

# Global motion based video super-resolution reconstruction using discrete wavelet transform

Wasnaa Witwit<sup>1</sup> · Yifan Zhao<sup>2</sup>  · Karl Jenkins<sup>3</sup> · Sri Addepalli<sup>2</sup>

Received: 4 July 2017 / Revised: 19 March 2018 / Accepted: 26 March 2018 /  
Published online: 11 April 2018  
© The Author(s) 2018

**Abstract** Different from the existing super-resolution (SR) reconstruction approaches working under either the frequency-domain or the spatial-domain, this paper proposes an improved video SR approach based on both frequency and spatial-domains to improve the spatial resolution and recover the noiseless high-frequency components of the observed noisy low-resolution video sequences with global motion. An iterative planar motion estimation algorithm followed by a structure-adaptive normalised convolution reconstruction method are applied to produce the estimated low-frequency sub-band. The discrete wavelet transform process is employed to decompose the input low-resolution reference frame into four sub-bands, and then the new edge-directed interpolation method is used to interpolate each of the high-frequency sub-bands. The novelty of this algorithm is the introduction and integration of a nonlinear soft thresholding process to filter the estimated high-frequency sub-bands in order to better preserve the edges and remove potential noise. Another novelty of this algorithm is to provide flexibility with various motion levels, noise levels, wavelet functions, and the number of used low-resolution frames. The performance of the proposed method has been tested on three well-known videos. Both visual and quantitative results demonstrate the high performance and improved flexibility of the proposed technique over the conventional interpolation and the state-of-the-art video SR techniques in the wavelet-domain.

**Keywords** Image enhancement · Image registration · Image reconstruction

---

✉ Yifan Zhao  
yifan.zhao@cranfield.ac.uk

<sup>1</sup> Through-Life Engineering Services Institute, School of Aerospace, Transport and Manufacturing, Cranfield University, Cranfield, Bedfordshire MK43 0AL, UK

<sup>2</sup> Through-Life Engineering Services Centre, School of Aerospace, Transport and Manufacturing, Cranfield University, Cranfield, Bedfordshire MK43 0AL, UK

<sup>3</sup> Centre for Computational Engineering Sciences, School of Aerospace, Transport and Manufacturing, Cranfield University, Cranfield, Bedfordshire MK43 0AL, UK

## 1 Introduction

High resolution (HR) images and videos are highly desirable, and strongly in demand for most electronic imaging applications not only for providing better visualisation but also for extracting additional information. However, HR images are not always available since the setup of high-resolution imaging can be expensive especially with the inherent physical limitations of the sensors, the optics manufacturing technology, the data storage and the sensor's communication bandwidth. Therefore, it is essential to find an effective way in image processing to increase the resolution level at a low-cost, without replacing the existing imaging system. To address this challenge, the concept of super-resolution (SR) has now been sought after. This technique aims to produce a single HR image, or HR video, from a set of different successive low-resolution (LR) images captured from the same scene in order to overcome the limitations and/or possibly ill-posed conditions of the imaging system [53]. Due to its wide applications, SR has been an active area of research over the last two decades for a variety of applications, such as satellite imaging [10, 21], medical imaging [19, 46], forensic imaging [29, 47] and video surveillance systems [20, 67].

Most SR methods consist of two main parts: I: image registration and II: image reconstruction. Image registration aims to estimate the motion between the LR images, while image reconstruction aims at combining the registered images to reconstruct the HR image. In image registration, the motion between the reference image and its neighbouring LR images is required to be estimated accurately to reconstruct the super-resolved image [45, 57]. When the camera is moving and the scene is stationary, global motion occurs. On the contrary, when the camera is fixed and the scene is moving, non-global (local) motion occurs. This paper primarily focuses on the first scenario.

### 1.1 Image registration

Image registration methods can be operated either in the spatial-domain or the frequency-domain. Frequency-domain methods are usually limited to global motion models, whereas spatial-domain methods usually allow more general motion models. In the frequency-domain, Vandewalle et al. [57] presented an image registration algorithm to accurately register a series of aliased images based on their low-frequencies, thereby aliasing its free-part. They used a planar motion model to estimate the shift and rotation between the images, particularly for the scenario when a set of images are captured in a short period of time with a small camera motion. Vandewalle's method performs better than the other frequency-domain registration methods, such as Marcel et al. [39] and Luchese and Cortelazzo [36]. The advantage of Vandewalle's method is that it is based on discarding the high-frequency components, where aliasing may have occurred, in order to be more robust. In the spatial domain, Keren et al. [30] developed an iterative planar motion estimation algorithm that uses different, down-sampled versions of the images to estimate the shift and rotation parameters based on the Taylor series expansions. The goal of this pyramidal scheme is to increase the accuracy for estimating large motion parameters. Keren's method and Vandewalle's method, have been well accepted to tackle global motions [57]. However, the existing sub-pixel registration methods become inaccurate when the motion is non-global. There are several recent approaches dealing with the general motion estimation in video SR. For example, Liu and Sun [35] used optical flow techniques to register multiple images with sub-pixel accuracy whereas Liao et al. [34] used an ensemble of optical flow models to reconstruct the original HR frames with rich high-frequency details.

## 1.2 Image reconstruction in spatial and frequency domain

Image reconstruction methods can also be classified into frequency domain-based and spatial domain-based approaches. The first frequency-domain-based SR approach was proposed by Tsai and Huang [55], which formulates the system equations that relate the HR image to the observed LR images by estimating the relative shifts between a sequence of down-sampled, aliased and noise-free LR images. This method was extended by Kim et al. [31] by proposing a weighted least squares solution based on the assumption that the blur and the noise characteristics are the same for all LR images. A major advantage of the frequency-domain-based SR methods is that they are usually theoretically simple and computationally inexpensive. However, these methods are insufficient to handle the real-world applications, as they are limited to only global translational motion and linear space-invariant blur during image acquisition process.

For the spatial-domain-based SR approaches, Non-uniform interpolation method [40, 56] is one of the most intuitive approaches with relatively low computational complexity. However, degradation models are applicable only if all LR images have the same blur and noise characteristics. Iterative back-projection (IBP) method [22, 43] can accommodate both global translational and rotational motions. However, the solution might not be unique due to the ill-posed nature of the SR problem and the selection of some parameters is usually difficult. Projection onto convex sets (POCS) method [17, 42] benefits from the utilisation of the efficient observation model and a proper priori information. The disadvantages, on the other hand, are the lack of a unique solution, a slow convergence rate and an expensive computational cost. Regularised-based SR methods include Maximum likelihood (ML) method [54] and Maximum a posteriori (MAP) method [48, 49]. The ML method only considers the relationship among the observed LR images and the original HR image without priori information while the MAP method considers both. An extension of this approach, called Hybrid ML/MAP-POCS method [16], were proposed to guarantee the single optimal solution. The spatial-domain-based SR methods can tackle the real-world applications better because they can accommodate both global and non-global motion models, linear space-variant blur and the noise during image acquisition process.

## 1.3 Wavelet-based image reconstruction

In addition to the frequency and spatial-based domains efforts have been made using the wavelet-domain. The wavelet-domain-based SR reconstruction approach is able to exploit both the spatial and frequency-domains, and integrate properties of both to reconstruct a HR image from observed LR images. The wavelet transform (WT) is an effective tool that divides an image into low and high-frequency sub-bands, each of which is examined independently with a resolution matched to its scale [9]. The mechanism behind the strategy of WT is that the features of the image at different scales can be separated, analysed and manipulated. Global features can be examined at coarse scales, while local features can be analysed at fine scales [40]. The attractive properties of WT, such as locality, multi-resolution, and compression make it effective for analysing real-world signals [7]. Discrete wavelet transform (DWT) is one of the recent wavelet transforms: it being employed as a powerful tool for many image and video processing applications to isolate and preserve the high-frequency components of the image. DWT decomposes the given image into one low-frequency sub-band and three high-frequency sub-bands using the property of dilations and translations by a single wavelet function called mother wavelet [18].

One of the challenges in SR is to preserve or recover the true edges of objects meanwhile compressing noise, which is usually difficult to be achieved simultaneously using frequency-based methods due to similar response of edges and noise in frequency band. WT offers an alternative solution to analyse true edges and noise separately. Manipulating wavelet coefficients in sub-bands containing high-pass frequency spatial information is the essential target of wavelet-based methods to solve the SR reconstruction problem. A common assumption of WT-based methods is that the LR image is the low-pass filtered sub-band produced by WT of the HR image [52]. The existing literature on WT-based methods is in both the single frame case and multi-frame (video) case. For the multi-frame case, Izadpanahi et al. [25] presented a SR technique using DWT and bicubic interpolation. They applied an illumination enhancement method based on singular value decomposition before the registration process of the LR frames to reduce the illumination inconsistencies between the frames. Anbarjafari et al. [3] proposed a SR technique for the LR video sequences using DWT and stationary wavelet transform (SWT). However, these available methods have limited performance for variety of noise levels, motion levels, wavelet functions, and the number of used frames.

#### 1.4 Other types of SR approaches

Recently, learning-based SR methods have emerged to further boost the efficiency of SR. These methods consist of two main parts: learning and recovering. In the learning part, a dictionary which contains a large number of LR and HR patch pairs is constructed. In the recovering part, the LR frame is divided into overlapped patches, and each patch searches its more similar LR patch from the dictionary. The HR frame is obtained by incorporating the corresponding HR patch to the LR frame. Takeda et al. [51] introduced a method based on the extension of steering kernel regression framework to 3-D signals for performing video de-noising, spatiotemporal upscaling and SR, without the need for explicit sub-pixel accuracy motion estimation. To generate better results, multi-dimensional kernel regression was applied. Yang et al. [65] proposed a sparse-coding method where LR and HR patch pairs in the dictionary share the same sparse representation. The sparse representation of a LR patch can be incorporated to the HR dictionary to obtain HR patch. Li et al. [33] introduced an adaptive subpixel-guided auto-regressive (AR) model in which key-frames are up-sampled by a sparse regression while non-key-frames are super-resolved by simultaneously exploiting the spatiotemporal correlations. Deep learning-based SR approaches and deep learning networks SR approaches [6, 8, 13, 14, 27, 28, 37, 38, 50, 59, 61–64] have been developed in recent years to improve SR results, and to better model complex image contents and details. For example, Dong et al. [14] proposed a SR convolution neural network (SRCNN) method to perform a sparse reconstruction. Jiang et al. [28] addressed the problem of learning the mapping functions (i.e. projection matrices) by introducing the non-local self-similarity and local geometry priors of the training data for fast SR. However, this type of methods is usually computationally costly and requires a large amount of training data.

#### 1.5 Focus of this study

This paper proposes a robust video super-resolution approach, based on a combination of the so-called discrete wavelet transform-new edge-directed interpolation and a soft-thresholding for increasing the spatial resolution and recovering the noiseless high-frequency details of the observed noisy LR video frames with global motion, which

integrates merits from the methods of image registration and reconstruction in both frequency-domain and spatial-domain. The application of the proposed SR technique is particularly useful when the camera is moving and the observed scene is stationary. One of the motivations of this technique is to provide flexibility for a variety of motion levels, noise levels, wavelet functions, and sufficient number of used LR frames since the existing wavelet-based SR methods have limited performance capabilities for these various factors and this potential has not yet been fully explored. The performance of this approach is tested on three well-known videos. The robustness of the proposed algorithm is then evaluated through an empirical test with various motion levels, noise levels, wavelet functions, and the number of used frames. Most of the existing wavelet-based SR methods have limited discussion on the above-mentioned factors.

## 2 Methods

### 2.1 Observation model

The observation model describes the relationship between the original referenced HR image and the observed LR images. The image acquisition process in the spatial domain involves warping, blurring, down-sampling, and noising to produce the LR images from the HR image, as shown in Fig. 1. The blurring step generates blurred images from the warped HR image based on the point spread function introduced by the camera. The down-sampling step produces down-sampled (aliased) LR images from the warped and blurred HR image, and the noise step represents the additive noise applied to each observed LR image. Let us assume that the HR image can be represented in the vector form as  $x = [x_1, x_2, \dots, x_{L_1N_1 \times L_2N_2}]^T$ , where  $L_1N_1 \times L_2N_2$  is the size of the HR image. Assume that  $L_1$  and  $L_2$  represent the down-sampling factors in the horizontal and vertical directions, respectively, and each LR image has the size of  $N_1 \times N_2$ . Let the LR image be denoted in the vector form by  $y_k = [y_{k1}, y_{k2}, \dots, y_{k(N_1 \times N_2)}]^T$ ,  $k = 1, 2, \dots, p$ , where  $p$  is the number of LR images. If each LR image is corrupted by additive Gaussian noise, the observation model can be represented as

$$y_k = DB_kM_kx + n_k \tag{1}$$

where  $M_k$  is the warp matrix of size  $L_1N_1L_2N_2 \times L_1N_1L_2N_2$ ,  $B_k$  is the camera blur matrix also with the same size,  $D$  is the down-sampling matrix of size  $N_1N_2 \times L_1N_1L_2N_2$ , and  $n_k$  represents the  $N_1N_2 \times 1$  noise vector. It is assumed that all LR images have the same blur and so the matrix  $B_k$  can be substituted by  $B$ . These operations can be incorporated into one matrix [41, 67] and can be expressed as

$$y_k = DBM_kx + n_k = H_kx + n_k \tag{2}$$

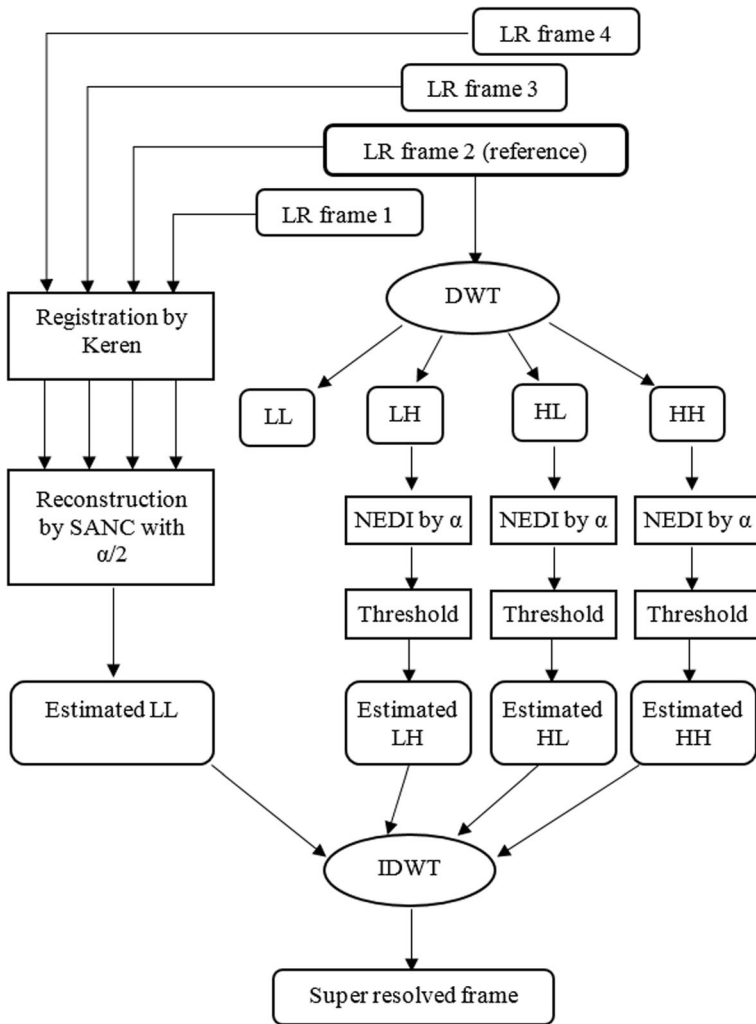


Fig. 1 The observation model

## 2.2 Proposed video super resolution technique

Recovering the missing high-frequency details of the given LR frames is the fundamental target of the video SR methods. The first step is sub-pixel image registration that aims to estimate the motion parameters between the reference frame and each of the neighbouring LR frames. When the camera is moving and the scene is stationary, global motion occurs including translation and rotation. In this work, Keren's method [30] is selected for global motion estimation which is one of the most accurate methods for sub-pixel image registration in the spatial-domain.

For image reconstruction, conventional interpolation methods (e.g., nearest neighbour, bilinear and bicubic) address the problem of reconstructing a HR image from the available LR image. However, these methods generally yield images with blurred edges and undesirable artefacts because they do not use any information pertinent to the edges in the original image. Therefore, the wavelet-based method is applied to preserve the high-frequency details (i.e. edges) and consequently construct the HR image from the given LR image. In the proposed technique, the discrete wavelet transform (DWT) is employed to isolate and preserve the high-frequency components of the image, and then the interpolation is applied to the high-frequency sub-bands. This is because the interpolation of isolated high-frequency components in the high-frequency sub-bands will preserve more edges of the image than using a direct interpolation. A number of DWT-based interpolation methods [2, 11, 12] have been developed to preserve the high-frequency components in the interpolated sub-bands. Nevertheless, the blurring effect from the employed interpolation method causes the potential loss of edges in these sub-bands. For example, the bicubic interpolation method produces blurring around the edges, even though the method is well accepted for resolution enhancement. Dual-tree complex wavelet transform (DT-CWT)-based interpolation methods [23, 26] are also being applied to address this problem by utilising an alternative interpolation method. Jagadeesh and Pragatheeswaran [26] used edge-directed interpolation EDI [1] as an alternative interpolation method to high-frequency sub-bands produced by DT-CWT. Later, this method was extended by Izadpanahi and Demirel [23] for video SR. Recently, the same authors applied new-edge directional interpolation NEDI [32] method to better preserve the edges of the interpolated high-frequency sub-bands generated by DT-CWT for local motion-based video SR [24]. However, none of these existing wavelet-based methods have tackled the problem of noisy high-frequency details corrupted by the limitations of imaging systems. For the current work, a combination of DWT, NEDI and an adaptive threshold is proposed not only for preserving the high-frequency details, but also for recovering the noiseless high-frequency information. One-level DWT process decomposes the input LR reference frame into four frequency sub-bands (LL, LH, HL and HH) in the frequency-domain. The high-frequency sub-bands (LH, HL and HH) are interpolated using the NEDI method with the scale factor,  $\alpha$ . The block diagram of the proposed video SR technique is illustrated by Fig. 2, which shows the combination of the so-called DWT-NEDI and an adaptive threshold with registration and reconstruction methods, in order to produce the estimated high and low-frequency sub-bands, respectively, for the inverse DWT. Generally, real video sequences are most commonly corrupted by noise such as additive Gaussian noise. Therefore, to better preserve the edges and remove potential noise in the estimated high-frequency sub-bands, a thresholding procedure that uses an adaptive threshold is applied to process the produced wavelet coefficients. Many types of thresholding functions have been introduced for the modification of estimated wavelet coefficients, such as Hard, Soft,



**Fig. 2** Block diagram of the proposed video SR technique

Semisoft and Garrote [5]. This paper employs a soft-thresholding technique proposed by Donoho [15] and extended by Zhang [66]. A universal threshold  $\tau$  for the considered sub-band can be calculated by

$$\tau = \sigma \sqrt{2 \log(N)/N} \tag{3}$$

where  $\sigma$  is the standard deviation of the sub-band, and  $N$  is the total number of pixels. The nonlinear soft-thresholding function is defined as

$$X_{out}(i, j) = \begin{cases} X_{in}(i, j) - \tau & X_{in}(i, j) > \tau \\ 0 & |X_{in}(i, j)| \leq \tau \\ X_{in}(i, j) + \tau & X_{in}(i, j) < -\tau \end{cases} \tag{4}$$



Equation (3) is chosen in the proposed method considering the prospect of automation in the proposed method and successful application of this equation in similar studies [66].

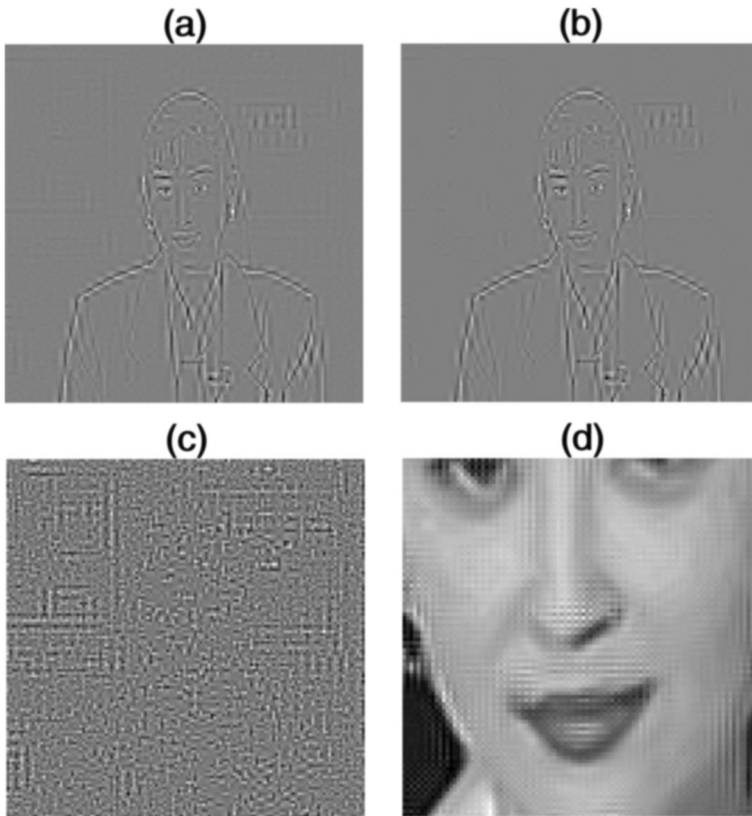
The rationale to include this thresholding process is that the energy of a signal is often concentrated on a few coefficients while the energy of noise is spread among all coefficients in the wavelet domain. Therefore, the nonlinear soft-thresholding tends to maintain few larger coefficients representing the signal and reduces noise coefficients to zero in the wavelet domain. The universal threshold is intuitively expected to uniformly remove the noise since the Gaussian noise still has the same variance over different scales in the transform domain [66]. On the other hand, in the spatial-domain, when the LR frames are precisely registered by Keren's method, the registered frames can be combined to reconstruct the missing high-frequency information and produce the low-frequency sub-band. In this work, structure-adaptive normalised convolution (SANC) reconstruction method [44] is applied, with half of the scale factor  $\alpha/2$ . This algorithm is used for fusion of irregularly sampled LR frames to recover the high-frequency details and generate the estimated LL sub-band, as the LL sub-band produced by the DWT does not contain any high-frequency information. Finally, inverse DWT (IDWT) process is applied to achieve a super-resolved frame by combining the estimated LL sub-band and processed high-frequency sub-bands.

The combination of DWT with NEDI aims to recover the edge details of directional high-frequency sub-bands and decrease the undesirable inter-directional interference in the SR process. This merit cannot be achieved using only the NEDI method, as indicated in the results section. The application of this soft-thresholding function is based on the hypothesis that the large coefficients in the high-frequency sub-bands reflect the true edges of objects, while the small coefficients reflect the noise, which is demonstrated by Fig. 3. Figure 3a shows the reconstruction image of high-frequency sub-bands only using IDWT without thresholding. Both true edges and noise can be clearly observed. Figure 3b shows the reconstruction image of high-frequency sub-bands only where the small coefficients are removed. It can be observed that the noise is significantly reduced, particularly in the background, while most of the true edges of the human body are preserved. The reconstruction image of high-frequency sub-bands can be acquired only where the large coefficients are removed and is as illustrated in Fig. 3c, which is dominated by noise and very few true edge information can be observed. To demonstrate the importance of this process, a region of the produced HR image using the proposed method without the soft-thresholding is shown in Fig. 3d, where the noise can be clearly observed.

The proposed technique can be summarised by the following steps:

1. Consider four consecutive frames from the LR video;
2. Estimate the motion parameters between the reference frame and each of the other LR frames using global motion estimation algorithm proposed by Keren;
3. Apply one-level DWT to decompose the input LR reference frame into four frequency sub-bands;
4. Apply the NEDI method to the LH, HL and HH high-frequency sub-bands with the scale factor of  $\alpha$ ;
5. Calculate the threshold  $\tau$  for each high-frequency sub-band;
6. Apply the nonlinear soft-thresholding process for each high-frequency sub-band to create the estimated  $\hat{LH}$ ,  $\hat{HL}$  and  $\hat{HH}$ ;
7. In the spatial-domain, employ SANC with half the scale factor  $\alpha/2$  to create the estimated  $\hat{LL}$ ;





**Fig. 3** An example to help justify the use of thresholding process. **a** the reconstruction image of high-frequency sub-bands only without thresholding; **b** the reconstruction image of high-frequency sub-bands only where the small coefficients are removed; **c** the reconstruction image of high-frequency sub-bands only where the large coefficients are removed; **d** the produced HR image without thresholding

8. Apply IDWT using  $(\hat{L}\hat{L}, \hat{L}\hat{H}, \hat{H}\hat{L}, \hat{H}\hat{H})$  to produce the output super-resolved frame.

### 3 Results

The proposed super-resolution technique was tested on three well-known video sequences, namely, "Mother & daughter", "Akiyo", and "Foreman". The files were downloaded from a public database [Xiph.org](http://Xiph.org). The proposed algorithm and other methods for comparison were implemented using Matlab 2015. The original high-resolution test videos were resized to  $512 \times 512$  pixels which are considered as the ground truth to evaluate the performance of the proposed approach. The frame rate of the test videos is 30 frames per second and each of the video sequences have 100 frames. Based on the observation model, the input LR video frames with the size of  $128 \times 128$  pixels were created as follows. Each original HR video frame is (1) blurred by a low-pass filter, (2) down-sampled in both the vertical and the horizontal directions by a scale factor of 1/4, and (3) added by a white Gaussian noise with a certain value of signal-to-noise ratio (SNR).

### 3.1 Visual and quantitative performance evaluation

This example aims to evaluate the overall performance of the proposed technique with a typical selection of parameters against other methods. Four shifted and rotated LR frames for each original HR frame were generated and down-sampled, and Gaussian noise was then added with a SNR value of 30 dB. The motion vectors were randomly produced with a standard deviation (STD) of 2 for shift and 1 for rotation. The wavelet function was chosen as db.9/7.

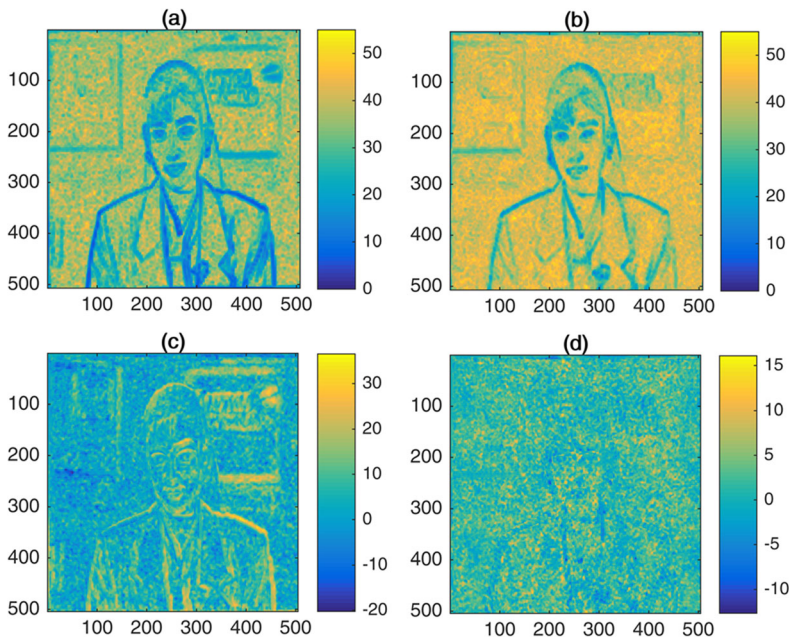
Figure 4 shows the super-resolved frames using the proposed method and other methods, selected arbitrarily from the video sequences *Akiyo*, *Mother & daughter*, and *Foreman*, respectively. It has been observed that the proposed technique produces the best visual quality in terms of preserving the edges, and removing the noise and aliasing artefacts in comparison to the other considered methods. The proposed technique preserves more information of the edges of the original HR video frame without smearing. For example, the edges of face in *Akiyo* produced by the proposed method are much cleaner in comparison with the images produced using other methods. Similar visual results have been observed for other edges in *Akiyo* and for other tested videos. Additionally, the noise and aliasing artefacts have been removed by the proposed method in comparison with the other methods. For example, the aliasing artefacts in *Mother* shoulders and hands have been removed substantially by the proposed technique as well as the Gaussian noise on the face of *Foreman*, while these noise and artefacts are clearly presented in the images produced by the other methods. From the motion estimation point of view, the aliasing high-frequency components due to down-sampling process appear to have different motions than the low-frequency components, and cause incorrect motion estimation [35]. Moreover, a larger noise level generates errors in motion estimation.



**Fig. 4** Results of produced HR images using different SR methods for a randomly selected frame of *Akiyo*, *Mother & daughter*, and *Foreman* video sequence respectively. Column 1: input LR frame, Column 2: Bicubic, Column 3: Vandwalle-SANC, Column 4: Keren-SANC, and Column 5: the proposed technique

To further investigate the improvement of the proposed method, Fig. 5 shows the local PSNR maps for different scenarios of the example *Akiyo*. The local PSNR map was calculated by a  $5 \times 5$  pixels window. Figure 5a shows the PSNR distribution between the raw HR image and the interpolated HR image produced by nearest neighbour method, which indicates the location of noise introduced by the degrading process. Regions of the human body have more noise (blue regions) introduced by the degrading process, while background regions have less (yellow regions). The blue regions tend to be areas with fine features (like boundary of the human body), while the yellow regions tend to have more coarse structures. Similar representation using the proposed method is shown in Fig. 5b, inspection of which proves that both background and fine features have been better recovered. To break down the contribution of each component, Figure 5c and d show the PSNR gain of the proposed method over NEDI and Keren-SANC, respectively. It has been observed that fine features are significantly improved in comparison to NEDI due to the consideration of adjacent frames, while the improvement of coarse structures is relatively small. Coarse structures have been significantly improved in comparison to Keren-SANC while the improvement of fine features is relatively small. All these observations clearly demonstrate that the proposed method improves the quality of both background and true edges, but other methods can only have one merit.

For quantitative evaluation of the experimental results, the nearest neighbour and bicubic interpolation methods, state-of-the-art resolution enhancement methods including NEDI [32], DASR [2], DWT-SWT [11], DWT-Dif [12], and state-of-the-art SR methods Keren-SANC and Vandewalle-SANC have been implemented to compare the performance of the proposed technique. The difference between the super-resolved images from different techniques can be small and sometimes it is difficult to be inspected visually. In this paper, as one of the most



**Fig. 5** An example to show the improvement distribution of the proposed technique. **a** The distribution of PSNR between the HR image using the nearest interpolation and the raw image; **b** the distribution of PSNR between the super-resolved image using the proposed method and the raw image; **c** the PSNR gain between the proposed method and NEDI; (d) the PSNR gain between the proposed method and Keren-SANC

commonly used objective fidelity criteria to evaluate image quality, the peak-signal-to-noise-ratio (PSNR) between the super-resolved image and the original HR image is used. It can be calculated by

$$PSNR = 10\log_{10}\left(\frac{L^2}{MSE}\right) \tag{5}$$

where  $L$  is the maximum fluctuation in the image. If the image is represented by 8-bit grayscale, the value of  $L$  will be 255.  $MSE$  represents the mean-square-error between the super- resolved image  $\hat{X}(i, j)$  and the original HR image  $X(i, j)$ . It can be calculated by

$$MSE = \frac{1}{W \times H} \sum_{i=1}^W \sum_{j=1}^H [\hat{X}(i, j) - X(i, j)]^2 \tag{6}$$

To complement the quantitative analysis, the structural similarity (SSIM) [58] image quality measure has also been applied. The SSIM index evaluates the visual effect of three characteristics of an image, namely, luminance, contrast and structure. It is based on the computation of these three components and is an inner product of them. It is defined by

$$SSIM = \frac{(2\mu_{\hat{X}}\mu_X + C_1)(2\sigma_{\hat{X}}\sigma_X + C_2)}{(\mu_{\hat{X}}^2 + \mu_X^2 + C_1)(\sigma_{\hat{X}}^2 + \sigma_X^2 + C_2)} \tag{7}$$

where  $\mu_{\hat{X}}, \mu_X$  are the local means for the images  $\hat{X}, X$ ,  $\sigma_{\hat{X}}, \sigma_X$  are corresponding standard deviations, and  $C_1, C_2$  are two constants to avoid the instability.

Table 1 shows the comparison of the averaged PSNR and SSIM results of 100 frames from the proposed method and other considered methods on the test videos. It is clearly shown that the proposed method achieves the highest average PSNR and SSIM values (31.48 dB, 30.57 dB and 23.88 dB for PSNR respectively; 0.90,0.91 and 0.84 for SSIM) for three tested videos. Achievement of this improved performance is due to the fact that the DWT-based SR reconstruction approach is more effective to recover the high-frequency details of the given LR frames, where the true edges are preserved and noise is removed benefiting from the nonlinear soft thresholding technique. Additionally, the combination of DWT and NEDI enables the

**Table 1** The averaged PSNR and SSIM values of 100 frames produced from different methods for three tested videos

SR methods	Mother& Daughter		Akiyo		Foreman	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Nearest	24.80	0.69	24.54	0.75	20.89	0.61
Bicubic	25.92	0.77	25.77	0.82	22.02	0.72
NEDI [32]	24.88	0.75	24.89	0.81	21.19	0.71
DASR [2]	23.76	0.62	23.83	0.69	20.16	0.55
DWT-Dif [12]	22.57	0.55	22.58	0.62	18.97	0.47
DWT-SWT [11]	23.10	0.57	23.09	0.65	19.83	0.50
Vandewalle-SANC	22.05	0.75	25.18	0.81	19.96	0.72
Keren- SANC	27.14	0.82	27.51	0.84	20.41	0.70
Proposed method	31.48	0.90	30.57	0.91	23.88	0.84

recovering of the edge details of directional high-frequency sub-bands and reduces the annoying inter-directional interference in the SR process. For the videos *Mother & daughter* and *Akiyo*, the proposed technique based on global motion of the entire frame is suitable for these videos and produces better PSNR results than the interpolation and classic SR methods (16% and 11% increment over Keren-SANC respectively). For the video *Foreman*, although PSNR result achieved by the proposed method is higher than the other considered methods (17% increment over Keren-SANC), the performance gain can be further increased by utilising local motion which divides each frame into multiple blocks and processes each block individually. Similar observation can be concluded based on SSIM values.

### 3.2 Performance for variety of noise levels

To demonstrate the robustness of the proposed method against noise benefiting from the adaptive thresholding process, four shifted and rotated LR frames for each original HR frame were generated and down-sampled, and the motion vectors were randomly produced with a standard deviation of 2 for shift and 1 for rotation. The wavelet function was chosen as db.9/7. The noise level was increased from 50 dB to 20 dB with the step of 5 dB. The first 10 frames from *Akiyo* video were tested by the proposed method and other different methods, and the results were averaged.

Table 2 shows the comparison of averaged PSNR results between different methods. The last column shows the PSNR increment over Keren-SANC in percentage by the proposed technique. It can be clearly seen that the proposed method consistently has the best performance for every noise level. Furthermore, the performance is even better for images corrupted by high noise level (15%, 21%, 18% increment for 30 dB, 25 dB and 20 dB respectively) than those with low noise level (10%, 8%, 5%, 3% increment for 50 dB, 45 dB, 40 dB and 35 dB respectively).

### 3.3 Performance for variety of wavelet functions

The above results from the proposed technique have been produced by the most widely used wavelet function db.9/7 in image SR applications. This section discusses the prospect of the proposed approach using other wavelet functions. Previously, our research showed that the selection of wavelet function can affect the performance of SR techniques [60]. In this experiment, the same parameters were chosen except that the noise level was fixed as 30 dB, and the wavelet function was set as a variable entity. Table 3 shows the averaged

**Table 2** The averaged PSNR results of 10 frames from Akiyo test video for each noise level, range from 20 dB to 50 dB with 5 dB step

SNR	Nearest	Bicubic	Keren-SANC	Proposed method	Increment
50 dB	25.20	26.50	28.18	31.13	10.47%
45 dB	23.86	24.85	28.46	30.82	8.29%
40 dB	24.98	26.29	30.05	31.64	5.29%
35 dB	24.18	25.20	30.49	31.30	2.65%
30 dB	23.84	24.86	26.50	30.37	14.60%
25 dB	23.69	24.85	24.80	30.08	21.29%
20 dB	23.42	24.93	24.31	28.63	17.78%

**Table 3** The averaged PSNR and SSIM values of 10 frames produced by the proposed technique for different wavelet functions

Wavelet functions	PSNR	SSIM
Db1	28.18	0.88
Db2	29.95	0.90
Sym16	30.83	0.92
Sym20	30.89	0.92
Coif1	30.39	0.91
Coif2	26.78	0.85
Bior4.4	30.72	0.92
Bior5.5	30.20	0.90
Bior6.8	30.81	0.92

PSNR and SSIM results of the first 10 frames for the *Akiyo* video sequence produced by the proposed method using nine wavelet functions, which include db1, db2, sym16, sym20, coif1, coif2, bior4.4 (db.9/7), bior5.5, and bior6.8. The results show that the proposed technique can perform well under other wavelet functions apart from db.9/7, even better than db.9/7. Not that bior4.4 is equal to db.9/7 [4]. In terms of both PSNR and SSIM values, the wavelet function with top 5 performances are sym20, sym16, bior6.8, bior4.4 and coif1 respectively, although the difference between them is not significant.

### 3.4 Performance for variety of motion levels

This section is dedicated to discussing the effectiveness of motion level (shift and rotation) on the performance of the proposed algorithm. In this experiment, the shift on both horizontal and vertical directions and rotation angle were randomly selected with the standard variation STD changing from 1 to 4 during generating the input LR frames from a HR frame. Four shifted and rotated LR frames for each original HR frame were generated and down-sampled. The wavelet function was chosen as db.9/7, and the noise level was fixed as 30 dB.

The averaged PSNR and SSIM results of the first 10 frames from *Akiyo* video produced by the proposed technique with different motion levels are shown in Table 4. It can be observed that the proposed method produces the higher PSNR and SSIM values when the motion level is relatively small. This is because of the fact that the estimation of a small motion is usually more accurate and leads to better reconstruction. When the motion level is large, the values of PSNR and SSIM drop as expected. This is because a large motion is more difficult to be measured accurately and errors in motion estimation prevents reconstructing the original HR

**Table 4** The averaged PSNR and SSIM values of 10 frames produced by the proposed technique with different motion levels

STD of Shift	STD of Rotation					
	1		2		4	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
1	30.74	0.92	29.44	0.90	28.09	0.87
2	29.16	0.90	29.00	0.89	27.97	0.88
4	27.35	0.86	27.13	0.86	26.90	0.85



**Table 5** The averaged PSNR and SSIM values of 10 frames produced by the proposed technique by sampling different number of frames

Number of sampled frames	PSNR	SSIM
4	29.81	0.91
8	30.06	0.92
16	29.98	0.92
32	30.07	0.92

frames correctly. From the considered smallest to largest motions, 12% and 7% decrease of PSNR and SSIM values respectively has been observed.

### 3.5 Performance for variety of number frames

This section aims to evaluate the effectiveness of the number of used frames on the performance. In all previous experiments, it was assumed that the shift and rotation parameters are randomly produced. But in real applications, the camera usually moves towards one direction which means the shifts change monotonously. The shifts include two motion vectors, horizontal shift  $\Delta_x$  and vertical shift  $\Delta_y$ . To simplify the process, in this experiment, the rotation angle was randomly selected with the standard deviation of 1, and only shifts are changed. The shifts were produced based on

$$\Delta_x(i) = \Delta_y(i) = \frac{i}{N} \quad (8)$$

where  $i(i = 1, 2, \dots, N)$  denotes the time index of LR frames and  $N$  denotes the total number of used frames. All other parameters are same as the previous experiment.

Table 5 shows the averaged PSNR and SSIM results of the first 10 frames for the *Akiyo* video sequence produced by the proposed technique using the number frames of 4, 8, 16 and 32 respectively. It has been observed that, as expected, higher PSNR and SSIM values were achieved with more number of sampled frames. However, the increment is only about 1% when the number changes from 4 to 32. This observation is because the used motion model is simple, and there is very limited extra contribution from 32 frames in comparison to 4 frames. A higher percentage of improvement could be achieved when the motion is more complicated or the motion is corrupted by more noise.

## 4 Conclusions

A robust video super-resolution reconstruction approach based on combining discrete wavelet transform, new edge-directed interpolation and the nonlinear soft-thresholding has been proposed in this paper for noisy LR video sequences with global motion to recover the noiseless high-frequency details and increase the spatial resolution, which integrates properties from methods of image registration and reconstruction. Firstly, an iterative planar motion estimation algorithm by Keren is used to estimate the motion parameters between a reference frame and its neighbouring LR frames in the spatial domain. The registered frames are combined by the SANC reconstruction method to output the estimated low-frequency sub-band. Secondly, the DWT is employed to decompose each input LR reference frame into four frequency sub-bands in the frequency-domain. The NEDI is employed to process each of



three high-frequency sub-bands, which are then filtered using the adaptive thresholding process to preserve the true edges and reduce the noise in the estimated high-frequency sub-bands. Finally, by combining the estimated low-frequency sub-band and three high-frequency sub-bands, a super-resolved frame is recovered through the invert-DWT process.

Subjective results show that this approach can better preserve the edges and remove potential noise in the estimated high-frequency sub-bands since a direct interpolation will blur the areas around edges. Three well-known videos (totally 100 frames for each) have been tested, and the quantitative results show that the superior performance of the proposed method. The proposed method tops the averaged PSNR and SSIM values (31.48 dB, 30.57 dB and 23.88 dB for PSNR respectively; 0.90, 0.91 and 0.84 for SSIM) for three videos respectively, and the averaged increment over KEREN-SANC is 16%, 11%, and 17% respectively. The performance against noise has also been analysed. Analysis based on the contribution of each component clearly demonstrates that the proposed method improves the quality of both background and true edges, but other methods usually can only have one merit.

One of the motivations of this paper is to address the limited performance capabilities of most of the existing wavelet-based SR methods for a variety of motion levels, noise levels, wavelet functions and adequate number of used frames, do empirical tests and analyse how these factors can affect the performance of the proposed method. The conclusions are:

- The proposed technique has produced 10%, 8%, 5% and 3% averaged increment of PSNR for an image corrupted by low level noise with the SNR value of 50 dB, 45 dB, 40 dB and 35 dB respectively. It has produced 15%, 21% and 18% averaged increment of PSNR for the image corrupted by high level noise with the SNR value of 30 dB, 25 dB and 20 dB respectively.
- The proposed technique can perform well using other wavelet functions apart from db.9/7, even better than db.9/7, although the difference between them is not significant.
- The performance of the proposed method is affected by the level of motion. Based on the considered smallest to largest motions, 12% and 7% decrease of PSNR and SSIM values respectively has been observed.
- If the motion is simple, the number of sampled frames has limited improvement on the performance due to the limited extra information. If the motion is complex and corrupted by high level of noise, significant improvement is expected using more frames.

A limitation of this method is that it can only be applied to video sequences with global motion. However, it can be extended to local motion by dividing the video frame into multiple blocks and then applying this method to each block.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

1. Allebach J, Wong PW Edge-directed interpolation. Proceedings of 3rd IEEE International Conference on Image Processing 3:707–710
2. Anbarjafari G, Demirel H (2010) Image super resolution based on interpolation of wavelet domain high frequency subbands and the spatial domain input image. ETRI J 32(3):390–394

3. Anbarjafari G, Izadpanahi S, Demirel H (2015) Video resolution enhancement by using discrete and stationary wavelet transforms with illumination compensation. *Signal, Image Video Process* 9:87–92
4. Antonini M, Barlaud M, Mathieu P, Daubechies I (1992) Image coding using wavelet transform. *IEEE Trans Image Process* 1(2):205–220
5. Bhandari AK, Soni V, Kumar A, Singh GK (Jul. 2014) Cuckoo search algorithm based satellite image contrast and brightness enhancement using DWT–SVD. *ISA Trans* 53(4):1286–1296
6. Chen Y, Pock T (2017) Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration. *IEEE Trans Pattern Anal Mach Intell* 39(6):1256–1272
7. Crouse MS, Nowak RD, Baraniuk RG (1998) Wavelet-based statistical signal processing using hidden markov. *IEEE Trans Signal Process* 46(4):886–902
8. Cui Z, Chang H, Shan S, Zhong B, Chen X (2014) Deep network cascade for image super-resolution. *ECCV* 5:49–64
9. Daubechies I (1992) Ten lectures on wavelets. Society for Industrial Applied Mathematics, Philadelphia
10. Demirel H, Anbarjafari G (2010) Satellite image resolution enhancement using complex wavelet transform. *IEEE Geosci Remote Sens Lett* 7(1):123–126
11. Demirel H, Anbarjafari G (2011) IMAGE resolution enhancement by using discrete and stationary wavelet decomposition. *IEEE Trans Image Process* 20(5):1458–1460
12. Demirel H, Anbarjafari G (2011) Discrete wavelet transform-based satellite image resolution enhancement. *IEEE Trans Geosci Remote Sens* 49(6):1997–2004
13. Dong C, Loy CC, He K, Tang X (2014) Learning a deep convolutional network for image super-resolution. *ECCV* 4:184–199
14. Dong C, Loy CC, He K, Tang X (2016) Image super-resolution using deep convolutional networks. *IEEE Trans Pattern Anal Mach Intell* 38(2):295–307
15. Donoho DL (1995) De-noising by soft-thresholding. *IEEE Trans Inf Theory* 41(3):613–627
16. Elad M, Feuer A (1997) Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images. *IEEE Trans Image Process* 6(12):1646–1658
17. Eren PE, Sezan MI, Tekalp AM (1997) Robust, object-based high-resolution image reconstruction from low-resolution video. *IEEE Trans Image Process* 6(10):1446–1451
18. Gonzalez RC, Woods RE (2007) Digital image processing. Prentice Hall, Englewood Cliffs
19. Greenspan H (2008) Super-resolution in medical imaging. *Comput J* 52(1):43–63
20. Huang S-C (2011) An advanced motion detection algorithm with video quality analysis for video surveillance images. *IEEE Transactions on Circuits and Systems for video Technology* 20(1):1–13
21. Iqbal MZ, Ghafoor A, Siddiqui AM (2013) Satellite image resolution enhancement using dual-tree complex wavelet transform and nonlocal means. *IEEE Geosci Remote Sens Lett* 10(3):451–455
22. Irani M, Peleg S (1991) Improving resolution by image registration. *CVGIP Graph Model Image Process* 53(3):231–239
23. Izadpanahi S, Demirel H (2012) Multi-frame super resolution using edge directed interpolation and complex wavelet transform. In: IET Conference on Image Processing (IPR 2012), pp A9–A9. <https://doi.org/10.1049/cp.2012.0447>
24. Izadpanahi S, Demirel H (2013) Motion based video super resolution using edge directed interpolation and complex wavelet transform. *Signal Process* 93(7):2076–2086
25. Izadpanahi S, Ozcinar C (2013) DWT based resolution enhancement of video sequences. no. 2000
26. Jagadeesh P, Pragatheeswaran J (2011) Image resolution enhancement based on edge directed interpolation using dual tree — Complex wavelet transform. In: 2011 International Conference on Recent Trends in Information Technology (ICRTIT), pp 759–763. <https://doi.org/10.1109/ICRTIT.2011.5972260>
27. Jiang J, Hu R, Han Z, Lu T (2014) Efficient single image super-resolution via graph-constrained least squares regression. *Multimed Tools Appl* 72(3):2573–2596. <https://doi.org/10.1007/s11042-013-1567-9>
28. Jiang J, Ma X, Chen C, Lu T, Wang Z, Ma J (2017) Single Image Super-Resolution via Locally Regularized Anchored Neighborhood Regression and Nonlocal Means. *IEEE Trans Multimed* 19(1):15–26
29. Kamenicky J, Bartos M, Flusser J, Mahdian B, Kotera J, Novozamsky A et al (2016) PIZZARO: Forensic analysis and restoration of image and video data. *Forensic Sci Int* 246:153–166
30. Keren D, Peleg S, Brada R (1988) Image sequence enhancement using sub-pixel displacements. In: Proceedings CVPR '88: The Computer Society Conference on Computer Vision and Pattern Recognition, pp 742–746. <https://doi.org/10.1109/CVPR.1988.196317>
31. Kim SP, Bose NK, Valenzuela HM (1990) Recursive reconstruction of high resolution image from noisy undersampled multiframes. *IEEE Trans Acoust* 38(6):1013–1027
32. Li X, Orchard MT (2001) New edge-directed interpolation. *IEEE Trans Image Process* 10(10):1521–1527
33. Li K, Zhu Y, Yang J, Jiang J (2016) Video super-resolution using an adaptive superpixel-guided autoregressive model. *Pattern Recogn* 51:59–71

34. Liao R, Tao X, Li R, Ma Z, Jia J (2015) Video Super-Resolution via Deep Draft-Ensemble Learning. In: 2015 I.E. International Conference on Computer Vision (ICCV), pp 531–539. <https://doi.org/10.1109/ICCV.2015.68>
35. Liu C, Sun D (2014) On Bayesian adaptive video super resolution. *IEEE Trans Pattern Anal Mach Intell* 36(2):346–360
36. Lucchese L, Cortelazzo GM (2000) A noise-robust frequency domain technique for estimating planar rotations. *IEEE Trans Signal Process* 48(6):1769–1786
37. Ma J, Zhao J, Tian J, Yuille AL, Tu Z (2014) Robust point matching via vector field consensus. *IEEE Trans Image Process* 23(4):1706–1721
38. Ma J, Qiu W, Zhao J, Ma Y, Yuille AL, Tu Z (2015) Robust L2E estimation of transformation for non-rigid registration. *IEEE Trans Signal Process* 63(5):1115–1129
39. Marcel B, Briot M, Murrieta R (1997) Calcul de Translation et Rotation par la Transformation de Fourier. *Traitement du Signal* 14(2):135–149
40. Nguyen N, Milanfar P (2000) A Wavelet-based interpolation-restoration method for superresolution (wavelet superresolution). *Circuits Syst Signal Process* 19(4):321–338
41. Park SC, Park MK, Kang MG (2003) Super-resolution image reconstruction: a technical overview. *IEEE Signal Process Mag* 20(3):21–36
42. Patti AJ, Sezan MI, Tekalp AM (1997) Superresolution video reconstruction with arbitrary sampling lattices and nonzero aperture time. *IEEE Trans Image Process* 6(8):1064–1076
43. Peleg S, Keren D, Schweitzer L (1987) Improving Image Resolution Using Subpixel Motion. *Pattern Recogn Lett* 5(3):223–226
44. Pham TQ, Van Vliet LJ, Schutte K (2006) Robust fusion of irregularly sampled data using adaptive normalized convolution. *EURASIP J Appl Signal Processing* 2006:1–12
45. Protter M, Elad M, Takeda H, Milanfar P (2009) Generalizing the non-local-means to super-resolution reconstruction. *IEEE Trans Image Process* 18(1):1958–1975
46. Robinson FS, Chiu SJ, Lo JY, Toth CA, Izatt JA (2010) Novel applications of super-resolution in medical imaging. In: Milanfar P (ed) *Super-Resolution Imaging*. CRC Press, Boca Raton, pp 383–412
47. Satiro J, Nasrollahi K, Correia PL, Moeslund TB (2015) Super-resolution of facial images in forensics scenarios. In: 2015 International Conference on Image Processing Theory, Tools and Applications (IPTA), pp 55–60. <https://doi.org/10.1109/IPTA.2015.7367096>
48. Schultz RR, Stevenson RL (1996) Extraction of high-resolution frames from video sequences. *IEEE Trans Image Process* 5(6):996–1011
49. Schultz RR, Stevenson RL (1994) A bayesian approach to image eXpansin for improved definition. *IEEE Trans Image Process* 3(3):233–242
50. Shi W, Caballero J, Huszar J, Totz J, Aitken AP, Bishop R, Rueckert D, Wang Z (2016) Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In: 2016 I.E. Conference on Computer Vision and Pattern Recognition (CVPR), pp 1874–1883. <https://doi.org/10.1109/CVPR.2016.207>
51. Takeda H, Milanfar P, Protter M, Elad M (2009) Super-Resolution without Explicit Subpixel Motion Estimation. *IEEE Trans Image Process* 18(9):1958–1975
52. Temizel A (2007) Image resolution enhancement using wavelet domain hidden markov tree and coefficient sign estimation. *Proc Int Conf Image Process* 5:V – 381–V – 384
53. Tian J, Ma K-K (2011) A survey on super-resolution imaging. *Signal, Image Video Process* 5(3):329–342
54. Tom BC, Katsaggelos AK, Galatsanos NP (1994) Reconstruction of a high resolution image from registration and restoration of low resolution images. In: *Proceedings - International Conference on Image Processing, ICIP*, 3, pp 553–557. <https://doi.org/10.1109/ICIP.1994.413745>
55. Tsai RY, Huang TS (1984) Multiframe image restoration and registration. In: *Advances in Computer Vision And Image Processing*, vol 1. JAI Press, London, pp 317–339
56. Ur H, Gross D (1992) Improved resolution from subpixel shifted pictures. *CVGIP Graph Model Image Process* 54(2):181–186
57. Vandewalle P, Süsstrunk S, Vetterl M (2006, 2006) A frequency domain approach to registration of aliased images with application to super-resolution. *EURASIP J Appl Signal Processing*:1–14
58. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Process* 13(4):600–612. <https://doi.org/10.1109/TIP.2003.819861>
59. Wang Z, Yang Y, Wang Z, Chang S, Han W, Yang J, Huang TS (2015) Self-tuned deep super resolution. *CVPR Workshops*:1–8
60. Witwit W, Zhao Y, Jenkins K, Zhao Y (2016) An optimal factor analysis approach to improve the wavelet-based image resolution enhancement techniques. *Global J Comp Sci Technol* 16(3F):1–11
61. Yan C, Zhang Y, Xu J, Dai F, Li L, Dai Q, Wu F (2014) A highly parallel framework for HEVC Coding unit partitioning tree decision on many-core processors. *IEEE Signal Process Lett* 21(5):573–576

62. Yan C, Zhang Y, Xu J, Dai F, Zhang J, Dai Q, Wu F (2014) Efficient parallel framework for HEVC motion estimation on many-core processors. *IEEE Trans Circuits Syst Video Technol* 24(12):2077–2089
63. Yan C, Xie H, Yang D, Yin J, Zhang Y, Dai Q (2018) Supervised hash coding with deep neural network for environment perception of intelligent vehicles. *IEEE Trans Intell Transp Syst* 19(1):284–295. <https://doi.org/10.1109/TITS.2017.2749965>
64. Yan C, Xie H, Liu S, Yin J, Zhang Y, Dai Q (2018) Effective uyghur language text detection in complex background images for traffic prompt identification. *IEEE Trans Intell Transp Syst* 19(1):220–229. <https://doi.org/10.1109/TITS.2017.2749977>
65. Yang J, Wright J, Huang TS, Ma Y (2010) Image super-resolution via sparse representation. *IEEE Trans Image Process* 19(11):2861–2873
66. Zhang X-P (2001) Thresholding neural network for adaptive noise reduction. *IEEE Trans Neural Netw* 12(3):567–584
67. Zhang L, Zhang H, Shen H, Li P (Mar. 2010) A super-resolution reconstruction algorithm for surveillance images. *Signal Process* 90(3):848–859



**Wasnaa Witwit** is currently pursuing the PhD at Cranfield University, United Kingdom. She worked as a research assistant at the Department of Physics of Babylon University, Iraq. She received her MSC degree in Resampling of Astronomical Images from the University of Babylon in 2002. Her research interests are in image processing, astronomical and satellite images analysis, interpolation, and video super resolution.



**Yifan Zhao** is a Lecturer in Image and Signal Processing and Degradation Assessment at Cranfield University. He received his PhD degree in Automatic Control and System Engineering from the University of Sheffield in 2007. He received the MSc and BEng degrees in Automatic Control Theory and Engineering from Beijing Institute of Technology, China in 2003 and 2000 respectively. His current research interests are computer vision, super resolution, thermography for Non-destructive Testing, and EEG data analysis.



**Karl Jenkins** is currently Head of the Centre for Computational Engineering Sciences within the School of Aerospace, Transport and Manufacturing at Cranfield University. He is also Programme Director for the MSc in Computational and Software Techniques in Engineering with specialist option in Digital Signal and Image Processing at Cranfield University. He received his PhD degree in Wave modelling from the University of Manchester in 1996. He was the Sir Arthur Marshal Research Fellow at Cambridge University between 1998 and 2004. His current research interests are computational engineering, high fidelity CFD simulations and high performance.



**Sri Addepalli** is a Research Fellow in Mechanical Component Degradation Analysis at Cranfield University. He was the recipient of the Prince of Wales Innovation Scholarships (PoWIS) awarded by the University of Wales in 2010. His research interests include passive and active thermography and material / component damage characterisation with specific expertise in composite materials. Sri holds an MPhil in Materials Technology from Swansea University, UK and a BEng in Mechanical Engineering from Anna University, India. He is also a CSWIP Certified Level 1 Thermography Inspector and a full time professional member of the Institute of Engineering Technology (MIET).