Three-Dimensional Video Super-Resolution Reconstruction Scheme Based on Histogram Matching and Recursive Bayesian Algorithms

GHALIB H. ALSHAMMRI [1,2], **AMANI K. SAMHA** [3], **WALID EL-SHAFAI** [4,5], **EMAD A. ELSHEIKH** [6,7], **EBRAHIM ABDEL HAMID** [6], **MOHAMED I. ABDO** [6], **MOHAMMED AMOON** [1,8], **AND FATHI E. ABD EL-SAMIE** [4,9]

[1] Department of Computer Science, Community College, King Saud University, Riyadh 11437, Saudi Arabia
[2] Deanship of Scientific Research, Saudi Electronic University, Riyadh 11673, Saudi Arabia
[3] Management Information System Department, College of Business Administration, King Saud University, Riyadh 28095, Saudi Arabia
[4] Department of Electronics and Electrical Communications Engineering, Faculty of Electronic Engineering, Menoufia University, Menouf 32952, Egypt
[5] Security Engineering Laboratory, Computer Science Department, Prince Sultan University, Riyadh 11586, Saudi Arabia
[6] Department of Industrial Electronics and Control Engineering, Faculty of Electronic Engineering, Menoufia University, Menouf 32952, Egypt
[7] Department of Computers and Systems Engineering, High Institute of Electronic Engineering, Ministry of Higher Education, Bilbis, Sharqiya 44621, Egypt
[8] Department of Computer Science and Engineering, Faculty of Electronic Engineering, Menoufia University, Menouf 32952, Egypt
[9] Department of Information Technology, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, Riyadh 11564, Saudi Arabia

Corresponding author: Walid El-Shafai (eng.waled.elshafai@gmail.com)

ABSTRACT Multimedia Super-Resolution (SR) reconstruction is an essential and mandatory process for different visualization functions. Recently, several schemes have been suggested for single- and multi-image SR reconstruction. This work presents an effective SR reconstruction scheme for visual quality and resolution enhancement of 3D Video (3DV) sequences. The idea behind the proposed 3DV SR reconstruction scheme is the utilization of a recursive Bayesian algorithm for improving the resolution of the degraded 3DV sequences with down-sampling, blurring, and noise effects. In addition, a significant stage of histogram matching based on a visual image with a better-distributed histogram is employed. The main aim of employing the histogram matching stage for enhancing the 3DV sequence is to introduce a dynamic range modification of each 3DV frame. Hence, it presents a 3DV sequence with an enhanced distribution of intensities. This improves the whole performance efficiency of the suggested scheme. The performance of the proposed SR reconstruction scheme is compared with that of the conventional bicubic interpolation scheme. Comparisons with recent and related SR reconstruction schemes are also introduced. Simulation results reveal that the proposed scheme achieves superior outcomes in terms of Structural Similarity (SSIM) index, local contrast, average gradient, Mean Square Error (MSE), edge intensity, entropy, and Peak Signal-to-Noise Ratio (PSNR) of the resulting 3DV frames.

INDEX TERMS 3DV SR, recursive Bayesian algorithm, bicubic interpolation, histogram matching, image quality enhancement.

I. INTRODUCTION

Multimedia Super-Resolution (SR) reconstruction schemes are used to enhance the resolution of a single image, a single video frame, multiple images, or a video stream. Hence, SR reconstruction schemes have been introduced to allow High-Resolution (HR) images to be created from numerous detected Low-Resolution (LR) images. The single-image-based and multiple-image-based SR reconstruction schemes

are built based on the number of input images. The input of a single-image SR reconstruction scheme is a single LR image. On the other hand, in the multiple-image SR reconstruction, different images of the same scene are utilized. SR reconstruction schemes can be employed in the spatial or a transform domain [1], [2]. In the spatial domain, the operations are directly performed on pixels. Different transforms are utilized, such as Fourier transform, wavelet transform, dual-tree complex wavelet transform, etc. The transform converts the input image into different components, and then mathematical operations are performed on them. After that, the

The associate editor coordinating the review of this manuscript and approving it for publication was Jiachen Yang .

image is transformed back into the spatial domain for display purposes.

According to the operation principles, SR reconstruction schemes may be categorized into three principal categories: learning-based, reconstruction-based, and interpolation-based schemes [3], [4]. The first and second categories are extremely popular among the existing SR reconstruction schemes. The reconstruction-based SR category is mainly concerned with multiple-image SR reconstruction schemes [5], [6].

The image SR reconstruction problem has been treated in the literature for decades as an inverse problem. In this type of inverse problems, multiple frames degraded with both blurring and noise, in addition to some registration and motion errors, are available. The objective is to get HR images of the LR degraded images. The registration process must be firstly performed, and then the degradation model is inverted either on a single-channel or multi-channel basis in the presence of blurring and noise effects. The ill-posedness of this problem is attributed to the inversion process of the channel degradation model in the presence of noise. This inversion may lead to noise enhancement [5], [6]. The solution of this problem begins with some attempts to adapt the Linear Minimum Mean Square Error (LMMSE) algorithm [7]. This solution is performed with Fourier transform and Toeplitz-to-circulant approximations [8], [9]. However, this solution requires estimating the Signal-to-Noise Ratio (SNR) of the image before the implementation of the SR reconstruction scheme. Besides, there is a need for approximating the auto-correlation matrix for the unavailable HR image from the degraded ones. To overcome the limitations of the LMMSE algorithm, entropy concepts have been adopted. In this approach, the HR image is reconstructed based on a cost function, and various constraints are set to maximize the entropy [10]. This approach is easy to implement compared to the inverse solution that requires diagonal matrices. However, it does not consider the image local activity levels [11].

The regularization algorithms are considered excellent solutions for ill-posed problems in image processing [12]. The regularization-based SR reconstruction is performed using multiple stages of multi-channel restoration, image fusion, and image interpolation [13]. Other trends have been presented for blind SR reconstruction of images based on the greatest common divisor algorithm. However, these trends depend on the assumption of co-prime blurring operators. In addition, the quality of the obtained images are very limited [14].

The process of video SR reconstruction aims to obtaining an HR video from one or multiple LR videos to enhance both temporal and spatial resolutions. The temporal correlations in the input LR frames can be exploited to obtain the best HR video quality. Most SR reconstruction schemes use the information from multiple frames to produce an upscaled frame [12]. These multi-frame schemes differ considerably from the complicated single-frame upscaling schemes that generate artificial information.

There are two main types of resolution in an imaging system: spatial and temporal. The spatial resolution depends on the spatial density of the sensor (or detector) array. Moreover, the blur effect is attributed to the point spread function of the sensors. On the other hand, the temporal resolution is affected by the camera exposure time and its frame rate [15]. Therefore, the aliasing effect on the spatial and temporal resolutions causes more degradations in the captured images and videos. The spatial aliasing is introduced in video frames or images once the lens cut-off frequency is higher than the detector cut-off frequency. On the other hand, temporal aliasing occurs in the video frames, when the camera frame rate is low and the captured objects move faster. The blur effect on videos and images that degrades their quality has different types like motion blur, defocus blur, detector blur, and optical blur. Hence, there is an urgent need for introducing efficient SR reconstruction schemes to mitigate these problems.

Therefore, the proposed work considers the blurring, down-sampling, and noise effects for efficiently improving the performance of the 3DV SR reconstruction scheme. The proposed 3DV SR reconstruction scheme assumes that the input degraded 3DV LR frames are blurred and down-sampled noisy forms of the 3DV HR frames. Consequently, this paper presents an efficient SR reconstruction scheme for visual quality and resolution enhancement of 3DV sequences. A simple iterative Bayesian algorithm is followed in the de-blurring and de-noising processes. In addition, a further stage of histogram matching is utilized to enhance the quality of the resulting 3DV frames from the recursive Bayesian algorithm to improve the whole efficiency of the suggested SR reconstruction scheme.

The major merits of the proposed 3DV SR reconstruction scheme are as follows:

- The Bayesian probabilistic model can efficiently formulate the nonlinear mappings in the temporal and spatial directions of the 3DV frames. It can improve the fidelity and visual quality of multiple video frames.
- The proposed scheme deals with the motion, blur, and noise effects. It can achieve good results in the presence of undetermined noise, unspecified blur kernel, and random motion.
- An efficient motion estimation algorithm is utilized to minimize the processing time of the SR reconstruction scheme.
- A non-uniformity correction algorithm based on histogram matching is used to enhance the quality of the reconstructed 3DV HR frames.

The remainder of the work is structured as follows. Section 2 demonstrates the literature review of various image and video SR reconstruction schemes. The proposed 3DV SR reconstruction scheme is discussed in Section 3. Section 4 shows the analysis of the proposed scheme. Finally, the concluding remarks and future research directions are given in Section 5.

II. RELATED WORK

The resolution of an image is one of the most important parameters for efficient image representation. The SR reconstruction schemes are used to enhance the resolution of a specific image extracted from multiple LR images or a single LR image. SR reconstruction schemes have several applications such as satellite imaging, medical imaging, High-Definition TV (HDTV), and surveillance [5], [11].

The SR reconstruction process from a single image is termed SISR [16]. The SISR schemes comprise reconstruction-, learning-, and interpolation-based schemes. Reconstruction-based schemes produce water-color artifacts. On the other hand, interpolation-based schemes, like bicubic interpolation, remain the main tool of digital zooming in consumer devices, but they yield ringing or blurring artifacts. The SISR can be implemented with the reconstruction-based and interpolation-based schemes. In the learning-based schemes, a mapping from LR patches to their corresponding HR patches is learned to utilize a set of HR-LR patch pairs. Most of these mappings are computation- and memory-intensive due to the creation of an LR-HR patch database through searching inside the same image. However, direct nonlinear regression mapping amongst HR and LR patches needs adequate time and memory requirements.

The image and video frame resolutions are limited due to sensor physical characteristics like detector density and size. In addition, image or video frame degradations are caused during the recording process. For example, motion blur is caused by limited shutter speed, optical distortion, aliasing effects, and noise. There are two approaches to produce an HR frame. The first one is by installing an HR sensor. Unfortunately, it is not reasonable due to the cost increase and the consumed power. The second one is by using a post-processing method to build the HR video frame. The more appealing approach in this research area is the resolution enhancement or SR reconstruction [15], [16].

Several schemes can be utilized to recover an HR image from one or multiple LR images. The standard interpolation techniques, such as nearest neighbor interpolation, have been used [17]. The resulting images from these techniques have blurred edges, but they do well in smoother regions. In the case of conventional multi-frame SR reconstruction schemes, several LR images of the same scene with different pixel shifts are taken as inputs, and the correspondence between HR and LR patches is learned from a database that consists of LR and HR image pairs. After that, this knowledge is exploited on a new LR image to rebuild the corresponding HR image.

In [18], the authors presented an efficient algorithm to obtain image patches with consistent structures and exact pixel values. Accordingly, it has flexibility and power to designate different image patterns. Asuni *et al.* [19] introduced an algorithm named Smooth Regression with Local Structure Prior (SRLSP) that was employed to forecast the expected HR pixel data from the LR patch. The main concept of this algorithm is to create an HR image from an LR one using a set of training samples.

Jin *et al.* [20] suggested the Bayesian and affine geometric distortion algorithm for realistic sequence SR reconstruction. They presented a design of a robust kernel for improved edge localization. Moreover, they introduced a sufficient procedure for initializing the optimization method, including misregistration uncertainties. The authors of [21] proposed an adaptive image prior model and established a Bayesian SR reconstruction algorithm. In this algorithm, the degree of correlation between pixels is modified adaptively by an adaptive norm, which is inferred based on the local image features. This algorithm avoids artifacts in the smoothed regions and also preserves edge details.

Katartzis *et al.* [22] proposed the Convolutional Sparse Coding (CSC) algorithm with joint Bayesian learning for SR reconstruction of images. The algorithm performance depends on the selection of certain parameters. In addition, a coupled Beta-Bernoulli procedure is utilized to nominate suitable Sparse Coding Maps (SCMs) and filters for both low- and high-resolution images. The sparse feature maps and filters for both HR and LR images are realized adaptively through the Bayesian learning methodology. Zhao *et al.* [23] suggested a video SR reconstruction scheme by filling the missing pixels with virtual-view or spatial interpolated pixels. This scheme achieved better performance than other ones.

Ge *et al.* [24] suggested a multi-frame SR reconstruction scheme with less processing time and lower complexity than those of other multi-frame SR reconstruction schemes. This scheme could reproduce adequate HR video streams with higher HR frame quality. Jin *et al.* [25] focused on improving a Bayesian method to combine SR reconstruction with video stream merging. An adaptive video SR reconstruction scheme based on a Bayesian probabilistic model via concurrently estimating the underlying blur kernel, motion, and noise level, while recovering the original HR frames, has been proposed [26]. This scheme delivers auspicious SR outcomes, and it is adapted to various blur kernels and noise levels. To investigate the effect of blur kernel and noise, the Cramer-Rao bounds have been used to perform a two-step analysis to show the impact of the blur kernel and noise level on the performance of the SR reconstruction scheme.

Chen *et al.* [27] suggested a multi-frame SR reconstruction scheme based on the Bidirectional Recurrent Convolutional Network (BRCN). The primary impact is for the utilization of 3D feedforward convolutions, in recurrent, and bidirectional networks for sufficient temporal modeling. This scheme has some limitations of computational complexity, lower reconstruction performance, and the need for large-scale video SR datasets. Liu *et al.* [28] introduced an image SR reconstruction scheme that has been combined with a multi-camera system to exploit the benefits of temporal and spatial correlations amongst the recorded streams. Three SR reconstruction schemes have been suggested: spatial-temporal, temporal-spatial, and mixed SR. The mixed SR scheme outperforms the

other schemes, because it has a flexible scaling factor, and it has a reduced computational cost as a unique SR reconstruction procedure is performed. Huang *et al.* [29] proposed a video SR reconstruction scheme with improved robustness to innovation outliers in real-time operation. An intuitive study was suggested for the proximal-point cost function representation of the Recursive Least Mean Squares (RLMS) gradient descent algorithm. This regularization provides faster convergence of the solution in the subspace related to innovations, while maintaining earlier estimated details.

Quevedo *et al.* [30] introduced an efficient configuration model for image and video SR reconstruction. Due to the weaknesses of the Basic Particle Swarm Optimization (BPSO), the genetic and ant colony algorithms have been utilized in the introduced model. Borsoi *et al.* [31] offered a new procedure for video SR reconstruction through a Generative Adversarial Network (GAN). GANs are Artificial Intelligence (AI) algorithms composed of two Neural Networks (NNs) to learn deep representations. It has been proved that GAN results for large upscaling factors are, by a substantial margin, more photo-realistic than reconstructions achieved with reference schemes.

Yu *et al.* [32] proposed a temporal adaptive network and examined various ways of image alignment involving a spatial alignment network to learn the temporal dynamics in order to improve video SR reconstruction. Video SR reconstruction can be launched from two aspects: the inter-frame temporal relation and the intra-frame spatial relation. The spatial alignment modules and the temporal adaptation are manifested to enhance the SR reconstruction quality. Gopan *et al.* [33] suggested depth SR reconstruction on RGB-D video streams with significant displacement 3D motion. This method is improved in two phases: merging of compensated depth images and motion compensation of depth images. For the compensation phase, a 3D Nearest-Neighbor Field (NNF) estimation method called RGB-D super-pixel matching was proposed, demonstrating that it is robust in the existence of large-displacement 3D motion. A deep Convolutional Neural Network (CNN) structure has been developed for the fusion phase to achieve an efficient performance. Liu *et al.* [34] suggested an approach for video SR reconstruction. This approach is the first very deep non-simultaneous, fully-recurrent CNN for video SR reconstruction. The temporal dependencies are modeled by late fusion depending on fully-recurrent convolutional layers and motion compensation. Deep residual learning is employed in the recurrent structure to increase the representation ability of large and complex motions.

Wang *et al.* [35] proposed an efficient algorithm to rebuild high-spatial-resolution video satellite images from a supervised perspective depending on CNNs in order to increase adaptability to video satellite imagery. It was confirmed that utilizing high-spatial-resolution static images for training is necessary for the reconstruction of video satellite motion images. Also, it was proved that the CNN-based SR reconstruction schemes are better than bicubic interpolation.

Li *et al.* [36] introduced a blind deconvolution and SR reconstruction scheme to work on LR video. To estimate the blur(s), the input frames are first up-sampled using the Non-Uniform Interpolation (NUI) SR reconstruction method, assuming that the blurs are identical or have slow variations over time. Then, the blurs are determined iteratively from some enhanced edges in the upsampled frames. After completion of blur estimation, the reconstructed frames are discarded, and a non-blind iterative SR reconstruction process is performed to obtain the final reconstructed frames using the estimated blur(s). A masking process is employed throughout each iteration of the final frame rebuilding to avoid artifacts produced from incorrect motion prediction, successively.

It is observed that several authors presented a lot of research activities on image SR reconstruction in the spatial domain. On the other hand, various modern image and video SR reconstruction schemes have been employed in transform domains. Commonly, there are few contributions in the literature on 3DV SR reconstruction schemes. Moreover, some of the introduced schemes have significant problems with the reconstruction quality of 3DV frames. Because the conventional video SR reconstruction schemes have not accomplished satisfactory subjective and objective quality levels in the presence of down-sampling, blurring, and noise effects, we have been motivated to present the proposed work. Therefore, considering the constraints of the conventional video and image SR reconstruction schemes, the significant contribution of this paper is the introduction of a hybrid 3DV SR reconstruction scheme for improving the resolution and visual quality of 3DV sequences. The proposed 3DV SR reconstruction scheme consists of the recursive Bayesian and the histogram matching algorithms. The first stage is presented to mitigate the effects of blurring, down-sampling, and noise. The second stage is introduced to improve the visual representation of the 3DV frames resulting from the first stage.

III. PROPOSED HYBRID 3DV SR RECONSTRUCTION SCHEME

As stated previously, the video SR reconstruction process can be carried out on either a single LR video frame or multiple LR video frames. Therefore, a video frame SR reconstruction scheme can improve the details in a single LR video frame to produce an enhanced HR video frame. Also, a video frame SR reconstruction scheme can be used for improving multiple LR frames to produce an enhanced HR video frame. The resolution is improved for a single LR frame with a video SR reconstruction scheme by enhancing the edges of the objects presented in the LR frame or replacing an LR patch with its corresponding HR patch through a patch redundancy algorithm. The proposed 3DV SR reconstruction scheme works on multiple LR frames, where enhancing high-frequency components gives the 3DV HR frame by integrating non-repeatable information within multiple LR frames. Therefore, the proposed 3DV SR reconstruction scheme has the main benefit that the resulting 3DV HR frame is generated from

multiple 3DV LR frames with less power consumption and lower computational cost.

In the proposed 3DV SR reconstruction scheme, the basic reconstruction idea of a 3DV HR frame from multiple 3DV LR frames is the exploitation of spatial, temporal, and inter-view correlations from the surrounding 3DV frames in the same scene or neighboring scenes. Figure 1 presents the basic idea for the proposed 3DV SR reconstruction scheme from multiple 3DV LR frames. If the 3DV LR frames have a minor shift, then the 3DV HR reconstructed frame will not contain any new information. By taking one frame as a reference, pixels are interleaved as a 3DV HR frame is generated. In addition, Fig. 1 shows that the subpixel motion provides complementary information among the LR frames that makes the SR reconstruction possible. The main steps of the proposed scheme can be summarized as registration or motion estimation, interpolation, deblurring or noise removal, and histogram matching, as shown in Fig. 2.

Figure 2 displays the proposed 3DV SR reconstruction scheme, which comprises the registration and restoration processes. First, sequences are registered from the input stream of LR 3DV frames to determine the motion relationship between a reference frame q_{LR_i} and the keyframe q_{LR_o}. The registration process estimates the motion of the blocks for every reference frame to the keyframe q_{LR_o}. Then, these 3DV LR input frames are projected into the coordinates of the keyframe, and after that, they are interpolated to recover the 3DV HR keyframe. The keyframe of the 3DV HR sequence is the projected frame of the HR blurred frame. After that, the authentic 3DV HR frame is obtained by deblurring the HR frame. Finally, the histogram matching is exploited to generate the final 3DV HR frame with better distribution of intensities.

It is assumed that the 3DV LR frames q_{LR_i}, with $i = \{-L, \ldots, L\}$, are obtained from the 3DV HR original frames q_{HR_i} as illustrated in (1), where V in video space indicates the down-sampling operator. E indicates the blur operator, and it is performed on the received video frames by convoluting with a blur kernel E and ψ_i denotes the additive system noise.

$$q_{LR_i} = V \times E \times q_{HR_i} + \psi_i \qquad (1)$$

In general, camera movements and scene details can be shifted horizontally, vertically, and arbitrarily rotated [24]. Thus, the q_{HR_i} differs from the q_{HR_o} by a shifting value $\Delta\delta_i$ and a revolving value i_q (frame i motion parameters). Therefore, the $F_{\Delta\delta_i}$ is the shifting function and R_{φ_i} is the revolving function from the q_{HR_i} to q_{HR_o}. The q_{HR_o} can be described as in (2). Thence, equation (1) can be reformulated as expressed in (3).

$$q_{HR_o} = R_{\varphi_i} \times F_{\Delta\delta_i} \times q_{HR_i} \qquad (2)$$

$$q_{LR_i} = V \times E \times R_{\varphi_i} \times F_{\Delta\delta_i} \times q_{HR_o} + \psi_i \qquad (3)$$

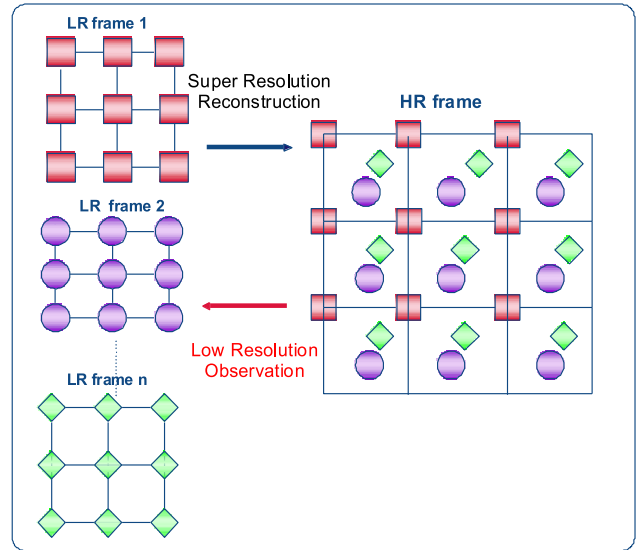After the blurring of the 3DV frame resulting from sampling at LR and vibration or camera motion, the 3DV frames

FIGURE 1. The basic idea for the proposed 3DV frame SR reconstruction from multiple 3DV LR frames.

are obtained. A camera usually moves low. Consequently, the motion blur does not significantly affect the resolution quality of the obtained 3DV frames. The sampling blur can considerably affect capturing a real scene with much details. If the motion blur and sampling blur parameters are required to be estimated simultaneously, afterwards, the estimation process for the blur kernel might be extraordinarily complex. We concentrate on the sampling blur and exclude the motion blur in the proposed scheme to facilitate the SR reconstruction process. Each LR pixel is an average color value of all HR pixels within the range of color values. Therefore, the sampling blur kernel can be determined by a mean filter with a size $h \times h$, and it can be expressed as:

$$E = \frac{1}{h \times h}
\begin{bmatrix}
1 & 1 & \cdots & 1 \\
1 & 1 & \cdots & 1 \\
\vdots & \vdots & \vdots & \vdots \\
1 & 1 & \cdots & 1
\end{bmatrix} \qquad (4)$$

The unknowns include the pixels of the present 3DV HR frame, and the set of parameters of noise and motion. The Bayesian Maximum A Posteriori (MAP) likelihood [24] is exploited in the proposed scheme to get the best possible solution to estimate the unknowns. The utilized Bayesian MAP algorithm has low complexity by solving for the unknown parameters for a block of pixels. The proposed model for estimation is as follows:

$$\{q_{HR_o}^*, \{\Delta_{\delta_i}^*\}, \{\psi_i^*\}, \{\varphi_i^*\}\}$$
$$= argmax_{q_{HR_o}, \{\Delta_{\delta_i}\}, \{\psi_i\}, \{\varphi_i\}}$$
$$\times p(q_{HR_o}, \{\Delta_{\delta_i}\}, \{\psi_i\}, \{\varphi_i\} \,|\{q_{LR_i}\}) \qquad (5)$$

where

$$p(q_{HR_o}, \{\Delta_{\delta_i}\}, \{\psi_i\}, \{\varphi_i\} \,\big|\{q_{LR_i}\})$$
$$= \prod_i p\left(q_{LR_i} \,\big| q_{LR_o}, \Delta\delta_i, \psi_i, \varphi_i\right)$$
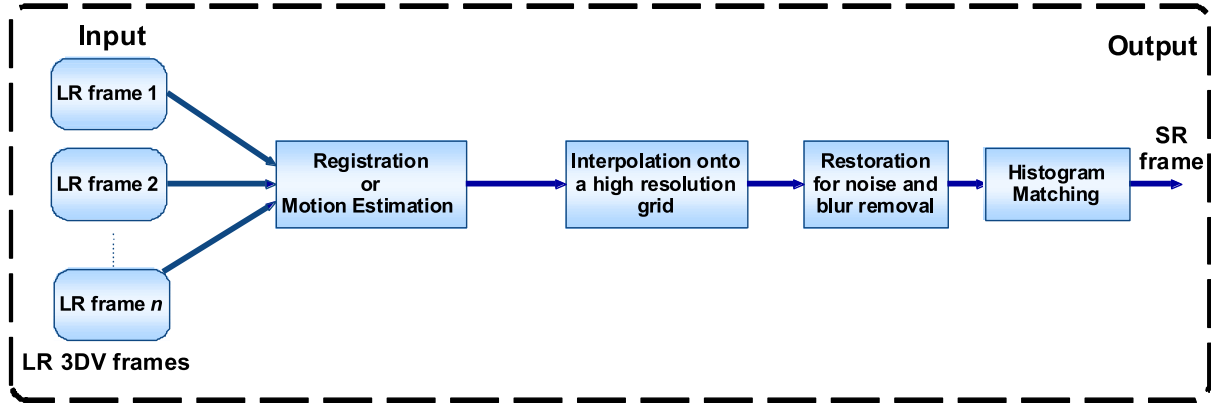
FIGURE 2. The basic steps of the suggested 3DV SR reconstruction scheme.

$$\times p\left(q_{HRo}\right) \prod_i p\left(\Delta\delta_i\right) \prod_i p\left(\psi_i\right) \prod_i p\left(\varphi_i\right) \qquad (6)$$

From equations (5) and (6), we have,

$$\{q^*_{HRo}, \{\Delta^*_{\delta_i}\}, \{\psi^*_i\}, \{\varphi^*_i\}\}$$

$$= argmax_{q_{HRo},\{\Delta_{\delta_i}\},\{\psi_i\},\{\varphi_i\}}$$

$$\times \{\prod_i p\left(q_{LR_i} \middle| q_{LR_o}, \Delta\delta_i, \psi_i, \varphi_i\right)$$

$$\times p\left(q_{HRo}\right) \prod_i p\left(\Delta\delta_i\right) \prod_i p\left(\psi_i\right) \prod_i p\left(\varphi_i\right)\}$$

$$= argmax_{q_{HRo},\{\Delta_{\delta_i}\},\{\psi_i\},\{\varphi_i\}}$$

$$\times \{\sum_i logp\left(q_{LR_i} \middle| q_{LR_o}, \Delta\delta_i, \psi_i, \varphi_i\right)$$

$$+ log(q_{HR_o}) + \sum_i logp(\Delta\delta_i)$$

$$+ \sum_i logp(\psi_i) + \sum_i logp(\varphi_i)\} \qquad (7)$$

where $p\left(q_{HR_i} \middle| q_{HR_o}, \Delta\delta_i, \psi_i, \varphi_i\right)$ refers to the probability distribution of estimated parameters, $\Delta\delta_i$, ψ_i, and φ_i for the likelihood of q_{HR_i}. The log_2 with a base two is used in all presented equations. An exponential distribution is presumed according to the probability given in (8).

$$p\left(q_{LR_i} \middle| q_{HR_o}, \Delta\delta_i, \psi_i, \varphi_i\right)$$

$$= \frac{1}{M_{\alpha i}} exp\{-\left\| q_{LR_i} \right.$$

$$= V \times E \times R_{\varphi_i} \times F_{\Delta\delta_i} \times q_{HR_o} + \psi_i\|\} \qquad (8)$$

where the state numbers for the parameter estimation input are G and $M_{\alpha i}$ as follows:

$$exp\{-\left\| q_{LR_i} = V \times E \times R_{\varphi_i} \times F_{\Delta\delta_i} \times q_{HR_o} + \psi_i\right\|\},$$

$$M_{\alpha i} = \sum_{n=1}^g exp\{-\left\| q_{LR_i} = V \times E \right.$$

$$\times R_{\varphi_i} \times F_{\Delta\delta_i} \times q_{HR_o} + \psi_i\|\} \qquad (9)$$

The parameters q_{HR_o}, $\Delta\delta_i$, ψ_i, and φ_i are optimal, if their gradients have the smallest values. Their exponential probability distributions are given as follows,

$$p\left(q_{HR_o}\right) = \frac{1}{M_\beta} exp\left\{-\left\| \nabla q_{HR_o} \right\|\right\} \qquad (10)$$

FIGURE 3. The original Poznan, Balloons, and Shark frames of the tested 3DV sequences.

$$p\left(\Delta_{\delta_i}\right) = \frac{1}{M_{\gamma_i}} exp\left\{-\left\| \nabla\Delta_{\delta_i} \right\|\right\} \qquad (11)$$

$$p\left(\psi_i\right) = \frac{1}{M_{\varepsilon_i}} exp\left\{-\left\| \nabla\psi_i \right\|\right\} \qquad (12)$$

$$p\left(\varphi_i\right) = \frac{1}{M_{\varepsilon_i}} exp\left\{-\left\| \nabla\varphi_i \right\|\right\} \qquad (13)$$

where M_β, M_{γ_i}, and M_{ε_i} indicate the normalized values for their likelihood distributions. ∇ is the gradient operator,

$$\left\| \nabla \, q_{HR_o} \right\| = \sum_{x,y} \left\| \nabla \, q_{HR_o}\left(x,y\right) \right\|$$

$$= \sum_{x,y} \left(\left| \frac{\partial}{\partial x} q_{HR_o} \right| + \left| \frac{\partial}{\partial y} q_{HR_o} \right|\right) \qquad (14)$$

With equations (6) to (14), equation (5) can be described as:

$$\{q^*_{HRo}, \{\Delta^*_{\delta_i}\}, \{\psi^*_i\}, \{\varphi^*_i\}\}$$

$$= argmax_{q_{HRo},\{\Delta_{\delta_i}\},\{\psi_i\},\{\varphi_i\}}$$

$$\times \{\sum_i \left\| q_{LR_i} - V \times E \times R_{\varphi_i} \times F_{\Delta_{\delta_i}} \times q_{HR_o} + \psi_i \right\|$$

$$\times \left\| \nabla \, q_{HR_o} \right\| + \sum_i \left\| \nabla \, \Delta\delta_i \right\| + \sum_i \left\| \nabla \, \psi_i \right\|$$

$$+ \sum_i \left\| \nabla \, \varphi_i \right\| + \sum_i logM_{\alpha i} + \sum_i log M_\beta$$

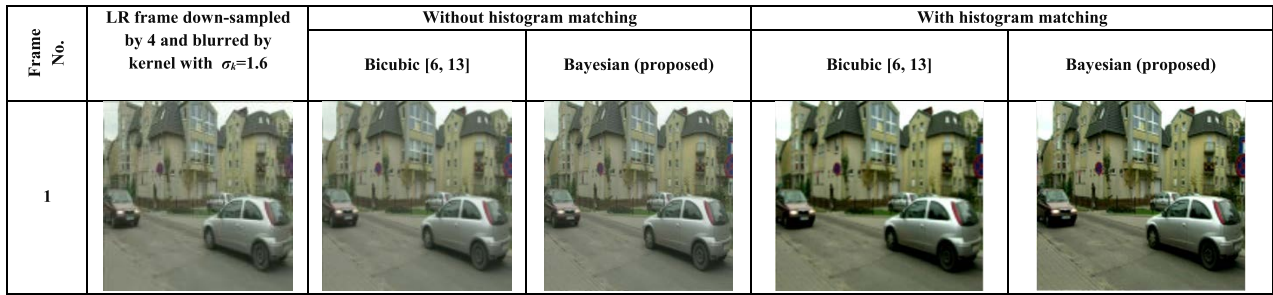| Frame No.	LR frame down-sampled by 4 and blurred by kernel with $\sigma_k=1.6$	Without histogram matching		With histogram matching	
		Bicubic [6, 13]	Bayesian (proposed)	Bicubic [6, 13]	Bayesian (proposed)
1					

FIGURE 4. The visual outcomes of the suggested scheme compared to the bicubic technique for frame 1 of the Poznan 3DV sequence by employing down-sampling by 4, and a blur kernel with $\sigma_k = 1.6$, without AWGN effect.
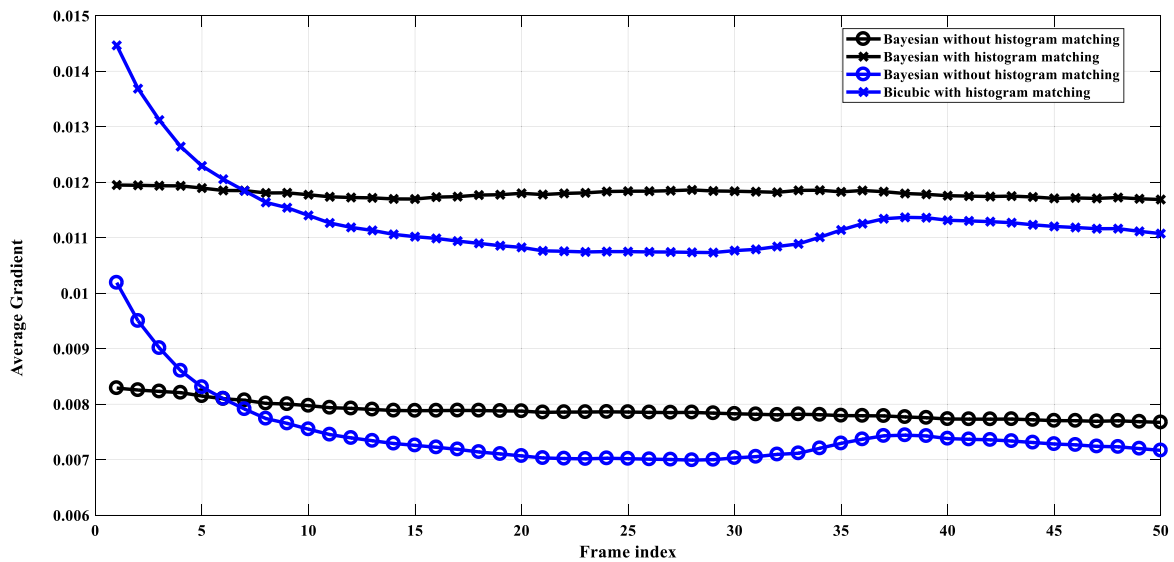

FIGURE 5. Objective comparison of average intensity outcomes of the suggested scheme compared to the conventional bicubic technique on the first 50 frames of the Poznan 3DV sequence for down-sampling by 4, and a blur kernel with $\sigma_k = 1.6$, without AWGN effect.

$$+ \sum_i \log M_{\gamma_i} + \sum_i \log M_{\varepsilon_i} \sum_i \log M_{\varepsilon_i}\} \qquad (15)$$

Because

$$\sum_i \log M_{\alpha i}, \sum_i \log M_{\beta}, \sum_i \log M_{\gamma i}, \text{ and } \sum_i \log M_{\varepsilon i}$$

are constants. Hence, the optimum condition for the proposed SR reconstruction scheme is,

$$\{q^*_{HRo}, \{\Delta^*_{\delta_i}\}, \{\psi^*_i\}, \{\varphi^*_i\}\}$$
$$= argmax_{q_{HRo}, \{\Delta_{\delta_i}\}, \{\psi_i\}, \{\varphi_i\}}$$
$$\times \{\sum_i \left\| q_{LR_i} - V \times E \times R_{\varphi_i} \times F_{\Delta_{\delta_i}} \times q_{HRo} + \psi_i \right\|$$
$$+ \left\| \nabla q_{HR_o} \right\| + \sum_i \left\| \nabla \Delta \delta_i \right\| + \sum_i \left\| \nabla \psi_i \right\|$$
$$+ \sum_i \left\| \nabla \varphi_i \right\|\} \qquad (16)$$

The last step in the suggested 3DV SR reconstruction scheme is histogram matching. Because the histogram shows the different gray level probabilities within the video frame, the histogram matching is a processing tool employed to enhance the visual representation of the resulting 3DV HR frames.

To perform the histogram matching process, we must choose a better-distributed histogram for a reference image in order to achieve a better distribution of intensities for the 3DV HR frames. If we alter the histogram of every frame in the 3DV SR video to be spread over the whole range of the reference image, then the SR video frames will be visually improved. This procedure is known as the histogram matching process. It can be implemented by selecting a reference image with large variance, such that the mean η and variance ζ of the SR video frame become dependent on their counterparts in the reference image. The steps of the employed histogram matching of the SR video frame to the reference image are explained below:

1. Calculate the mean of the authentic SR video frame $B(i, j)$.

$$\hat{\eta}_B = \sum_{i=1}^{K} \sum_{j=1}^{H} B(i, j) \qquad (17)$$
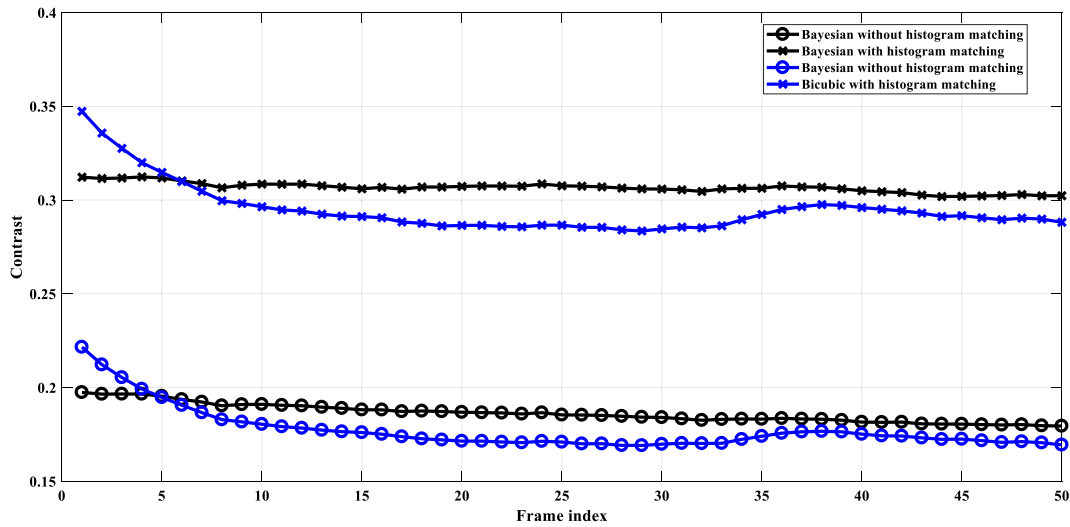
where K and H are the SR video frame dimensions.

FIGURE 6. Objective comparison contrast outcomes of the suggested scheme compared to the conventional bicubic technique on the first 50 frames of the 3DV Poznan sequence for down-sampling by 4, and a blur kernel with $\sigma_k = 1.6$, without AWGN effect.
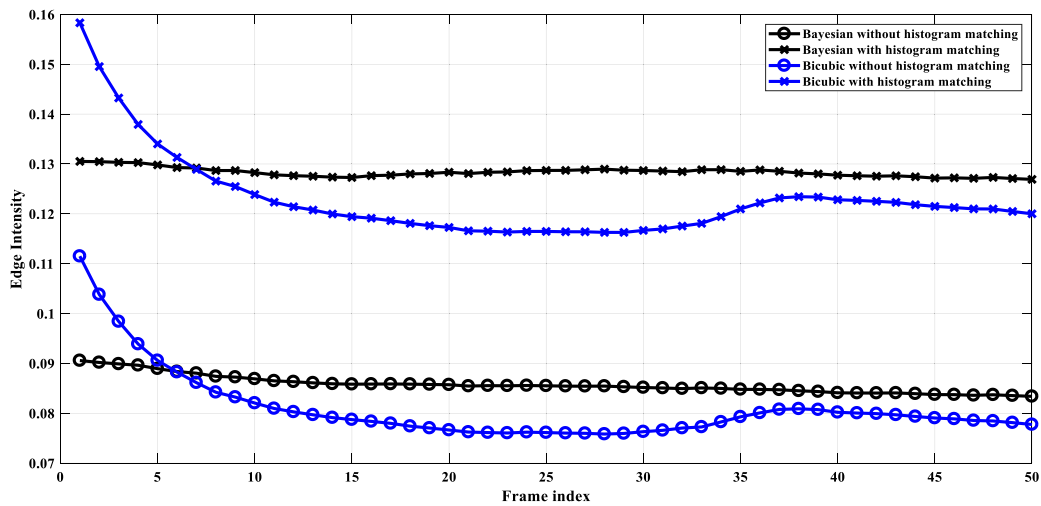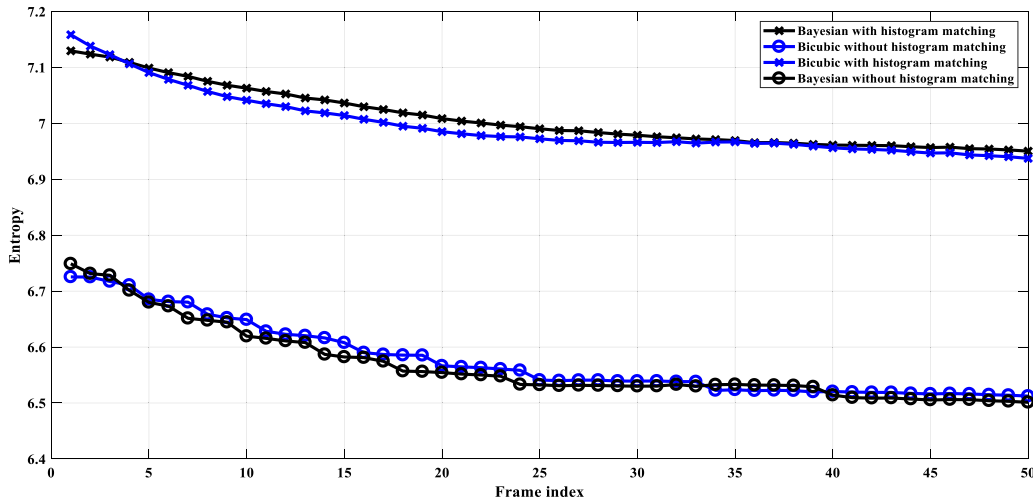

FIGURE 7. Objective comparison edge intensity outcomes of the suggested scheme compared to the conventional bicubic technique on the first 50 frames of the 3DV Poznan sequence for down-sampling by 4, and a blur kernel with $\sigma_k = 1.6$, without AWGN effect.

2. Calculate the reference image $s(w, g)$ mean.

$$\hat{\eta}_s = \sum_{w=1}^{W} \sum_{g=1}^{G} s(w, g) \tag{18}$$

where W and G are the reference image dimensions.

3. Estimate the SR video frame standard deviation ζ_{std1}.

$$\zeta_{std1} = \sqrt{\frac{1}{K \times H} \sum_{i=1}^{K} \sum_{j=1}^{H} (B(i,j) - \hat{\eta}_B)^2} \tag{19}$$

4. Estimate the reference image standard deviation ζ_{std2}.

$$\zeta_{std2} = \sqrt{\frac{1}{W \times G} \sum_{w=1}^{W} \sum_{g=1}^{G} (s(w, g) - \hat{\eta}_s)^2} \tag{20}$$

5. Determine the multiplicative correction factor G_f by dividing the reference image standard deviation by the SR video frame standard deviation.

$$G_f = \frac{\zeta_{std2}}{\zeta_{std1}} \tag{21}$$

6. Calculate the additive correction factor f_{MC}.

$$f_{MC} = \hat{\eta}_s - G_f \times \hat{\eta}_B \tag{22}$$

FIGURE 8. Objective comparison entropy outcomes of the suggested scheme compared to the conventional bicubic technique on the first 50 frames of the 3DV Poznan sequence for down-sampling by 4, and a blur kernel with $\sigma_k = 1.6$, without AWGN effect.
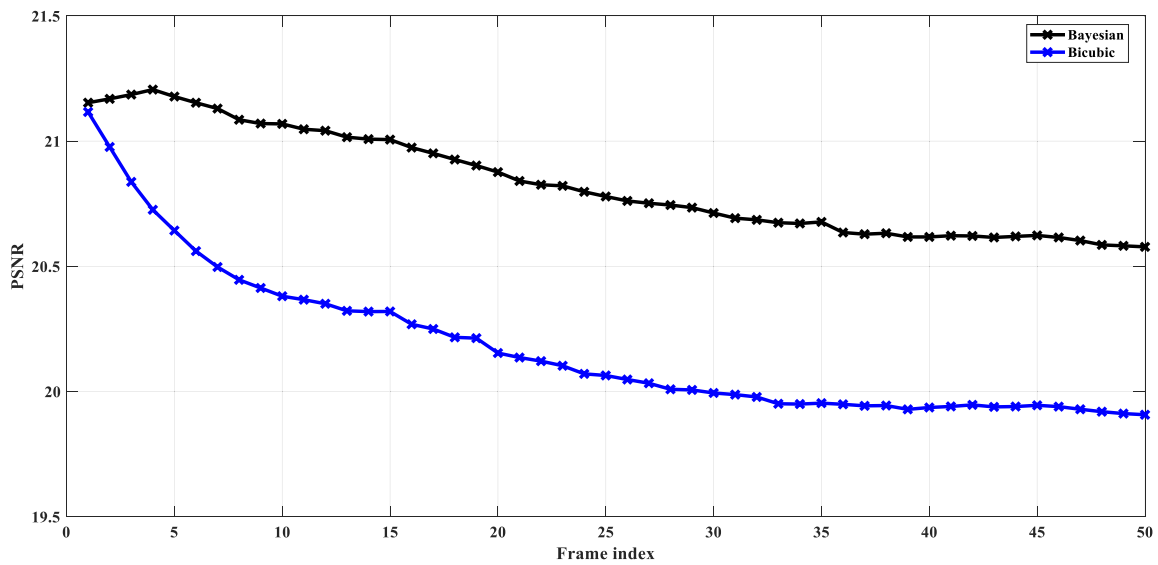
FIGURE 9. Objective comparison PSNR outcomes of the suggested scheme compared to the conventional bicubic technique on the first 50 frames of the 3DV Poznan sequence for down-sampling by 4, and a blur kernel with $\sigma_k = 1.6$, without AWGN effect.

FIGURE 10. Visual outcomes of the suggested scheme compared to the bicubic technique for frame 25 of the Balloons 3DV sequence without down-sampling, and with a blur kernel with $\sigma_k = 1.2$, and AWGN effect with $\sigma_n = 0.01$.

7. Determine the histogram-matched video frame.

$$B_{HM}(i, j) = f_{MC} + B(i, j) \times G_f \qquad (23)$$

IV. RESULTS AND COMPARATIVE ANALYSIS

To evaluate the suggested 3DV SR reconstruction scheme, numerous analyses and tests on the most common and
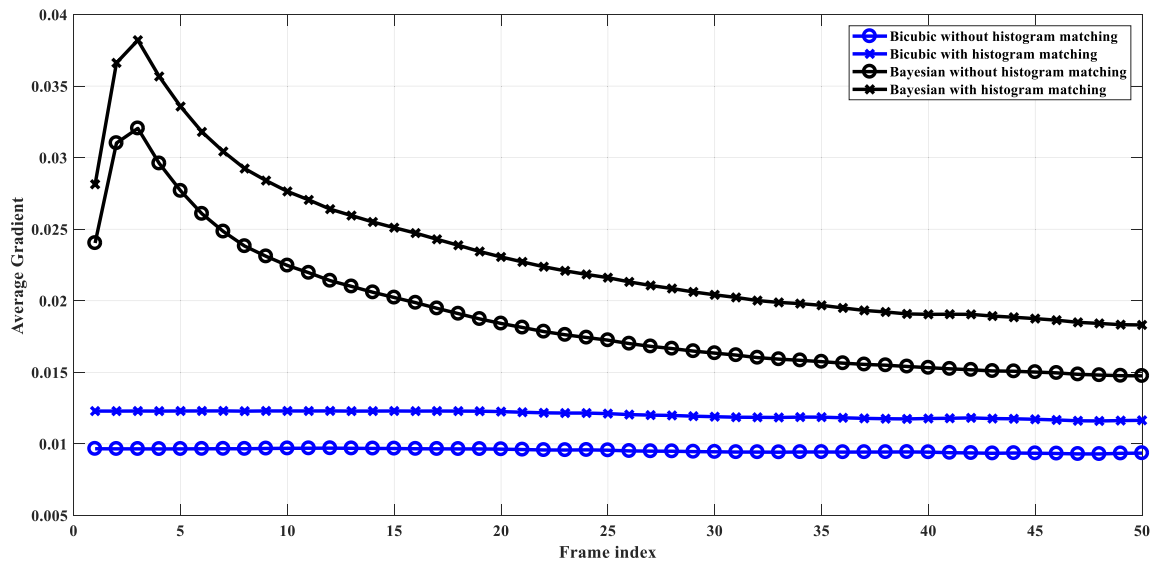
FIGURE 11. Objective comparison of average intensity outcomes of the suggested scheme compared to the bicubic technique on the first 50 frames for the Balloons 3DV sequence without down-sampling, and with a blur kernel with $\sigma_k = 1.2$, and AWGN effect with $\sigma_n = 0.01$.
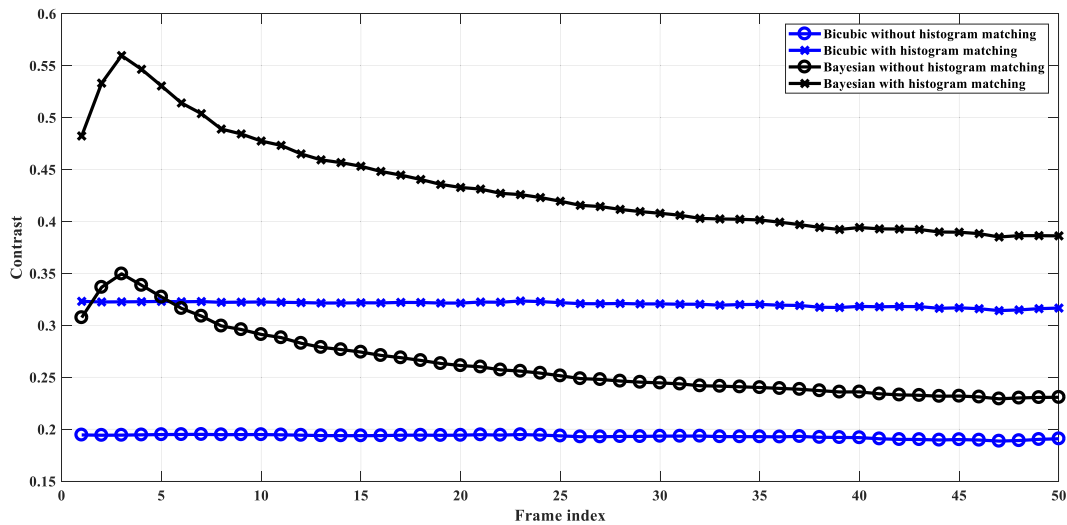

FIGURE 12. Objective comparison contrast outcomes of the suggested scheme compared to the bicubic technique on the first 50 frames for the Balloons 3DV sequence without down-sampling, and with a blur kernel with $\sigma_k = 1.2$, and AWGN effect with $\sigma_n = 0.01$.

standard 3DV (Poznan, Balloons, and Shark) 1920 × 1088 streams have been carried out [29]. Different evaluation metrics are used to assess performance on the utilized 3DV frames after the proposed processing operations and then compare the results with the original video frames. In this paper, the average gradient, local contrast, MSE, edge intensity, PSNR, SSIM index, and entropy quality metrics [16], [17] are utilized to assess the accomplishment of the suggested 3DV SR reconstruction scheme compared to the traditional schemes. Further explanations, descriptions, and mathematical representations of the utilized assessment metrics used in this section could be found in [7], [13], [16].

In our simulation experiments, to create the LR 3DV streams, the HR 3DV streams are downsampled by 4. After that, the proposed SR reconstruction scheme is employed to upsample the LR 3DV streams by 4. In the simulation experiments, we tested different consecutive and non-consecutive frames to assess the proposed 3DV SR scheme, and we only presented samples of the 3DV frames to prove the effectiveness of the suggested scheme. A size of 16 × 16 is chosen for frame blocks in the motion estimation process. The suggested SR reconstruction scheme is compared to the conventional bicubic technique and the more recent related works.
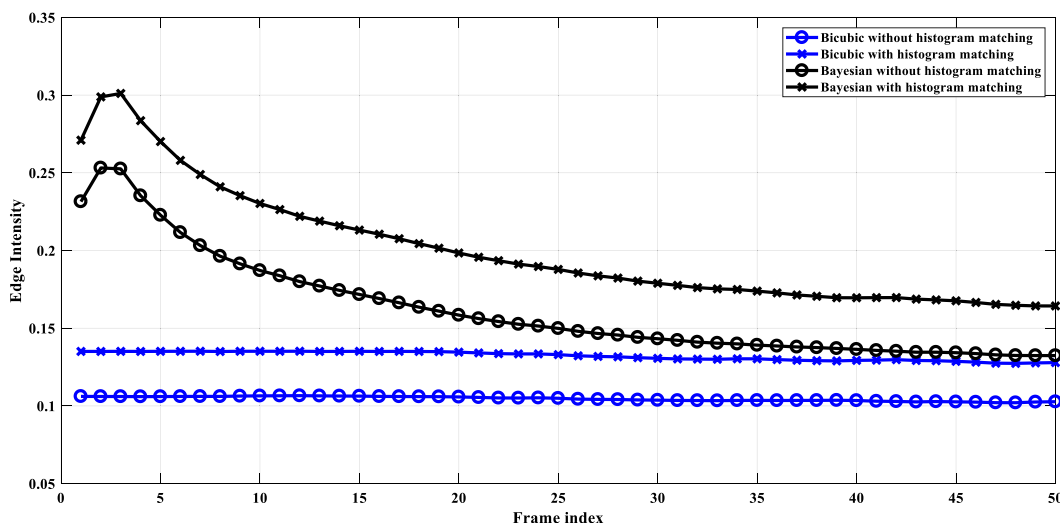
FIGURE 13. Objective comparison edge intensity outcomes of the suggested scheme compared to the bicubic technique on the first 50 frames for the Balloons 3DV sequence without down-sampling, and with a blur kernel with $\sigma_k = 1.2$, and AWGN effect with $\sigma_n = 0.01$.
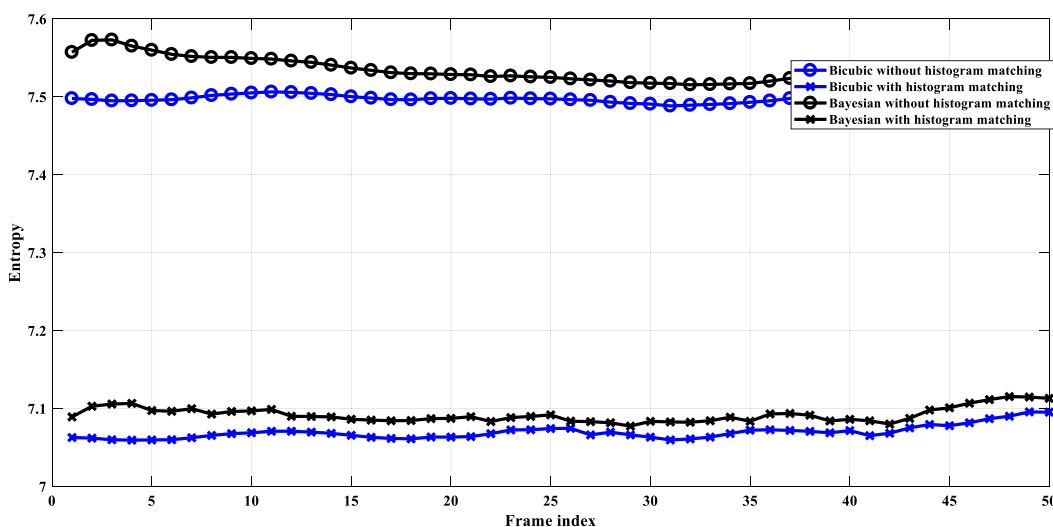

FIGURE 14. Objective comparison entropy outcomes of the suggested scheme compared to bicubic technique on the first 50 frames for the Balloons 3DV sequence without down-sampling, and with a blur kernel with $\sigma_k = 1.2$, and AWGN effect with $\sigma_n = 0.01$.

The LR 3DV frames are produced in the experimental work by employing the down-sampling process on the tested 3DV HR frames with a factor $L = 4$, blurring effect, and Gaussian noise effect. These processes are considered to simulate the acquisition process of the 3DV frames using a low-quality camera. The proposed SR reconstruction scheme is evaluated by presenting more objective and subjective outcomes. The visual results of the reconstructed 3DV frames with the proposed scheme are firstly compared to those of the bicubic technique. Then, the proposed SR reconstruction scheme is compared to different traditional SR reconstruction schemes. Moreover, more objective results of the local contrast, average gradient, edge intensity, PSNR, SSIM index, correlation,

and entropy metrics between the reconstructed frames and the original HR frames are presented to carefully assess the proposed 3DV SR scheme compared to the traditional ones. In the introduced objective outcomes for all tested 3DV frames, we assess the efficiency of the investigated SR reconstruction schemes on the first 50 frames for each tested 3DV sequence. The samples of the original 3DV frames of the tested streams in the introduced simulation results are given in Fig. 3.

Due to space limitations, only a single frame from the shown frames in Fig. 3 is considered in the subjective results for each tested 3DV sequence. Figure 4 illustrates the results of the proposed 3DV SR reconstruction scheme
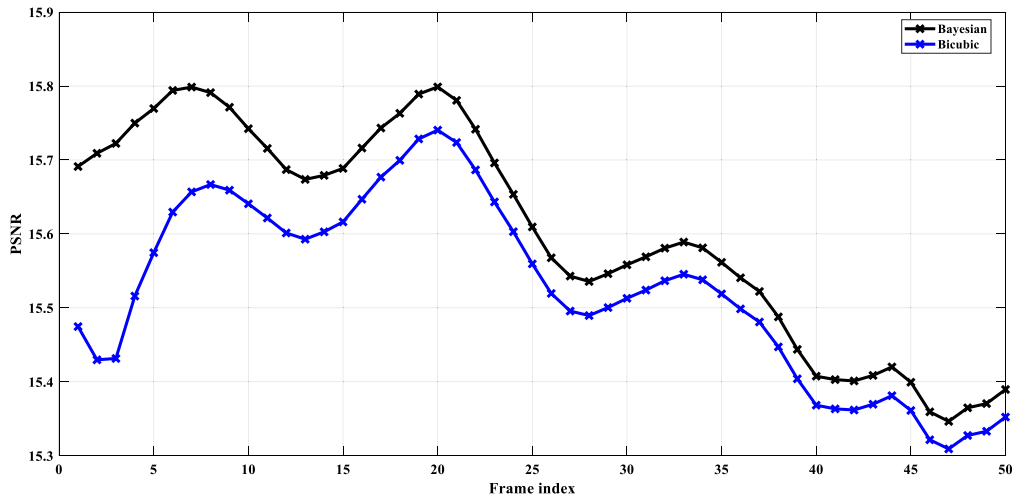
FIGURE 15. Objective comparison PSNR outcomes of the suggested scheme compared to the bicubic technique on the first 50 frames for the Balloons 3DV sequence without down-sampling, and with a blur kernel with $\sigma_k = 1.2$, and AWGN effect with $\sigma_n = 0.01$.

Frame No.	Low resolution frames down-sampled by 4 and blurred by $\sigma_k = 1.6$	Without histogram matching		With histogram matching	
		Bicubic [6, 13]	Bayesian (proposed)	Bicubic [6, 13]	Bayesian (proposed)
50					

FIGURE 16. Visual outcomes of the suggested approach compared to the bicubic technique for frame 50 of the Shark 3DV frame for down-sampling by 4, and a blur kernel with $\sigma_k = 1.6$, without AWGN effect.

FIGURE 17. Objective comparison of average intensity outcomes of the suggested scheme compared to the bicubic technique on the first 50 frames of the 3DV Shark sequence with down-sampling by 4, a blur kernel with $\sigma_k = 1.6$, and without AWGN effect.

compared to the conventional bicubic technique [6], [13] for frame 1 of the tested Poznan 3DV stream with down-sampling by 4, a standard deviation of the blur kernel $\sigma_k = 1.6$, without Additive White Gaussian Noise (AWGN) effect.

FIGURE 18. Objective comparison of contrast outcomes of the suggested scheme compared to the bicubic technique on the first 50 frames of the 3DV Shark sequence for down-sampling by 4, and a blur kernel with $\sigma_k = 1.6$, without AWGN effect.
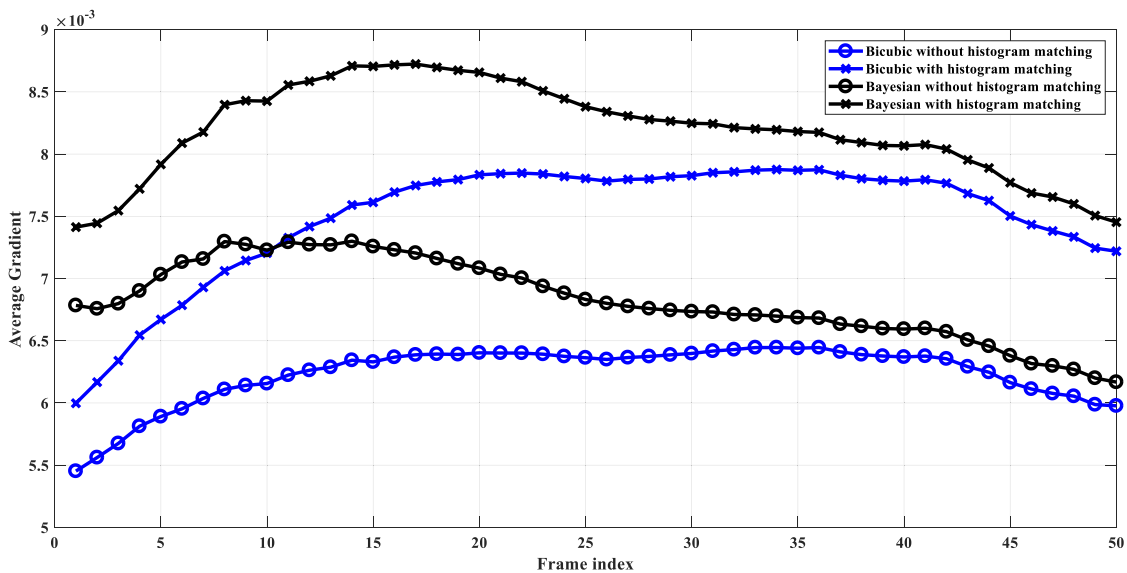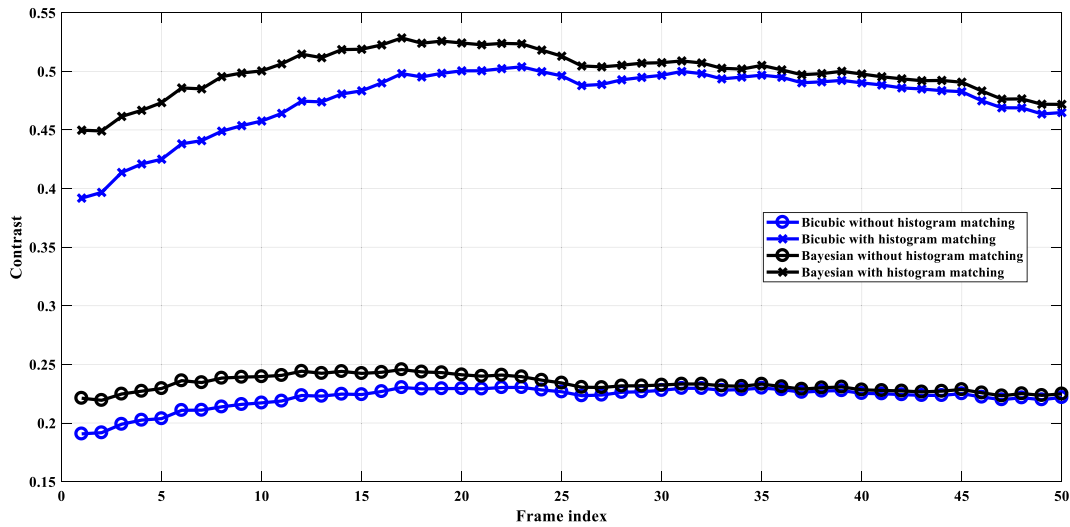

FIGURE 19. Objective comparison of edge intensity outcomes of the suggested scheme compared to the bicubic technique on the first 50 frames of the 3DV Shark sequence for down-sampling by 4, and a blur kernel with $\sigma_k = 1.6$, without AWGN effect.
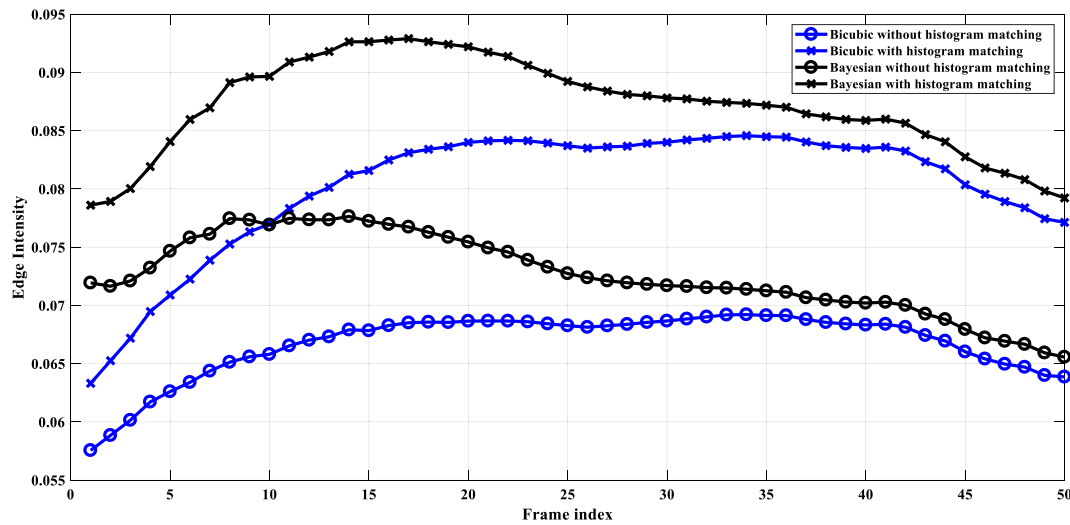
TABLE 1. Objective comparison outcomes for the sample first frame of the Shark 3DV stream without histogram matching, by employing down-sampling by 4, with blurring effect with different standard deviation values, and AWGN effect with different noise variances.

	Before histogram matching																	
AGWN	$\sigma_n = 0.01$						$\sigma_n = 0.03$						$\sigma_n = 0.05$					
Blur kernel	$\sigma_k = 1.2$		$\sigma_k = 1.6$		$\sigma_k = 2.0$		$\sigma_k = 1.2$		$\sigma_k = 1.6$		$\sigma_k = 2.0$		$\sigma_k = 1.2$		$\sigma_k = 1.6$		$\sigma_k = 2.0$	
Method	Bay.	Bi.	Bay.	Bi.	Bay.	Bi.	Bay.	Bi.	Bay.	Bi.	Bay.	Bi.	Bay.	Bi.	Bay.	Bi.	Bay.	Bi.
Entropy	7.5872	7.5525	7.5883	7.5516	7.5902	7.5515	7.6507	7.5525	7.6512	7.5516	7.6503	7.5515	7.7128	7.5525	7.7072	7.5516	7.7047	7.5515
Average gradient	0.0094	0.0061	0.0089	0.0055	0.0089	0.0050	0.0136	0.0061	0.0133	0.0055	0.0131	0.0050	0.0172	0.0061	0.0170	0.0055	0.0166	0.0050
Local contrast	0.2698	0.2118	0.2602	0.1977	0.2579	0.1865	0.3561	0.2118	0.3525	0.1977	0.3484	0.1865	0.4454	0.2118	0.4383	0.1977	0.4325	0.1865
Edge intensity	0.1018	0.0672	0.0969	0.0607	0.0959	0.0553	0.1464	0.0672	0.1436	0.0607	0.1412	0.0553	0.1861	0.0672	0.1835	0.0607	0.1794	0.0553
RMSE	0.0368	0.0338	0.0368	0.0346	0.0374	0.0358	0.0432	0.0338	0.0434	0.0346	0.0446	0.0358	0.0517	0.0338	0.0523	0.0346	0.0525	0.0358
PSNR	28.676	29.426	28.68	29.22	28.541	28.93	27.297	29.426	27.246	29.223	27.021	28.93	25.736	29.426	25.638	29.223	25.602	28.93
SSIM	0.8022	0.852	0.796	0.844	0.7823	0.8352	0.6705	0.8517	0.6619	0.8440	0.6493	0.8352	0.5615	0.8517	0.5497	0.8440	0.5471	0.8352

Figures 5 to 9 show the objective comparison results of the average gradient, contrast, edge intensity, entropy, and PSNR for the first 50 frames of the Poznan 3DV sequence by employing down-sampling by 4, a standard deviation of the blur kernel $\sigma_k = 1.6$, without AWGN effect. From the presented objective and subjective results in Figs. 4 to 9, it is clear that there is a great similarity between the reconstructed and the original 3DV frames with the suggested
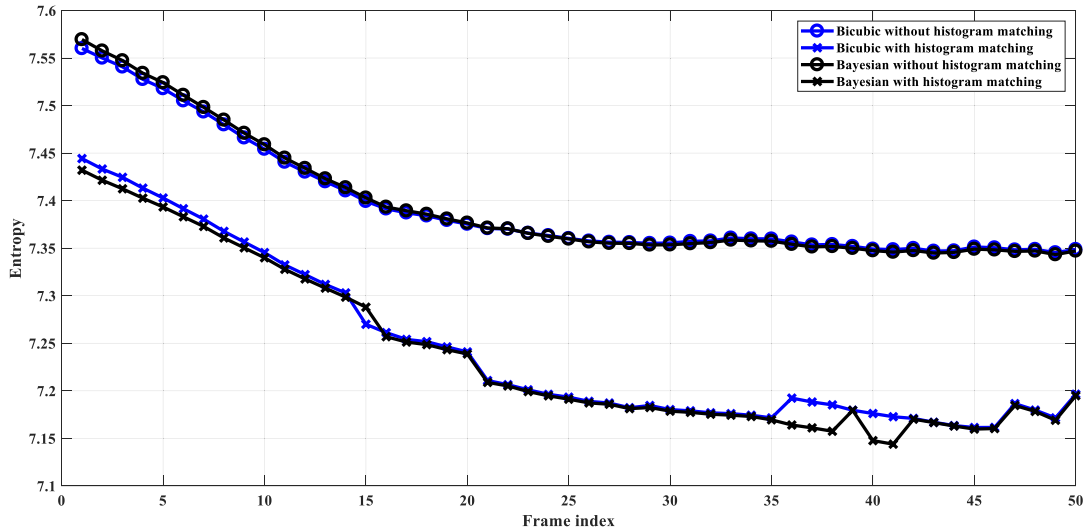
FIGURE 20. Objective comparison of entropy outcomes of the suggested scheme compared to the bicubic technique on the first 50 frames of the 3DV Shark sequence for down-sampling by 4, with a blur kernel with $\sigma_k = 1.6$, and AWGN effect.
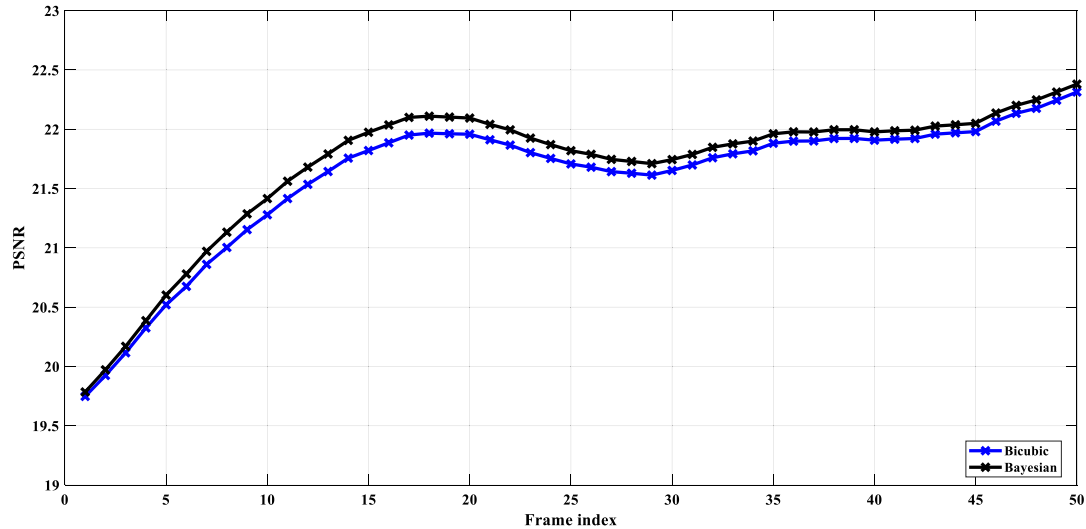

FIGURE 21. Objective comparison of PSNR outcomes of the suggested scheme compared to the bicubic technique on the first 50 frames for the 3DV Shark sequence, for down-sampling by 4, and a blur kernel with $\sigma_k = 1.6$, without AWGN effect.

TABLE 2. Objective comparison outcomes for the sample first frame of the Shark 3DV stream with histogram matching, for down-sampling by 4, with blurring effect with different standard deviation values, and AWGN effect with different noise variances.

	After histogram matching																	
AGWN	$\sigma_n = 0.01$						$\sigma_n = 0.03$						$\sigma_n = 0.05$					
Blur kernel	$\sigma_k = 1.2$		$\sigma_k = 1.6$		$\sigma_k = 2.0$		$\sigma_k = 1.2$		$\sigma_k = 1.6$		$\sigma_k = 2.0$		$\sigma_k = 1.2$		$\sigma_k = 1.6$		$\sigma_k = 2.0$	
Scheme	Bay.	Bi.	Bay.	Bi.	Bay.	Bi.	Bay.	Bi.	Bay.	Bi.	Bay.	Bi.	Bay.	Bi.	Bay.	Bi.	Bay.	Bi.
Entropy	7.5872	7.5525	7.5883	7.5516	7.5902	7.5515	7.6507	7.5525	7.6512	7.5516	7.6502	7.5515	7.7126	7.5525	7.7070	7.5516	7.7044	7.5515
Average gradient	0.0096	0.0062	0.0092	0.0056	0.0091	0.0051	0.0138	0.0062	0.0135	0.0056	0.0134	0.0051	0.0173	0.0062	0.0171	0.0056	0.0168	0.0051
Local contrast	0.2816	0.2210	0.2726	0.2073	0.2710	0.1965	0.3680	0.2210	0.3663	0.2073	0.3647	0.1965	0.4543	0.2210	0.4473	0.2073	0.4452	0.1965
Edge intensity	0.1039	0.0687	0.0991	0.0622	0.0983	0.0569	0.1484	0.0687	0.1461	0.0622	0.1441	0.0569	0.1869	0.0687	0.1848	0.0622	0.1815	0.0569
RMSE	0.0334	0.0298	0.0333	0.0307	0.0339	0.0339	0.0404	0.0298	0.0406	0.0307	0.0419	0.0320	0.0493	0.0298	0.0500	0.0307	0.0503	0.0320
PSNR	29.5378	30.5144	29.5378	30.2690	29.3878	29.9082	27.8760	30.5144	27.8200	30.2690	27.5563	29.9082	26.1470	30.5144	26.0244	30.2690	25.9679	29.9082
SSIM	0.8166	0.8689	0.8103	0.8610	0.7957	0.8516	0.6809	0.8689	0.6711	0.8610	0.6571	0.8516	0.5718	0.8689	0.5586	0.8610	0.5541	0.8516

scheme compared to the conventional bicubic technique. So, the suggested 3DV SR reconstruction scheme introduces better-reconstructed frames than those of the bicubic technique. Furthermore, the suggested scheme achieves superior objective results of average gradient, contrast, edge intensity, entropy, and PSNR values for the tested 3DV Poznan frames compared to the bicubic technique. Also, the significant effect of using the histogram matching is observed in improving the quality of the reconstructed 3DV frames.

TABLE 3. Results of average PSNR (dB), SSIM, and CPU time (sec) assessment outcomes of the suggested 3DV SR scheme and the traditional schemes on the 3DV Shark frames.

SR Scheme	SSIM	PSNR	CPU time (sec)
[16]	0.8684	31.24	2.475
[23]	0.8499	32.61	2.894
[30]	0.8992	31.62	4.647
[36]	0.9103	32.47	5.206
[37]	0.8961	31.90	4.973
Proposed	0.9075	32.52	2.867

Figure 10 demonstrates the outcomes of the suggested 3DV SR reconstruction scheme compared to the conventional bicubic technique [6], [13] for frame 25 of the tested Balloons 3DV stream without employing down-sampling, with a standard deviation of blur kernel $\sigma_k = 1.6$, and AWGN effect ($\sigma_n = 0.01$). Figures 11 to 15 show the objective comparison results of the average gradient, contrast, edge intensity, entropy, and PSNR for the first 50 frames of the Balloons 3DV sequence without employing down-sampling, with a standard deviation of blur kernel $\sigma_k = 1.6$, and AWGN effect ($\sigma_n = 0.01$).

It is observed from the quantitative and visual results shown in figures 10 to 15 that the proposed 3DV SR reconstruction scheme presents improved reconstructed frames compared to those of the bicubic technique. Also, it is noticed that there is a significant effect of the histogram matching process employed in the proposed scheme for obtaining the reconstructed 3DV frames with recommended and superior quality. Moreover, it is detected that there is a high level of similarity between the reconstructed and original 3DV frames with the suggested 3DV SR scheme compared to the conventional Bicubic technique. Additionally, it is observed that the proposed scheme accomplishes greater and superior values of entropy, average gradient, edge intensity, contrast, and PSNR values for the utilized 3DV Balloons frames compared to those of the bicubic technique, which confirms the pre-eminence of the suggested 3DV SR reconstruction scheme. Figure 16 clarifies the outcomes of the suggested hybrid 3DV SR reconstruction scheme compared to the conventional bicubic technique [6], [13] for the frame 50 of the tested Shark 3DV stream with down-sampling by 4, standard deviation of blur kernel $\sigma_k = 1.6$, without AWGN effect. Figures 17 to 21 show the objective comparison results of the average gradient, contrast, edge intensity, entropy, and PSNR for the first 50 frames of the Shark 3DV sequence with down-sampling by 4, standard deviation of blur kernel $\sigma_k = 1.6$, without AWGN effect.

From the presented objective and subjective results in Figs. 16 to 21, it is clear that there is a great similarity between the reconstructed and original 3DV frames with the suggested scheme compared to the conventional bicubic technique.

So, the suggested 3DV SR reconstruction scheme introduces better-reconstructed frames than those of the bicubic technique. Furthermore, the suggested scheme achieves superior objective results of average gradient, contrast, edge intensity, entropy, and PSNR values for the tested 3DV Shark frames compared to those of the bicubic technique. Also, it is observed that there is a great effect of the histogram matching for improving the quality of the reconstructed 3DV frames.

Tables 1 and 2 present the objective comparison results of the local contrast, entropy, edge intensity, MSE, average gradient, PSNR, and SSIM for the first frame of the Shark 3DV sequence with and without histogram matching, for down-sampling by 4, a blurring effect with different standard deviation values, and AWGN effect with various noise variances. It is observed that the suggested 3DV SR reconstruction scheme with histogram matching achieves superior outcomes in terms of edge intensity, average gradient, MSE, PSNR, local contrast, SSIM index, and entropy metrics for the resulting 3DV frames compared to the case of not employing histogram matching.

The whole introduced subjective and objective outcomes demonstrate that the suggested 3DV SR reconstruction scheme is more appreciated and recommended for 3DV communication applications. It has more superior visual and objective results than those of the traditional schemes that do not use the Bayesian and histogram matching algorithms, especially in mitigating the effect of blurring, noise, and down-sampling effects. Furthermore, it is noticed that the suggested 3DV SR reconstruction scheme gives preferable outcomes for various well-known 3DV streams that have different spatial and temporal features.

Furthermore, to highlight the efficacy of the suggested hybrid SR reconstruction scheme for reliable reconstruction of 3DV frames, different comparisons are implemented to assess its outcomes to be compared with those of the previous schemes [16], [23], [30], [36], [37]. The comparisons have been implemented on the 3DV Shark stream in the presence of down-sampling, blurring, and AWGN effects.

Table 3 reveals the average CPU time, SSIM, and PSNR values of the suggested 3DV SR scheme compared to those of the previous SR schemes in [16], [23], [30], [36], [37]. We notice that the suggested 3DV SR reconstruction scheme surpasses the recent previous schemes by presenting better average values of SSIM and PSNR. It also offers lower computational times compared to those of the other related studies. Consequently, it is seen that by deploying the histogram matching besides the Bayesian model, the 3DV SR reconstruction quality can be improved with good SSIM and PSNR values compared to those of the conventional schemes. In addition, there is a great benefit provided by the proposed SR reconstruction scheme in terms of computational complexity.

V. CONCLUSION AND FUTURE RESEARCH DIRECTIONS

An efficient SR reconstruction scheme for 3DV communication has been suggested in this paper. The proposed scheme is based on recursive probabilistic Bayesian and histogram matching algorithms for recovering HR 3DV frames with good features. It is observed that the proposed hybrid 3DV SR scheme has higher robustness in the cases of unknown noise levels, unknown blur kernels, and arbitrary motion compared to the contemporary schemes available in the literature. Furthermore, the introduced simulation results on different standard 3DV sequences with different spatial and temporal features using different quality metrics demonstrated that the proposed hybrid 3DV SR reconstruction scheme outperforms many existing SR reconstruction schemes. Furthermore, the histogram matching algorithm proved its efficiency in the proposed scheme to generate 3DV sequences with better distributions of intensities. As a result, the suggested scheme has demonstrated favorable outcomes, and it is feasible for reliable 3DV SR reconstruction applications. In the future, deep-learning-based enhancement steps can be incorporated for improving the resolutions of 3DV frames. In addition, advanced segmentation techniques can be utilized to improve the region of interest in the degraded 3DV frames.

REFERENCES

[1] Y. Luo, L. Zhou, S. Wang, and Z. Wang, "Video satellite imagery super resolution via convolutional neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 12, pp. 2398–2402, Dec. 2017.

[2] E. Faramarzi, D. Rajan, F. C. A. Fernandes, and M. P. Christensen, "Blind super resolution of real-life video sequences," *IEEE Trans. Image Process.*, vol. 25, no. 4, pp. 1544–1555, Apr. 2016.

[3] R. Timofte, R. Rothe, and L. Van Gool, "Seven ways to improve example-based single image super resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 1865–1873.

[4] J. Wang, S. Zhu, and Y. Gong, "Resolution enhancement based on learning the sparse association of image patches," *Pattern Recognit. Lett.*, vol. 31, no. 1, pp. 1–10, Jan. 2010.

[5] H. Liu and D. Huang, "A short survey of image super resolution algorithms," *J. Comput. Sci. Technol. Updates*, vol. 2, no. 2, pp. 19–29, Dec. 2015.

[6] E. Mandanici, L. Tavasci, F. Corsini, and S. Gandolfi, "A multi-image super-resolution algorithm applied to thermal imagery," *Appl. Geomatics*, vol. 11, no. 3, pp. 215–228, Sep. 2019.

[7] J. Tian and K.-K. Ma, "A survey on super-resolution imaging," *Signal, Image Video Process.*, vol. 5, no. 3, pp. 329–342, 2011.

[8] J. Sun, J. Sun, Z. Xu, and H.-Y. Shum, "Gradient profile prior and its applications in image super-resolution and enhancement," *IEEE Trans. Image Process.*, vol. 20, no. 6, pp. 1529–1542, Jun. 2011.

[9] S. E. El-Khamy, M. M. Hadhoud, M. I. Dessouky, B. M. Salam, and F. E. A. El-Samie, "Wavelet fusion: A tool to break the limits on LMMSE image super-resolution," *Int. J. Wavelets, Multiresolution Inf. Process.*, vol. 4, no. 1, pp. 105–118, Mar. 2006.

[10] M. D. Robinson, C. A. Toth, J. Y. Lo, and S. Farsiu, "Efficient Fourier-wavelet super-resolution," *IEEE Trans. Image Process.*, vol. 19, no. 10, pp. 2669–2681, Oct. 2010.

[11] H. Huang, R. He, Z. Sun, and T. Tan, "Wavelet-SRNet: A wavelet-based CNN for multi-scale face super resolution," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 1689–1697.

[12] X. Li, Y. Hu, X. Gao, D. Tao, and B. Ning, "A multi-frame image super-resolution method," *Signal Process.*, vol. 90, no. 2, pp. 405–414, Feb. 2010.

[13] F. El-Samie, M. Hadhoud, and S. El-Khamy, *Image Super-Resolution and Applications*. Boca Raton, FL, USA: CRC Press, 2012.

[14] Z. Chen and S. Haykin, "On different facets of regularization theory," *Neural Comput.*, vol. 14, no. 12, pp. 2791–2846, Dec. 2002.

[15] S. Cheol Park, M. Kyu Park, and M. Gi Kang, "Super-resolution image reconstruction: A technical overview," *IEEE Signal Process. Mag.*, vol. 20, no. 3, pp. 21–36, May 2003.

[16] H. Ji and C. Fermuller, "Robust wavelet-based super-resolution reconstruction: Theory and algorithm," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 4, pp. 649–660, Apr. 2009.

[17] T. Hilker, M. A. Wulder, N. C. Coops, J. Linke, G. McDermid, J. G. Masek, F. Gao, and J. C. White, "A new data fusion model for high spatial- and temporal-resolution mapping of forest disturbance based on Landsat and MODIS," *Remote Sens. Environ.*, vol. 113, no. 8, pp. 1613–1627, Aug. 2009.

[18] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4681–4690.

[19] N. Asuni and A. Giachetti, "Accuracy improvements and artifacts removal in edge based image interpolation," *VISAPP*, vol. 1, no. 8, pp. 58–65, 2008.

[20] J. Jiang, R. Hu, Z. Wang, Z. Han, and S. Dong, "Manifold regularized sparse support regression for single image super-resolution," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2013, pp. 1429–1433.

[21] J. Jiang, C. Chen, J. Ma, Z. Wang, Z. Wang, and R. Hu, "SRLSP: A face image super-resolution algorithm using smooth regression with local structure prior," *IEEE Trans. Multimedia*, vol. 19, no. 1, pp. 27–40, Jan. 2017.

[22] A. Katartzis and M. Petrou, "Robust Bayesian estimation and normalized convolution for super-resolution image reconstruction," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1–7.

[23] S. Zhao, R. Jin, X. Xu, E. Song, and C.-C. Hung, "A variational Bayesian superresolution approach using adaptive image prior model," *Math. Problems Eng.*, vol. 2015, pp. 1–13, Jan. 2015.

[24] Q. Ge, W. Shao, and L. Wang, "Joint Bayesian convolutional sparse coding for image super-resolution," *PLoS ONE*, vol. 13, no. 9, Sep. 2018, Art. no. e0201463.

[25] Z. Jin, T. Tillo, C. Yao, J. Xiao, and Y. Zhao, "Virtual-View-Assisted video super-resolution and enhancement," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 3, pp. 467–478, Mar. 2016.

[26] C. Bui-Thu, T. Do-Hong, T. Le-Tien, and H. Nguyen-Duc, "An efficient approach based on Bayesian MAP for video super-resolution," in *Proc. IEEE Int. Conf. Adv. Technol. Commun. (ATC)*, Oct. 2014, pp. 522–527.

[27] J. Chen, J. L. Nunez-Yanez, and A. Achim, "Bayesian video super-resolution with heavy-tailed prior models," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 6, pp. 905–914, Jun. 2014.

[28] C. Liu and D. Sun, "On Bayesian adaptive video super resolution," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 2, pp. 346–360, Feb. 2014.

[29] Y. Huang, W. Wang, and L. Wang, "Bidirectional recurrent convolutional networks for multi-frame super-resolution," in *Proc. Adv. Neural Inf. Process. Syst.*, 2015, pp. 235–243.

[30] E. Quevedo, J. de la Cruz, G. M. Callico, F. Tobajas, and R. Sarmiento, "Video enhancement using spatial and temporal super-resolution from a multi-camera system," *IEEE Trans. Consum. Electron.*, vol. 60, no. 3, pp. 420–428, Aug. 2014.

[31] R. A. Borsoi, G. H. Costa, and J. C. M. Bermudez, "A new adaptive video super-resolution algorithm with improved robustness to innovations," *IEEE Trans. Image Process.*, vol. 28, no. 2, pp. 673–686, Feb. 2019.

[32] W. Yu and M. Zhang, "A mixed particle swarm optimization algorithm's application in image/video super-resolution reconstruction," in *Proc. 2nd IEEE Int. Conf. Image, Vis. Comput. (ICIVC)*, Jun. 2017, pp. 526–530.

[33] K. Gopan and G. S. Kumar, "Video super resolution with generative adversarial network," in *Proc. 2nd Int. Conf. Trends Electron. Informat. (ICOEI)*, May 2018, pp. 1489–1493.

[34] D. Liu, Z. Wang, Y. Fan, X. Liu, Z. Wang, S. Chang, X. Wang, and T. S. Huang, "Learning temporal dynamics for video super-resolution: A deep learning approach," *IEEE Trans. Image Process.*, vol. 27, no. 7, pp. 3432–3445, Jul. 2018.

[35] Y. Wang, J. Zhang, Z. Liu, Q. Wu, Z. Zhang, and Y. Jia, "Depth super-resolution on RGB-D video sequences with large displacement 3D motion," *IEEE Trans. Image Process.*, vol. 27, no. 7, pp. 3571–3585, Jul. 2018.

[36] D. Li, Y. Liu, and Z. Wang, "Video super-resolution using non-simultaneous fully recurrent convolutional network," *IEEE Trans. Image Process.*, vol. 28, no. 3, pp. 1342–1355, Mar. 2019.

[37] Y. Xie, J. Xiao, T. Tillo, Y. Wei, and Y. Zhao, "3D video super-resolution using fully convolutional neural networks," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jul. 2016, pp. 1–6.

GHALIB H. ALSHAMMRI received the Master of Computer Science degree with a focus on intelligent systems from the University of Wollongong, Australia, in 2009, and the Ph.D. degree with particular interests in artificial intelligence and nano molecular communication from the Electrical and Computer Engineering Department, Stevens Institute of Technology, Hoboken, NJ, USA, in May 2019. His dissertation is titled "Artificial Intelligence Techniques for Diffusion-Based Bio-Molecular Nano Communication Networks." He has taught many computer science and engineering courses at King Saud University (since 2003) and the Steven Institute of Technology (since Spring 2017) for graduate and undergraduate students. He is currently an Assistant Professor in artificial intelligence and data science at the Community College, King Saud University, and a Scientific Research Consultant at Saudi Electronic University. He is also the Co-Founder and Board Member at the Saudi Society for Data Science (SSDS). He has published several scientific papers and a provisional patent. His research interests include artificial intelligence, robotics, machine learning, fuzzy techniques, and the Internet of Bio-Things.

AMANI K. SAMHA received the Master of Information Technology degree from Australian National University, Canberra, Australia, in July 2011, and the Ph.D. degree with particular interests in data mining from the Information Technology Department, Queensland University of Technology, Australia, in July 2016. Her dissertation is titled "Aspect-Based Opinion Mining from Customer Reviews." She has taught several information technology, computer science, and information system courses at King Saud University (since 2016) for graduate and undergraduate students. She was the Vice Dean for quality and development, in 2018. She is currently an Assistant Professor in artificial intelligence and data science at the College of Business Administration, Management Information Systems and Digital Transformation, King Saud University. She is also an Information Security Consultant at the Ministry of Education. She is also a Consultant at the National Committee for Labour Committees in Saudi Arabia (Saudi Arabia L20). Her research interests include artificial intelligence, machine learning, fuzzy techniques, and the Internet of Things.

WALID EL-SHAFAI was born in Alexandria, Egypt. He received the B.Sc. degree (Hons.) in electronics and electrical communications engineering from the Faculty of Electronic Engineering (FEE), Menoufia University, Menouf, Egypt, in 2008, the M.Sc. degree from the Egypt-Japan University of Science and Technology (E-JUST), in 2012, and the Ph.D. degree from the Faculty of Electronic Engineering, Menoufia University, in 2019. He is currently working as a Lecturer and an Assistant Professor with ECE Department, FEE, Menoufia University. Since January 2021, he has been with the Security Engineering Laboratory (SEL), Prince Sultan University (PSU), Riyadh, Saudi Arabia, as a Postdoctoral Research Fellow. His research interests include wireless mobile and multimedia communications systems, image and video processing, efficient 2-D video/3-D multi-view video coding, multi-view video plus depth coding, 3-D multi-view video coding and transmission, quality of service and experience, digital communication techniques, cognitive radio networks, adaptive filter design, 3-D video watermarking, steganography, encryption, error resilience and concealment algorithms for H.264/AVC, H.264/MVC, and H.265/HEVC video codecs standards, cognitive cryptography, medical image processing, speech processing, security algorithms, software defined networks, the Internet of Things, medical diagnosis applications, FPGA implementations for signal processing algorithms and communication systems, cancellable biometrics, pattern recognition, image and video magnification, artificial intelligence for signal processing algorithms and communication systems, modulation identification and classification, image and video super-resolution, denoising, cybersecurity applications, and deep learning in signal processing and communication system applications. He serves as a reviewer in several international journals.

EMAD A. ELSHEIKH received the B.Sc. (Hons.), M.Sc., and Ph.D. degrees from the Faculty of Electronic Engineering, Menoufia University, Menouf, Egypt, in 2003, 2009, and 2016, respectively. He has been the Teaching Staff with the Department of Industrial Electronics and Control, Faculty of Electronic Engineering, Menoufia University, since 2016. His current research interests include intelligent control, visual servoing, robotics, real-time object tracking, and super-resolution reconstruction of images.

EBRAHIM ABDEL HAMID received the B.Sc., M.Sc., and Ph.D. degrees in electronic engineering and automatic control engineering, in 1999, 2007, and 2015, respectively. He is currently a Lecturer with the Department of Industrial Electronic and Control Engineering, Faculty of Electronic Engineering, Menoufia University. His research interests include intelligent control systems, computer controlled systems, embedded system design for control systems based on FPGA, and super-resolution reconstruction of images.

MOHAMED I. ABDO received the B.Sc., M.Sc., and Ph.D. degrees in electronic/automatic control engineering from Menoufia University, Egypt, in 2000, 2009, and 2017, respectively. Since 2017, he has been a Lecturer with the Department of Industrial Electronics and Control Engineering, Menoufia University. His research interests include intelligent control systems, fault-tolerant control systems, robust control, intelligent instrumentation, fuzzy control, system identification, and image processing.

MOHAMMED AMOON received the B.Sc. degree in electronic engineering and the M.Sc. and Ph.D. degrees in computer science and engineering from Menoufia University, in 1996, 2001, and 2006, respectively. He is currently a Professor of computer science and engineering at the Department of Computer Science and Engineering, Menoufia University. He is also a Professor of computer science with the Department of Computer Science, King Saud University. His research interests include agent-based systems, fault tolerance techniques, scheduling algorithms, green computing, distributed computing, grid computing, cloud computing, fog computing, and the Internet of Things (IoT).

FATHI E. ABD EL-SAMIE received the B.Sc. (Hons.), M.Sc., and Ph.D. degrees from the Faculty of Electronic Engineering, Menoufia University, Menouf, Egypt, in 1998, 2001, and 2005, respectively. Since 2005, he has been a Teaching Staff Member with the Department of Electronics and Electrical Communications, Faculty of Electronic Engineering, Menoufia University. His current research interests include image enhancement, image restoration, image interpolation, super-resolution reconstruction of images, data hiding, multimedia communications, medical image processing, optical signal processing, and digital communications. He was a recipient of the Most Cited Paper Award from the Digital Signal Processing journal, in 2008.

• • •