 1080. The multiple sequences with EV shifted by ±(0.5, 0.7, 1.0, 2.0, 3.0) are well-aligned. In each image sequence, we first select three images with different exposures and then crop 520 × 400 size images from different positions of the three images as input. The size of the overlap area between the input adjacent images is 520 × 60.

2) IMFs Comparison Description: In the VETHDR-Nikon dataset [36], the overlapping areas in each pair of images are well-aligned. However, in actual situations, since the shooting angles of the two images to be stitched are different, the overlapping areas are always not well-aligned. In order to simulate the real situation of the overlapping area, Fig. [2] shows
how to process the input image. Specifically, the $N_c$ columns of pixels on the left and $N_c$ rows of pixels on the bottom of the image (a) are cut off, the $N_c$ columns of pixels on the right and $N_c$ rows of pixels on the top of the image (b) are also cut off. The remaining areas are used as the simulated overlap area. The IMFs are first estimated by the simulated overlap areas and then used to correct the brightness of image (a) or (b).

Fig. 2. An example of simulating overlapping areas. (a) and (b) are the input image pairs, (c) and (d) show the way to simulate the overlapping area of (a) and (b), respectively.

Moreover, in order to test the IMFs more comprehensively, each pair of images in the VETHDR-Nikon dataset will be tested twice. Specifically, use the dark image as the reference to correct the bright image, and then use the bright image as the reference to correct the dark image. So each IMF estimation algorithm will be tested 990 times on the VETHDR-Nikon dataset.

In terms of quantitative comparison, we choose PSNR, SSIM, FSIM, iCID and running time as the evaluation metrics. All results are tested on a laptop with Intel Core i7-9750H CPU 2. 59GHz, 32. 0 GB memory and Matlab R2019a installed. The all evaluation metrics are explained as follows:

- **PSNR**: Peak signal-to-noise ratio, the higher score comes with the better performance.
- **SSIM**: Structural similarity [37], the higher score comes with the better performance.
- **FSIM**: Feature similarity [38], the higher score comes with the better performance.
- **iCID**: Improved color image difference [39], the lower score comes with the better performance.
- **Time**: The time it takes for the IMFs algorithm to map a image, it does not include the time to read and save the image.

Table I summarizes the IMF algorithms involved in the experiments.

### B. Ablation Study

In this section, on the one hand, we show that for the empty-bins problem in WHA, employing linear interpolation is a simple and straightforward method. On the other hand, we explore the variation of the accuracy of the WHA algorithm under different degrees of dislocation through experiments, which proves that the WHA algorithm is more robust and accurate than the traditional method.

1) **Effectiveness of Interpolation Algorithm**: Experiments on different models are conducted to validate the interpolation algorithm as a simple and straightforward method.

The HHM algorithm [4] proposes an interesting way to correct the IMF, which can be adopted in the proposed algorithm to deal with the empty-bins problem while correcting the IMF curve. This method of applying the correction method in the HHM to the proposed WHA is abbreviated as WHA-HHM. The test results are shown in Table II. When the correction method in the HHM algorithm is adopted to correct the WHA, the performance of the algorithm is degraded a lot. This implies that the correction method in the HHM algorithm [4] is not applicable to the proposed WHA.

### TABLE I

| Method | Summary Description |
|--------|---------------------|
| CHM [2] | using cumulative histogram |
| GC [18] | using geometric correspondence between pixels |
| HHM [1] | using hybrid histogram matching algorithm |
| PCRF [15] | using prior information about camera response function |
| TG [29] | using histogram simulated by truncated Gaussians |
| AM [27] | using affine map |
| 3MS [30] | using monotonic splines and three-channel correlation |
| GPS [31] | using gradient preserving spline |
| MV [25] | using the mean and variance of the images |
| WHA | using weighted histogram averaging algorithm |

### TABLE II

| Method | PSNR↑ | SSIM↑ | FSIM↑ | iCID(%)↓ | Time(s)↓ |
|--------|-------|-------|-------|---------|---------|
| WHA-HHM | 30.91 | 0.9033 | 0.9698 | 7.79 | 0.09 |
| WHA | 34.38 | 0.9153 | 0.9815 | 4.77 | 0.08 |

2) **Robustness and Accuracy Analysis**: In order to verify that the WHA algorithm can inherit the high robustness of the CHM algorithm [2] and the high accuracy of the GC algorithm [18], the three algorithms are compared in different situations quantitatively. Specifically, when using the simulated overlap area to estimate the IMFs, the value of $N_c$ is changed to simulate the different degrees of misalignment between the overlap areas. The larger $N_c$, the greater the degree of misalignment in the overlapping area. Note that when $N_c = 0$, because the empty value problem in the GC algorithm will not affect the accuracy, there is no need to use the corrected process in section II-B2. When $N_c \neq 0$, the corrected process will be used.

The quantitative comparison results are shown in Table III. It can be seen that when the overlapping area is well aligned ($N_c = 0$), the accuracy of GC is the highest, and the accuracy of CHM is the lowest, but the accuracy of the WHA is very close to that of GC. This shows that the WHA
Quantitative comparison among the CHM, GC and the proposed WHA in different numbers of offset pixels. The best results are shown in bold, and the second-best results are shown in red. The results come from the VETHDR-Nikon dataset.

| NC | PSNR↑ | SSIM↑ | FSIM↑ | iCID(%)↓ |
|----|-------|-------|-------|---------|
| 0  | WHA   | CHM   | GC    | WHA     | CHM     | GC     | WHA   | CHM   | GC    | WHA   | CHM   | GC    | W        | C        | G        |
|    |       |       |       |         |         |        |       |       |       |       |       |       |         |         |         |
| 2  | 34.54 | 32.49 | 34.83 | 0.9156  | 0.9032  | 0.9210 | 0.9817 | 0.9743 | 0.9814 | 4.70  | 7.04  | 4.45  | 4.70    | 7.04    | 4.45    |
| 4  | 34.51 | 32.46 | 30.45 | 0.9156  | 0.9031  | 0.8927 | 0.9817 | 0.9743 | 0.9619 | 4.71  | 7.05  | 8.54  | 4.70    | 7.04    | 8.54    |
| 6  | 34.47 | 32.43 | 29.54 | 0.9155  | 0.9031  | 0.8849 | 0.9816 | 0.9742 | 0.9546 | 4.73  | 7.06  | 9.70  | 4.75    | 7.08    | 10.83   |
| 8  | 34.43 | 32.40 | 28.87 | 0.9154  | 0.9030  | 0.8774 | 0.9816 | 0.9741 | 0.9467 | 4.75  | 7.08  | 10.83 | 4.75    | 7.08    | 10.83   |
| 10 | 34.38 | 32.36 | 28.17 | 0.9153  | 0.9029  | 0.8693 | 0.9815 | 0.9740 | 0.9385 | 4.77  | 7.10  | 11.98 | 4.79    | 7.13    | 12.99   |
| 12 | 34.32 | 32.31 | 27.64 | 0.9152  | 0.9027  | 0.8622 | 0.9814 | 0.9739 | 0.9314 | 4.79  | 7.13  | 12.99 | 4.82    | 7.16    | 13.79   |
| 14 | 34.25 | 32.25 | 27.24 | 0.9150  | 0.9025  | 0.8571 | 0.9812 | 0.9738 | 0.9258 | 4.82  | 7.16  | 13.79 | 4.86    | 7.19    | 14.64   |
| 16 | 34.18 | 32.20 | 26.87 | 0.9148  | 0.9024  | 0.8514 | 0.9811 | 0.9737 | 0.9196 | 4.86  | 7.19  | 14.64 | 4.86    | 7.19    | 14.64   |

Fig. 3. Qualitative comparison among the existing IMFs’ algorithms and the proposed WHA. (a) Original Images, (b) Ground truth, (c) the CHM [2], (d) the HHM [4], (e) the PCRF [16], (f) the GPS [31], (g) the proposed WHA. The inputs are from the VETHDR-Nikon dataset [36].
has inherited the high precision of the GC well. When the degree of misalignment in the overlapping area becomes larger \((N_c \uparrow)\), the sensitivity of GC to misalignment causes the accuracy to drop quickly, but the accuracy of the WHA and CHM only drops a little bit. This shows that the WHA also inherits the robustness of the CHM against dislocation. As the WHA inherits the advantages of the GC and CHM, the WHA maintains the highest accuracy in the case of non-alignment \((N_c \neq 0)\).

C. Comparison with Existing Intensity Mapping Algorithms

The proposed WHA algorithm is compared with nine existing intensity mapping algorithms in [2], [4], [16], [18], [25], [27], [29]–[31] qualitatively and quantitatively on the VETHDR-Nikon dataset [36]. Table I summarizes all these algorithms.

\(N_c\) is set to 10 in order to verify the accuracy and robustness of all the IMF estimation algorithms simultaneously. The results of 5 evaluation metrics are reported in Table IV. With the second fastest running speed, our method achieves the best performance in all the PSNR, SSIM, FSIM, and iCID.

For the qualitative comparison, several representative methods were selected. The visualization results and their detailed parts are shown in Fig. 3. In the first row, the color of the door is mapped incorrectly by the algorithms in (d), (e) and (f). In the second row, the luminance of images in (d) and (f) is different from the ground-truth, and the color of images in (c),
TABLE IV
QUANTITATIVE COMPARISON AMONG THE EXISTING INTENSITY MAPPING ALGORITHMS. THE BEST RESULTS ARE SHOWN IN BOLD, AND THE SECOND-BEST RESULTS ARE SHOWN IN RED.

| Method   | PSNR↑  | SSIM↑  | FSIM↑   | iCID(%)↓ | Time(s)↓ |
|----------|--------|--------|---------|----------|----------|
| HHM [4]  | 32.36  | 0.9029 | 0.9740  | 7.10     | 0.9740   |
| PCRF [17]| 29.34  | 0.8897 | 0.9391  | 11.35    | 12.52    |
| TG [27]  | 25.68  | 0.8413 | 0.8882  | 18.49    | 0.54     |
| AM [27]  | 30.96  | 0.9857 | 0.9455  | 11.17    | 7.70     |
| 3MS [30] | 28.14  | 0.8964 | 0.9470  | 10.28    | 0.66     |
| GPS [31] | 28.17  | 0.8943 | 0.9643  | 9.36     | 151.09   |
| MV [25]  | 32.36  | 0.9029 | 0.9740  | 7.10     | 1.77     |
| GC [13]  | 30.96  | 0.9857 | 0.9455  | 11.17    | 7.70     |
| WHA [4]  | 34.38  | 0.9153 | 0.9815  | 4.77     | 0.08     |

(e) and (f) is abnormal. In a nutshell, the results in (d), (e) and (f) are likely to have abnormal color and luminance, the results in (e) sometimes have abnormal color when mapped from lighter images. Different from these methods, the proposed algorithm is robust to most scenes, providing accurate results. The overall comparison demonstrates the effectiveness of the proposed WHA algorithm.

D. Comparison of Different Stitching Methods

This subsection compares the proposed HDR stitching method with traditional LDR stitching algorithms qualitatively. By using the traditional methods including the CHM [2], the HHM [4], the GPS [31], and the 3MS [30], all the input images are aligned with the image with the medium exposure such that the panoramic images have the least overexposed or underexposed area.

As shown in figure 4, there are two main problems for the traditional LDR stitching algorithms. Due to the inaccuracy of the IMF estimation algorithm, panoramic images could suffer from color shift, and the brightness transition at the splicing position might be unnatural. Meanwhile, fine details in the darkest or brightest regions are not preserved well. On the other hand, the proposed HDR panorama stitching algorithm uses the WHA with high precision, so there is no color shift problem. Furthermore, the stitching algorithm first generates three LDR panoramic images with different exposures, and then uses the MEF algorithm in [20] to merge them into an HDR panoramic image. The resultant HDR panoramic image usually contains scene information as much as possible, and the global brightness is also more comfortable. For example, the sky and grass in the first column are very clear, the overall light and shade in the second column are moderate, and the third column can clearly show the scene outside the train and the tunnel simultaneously.

E. Further Application to Ghost Removal

The stitched image could suffer from ghosting artifacts due to moving objects in the overlapping areas. In this subsection, the proposed IMF estimation method is applied to study ghost removal of differently exposed LDR images.

There are many ghost removal algorithms for differently exposed images [32], [40]–[42]. The proposed IMF estimation algorithm is adopted to improve the ghost removal algorithm in [40]. One of the differently exposed images is selected as the reference image. Pixels in all other images are classified into consistent and inconsistent pixels. All the inconsistent pixels in each image are corrected by using the correlation between the reference image and other images. The IMFs which are estimated by the algorithm in [2] are used to detect all the inconsistent pixels and correct them. The estimation algorithm in [2] is replaced by the proposed algorithm. As shown in Figure 5, there are serious color distortion by using the estimation method in [2] which can be shown more clearly by zooming in those images in the column (b). Fortunately, the color distortion is significantly reduced by the proposed estimation method.

VI. CONCLUSION REMARKS AND DISCUSSION

In this paper, a novel intensity mapping estimation algorithm is proposed by introducing a new concept of weighted histogram averaging (WHA), and it is applied to study image stitching. Instead of only correcting the color and brightness difference as in the existing image stitching algorithms, a set of differently exposed low dynamic range (LDR) panorama images is first synthesized and then fused via a multi-scale exposure fusion algorithm to produce the desired HDR panorama images. Details in the brightest and darkest regions of the input images are preserved in the HDR panorama images much better than traditional LDR stitching algorithms.
The proposed intensity mapping estimation algorithm can be applied to estimate optical flows and disparity of differently exposed images. The stitched image can be further improved by using existing detail enhancement algorithms such as [43] and the neural augmentation in [28]. All these problems will be studied in our future research.

REFERENCES

[1] M. Brown, D. G. Lowe et al., “Recognising panoramas,” in ICCV, vol. 3, 2003, p. 1218.
[2] M. D. Grossberg and S. K. Nayar, “Determining the camera response from images: What is knowable?”, IEEE Transactions on pattern analysis and machine intelligence, vol. 25, no. 11, pp. 1455–1467, 2003.
[3] A. Eden, M. Uyttendaele, and R. Szeliski, “Seamless image stitching of scenes with large motions and exposure differences,” in 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’06), vol. 2, IEEE, 2006, pp. 2498–2505.
[4] C. Ding and Z. Ma, “Multi-camera color correction via hybrid histogram matching,” IEEE Transactions on Circuits and Systems for Video Technology, 2020.
[5] F. Bellavia and C. Colombo, “Dissecting and reassembling color correction algorithms for image stitching,” IEEE Transactions on Image Processing, vol. 27, no. 2, pp. 735–748, 2018.
[6] P. Osiak, “Implementation of hdr panorama stitching algorithm,” in Proceedings of the 10th CESC Conference. Citeseer, 2004, pp. 24–26.
[7] Y. Li, F. Poclich, K. Black, C. Piatig, Q. Cogal, and Y. Leblebici, “Multi-camera platform for panoramic real-time hdr video construction and rendering,” Journal of Real-Time Image Processing, vol. 12, no. 4, pp. 697–708, 2016.
[8] “Google pixel phone: panorama camera modes,” https://www.youtube.com/watch?v=HzD7jY3P7PK, 2017.
[9] T. Hanvey, Z. Li, W. Young, and G. Seet, “Accurate inu factor using switched linear systems for vio,” IEEE Transactions on Industrial Electronics, vol. 68, no. 8, pp. 7199–7208, 2021.
[10] J. Cai, S. Gu, and L. Zhang, “Learning a deep single image contrast enhancer from multi-exposure images,” IEEE Transactions on Image Processing, vol. 27, no. 4, pp. 2049–2062, 2018.
[11] E. Reinhard, J. Debevec, and J. Malik, “Recovering high dynamic range radiance and the neural augmentation in,” in 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE, 2010, pp. 263–270.
[12] E. Reinhard, M. Adhikar, B. Gooch, and P. Shirley, “Color transfer between images,” IEEE Computer graphics and applications, vol. 21, no. 5, pp. 34–41, 2001.
[13] M. Brown and D. G. Lowe, “Automatic panoramic image stitching using invariant features,” International journal of computer vision, vol. 74, no. 1, pp. 59–73, 2007.
[14] Y. Hwang, J.-Y. Lee, I. So Kweon, and S. Joo Kim, “Color transfer using probabilistic moving least squares,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2014, pp. 3342–3349.
[15] F. Kou, Z. Li, Y. Yang, and S. Wu, “Single image brightening via multi-scale exposure fusion with hybrid learning,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 31, no. 4, pp. 1425–1435, 2020.
[16] M. Oliveira, A. D. Sappa, and V. Santos, “A probabilistic approach for color correction in image mosaicking applications,” IEEE Transactions on Image Processing, vol. 24, no. 2, pp. 508–523, 2014.
[17] H. Faridul, J. Stauder, J. Kervec, and A. Tréméau, “Approximate cross channel color mapping from sparse color correspondences,” in Proceedings of the IEEE international conference on computer vision workshops, 2013, pp. 860–867.
[18] F. Bellavia and C. Colombo, “Dissecting and reassembling color correction algorithms for image stitching,” IEEE Transactions on Image Processing, vol. 27, no. 2, pp. 735–748, 2017.
[19] J. Zheng, Z. Li, Z. Zhu, S. Wu, and S. Rahardja, “Robust alignment for panoramic stitching via an exact rank constraint,” IEEE Transactions on Image Processing, vol. 28, no. 10, pp. 4730–4745, 2019.
[20] T. T. Liao and N. Li, “Single-perspective warps in natural image stitching,” IEEE Transactions on Image Processing, vol. 29, pp. 724–735, 2020.
[21] X. Zhao, J. Liu, X. Wu, W. Chen, F. Guo, and Z. Li, “Probabilistic spatial distribution prior based attentional keypoints matching network,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 32, no. 3, pp. 1313–1327, 2022.
[22] C. Zheng, Z. Li, Y. Yang, and S. Wu, “Color transfer using probabilistic moving least squares,” in Proceedings of the IEEE international conference on computer vision workshops, 2013, pp. 860–867.
[23] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” IEEE Trans. on Image Processing, vol. 13, no. 4, pp. 600–612, 2004.
[24] L. Zhang, L. Zhang, X. Mou, and D. Zhang, “Fism: A feature similarity index for image quality assessment,” IEEE Transactions on Image Processing, vol. 20, no. 8, pp. 2378–2386, 2011.
[25] J. Preiss, F. Fernandes, and P. Urban, “Color-image quality assessment: From prediction to optimization,” IEEE Transactions on Image Processing, vol. 23, no. 3, pp. 1366–1378, 2014.
[26] Z. Li, J. Zheng, Z. Zhu, and S. Wu, “Selectively detail-enhanced fusion of differently exposed images with moving objects,” IEEE Transactions on Image Processing, vol. 23, no. 10, pp. 4372–4382, 2014.
[27] K. Myszkowski, K. P. Singh, and R. V. Babu, “Colour mapping: A review of recent methods, approaches for automatic multi-view image and video stitching,” in 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE, 2010, pp. 263–270.