# Multiview High Dynamic Range Image Synthesis Using Fuzzy Broad Learning System

Hongbin Guo, Bin Sheng, Ping Li, *Member, IEEE*, and C. L. Philip Chen, *Fellow, IEEE*

***Abstract*—Compared with the normal low dynamic range (LDR) images, the high dynamic range (HDR) images provide more dynamic range and image details. Although the existing techniques for generating the HDR images have a good effect for static scenes, they usually produce artifacts on the HDR images for dynamic scenes. In recent years, some learning-based approaches are used to synthesize the HDR images and obtain good results. However, there are also many problems, including the deficiency of explaining and the time-consuming training process. In this article, we propose a novel approach to synthesize multiview HDR images through fuzzy broad learning system (FBLS). We use a set of multiview LDR images with different exposure as input and transfer corresponding Takagi–Sugeno (TS) fuzzy subsystems; then, the structure is expanded in a wide sense in the "enhancement groups" which transfer from the TS fuzzy rules with nonlinear transformation. After integrating fuzzy subsystems and enhancement groups with the trained-well weight, the HDR image is generated. In FBLS, applying the incremental learning algorithm and the pseudoinverse method to compute the weights can greatly reduce the training time. In addition, the fuzzy system has better interpretability. In the learning process, IF-THEN fuzzy rules can effectively help the model to detect the artifacts and reject them in the final HDR result. These advantages solve the problem of existing deep-learning methods. Furthermore, we set up a new dataset of multiview LDR images with corresponding HDR ground truth to train our system. Our experimental results show that our system can synthesize high-quality multiview HDR images, which has a higher training speed than other learning methods.**

***Index Terms*—Fuzzy broad learning system (FBLS), high dynamic range (HDR) image, multiview synthesis.**

## I. Introduction

With the rapid demand of information contained in images [2]–[5], the high dynamic range (HDR) images acquisition technique is growing at an explosive speed. The current methods to obtain the HDR images are usually divided into two categories: 1) directly capturing with a professional camera and 2) merging from a set of different exposure low dynamic range (LDR) images which take from an ordinary digital camera. The former approach can directly obtain the high-quality HDR images in dynamic scenes using a professional camera with unique equipment. Nayar and Mitsunaga [6] placed an optical mask adjacent to a conventional image detector array to sample the spatial and exposure dimensions of image irradiance simultaneously. McGuire *et al.* [7] used multiple sensors to capture from the same viewpoint but have different image sensors and image parameters simultaneously. Hasinoff *et al.* [8] proposed the noise-optimal capture to capture the HDR or reduce noise using the SNR advantage of high ISO settings. Tocci *et al.* [9] presented an optical architecture for HDR imaging that allows simultaneous capture of high-, medium-, and low-exposure images on three sensors. However, these expensive professional hardwares result that the former methods could not popularize to the public easily. The second method, such as Debevec and Malik [10] and Mann and Picard [11] generated well HDR image from LDR images in the static scenes. However, if the scenes are dynamic or the camera is hand-held, these methods produced artifacts like ghosting during the process of aligning in the final HDR image. Therefore, "deghosting" is an important work in the process of synthesizing an HDR image. With the development of the HDR technology, patch-based approaches of Sen *et al.* [12] and Hu *et al.* [13], learning-based approaches of Kalantari and Ramamoorthi [14] and Eilertsen *et al.* [15], and other approaches of [16]–[21] all obtain good progress for ghost-free HDR image.

However, the existing methods focus on dynamic scenes which are moving or the slight shake of the hand-held camera. They neglect the change of views caused by the movement of the person holding the camera, which produces the multiview LDR images, for example, as shown in Fig. 1. Thus, we start to work on multiview HDR synthesis. The essential of synthesizing multiview HDR is to reduce artifacts produced in the process of aligning. Furthermore, we find that the artifacts can

Manuscript received March 18, 2019; revised June 24, 2019; accepted August 5, 2019. Date of publication August 30, 2019; date of current version April 15, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant 61872241, Grant 61572316, Grant 61751202, Grant 61751205, Grant 61572540, Grant U1813203, and Grant U1801262, in part by the Macau Science and Technology Development Fund under Grant 0027/2018/A1, Grant 079/2017/A2, Grant 024/2015/AMJ, and Grant 0119/2018/A3, in part by the National Key Research and Development Program of China under Grant 2017YFE0104000 and Grant 2016YFC1300302, and in part by the Science and Technology Commission of Shanghai Municipality under Grant 18410750700, Grant 17411952600, and Grant 16DZ0501100. This article was recommended by Associate Editor C.-F. Juang. *(Corresponding author: Bin Sheng.)*

H. Guo and B. Sheng are with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: shengbin@sjtu.edu.cn).

P. Li is with the Department of Computing, Hong Kong Polytechnic University, Hong Kong (e-mail: p.li@polyu.edu.hk).

C. L. P. Chen is with the School of Computer Science and Engineering, South China University of Technology, Guangzhou 510006, China, also with the Navigation College, Dalian Maritime University, Dalian 116026, China, and also with the Faculty of Science and Technology, University of Macau, Macau 999078, China (e-mail: philip.chen@ieee.org).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TCYB.2019.2934823.

Digital Object Identifier 10.1109/TCYB.2019.2934823


Fig. 1. Our method uses three different exposure LDR images of the multiview scene (shown on the bottom) as input, and output a high-quality HDR image (shown on the top) via FBLS. Note that we take the middle exposure LDR image as a reference, then use optical flow method of Liu [1] to align images with low and high exposures to the reference, respectively, to produce the alignment LDR images. Then, use alignment LDR images to synthesize HDR image which contains all of the three LDR images' lighting information through the FBLS.

be reduced significantly during the process of synthesizing by detecting artifact regions, which reduces the bad effect in the final HDR result. Thus, we use the learning method to imitate this complex process. Specifically, we use the fuzzy broad learning system (FBLS) [22] as our learning model, since IF-THEN fuzzy rules can effectively help the model to detect the artifacts and reject them in the final HDR result. This model not only solves the shortcomings of deep learning's long training time but also is interpretable compared with the traditional deep-learning model.

Generally, the process of obtaining the final HDR images can be divided into three steps: 1) align multiview LDR images to the reference, respectively; 2) synthesize the aligned LDR images into an HDR image; and 3) display HDR images after tonemapping. In this article, we take three multiview LDR images with high, medium, and low exposures and use the method of Liu [1] to align multiview LDR images. The LDR images with high and low exposures are aligned to the medium exposure (reference), respectively, to obtain three aligned images. In the second step, to eliminate the artifacts produced during the alignment, we adopt the FBLS to synthesize an HDR image. In this system, we use the acquired three aligned LDR images as input. We extract features through IF-THEN fuzzy rules to transfer corresponding Takagi–Sugeno (TS) fuzzy subsystems. Simultaneously, the structure is expanded in a wide sense in the "enhancement groups" which transfer from these extract features with nonlinear transformation to preserve the characteristic of inputs. Therefore, the structure of the system becomes wider instead

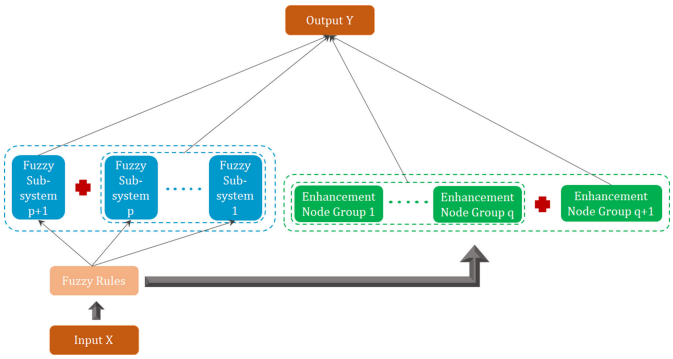



Fig. 2. Structure of increasing fuzzy subsystems and enhancement groups.

of deeper. Due to the FBLS's flat network, using pseudoinverse to compute the output-layer weights is a very convenient approach to reduce the training time. Finally, we connect all the fuzzy subsystems and enhancement groups to the output layer with a target weight. In the last step, we need tonemapper since the HDR images could not be displayed on a normal monitor which grayscale is 8 bits. Many powerful tonemapping techniques have been implemented and we choose Photoshop CS6 to tone map our HDR image. In addition, we retrain the model using the incremental learning algorithm through increasing the fuzzy subsystems and enhancement groups. It proves that our method is efficient and high quality. The experimental results demonstrate that our system is correct and efficient. In summary, our approach has the following contributions.

1) *Multiview HDR Image Synthesis:* This article proposes a novel research direction on the multiview HDR image synthesis which uses three multiview LDR images with different exposures. Most existing HDR datasets are captured from static scenes. Some of the datasets lack ground-truth images or have a small number of scenes with only rigid motion. In other datasets, the dynamic scenes are caused by the movement of the subject, while the views of different LDR images are the same. As a result, we create the multiview HDR dataset.
2) *Using FBLS to Synthesize HDR Image:* We adopt FBLS as our learning model. In the learning process, IF-THEN fuzzy rules can effectively help the model to detect the artifacts and reject them in the final HDR result. Its efficient, fast, and interpretable characteristic produces high-quality HDR result. Due to the flat structure of the FBLS, the coefficients in the consequent part of fuzzy rules in every fuzzy subsystem and the weights connecting the final output layer with the outputs of the enhancement layer can be computed using pseudoinverse. As a result, compared with the deep-learning-based methods, our method has fast computation nature.
3) *Incremental Learning Algorithm for FBLS:* Our system greatly reduces time as a result of incremental learning, which is a fast remodeling via increasing fuzzy subsystems and enhancement node groups instead of retraining the entire system compared to the other learning methods (see Fig. 2). When we increase the number

of fuzzy subsystems and enhancement node groups, the new weight matrix can be computed via the formerly calculated weights. Thus, we do not have to retrain the entire model, which saves lots of computational time to update the model.

## II. Related Work

HDR images have received extensive research over the past two decades and acquired a lot of valuable progress. Here, we will introduce the related work about aligning LDR images, rejecting moving object, and FBLS, respectively.

### A. Align LDR Images

Ward [23] and Tomaszewska and Mantiuk [24] aligned the LDR images through the automatic approaches. Ward [23] used an inexpensive shift and difference operations over each image to translate. Tomaszewska and Mantiuk [24] used the SIFT algorithm to search for key-points which constitute homography matrices. However, these simpler approaches are unable to use in the dynamic scene. There are also massive alignment algorithms based on optical flow. Bogoni [25] used local unconstrained motion estimation to align the LDR images. Kang *et al.* [26] used optical flow to compute a dense motion field that formed a local correction to the global transformation and then merged HDR images by rejecting the pixel which was not corresponding to the reference. Mangiat and Gibson [27] performed simpler block-based motion estimation and refined the motion vectors using color similarity in the adjacent frames. In [28], HDR images were aligned with the energy-based optical flow which minimizes an energy function of the data term on the gradient constancy assumption and smoothness term. Hu *et al.* [13] presented a displacement estimation method based on generating a perfectly aligned image which successfully deals with large saturated regions in the reference image. However, the state-of-the-art alignment algorithms using optical flow in the challenging scenes also produce artifacts on the final HDR image.

### B. Reject Moving Object

Rejecting moving object is another important work to synthesize the HDR images. These approaches identify and reject the moving object to avoid generating ghost. Liu and El Gamal [29] used advanced CMOS image sensors to capture multiple images within a normal exposure time to reject ghost. Grosch [30] predicted the pixel color from one image to another using camera response function and the difference between the two colors which indicated object motion. Gallo *et al.* [16] and Raman and Chaudhuri [31] did the similar approach with Grosch [30]. Khan *et al.* [32] and Heo *et al.* [17] did not require explicit object detection and motion estimation. Khan *et al.* [32] iteratively weighted each pixel according to its probability of belonging to the background. Heo *et al.* [17] utilized the global intensity transfer functions which obtained from the joint probability density functions between different exposure images and weighted each exposure according to the Gaussian-weighted distance to a reference value. Jacobs *et al.* [33] detected motions which used the difference in local entropy between different exposure images. Jinno and Okuda [34] used Markov random field [35], [36] to estimate displacements, occlusion, as well as saturated regions, and reject motion blur. Sidibe *et al.* [37] proposed a detection for ghosting region via an order relation between pixel values in consecutive images.

Min *et al.* [38] proposed a histogram-based ghost removal method, in which object motion and background change between two exposures were detected using multilevel thresholding of the intensity histogram. Pece and Kautz [39] used the median threshold bitmap algorithm to generate bitmaps for each exposure image and detected movement when the value changes in a pixel. Wu *et al.* [40] proposed three criteria of monotonous, pixel error, and color error to detect the moving objects. Zhang and Cham [41] utilized the gradient direction changes to reveal object movement. Sen *et al.* [12] proposed a novel patch-based energy-minimization formulation that integrates alignment and reconstruction in a joint optimization through an HDR image synthesis equation. Based on the Sen's patch-match method, Kalantari *et al.* [42] proposed the HDR video generation using alternate exposures in 2013 and used deep learning to reject artifact in 2017 [14] which was the first learning method in the HDR field. Granados *et al.* [43] proposed a method that models the noise distribution of color values and used a Markov random field to reconstruct irradiance from pixels that were likely to correspond to the same static scene object. Lee *et al.* [19] assumed that irradiance maps were linearly related to LDR image exposures and formulated ghost region detection as a rank minimization problem by restricting moving objects. Oh *et al.* [44] improved the Lee's method that handled moving objects with large overlapping area.

### C. FBLS

Before understanding fuzzy broad learning, we first introduce the deep learning and broad learning, respectively. Deep learning is a new field in machine-learning research [45]–[48], and it has a wide range of applications in image processing, speech recognition, and large-scale data processing [49]–[52]. Deep belief network (DBN) was one of the earliest neural networks. Hinton *et al.* [53] proposed the restricted Boltzmann machines which can be stacked and trained in a greedy manner to form a DBN. The convolutional neural network (CNN) [54] is the most popular neural network. Kalantari and Ramamoorthi [14] used CNN as a learning model and compared three different architectures which output the estimated HDR image, blending weights, and refined aligned to merge HDR images, respectively. Although they presented high-quality results, they suffered too long training time. The single-layer feedforward neural (SLFN) network has been widely used in many fields for its good learning ability [55], [56]. However, it has slow training speed, which easily falls into local minima and could not reach the global minimum. The random vector functional-link neural network (RVFLNN) [56], [57] effectively eliminated the drawback of the long training process. However, it could not do well on large-scale data and time variety. Chen and Wan [58] proposed

a fast learning algorithm which found optimal weights of the flat neural networks. Based on the work in [58], the broad learning system (BLS) [59], [60] was proposed. The feature nodes and the enhancement nodes from the input data were integrated into the output and the incremental learning algorithm can remodel the architecture through increasing input data, feature nodes, and enhancement nodes without retraining original architecture.

In the study of past fuzzy systems, the network based on a set of IF-THEN fuzzy rules of fuzzy system with the learning and connecting structure of neural network, called fuzzy neural network (FNN) achieved a lot of excellent results. However, the FNN follows the neural network's training method to train the parameters in the fuzzy rules and cost much time. Due to the massive data and big dimension, fuzzy rules' accuracy choice is a problem. To solve these problems, many improved approaches have been proposed recently. Wang *et al.* [61] proposed a novel hierarchical hybrid FNN which the fuzzy subsystems aggregates several discrete input variables into an intermediate variable and neural network rest consist of continuous input variables and intermediate variables. Rong *et al.* [62] proposed an online sequential fuzzy extreme learning machine (OS-Fuzzy-ELM), its learning can be done with the input data coming in a one-by-one mode or a chunk-by-chunk mode with and randomly assigned all the antecedent parameters of the membership functions to cut down the learning time. Sun *et al.* [63] proposed a neuro-fuzzy inference system which grouped the data by the $k$-means clustering method and the membership of arbitrary input for each fuzzy rule was derived through an ELM. However, they only consider one fuzzy system in their models. Recently, Feng and Chen [22] proposed a fuzzy learning model based on the BLS which includes fuzzy subsystems and enhancement groups.

## III. Approach Framework

In our proposed method, we use a set of multiview LDR images with different exposures $(L_1, L_2, L_3)$ to synthesize high-quality HDR images $(H)$ with little ghost via FBLS (see Algorithm 1). In the alignment stage, we regard the middle-exposure image $(L_2)$ as the reference; then, align low-exposure image $(L_1)$ and high-exposure image $(L_3)$ to the reference (the middle exposure) using optical-flow method of Liu [1], respectively. The new generated aligned images with different exposures are denoted as $I = \{I_1, I_2, I_3\}$. The most advanced algorithm which used optical flow to align the LDR images could not work accurately and produce ghosting artifacts in some complex motions. To reduce the influences of this phenomenon on the final HDR images, we proposed to generate HDR image using FBLS. In the training step, we take the aligned images $I_1, I_2$, and $I_3$ and the ground truth as input and output the connection weight. In the testing step, we input the aligned image $I_1, I_2, I_3$ and output our HDR result. The process of synthesizing multiview HDR image is shown in Fig. 3. In some cases, if the quality of the final HDR image quality could not reach our expectation, we need incremental learning to proceed to refine our learning model. We describe the HDR synthesis system in Section IV and interpret the incremental learning in Section V.

**Algorithm 1** Synthesis HDR Image Using FBLS

**Input:** Aligned Images $I$, fuzzy rules $K_i$, fuzzy subsystems $p$, enhancement node groups $q$
**Output:** Final HDR Image

```
Take X = I;
Random the parameter α_kt^i in [0, 1];
for i = 1; i <= p do
    for s = 1; s <= n do
        Calculate Z_si using Eq. (4);
        Calculate F_si using Eq. (9);
    end for
    Obtain the Z_i using Eq. (5);
    Obtain the F_i using Eq. (10);
end for
Obtain Z^p using Eq. (6);
for j = 1; j <= q do
    Randomly generated ω_j, β_j;
    Calculate H_j using Eq. (7);
end for
Obtain the enhancement node groups H^q using Eq. (8);
Obtain the fuzzy subsystems F^p using Eq. (11);
Use trained well weight W to generate final HDR image with Eq. (12);
```

## IV. Generate HDR Image

### A. Preprocessing

If the format of the multiview LDR images is not RAW, we use camera response curve in [10] to linearize them. Then, we use gamma curve ($\gamma = 2.2$) on these linear images to obtain LDR images $(L_1, L_2, L_3)$ suitable for our method. This step makes the LDR images closer to the real. Before aligning, we need to adjust exposure of the reference image $(L_2)$ to high-exposure image $L_3$ and adjust low-exposure image $L_1$ to the reference image $L_2$, respectively, since the optical-flow methods require brightness constancy. Formally, it is denoted as: $L^{1,2} = \text{clip}(L_1\text{exposure}(L_1, L_2)^{(1/\gamma)})$ and $L^{2,3} = \text{clip}(L_2\text{expourse}(L_2, L_3)^{(1/\gamma)})$, where exposure() is the exposure ratio between the reference and the low or high exposure. Then, we use optical-flow method in [1] to compute the flow between $L_1$ and $L^{1,2}$, and the flow between $L_3$ and $L^{2,3}$ to obtain the aligned images $I = \{I_1, I_2, I_3\}$.

### B. HDR Synthesis

In this step, we use the aligned images $I$ as input of FBLS $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)^T \in \mathbb{R}^{n \times m}$ and output the HDR image. At first, we use the first-order TS fuzzy model to map the input $\mathbf{x}_s = (x_{s1}, x_{s2}, \ldots, x_{sm})$ to the $i$th fuzzy system. In the first-order TS fuzzy model, the function of $x_{st}$ is the first-order polynomial, where $t = 1, 2, \ldots, m$. This can be defined as

$$z_{sk}^i = \sum_{t=1}^{m} \alpha_{kt}^i x_{st} \tag{1}$$

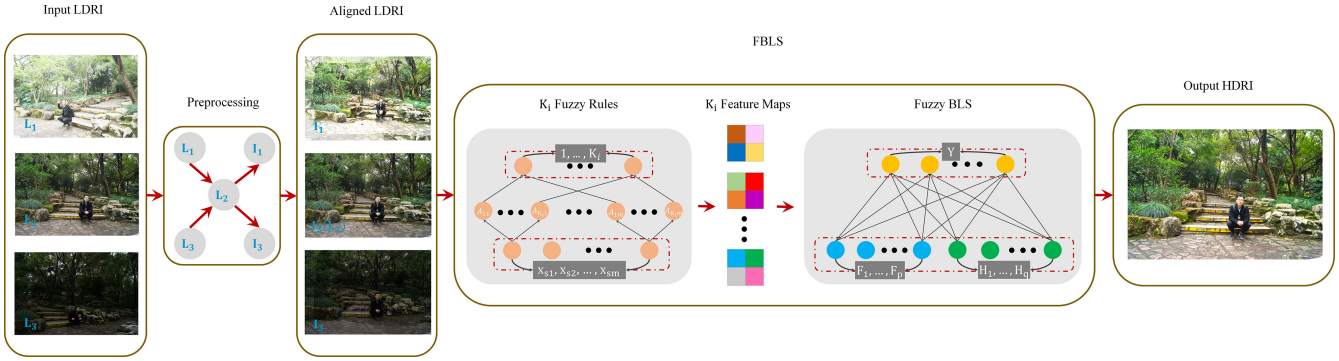



Fig. 3. Overview of the FBLS-HDR synthesis process. We take three multiview LDR images with different exposures $L_1, L_2, L_3$ as the input data. After preprocessing, we obtain the aligned LDR images. In this step, we use optical flow as our aligned method which use high- and low-exposure LDR image to align with middle exposure (reference), respectively, to generate aligned LDR images $I_1, I_2, I_3$. Then, we take $I_1, I_2, I_3$ as input of FBLS and extract features through $K_i$ fuzzy rules to transform fuzzy subsystems $F_i$. Simultaneously, the structure is expanded in a wide sense in the enhancement groups $H_i$ which transfer from these extract features with nonlinear transformation to preserve the characteristic of inputs. Finally, we connect all the fuzzy subsystems and enhancement groups to the output layer with a target weight.

where $\alpha_{kt}^i$ is a parameter generated randomly and $k = 1, 2, \ldots, K_i$ is the number of fuzzy rule of the $i$th fuzzy system. We adopt the Gaussian membership function as our membership function, this can be denoted as

$$\mu_{kt}^i(x_{st}) = e^{-\left(\frac{x_{st}^i - c_{kt}^i}{\sigma_{kt}^i}\right)^2} \tag{2}$$

where $c_{kt}^i$ are the centers of the Gaussian membership functions and the width $\sigma_{kt}^i$ is a fixed value. We use classic $k$-means method on the training data to obtain $K_i$ clustering centers. The values of $c_{kt}^i$ are decided by these clustering centers. Then, the weighted activation level for each rule can be computed as

$$\omega_{sk}^i = \frac{\prod_{t=1}^m \mu_{kt}^i(x_{st})}{\sum_{k=1}^{K_i} \prod_{t=1}^m \mu_{kt}^i(x_{st})}. \tag{3}$$

After the above calculation, the intermediate output $\mathbf{Z}_{si}$ for the $s$th training sample of the $i$th fuzzy subsystem can be denoted as

$$\mathbf{Z}_{si} = \left(\omega_{s1}^i z_{s1}^i, \omega_{s2}^i z_{s2}^i, \ldots, \omega_{sK_i}^i z_{sK_i}^i\right) \tag{4}$$

the intermediate output $\mathbf{Z}_i$ for all the training samples of the $i$th fuzzy subsystem can be denoted as

$$\mathbf{Z}_i = (\mathbf{Z}_{1i}, \mathbf{Z}_{2i}, \ldots, \mathbf{Z}_{ni}), i = 1, 2, \ldots, p. \tag{5}$$

Then, the intermediate output $\mathbf{Z}^p$ of $p$ fuzzy subsystems can be denoted as

$$\mathbf{Z}^p = \left(\mathbf{Z}_1, \mathbf{Z}_2, \ldots, \mathbf{Z}_p\right) \in \mathbb{R}^{n \times (K_1 + K_2 + \cdots + K_p)}. \tag{6}$$

Then, we enhance the intermediate output to the enhancement node groups, this can be defined as

$$\mathbf{H}_j = \psi\left(\mathbf{Z}^p \boldsymbol{\omega}_j + \boldsymbol{\beta}_j\right), \quad j = 1, 2, \ldots, q \tag{7}$$

where $\mathbf{H}_j$ is the enhancement groups transformed from $\mathbf{Z}^p$, which can preserve the characteristic of inputs. $\boldsymbol{\omega}_j$ and $\boldsymbol{\beta}_j$ are the weight and bias randomly generated from [0,1] with proper dimensions which transform $\mathbf{Z}^p$ to $\mathbf{H}_j$. $\psi(\cdot)$ is an activation function which use the Sigmoid function. And all the enhancement node groups are denoted as

$$\mathbf{H}^q = \left(\mathbf{H}_1, \mathbf{H}_2, \ldots, \mathbf{H}_q\right). \tag{8}$$

The output vector $\mathbf{F}_{si}$ for the $s$th training sample of the $i$th fuzzy subsystem consists of $z_{sk}^i$. To avoid the computational complexity, we do not compute the coefficient $\alpha_{kt}^i$ in pseudoinverse explained in (13), but introduce a new parameter $\lambda_{kc}^i$. It can be defined as

$$\begin{aligned}
\mathbf{F}_{si} &= \left(\sum_{k=1}^{K_i} \lambda_{k1}^i \omega_{sk}^i z_{sk}^i, \ldots, \sum_{k=1}^{K_i} \lambda_{kC}^i \omega_{sk}^i z_{sk}^i\right) \\
&= \left(\sum_{k=1}^{K_i} \lambda_{k1}^i \omega_{sk}^i \left(\sum_{t=1}^m \alpha_{kt}^i x_{st}\right)\right), \ldots \\
&\quad \left(\sum_{k=1}^{K_i} \lambda_{k1}^i \omega_{sk}^i \left(\sum_{t=1}^m \alpha_{kC}^i x_{st}\right)\right) \\
&= \sum_{t=1}^m \alpha_{kt}^i x_{st} \left(\omega_{s1}^i, \ldots, \omega_{sK_i}^i\right) \begin{pmatrix} \lambda_{11}^i & \cdots & \lambda_{1C}^i \\ \vdots & & \vdots \\ \lambda_{K_i1}^i & \cdots & \lambda_{K_iC}^i \end{pmatrix}
\end{aligned} \tag{9}$$

where $c = 1, 2, \ldots, C$. The output matrix $\mathbf{F}_i$ for all the training samples of the $i$th fuzzy subsystem is denoted as

$$\mathbf{F}_i = (\mathbf{F}_{1i}, \mathbf{F}_{2i}, \ldots, \mathbf{F}_{ni}) = \mathbf{D}\Omega^i\boldsymbol{\lambda}^i \tag{10}$$

where $\mathbf{D} = \text{diag}\{\sum_{t=1}^m \alpha_{kt}^i x_{1t}, \ldots, \sum_{t=1}^m \alpha_{kt}^i x_{nt}\}$, and

$$\Omega^i = \begin{pmatrix} \omega_{11}^i & \cdots & \omega_{1K_i}^i \\ \vdots & \cdots & \vdots \\ \omega_{n1}^i & \cdots & \omega_{nK_i}^i \end{pmatrix}, \quad \boldsymbol{\lambda}^i = \begin{pmatrix} \lambda_{11}^i & \cdots & \lambda_{1C}^i \\ \vdots & \cdots & \vdots \\ \lambda_{K_i1}^i & \cdots & \lambda_{K_iC}^i \end{pmatrix}.$$

Let $\mathbf{F}^p$ denote the aggregative output of $p$ fuzzy subsystems, written as

$$\mathbf{F}^p = \sum_{i=1}^p \mathbf{F}_i = \sum_{i=1}^p \mathbf{D}\Omega^i\boldsymbol{\lambda}^i = \mathbf{D}\left(\Omega^1, \ldots, \Omega^p\right) \begin{pmatrix} \boldsymbol{\lambda}^1 \\ \vdots \\ \boldsymbol{\lambda}^p \end{pmatrix} = \mathbf{D}\Omega\boldsymbol{\Lambda} \tag{11}$$

where $\Omega = (\Omega^1, \ldots, \Omega^p)$ and we denote $((\boldsymbol{\lambda}^1)^T, \ldots, (\boldsymbol{\lambda}^p)^T)^T$ as $\Lambda$.

Finally, we connect the fuzzy subsystems $\mathbf{F}^p$ and the enhancement node groups $\mathbf{H}^q$ to the output $\mathbf{Y}$. The weights

connecting to the output of $\mathbf{F}^p$ and $\mathbf{H}^q$ are $\mathbf{W}_f$ and $\mathbf{W}_h$, respectively. Hence, the FBLS model can be denoted as

$$\begin{aligned}\mathbf{Y} &= \mathbf{F}^p\mathbf{W}_f + \mathbf{H}^q\mathbf{W}_h \\ &= \mathbf{D}\Omega\Lambda + \mathbf{H}^q\mathbf{W}_h \\ &= \left(\mathbf{D}\Omega, \mathbf{H}^q\right)\begin{pmatrix}\Lambda \\ \mathbf{W}_h\end{pmatrix} \\ &= \mathbf{A}\mathbf{W}\end{aligned} \tag{12}$$

where $\mathbf{A} = (\mathbf{D}\Omega, \mathbf{H}^q)$ and the values of the weights $\mathbf{W}_f$ are set to be 1. $\mathbf{W} = \begin{pmatrix}\Lambda \\ \mathbf{W}_h\end{pmatrix}$ is the final connecting weight matrix of the FBLS. The weight matrix $\mathbf{W}$ can be computed using the training targets $\mathbf{Y}$ : $\mathbf{W} = \mathbf{A}^+\mathbf{Y}$.

The pseudoinverse matrix $\mathbf{A}^+$ could be obtained by the optimization problem

$$\mathbf{A}^+ = \arg\min_{\mathbf{W}} \left\|\mathbf{A}\mathbf{W} - \mathbf{Y}\right\|_2^2 + \lambda\left\|\mathbf{W}\right\|_1. \tag{13}$$

We use an improved least square to obtain the connecting weight $\mathbf{W}$ with smallest training errors, and $\lambda$ denote the further constraints on the sum of the squared weights in [59]. The first terms is an $l_2$-norm regularization which denotes the training errors. The second term is an $l_1$-norm regularization which prevents our model from overfitting. Obviously, we could obtain $\mathbf{W} = (\lambda\mathbf{E} + \mathbf{A}\mathbf{A}^T)^{-1}\mathbf{A}^T\mathbf{Y}$, where $\mathbf{E}$ is a unit matrix. If $\lambda \to 0$, we have : $\mathbf{A}^+ = \lim_{\lambda\to 0}(\lambda\mathbf{E} + \mathbf{A}\mathbf{A}^T)^{-1}\mathbf{A}^T$. Overall, in this step, we can synthesize the HDR image using the trained-well weight matrix $\mathbf{W}$ in the FBLS system.

## V. Optimization of Training Result

In other deep-learning models, if the learning effect does not work well as expected, they will increase the number of the filter or increase the number of the layer, which needs to retrain the new network for a long time. In our FBLS, we can increase additional enhancement node groups and fuzzy subsystems to reconstruct the model using an incremental learning algorithm without the process of retraining the entire system (see Algorithm 2). It has better performance. We use PSNR to evaluate the accuracy of our HDR results. If the PSNR value of the final HDR images is less than 40, we need to increase the enhancement node group and fuzzy subsystem to raise our HDR result's quality. We denote the additional enhancement node group as $\mathbf{H}_{q+1} = \psi(\mathbf{Z}^p \cdot \boldsymbol{\omega}_{q+1} + \boldsymbol{\beta}_{q+1})$, where $\boldsymbol{\omega}_{q+1}$, and $\boldsymbol{\beta}_{q+1}$ are the weight and bias generated randomly from the fuzzy subsystems to the additional enhancement node group with proper dimensions. And the new matrix after increasing additional enhancement node group is denoted as $\mathbf{A}^{q+1}$

$$\mathbf{A}^{q+1} = \left[\mathbf{A}|\mathbf{H}_{q+1}\right]. \tag{14}$$

Then, we calculate the pseudoinverse of the new matrix $(\mathbf{A}^{q+1})$ as

$$\left(\mathbf{A}^{q+1}\right)^+ = \begin{bmatrix}\mathbf{A}^+ - \left(\mathbf{A}^+\mathbf{H}_{q+1}\right)\mathbf{B}^T \\ \mathbf{B}^T\end{bmatrix} \tag{15}$$

where $\mathbf{B}$ is explained detailedly in [59].

**Algorithm 2** Incremental Learning Algorithm

**Input:** Training samples $(X, Y)$, original connecting weight
**Output:** Dynamic updated connecting weight

```
repeat
    if increase the enhancement group then
        Randomly generated ω_{q+1} and β_{q+1};
        Calculate new enhancement node group H_{q+1} = ψ(Z^p ω_{q+1} + β_{q+1});
        Set the new matrix A_{q+1} using Eq. (14);
        Update (A^{q+1})^+ using Eq. (15);
        Update connecting weight W^{q+1} using Eq. (16);
    end if
    if increase the fuzzy subsystem then
        Calculate new fuzzy subsystem F_{p+1} = DΩ^{p+1}λ^{p+1};
        Set the new matrix A_{p+1};
        Update (A_{p+1})^+ using Eq. (18);
        Update connecting weight W^{p+1} using Eq. (19);
    end if
until The PSNR value is satisfied
Update connecting weight;
```

Finally, the dynamic updating weight $\mathbf{W}^{q+1}$ can be calculated by

$$\mathbf{W}^{q+1} = \begin{bmatrix}\mathbf{W} - \left(\mathbf{A}^+\mathbf{H}_{q+1}\right)\mathbf{B}^T\mathbf{Y} \\ \mathbf{B}^T\mathbf{Y}\end{bmatrix}. \tag{16}$$

Only increasing the enhancement nodes may also not satisfy our quality requirements, because the feature mapping of the fuzzy subsystem we extract may not completely include the entire features of the input data. Therefore, we can increase the number of fuzzy subsystems to improve the quality. We denote the additional fuzzy subsystem as $\mathbf{F}_{p+1}$

$$\mathbf{F}_{p+1} = \mathbf{D}\Omega^{p+1}\lambda^{p+1}. \tag{17}$$

We denoted the new matrix as $\mathbf{A}^{p+1} = [\mathbf{A}|\mathbf{D}\Omega^{p+1}]$. Then, we calculate the pseudoinverse $(\mathbf{A}^{p+1})^+$ similar to (15) and the dynamic updating weight $\mathbf{W}^{p+1}$ similar to (16)

$$\left(\mathbf{A}^{p+1}\right)^+ = \begin{bmatrix}\mathbf{A}^+ - \left(\mathbf{A}^+\mathbf{D}\Omega^{p+1}\right)\mathbf{B}^T \\ \mathbf{B}^T\end{bmatrix} \tag{18}$$

$$\mathbf{W}^{p+1} = \begin{bmatrix}\mathbf{W} - \left(\mathbf{A}^+\mathbf{D}\Omega^{p+1}\right)\mathbf{B}^T\mathbf{Y} \\ \mathbf{B}^T\mathbf{Y}\end{bmatrix}. \tag{19}$$

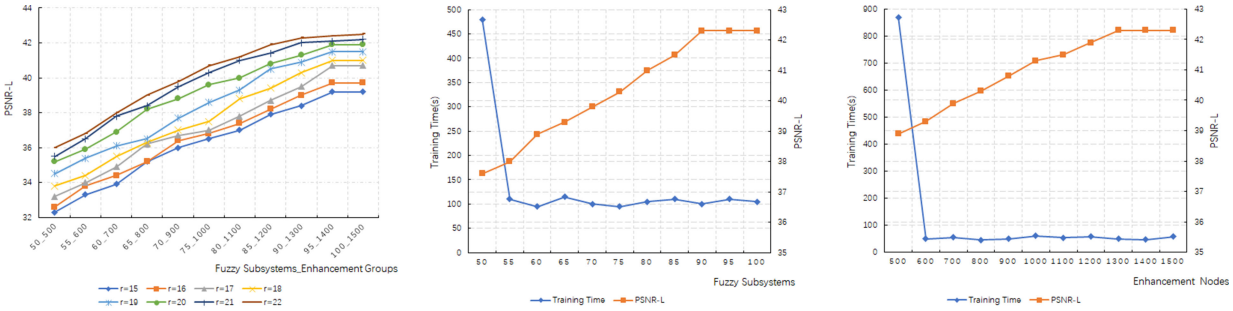
Overall, due to our model's flat structure, we can increase the fuzzy subsystems and enhancement node groups through pseudoinverse to avoid retraining the entire model.

## VI. Experimental Results

In this section, in order to prove that our proposed method is efficient and generates a high-quality result, we will show the experiment in four parts, including datasets and implementation, evaluation, comparisons, and running time.

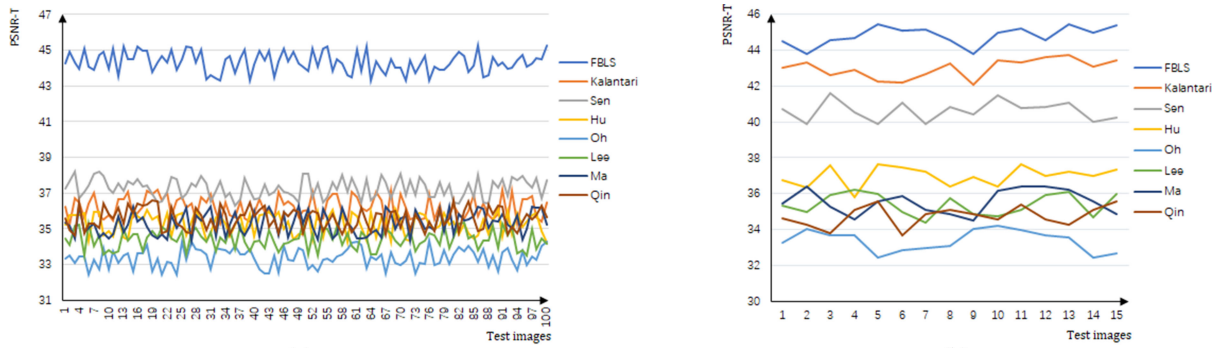
### A. Datasets and Implementation Details

*1) Datasets:* In order to train our FBLS to have a better performance without artifacts, we need a large dataset which




(a) (b) (c)

Fig. 4. Parameter analysis experiment. We keep the remaining parameters in our defined value and adjust some parameters in each experiment. (a) Experiment of increasing the fuzzy rules and fuzzy subsystem-enhancement group. (b) Experiment of increasing the fuzzy subsystems from 50 to 100 to observe the PSNR-L and the training time simultaneously. (c) Experiment of increasing the enhancement nodes from 500 to 1500 to observe the PSNR-L and the training time simultaneously.




(a) (b)

Fig. 5. Compare the values of PSNR-T with the state-of-the-art methods (Hu *et al.* [13], Sen *et al.* [12], Qin *et al.* [20], Ma *et al.* [21], Lee *et al.* [19], Oh *et al.* [44], and Kalantari and Ramamoorthi [14]) on our multiview test set and Kalantari and Ramamoorthi [14] test set (deep-learning method), respectively. (a) Test on 100 scenes of our multiview test set. (b) Test on 15 scenes of Kalantari test set.

TABLE I
QUANTITATIVE RESULTS OF THREE METRIC METHOD

| Kalantari's Dataset | Method | | | | |
|---|---|---|---|---|---|
| | Sen [12] | Hu [13] | Oh [44] | Kalantari [14] | Ours |
| PSNR-T | 40.70 | 35.53 | 32.27 | 42.72 | 43.24 |
| PSNR-L | 37.98 | 30.81 | 34.37 | 41.20 | 42.33 |
| HDR-VDP-2 | 63.85 | 60.74 | 61.28 | 64.03 | 63.51 |

consists of the multiview LDR images with different exposure and the corresponding ground truth. However, there is no such existing dataset. Therefore, we create a new dataset of multiview HDR synthesis which includes a set of 100 training scenes and 100 test scenes, each set of the scene contains three different exposure multiview LDR images in .tif format with its corresponding ground-truth image in .hdr format. The training scenes contain indoor, outdoor, sunny day, and cloudy day. The resolution of these images is $1500 \times 1000$ and the exposure biases of the LDR images are $(-2.0, 0.0, 2.0)$. Note that we use the approach of Kalantari and Ramamoorthi [14] to generate the corresponding ground-truth HDR image.

*2) Implementation Details:* We implement our approach using MATLAB on a laptop with 4.00-GHz Intel i7 CPU, 32-GB memory. We use the set of 100 scenes as our training data. We set the randomly generated weights in the enhancement node groups $\boldsymbol{\omega}_j$ and $\boldsymbol{\beta}_j$ within the interval of $(-1, 1)$, the

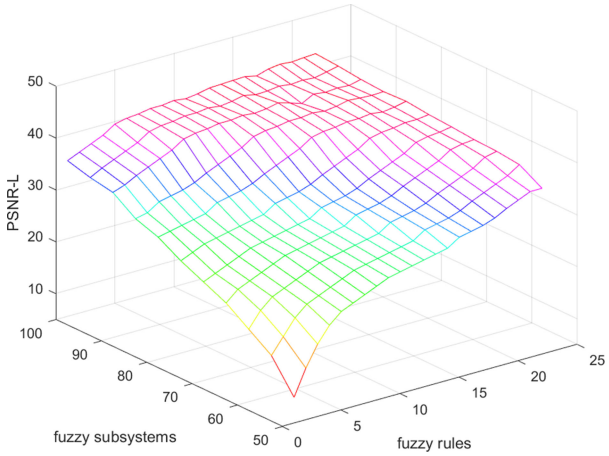



Fig. 6. With the fuzzy subsystems and fuzzy rules increasing, the values of PSNR-L are on the rise on the whole. It confirms the stability of our system.

coefficients $\alpha_{kt}^i$ are initialized randomly in $(0, 1)$. In addition, we set the parameter $\lambda$ for ridge regression in (13) as $10^{-6}$ and set the $\sigma_{kt}^i$ in (2) as 1. The activation function which transforms the enhancement node groups is the nonlinear sigmoid functions. For setting the number of fuzzy subsystems, enhancement node groups, and fuzzy rules, in the beginning, we set the number of fuzzy rules as 15, fuzzy subsystems as 50, and enhancement nodes as 500; then, we, respectively,

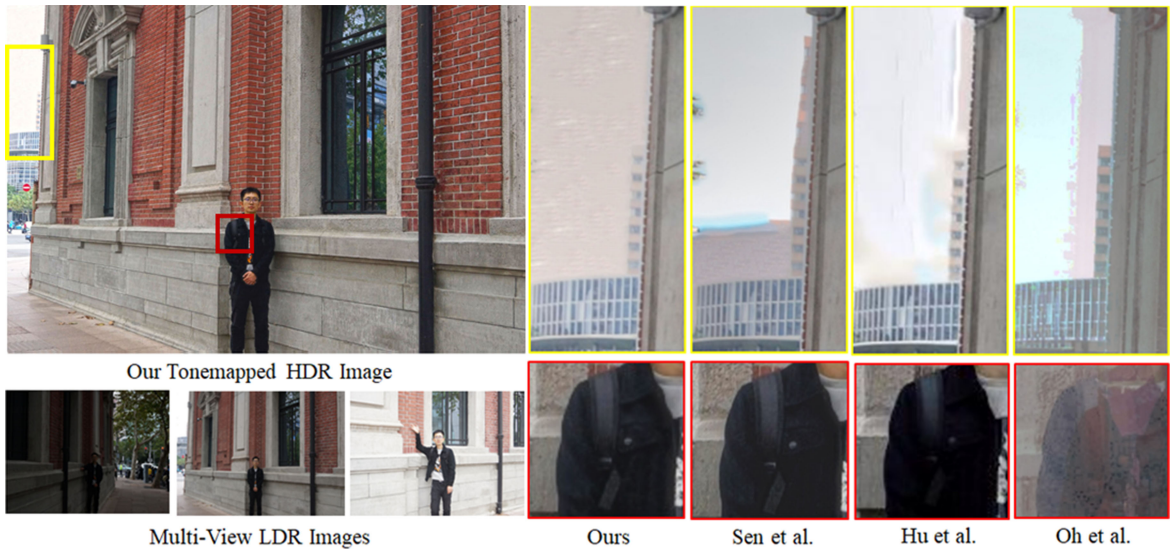



Fig. 7. Compare the HDR results on *man1* of our multiview dataset with the state-of-the-art methods of Sen *et al.* [12], Hu *et al.* [13], and Oh *et al.* [44].

TABLE II
COMPARE RUNTIME IN SECONDS

| Scenes | Image Size | Sen et al. [12] | Hu et al. [13] | Lee et al. [19] | Oh et al. [44] | Kalantari et al. [14] | FBLS |
|---|---|---|---|---|---|---|---|
| | | MATLAB+Mex | MATLAB+Mex | MATLAB+Mex | MATLAB | MATLAB+Mex | MATLAB |
| *man1* | 1500×1000 | 210.28 | 314.96 | 106.91 | 120.93 | 113.78 | 122.41 |
| *man2* | 1500×1000 | 237.52 | 332.34 | 115.35 | 108.41 | 106.12 | 131.53 |
| *man3* | 1500×1000 | 242.19 | 340.13 | 136.39 | 140.03 | 124.65 | 103.96 |
| *man4* | 1500×1000 | 205.57 | 338.70 | 121.27 | 129.14 | 119.76 | 116.72 |
| *arch* | 699×1024 | 95.38 | 154.96 | 56.24 | 70.33 | 52.04 | 60.39 |
| *forrest* | 1024×683 | 159.28 | 213.27 | 82.46 | 91.40 | 83.20 | 79.15 |

increase the amount of fuzzy subsystems from 50 to 100, and enhancement nodes from 500 to 1500 to retrain our model. Simultaneously, we observe the variety of PSNR-T (computed using tonemapped outputs and ground truth) and PSNR-L (computed using linear images and ground truth) to obtain the optimal model parameters. Then, we increase the number of fuzzy rules from 15 to 25 to adjust our model while keeping the number of fuzzy subsystems and enhancement nodes are constant. Finally, we set the number of fuzzy subsystems as 90, the number of enhancement groups as 1300, and the number of fuzzy rules as 22, which can obtain the highest quality HDR image and cost the least time relatively. If we keep on increasing the number of fuzzy subsystems, enhancement nodes, and fuzzy rules, the quality of the result has little improved and also cost much time. This parameter analysis experiment is shown in Fig. 4. From Fig. 4(b) and (c), we can see that the training time has a large decline when increasing the number. It proves that we do not need to retrain the entire model when increasing the fuzzy subsystems and enhancement nodes.

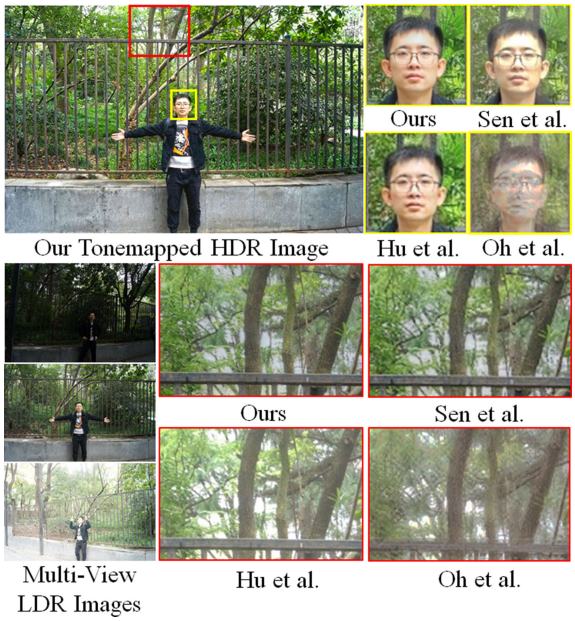



Fig. 8. Compare HDR results on *man2* of our multiview dataset with the state-of-the-art methods of Sen *et al.* [12], Hu *et al.* [13], and Oh *et al.* [44].

## B. Evaluations

In this section, we compare our method with the state-of-the-art methods, including two-based patch method [12], [13]; a deep learning CNN method [14]; and a motion rejection method [44]. All of the results are implemented by the authors' improved code. In order to prove that our method has better performance, we make quantitative evaluation using PSNR-T, PSNR-L, and HDR-VDP-2 [64] which is a calibrated visual metric for visibility and quality prediction for HDR images on 15 scenes of the Kalantari's test set [14]. Note that there


Fig. 9. Compare HDR results on *man3* of the Kalantari's dataset with the state-of-the-art methods of Kalantari and Ramamoorthi [14], Eilertsen *et al.* [15], Sen *et al.* [12], Hu *et al.* [13], and Oh *et al.* [44].


Fig. 10. Compare HDR results on *man4* of the Kalantari's dataset with the state-of-the-art methods of Kalantari and Ramamoorthi [14], Eilertsen *et al.* [15], Sen *et al.* [12], Hu *et al.* [13], and Oh *et al.* [44].

is no training data used in the test data. The higher average value of three kinds of metric methods denotes which is more similar to the ground truth. From Table I, we can see that our values of PSNR-T and PSNT-L are higher than the state-of-the-art methods, which means our proposed method better restores the real scene detail. The Kalantari's method obtains the highest HDR-VDP-2 score. We also compare the values of PSNR-T on 100 scenes of our multiview test set and 15 scenes of the Kalantari's test set with the state-of-the-art methods (Hu *et al.* [13], Sen *et al.* [12], Qin *et al.* [20], Ma *et al.* [21], Lee *et al.* [19], Oh *et al.* [44], and Kalantari and Ramamoorthi [14]) in Fig. 5. The row coordinates are different test images. From the figure we can see, our FBLS has better performance on both two test sets. It indicates that our proposed method can produce high-quality HDR results not only on the multiview but on the traditional dynamic scenes. However, other approaches are not able to apply in our multiview scenes. In addition, we evaluate the stability of our model. We set the enhancement nodes as 1500 and test our model with the increasing of fuzzy rules and fuzzy subsystems. From Fig. 6, we can see the values of PSNR-L are rising generally. It shows that our model is stable.

### C. Comparisons

We compare our method on our multiview test data with the state-of-the-art methods of Sen *et al.* [12], Hu *et al.* [13], and Oh *et al.* [44] which is shown in Figs. 7 and 8. In the left of each figure is the three multiview LDR images with different exposures and the middle is the reference. Fig. 7 shows a man on the street on a cloudy day. In the red block, Oh *et al.* [44] are not able to avoid alignment artifacts caused by the significant multiview motions. And in the dark region, Hu *et al.* [13] are not able to recover the detail of the black clothes. In the yellow block of the building in the distance, Oh *et al.* [44] still produce a lot of alignment artifacts, Hu *et al.* [13] produce blur and Sen *et al.* [12] generate both alignment artifacts as well as blur. Our method can synthesize a blur-free and artifact-free high-quality HDR result. Fig. 8 shows a man in the green park on a cloudy day. In the red block, other approaches are not able to avoid alignment artifacts (Oh *et al.* [44]) could not recover the highlight (Sen *et al.* [12]) and produce blur in the entire image (Hu *et al.* [13]). In the yellow block, the approach of Hu *et al.* [13] could not recover the highlight of the sky behind the leaves. The approach of Sen *et al.* [12] produces blur around the tree and the approach of Oh *et al.* [44]

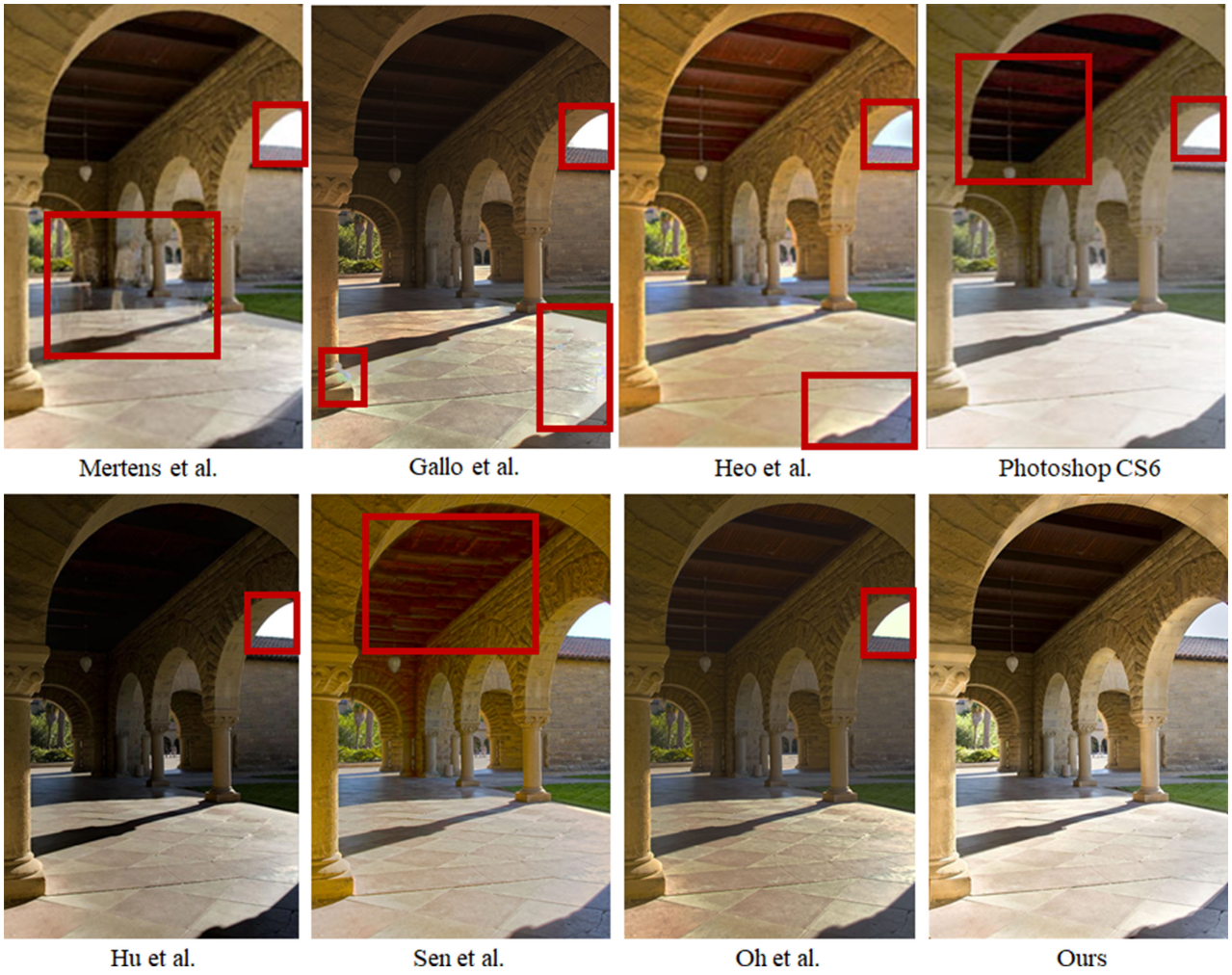



Fig. 11. Compare HDR results on *arch* of the Gallo's [16] dataset with the state-of-the-art methods of Mertens *et al.* [18], Gallo *et al.* [16], Heo *et al.* [17], Photoshop CS6, Hu *et al.* [13], Sen *et al.* [12], and Oh *et al.* [44].

generates serious artifacts. In contrast, our result has better performance.

We compare our method on the Kalantari's dataset [14] with the state-of-the-art methods of Kalantari and Ramamoorthi [14], Sen *et al.* [12], Hu *et al.* [13], Oh *et al.* [44], and Eilertsen *et al.* [15], which is shown in Figs. 9 and 10. In Fig. 9, because of the significant foreground motions of the people, the patch-based approaches (Sen *et al.* [12] and Hu *et al.* [13]) produce a lot of artifacts around the moving region. In contrast, the learning-based approaches (Kalantari and Ramamoorthi [14], Eilertsen *et al.* [15] and ours) produce better results. However, the result of Kalantari and Ramamoorthi [14] also has some artifacts and the result of Eilertsen *et al.* [15] produces saturated highlights and blur. Our approach is able to produce a high-quality HDR image with little artifacts and highlights. Similarly, in Fig. 10, in the highly saturated highlight regions, the result of Kalantari and Ramamoorthi [14], Sen *et al.* [12], Hu *et al.* [13], and Oh *et al.* [44] produces a lot of artifacts. The result of Eilertsen *et al.* [15] is better but still has a little blur. Our results are able to hallucinate plausible details in the saturated regions.

We compare our method on the Gallo's dataset [16] with the state-of-the-art methods shown in Figs. 11 and 12. The ArchSequence in Fig. 11 consists of five different exposure LDR images, we choose three of them to test our proposed method with exposure value $(-2.0, 0.0, 2.0)$. In Fig. 11, the result of Mertens *et al.* [18] has an obvious ghost of moving person, because they do not align the dynamic scenes and just handle the static scenes. In the lower-left corner and lower-right corner of Gallo *et al.* [16], their result produces blur and saturated highlights, respectively. The result of Heo *et al.* [17] and Photoshop CS6 generated the halo artifacts and blending artifacts, respectively. The patch-based result of Hu *et al.* [13] and Sen *et al.* [12] produced blur and noise in the dark regions since they preserved information from the reference heavily. These results all reject ghost successfully in addition to Mertens *et al.* [18]. The result of Oh *et al.* [44] and most of these results are not able to recover the sky. Our method shows more details and high quality of the scenes especially. In Fig. 12, we compare our result with Gallo *et al.* [16], Hu *et al.* [65], and Hu *et al.* [13]. In the red block which is a bright region, our result and Gallo's have better performance, however, Hu *et al.* [65] and Hu *et al.* [13] produce highlight.

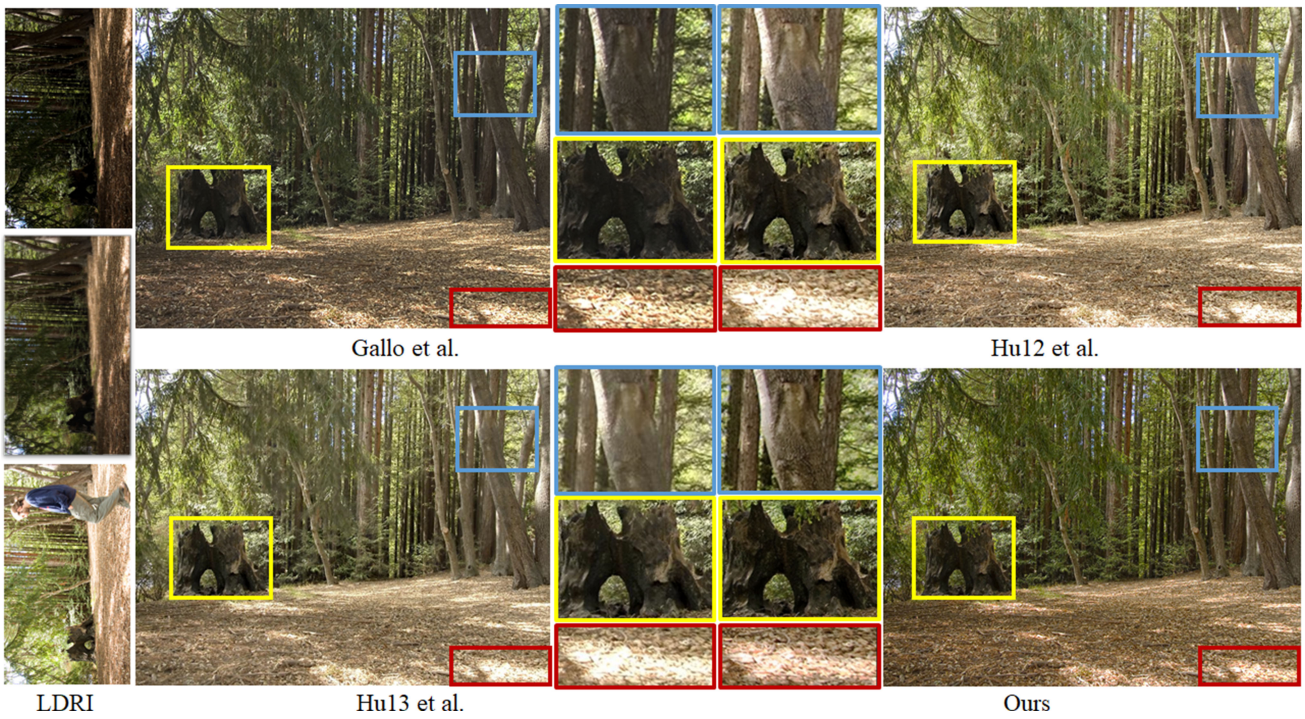



Fig. 12. Compare HDR results on *forrest* of the Gallo's [16] dataset with the state-of-the-art methods of Gallo *et al.* [16], Hu12 *et al.* [65], and Hu13 *et al.* [13].

On the contrary, in the yellow block which is a dark region, Gallo *et al.* [16] did not restitute detail of the scene and the tree stump contained some blur. Moreover, compared to our method in the blue block, other methods have different degrees of blur.

### D. Running Time

In order to prove our proposed approach is efficient, we compare the execution time with Hu *et al.* [13], Sen *et al.* [12], Oh *et al.* [44], Kalantari and Ramamoorthi [14], and Lee *et al.* [19] on six scenes from different dataset in Table II. From Table II, we can see our running time is much less than Sen *et al.* [12] and Hu *et al.* [13], and close to the running time of other methods. In addition, the greatest advantage of FBLS is the fast training. We compare the training time with Kalantari and Ramamoorthi [14] which cost nearly three days to train their CNN network on the Kalantari's dataset, while our FBLS only cost 10 min. We first input 30 sets of images as the training set, then add 10 sets each time and record the training time. With the increase of the training images, Kalantari's (the traditional deep-learning method) training time has significant growth. However, due to our model is a flat network, our training time is much less than the deep learning.

## VII. Conclusion

In this article, we propose a novel model in the HDR synthesis filed which synthesizes an HDR image using a set of multiview LDR images. For generating high-quality HDR results, we use the FBLS as our learning model and create the dataset of multiview LDR images with its corresponding ground truth. The training time of our system is much less than the deep neural network since our model is a flat network consisting of fuzzy subsystems and enhancement groups. In addition, we present the incremental learning algorithm for this model. We can increase the enhancement groups and fuzzy subsystems to optimize our model instead of retraining the entire network. We compare massive substantial experiments with the state-of-the-art methods to prove our results contain more details and less ghosting.

However, there are some limitations in this article. Due to the restriction of the model, we could not support multiple input LDRs for more than three images. More LDR images can contain more complete details of the scene. In addition, there are many matrix operations in the computation of weights, which requires lots of memory. Therefore, in the future, we will attempt to optimize our model to support more LDR inputs. And, we will try to reduce the space complexity for the computation of weights.

## References

[1] C. Liu, "Beyond pixels: Exploring new representations and applications for motion analysis," Ph.D. dissertation, Elect. Eng. Comput. Sci., Massachusetts Inst. Technol., Cambridge, MA, USA, 2009.
[2] Y. Tao, Y. Shen, B. Sheng, P. Li, and R. W. H. Lau, "Video decolorization using visual proximity coherence optimization," *IEEE Trans. Cybern.*, vol. 48, no. 5, pp. 1406–1419, May 2018.
[3] B. Zhang, B. Sheng, P. Li, and T.-Y. Lee, "Depth of field rendering using multilayer-neighborhood optimization," *IEEE Trans. Vis. Comput. Graphics*, to be published.
[4] B. Sheng *et al.*, "Retinal vessel segmentation using minimum spanning superpixel tree detector," *IEEE Trans. Cybern.*, vol. 49, no. 7, pp. 2707–2719, Jul. 2019.
[5] B. Sheng, P. Li, Y. Jin, P. Tan, and T.-Y. Lee, "Intrinsic image decomposition with step and drift shading separation," *IEEE Trans. Vis. Comput. Graphics*, to be published.
[6] S. K. Nayar and T. Mitsunaga, "High dynamic range imaging: Spatially varying pixel exposures," in *Proc. IEEE CVPR*, vol. 1, 2000, pp. 472–479.

[7] M. McGuire, W. Matusik, H. Pfister, B. Chen, J. F. Hughes, and S. K. Nayar, "Optical splitting trees for high-precision monocular imaging," *IEEE Comput. Graph. Appl.*, vol. 27, no. 2, pp. 32–42, Mar./Apr. 2007.

[8] S. W. Hasinoff, F. Durand, and W. T. Freeman, "Noise-optimal capture for high dynamic range photography," in *Proc. IEEE CVPR*, 2010, pp. 553–560.

[9] M. D. Tocci, C. Kiser, N. Tocci, and P. Sen, "A versatile HDR video production system," *ACM Trans. Graph.*, vol. 30, no. 4, p. 41, 2011.

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

[11] S. Mann and R. W. Picard, "Being 'undigital' with digital cameras: Extending dynamic range by combining differently exposed pictures," Singapore-MIT Alliance, Massachusetts Inst. Technol., Cambridge, MA, USA, Rep. 323, 1994.

[12] 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, p. 203, 2012.

[13] J. Hu, O. Gallo, K. Pulli, and X. Sun, "HDR deghosting: How to deal with saturation?" in *Proc. IEEE CVPR*, 2013, pp. 1163–1170.

[14] N. K. Kalantari and R. Ramamoorthi, "Deep high dynamic range imaging of dynamic scenes," *ACM Trans. Graph.*, vol. 36, no. 4, p. 144, 2017.

[15] G. Eilertsen, J. Kronander, G. Denes, R. K. Mantiuk, and J. Unger, "HDR image reconstruction from a single exposure using deep CNNs," *ACM Trans. Graph.*, vol. 36, no. 6, p. 178, 2017.

[16] O. Gallo, N. Gelfandz, W.-C. Chen, M. Tico, and K. Pulli, "Artifact-free high dynamic range imaging," in *Proc. IEEE Int. Conf. Comput. Photography*, 2009, pp. 1–7.

[17] Y. S. Heo, K. M. Lee, S. U. Lee, Y. Moon, and J. Cha, "Ghost-free high dynamic range imaging," in *Proc. Asian Conf. Comput. Vis.*, 2011, pp. 486–500.

[18] 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.

[19] C. Lee, Y. Li, and V. Monga, "Ghost-free high dynamic range imaging via rank minimization," *IEEE Signal Process. Lett.*, vol. 21, no. 9, pp. 1045–1049, Sep. 2014.

[20] X. Qin, J. Shen, X. Mao, X. Li, and Y. Jia, "Robust match fusion using optimization," *IEEE Trans. Cybern.*, vol. 45, no. 8, pp. 1549–1560, Aug. 2015.

[21] K. Ma, H. Li, H. Yong, Z. Wang, D. Meng, and L. Zhang, "Robust multi-exposure image fusion: A structural patch decomposition approach," *IEEE Trans. Image Process.*, vol. 26, no. 5, pp. 2519–2532, May 2017.

[22] S. Feng and C. L. P. Chen, "Fuzzy broad learning system: A novel neuro-fuzzy model for regression and classification," *IEEE Trans. Cybern.*, to be published.

[23] G. Ward, "Fast, robust image registration for compositing high dynamic range photographs from hand-held exposures," *J. Graph. Tools*, vol. 8, no. 2, pp. 17–30, 2003.

[24] A. W. Tomaszewska and R. Mantiuk, "Image registration for multi-exposure high dynamic range image acquisition," in *Proc. Int. Conf. Central Europe Comput. Graph. Visual. Comput. Vis.*, 2007, pp. 49–56.

[25] L. Bogoni, "Extending dynamic range of monochrome and color images through fusion," in *Proc. Int. Conf. Pattern Recognit.*, vol. 3, 2000, pp. 7–12.

[26] 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.

[27] S. Mangiat and J. Gibson, "High dynamic range video with ghost removal," in *Proc. SPIE*, vol. 7798, no. 1, 2010, Art. no. 779812.

[28] H. Zimmer, A. Bruhn, and J. Weickert, "Freehand HDR imaging of moving scenes with simultaneous resolution enhancement," *Comput. Graph. Forum*, vol. 30, no. 2, pp. 405–414, 2011.

[29] X. Liu and A. El Gamal, "Synthesis of high dynamic range motion blur free image from multiple captures," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 50, no. 4, pp. 530–539, Apr. 2003.

[30] T. Grosch, "Fast and robust high dynamic range image generation with camera and object movement," in *Proc. Vis. Model. Visual.*, 2006, pp. 1–9.

[31] S. Raman and S. Chaudhuri, "Reconstruction of high contrast images for dynamic scenes," *Vis. Comput.*, vol. 27, no. 12, pp. 1099–1114, 2011.

[32] E. A. Khan, A. O. Akyuz, and E. Reinhard, "Ghost removal in high dynamic range images," in *Proc. IEEE ICIP*, 2006, pp. 2005–2008.

[33] K. Jacobs, C. Loscos, and G. Ward, "Automatic high-dynamic range image generation for dynamic scenes," *IEEE Comput. Graph. Appl.*, vol. 28, no. 2, pp. 84–93, Mar./Apr. 2008.

[34] T. Jinno and M. Okuda, "Motion blur free HDR image acquisition using multiple exposures," in *Proc. IEEE ICIP*, 2008, pp. 1304–1307.

[35] Y. Nie, H. Sun, P. Li, C. Xiao, and K.-L. Ma, "Object movements synopsis via part assembling and stitching," *IEEE Trans. Vis. Comput. Graphics*, vol. 20, no. 9, pp. 1303–1315, Sep. 2014.

[36] Y. Nie, C. Xiao, H. Sun, and P. Li, "Compact video synopsis via global spatiotemporal optimization," *IEEE Trans. Vis. Comput. Graphics*, vol. 19, no. 10, pp. 1664–1676, Oct. 2013.

[37] D. Sidibe, W. Puech, and O. Strauss, "Ghost detection and removal in high dynamic range images," in *Proc. Eur. Signal Process. Conf.*, 2009, pp. 2240–2244.

[38] T.-H. Min, R.-H. Park, and S. Chang, "Histogram based ghost removal in high dynamic range images," in *Proc. IEEE ICME*, 2009, pp. 530–533.

[39] F. Pece and J. Kautz, "Bitmap movement detection: HDR for dynamic scenes," in *Proc. Conf. Vis. Media Prod.*, 2010, pp. 1–8.

[40] S. Wu, S. Xie, S. Rahardja, and Z. Li, "A robust and fast anti-ghosting algorithm for high dynamic range imaging," in *Proc. IEEE ICIP*, 2010, pp. 397–400.

[41] W. Zhang and W.-K. Cham, "Gradient-directed composition of multi-exposure images," in *Proc. IEEE CVPR*, 2010, pp. 530–536.

[42] 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, pp. 1–8, 2013.

[43] M. Granados, K. I. Kim, J. Tompkin, and C. Theobalt, "Automatic noise modeling for ghost-free HDR reconstruction," *ACM Trans. Graph.*, vol. 32, no. 6, pp. 1–10, 2013.

[44] T.-H. Oh, J.-Y. Lee, Y.-W. Tai, and I. S. Kweon, "Robust high dynamic range imaging by rank minimization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 6, pp. 1219–1232, Jun. 2015.

[45] D. Lin, R. Zhang, Y. Ji, P. Li, and H. Huang, "SCN: Switchable context network for semantic segmentation of RGB-D images," *IEEE Trans. Cybern.*, to be published.

[46] A. Kamel, B. Sheng, P. Yang, P. Li, R. Shen, and D. D. Feng, "Deep convolutional neural networks for human action recognition using depth maps and postures," *IEEE Trans. Syst., Man, Cybern., Syst.*, to be published.

[47] A. Kamel, B. Sheng, P. Li, J. Kim, and D. D. Feng, "Efficient body motion quantification and similarity evaluation using 3-D joints skeleton coordinates," *IEEE Trans. Syst., Man, Cybern., Syst.*, to be published.

[48] B. Sheng, P. Li, C. Gao, and K.-L. Ma, "Deep neural representation guided face sketch synthesis," *IEEE Trans. Vis. Comput. Graphics*, to be published.

[49] E. Shelhamer, J. Long, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 4, pp. 640–651, Apr. 2017.

[50] O. Abdel-Hamid, A.-R. Mohamed, H. Jiang, L. Deng, G. Penn, and D. Yu, "Convolutional neural networks for speech recognition," *IEEE Trans. Audio, Speech, Language Process.*, vol. 22, no. 10, pp. 1533–1545, Oct. 2014.

[51] M. Gong, J. Zhao, J. Liu, Q. Miao, and L. Jiao, "Change detection in synthetic aperture radar images based on deep neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 1, pp. 125–138, Jan. 2016.

[52] W. Hou, X. Gao, D. Tao, and X. Li, "Blind image quality assessment via deep learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 6, pp. 1275–1286, Jun. 2015.

[53] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, 2006.

[54] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.

[55] M. Leshno, V. Y. Lin, A. Pinkus, and S. Schocken, "Multilayer feedforward networks with a nonpolynomial activation function can approximate any function," *Neural Netw.*, vol. 6, no. 6, pp. 861–867, 1993.

[56] Y.-H. Pao and Y. Takefuji, "Functional-link net computing: Theory, system architecture, and functionalities," *Computer*, vol. 25, no. 5, pp. 76–79, May 1992.

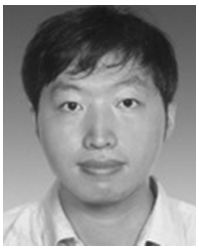
[57] Y.-H. Pao, G.-H. Park, and D. J. Sobajic, "Learning and generalization characteristics of the random vector functional-link net," *Neurocomputing*, vol. 6, no. 2, pp. 163–180, 1994.

[58] C. L. P. Chen and J. Z. Wan, "A rapid learning and dynamic stepwise updating algorithm for flat neural networks and the application to time-series prediction," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 29, no. 1, pp. 62–72, Feb. 1999.

[59] C. L. P. Chen and Z. Liu, "Broad learning system: An effective and efficient incremental learning system without the need for deep architecture," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 1, pp. 10–24, Jan. 2018.
[60] Z. Chen, T. Gao, B. Sheng, P. Li, and C. L. P. Chen, "Outdoor shadow estimating using multiclass geometric decomposition based on BLS," *IEEE Trans. Cybern.*, to be published.
[61] D. Wang, X.-J. Zeng, and J. A. Keane, "Hierarchical hybrid fuzzy-neural networks for approximation with mixed input variables," *Neurocomputing*, vol. 70, nos. 16–18, pp. 3019–3033, 2007.
[62] H.-J. Rong, G.-B. Huang, N. Sundararajan, and P. Saratchandran, "Online sequential fuzzy extreme learning machine for function approximation and classification problems," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 39, no. 4, pp. 1067–1072, Aug. 2009.
[63] Z.-L. Sun, K.-F. Au, and T.-M. Choi, "A neuro-fuzzy inference system through integration of fuzzy logic and extreme learning machines," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 37, no. 5, pp. 1321–1331, Oct. 2007.
[64] R. Mantiuk, K. J. Kim, A. G. Rempel, and W. Heidrich, "HDR-VDP-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions," *ACM Trans. Graph.*, vol. 30, no. 4, pp. 1–14, 2011.
[65] J. Hu, O. Gallo, and K. Pulli, "Exposure stacks of live scenes with hand-held cameras," in *Proc. Eur. Conf. Comput. Vis.*, 2012, pp. 499–512.

**Hongbin Guo** received the B.Eng. degree in intelligence science and technology from Sun Yat-sen University, Guangzhou, China. He is currently pursuing the M.Eng. degree in computer science with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, China.

His current research interests include high dynamic range image synthesis, fuzzy broad learning system, and computer vision.



**Bin Sheng** received the Ph.D. degree in computer science and engineering from the Chinese University of Hong Kong, Hong Kong.

He is currently an Associate Professor with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, China. His current research interests include virtual reality and computer graphics.

**Ping Li** (M'14) received the Ph.D. degree in computer science and engineering from the Chinese University of Hong Kong, Hong Kong.

He is currently a Research Assistant Professor with Hong Kong Polytechnic University, Hong Kong. He has one image/video processing national invention patent, and has excellent research project reported worldwide by *ACM TechNews*. His current research interests include image/video stylization, artistic rendering and synthesis, and creative media.

**C. L. Philip Chen** (S'88–M'88–SM'94–F'07) received the M.S. degree in electrical engineering from the University of Michigan, Ann Arbor, MI, USA, in 1985, and the Ph.D. degree in electrical engineering from Purdue University, West Lafayette, IN, USA, in 1988.

He is currently a Chair Professor and the Dean of the School of Computer Science and Engineering, South China University of Technology, Guangzhou, China. Being a Program Evaluator of the Accreditation Board of Engineering and Technology Education in the USA, for computer engineering, electrical engineering, and software engineering programs, he successfully architects the University of Macau's Engineering and Computer Science programs receiving accreditations from Washington/Seoul Accord through Hong Kong Institute of Engineers (HKIE), of which is considered as his utmost contribution in engineering/computer science education for Macau as the former Dean of the Faculty of Science and Technology. His current research interests include systems, cybernetics, and computational intelligence.

Dr. Chen was a recipient of the IEEE Norbert Wiener Award in 2018 for his contribution in systems and cybernetics, and machine learning, the 2018 Highly Cited Researcher Award in Computer Science by Clarivate Analytics, the 2016 Outstanding Electrical and Computer Engineers Award from his alma mater, Purdue University, after he graduated from the University of Michigan at Ann Arbor in 1985. He was the IEEE Systems, Man, and Cybernetics Society President from 2012 to 2013. He is currently the Editor-in-Chief of the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS: SYSTEMS, an Associate Editor of the IEEE TRANSACTIONS ON FUZZY SYSTEMS, and the IEEE TRANSACTIONS ON CYBERNETICS. He is currently the Vice President of the Chinese Association of Automation (CAA). He is a fellow of AAAS, IAPR, CAA, and HKIE; a member of Academia Europaea, the European Academy of Sciences and Arts, and the International Academy of Systems and Cybernetics Science.