



Dual-tree complex wavelet transform and super-resolution based video inpainting application to object removal and error concealment

ISSN 2468-2322

Received on 10th June 2019

Revised on 19th June 2020

Accepted on 14th July 2020

doi: 10.1049/trit.2019.0045

www.ietdl.org

Gajanan Tudavekar^{1,2} ✉, Sanjay R. Patil³, Santosh S. Saraf⁴

¹Department of Electronics & Communication Engineering, Angadi Institute of Technology and Management, Belagavi, Karnataka, India

²Visvesvaraya Technological University, Belagavi, Karnataka, India

³Department of Electronics & Telecommunication Engineering, Marathwada Institute of Technology, Aurangabad, Maharashtra, India

⁴Department of Electronics & Communication Engineering, KLS Gogte Institute of Technology, Belagavi, Karnataka, India

✉ E-mail: askgajanan@gmail.com

Abstract: Video inpainting is a technique that fills in the missing regions or gaps in a video by using its known pixels. The existing video inpainting algorithms are computationally expensive and introduce seam in the target region that arises due to variation in brightness or contrast of the patches. To overcome these drawbacks, the authors propose a novel two-stage framework. In the first step, sub-bands of wavelets of a low-resolution image are obtained using the dual-tree complex wavelet transform. Criminisi algorithm and auto-regression technique are then applied to these sub-bands to inpaint the missing regions. The fuzzy logic-based histogram equalisation is used to further enhance the image by preserving the image brightness and improve the local contrast. In the second step, the image is enhanced using super-resolution technique. The process of down-sampling, inpainting and subsequently enhancing the video using the super-resolution technique reduces the video inpainting time. The framework is tested on video sequences by comparing and analysing the structural similarity index matrix, peak-signal-to-noise ratio, visual information fidelity in pixel domain and execution time with the state-of-the-art algorithms. The experimental analysis gives visually pleasing results for object removal and error concealment.

1 Introduction

Video inpainting is a method to fill the missing region or gaps, recover objects in a video. The removal of undesired objects in the video can create missing areas [1–3]. The objective of video inpainting is to fill the missing areas so the results are visually convincing both in time and space. The inpainting techniques are classified as partial differential equation (PDE) based approaches and exemplar-based techniques. PDE-based approaches construct the diffusion PDE based on the edge information into the required region. Based on the work of Bertalmio *et al.* [4], Chan and Shen [5, 6] proposed PDE-based models and total variation (TV) model to handle the non-texture image inpainting problems. The drawback of these approaches is that blur gets introduced for the vast region of the target. Granados *et al.* [7] have proposed a homography registration based inpainting method that fills large holes in videos. The exemplar-based approach is proposed by the Liu and Caselles [8] and Xia *et al.* [9] for removing objects from video and filling the gaps by dividing the frame. For the restoration of the entirely damaged object, Ghanbari and Soryani [10] proposed a contour-based method for video inpainting. In this work, moving objects are separated, and then using a contour-based comparison and patch-based inpainting method, the damaged regions are inpainted. Tremendous work has been done in the past years on video inpainting, but issues remain when the hole to be filled is largely filled, and high computational time in general required for video inpainting. To address these issues, a super-resolution (SR) based video inpainting framework is proposed. SR is the technique to reconstruct high-resolution [11–13] image from one or more low-resolution (LR) images. Caselles and co-workers [4] have developed the first PDE-based inpainting algorithm. The algorithm iteratively updates the output image by adding objective information gained from the surrounding geometric structure. However, this algorithm is not

suitable for filling large missing areas. Also, due to the blurring artefacts generated by the diffusion process, the algorithm may cause distractions to the viewers of inpainting results. Also, the algorithm lacks explicit treatment of the pixels on edges. Later, inspired by Bertalmio's algorithm, Chan and Shen [6] proposed the TV inpainting model, which is closely connected to the classic TV denoising model of Rudin *et al.* [14]. To overcome this problem, Le Sun's *et al.* [15] proposed a novel Hue-saturation-intensity (HSI) mixed denoising method based on 3D spectral-spatial cross-TV. Amrani *et al.* [16] proposed a novel PDE-based inpainting algorithm to compress hyperspectral images. The method inpaints the known data in the spatial and spectral dimensions separately. Then, it applies a prediction model to the final inpainting solution to obtain a representation much closer to the original image. Newson *et al.* [17, 18] optimised a global patch-based function and thus made a significant improvement, especially in motion preservation, by incorporating the optical flow in several stages of the algorithm. In the paper authored by Huang and Tang [19], when restoring the damaged background, the information is directly replicated to fill in the damaged region from the constructed panoramic background image. When restoring the damaged foreground, it seeks the best state matching frame by comparing the next running foreground state of the damaged frame and searches the best matching patch in this frame. Meanwhile, the algorithm also improves the way of searching exemplar patch, the principle of matching cost and way of confidence term updating. Newson *et al.* [20] proposed a method combining the patch-based method and hierarchical pyramid decomposition. This method uses multiscale scheme inpainting with the help of k-nearest neighbours searching and energy minimum constraints. However, structural connectivity is not yet complete. Janardhana Rao *et al.* [21] recognised that the priority of the patch falls to a low value due to the rapid decrease in confidence term at lower iterations. To overcome this effect,

regularisation factor is added to the confidence term. The regularisation factor is used to reduce the smoothness of the confidence curve. In recent years, to improve the accuracy of the video inpainting, deep learning based image inpainting methods use SR. These methods produce accurate results but the runtime is more because these models need to be trained using external training data [22–24].

In this paper, a novel video inpainting framework that combines dual-tree complex wavelet transform (DTCWT), auto-regression (AR), fuzzy enhancement and SR is proposed. The process of down-sampling, inpainting and subsequently enhancing the video using SR technique reduces the video inpainting time. Also, the use of AR avoids the discontinuities that arise due to variation in brightness or contrast of patches.

The contributions of this paper are as below:

- *Patch priority selection*: Proposed a framework to overcome the drawback of the Criminisi *et al.* [25] algorithm, i.e. sensitive to the parameter setting of the inpainting method. To overcome this, we have inpainted input image several times with different configurations.
- *Reduction in seam*: In the exiting video inpainting methods, seam arises due to change in brightness of patches. To overcome this, AR model is used to select appropriate patches to be copied into the target region.
- *Reduction in execution time*: Existing video inpainting methods are computationally expensive. A modified SR algorithm is proposed to reduce the execution time.

2 Proposed method

The proposed method is the combination of two sequential operations, as shown in Fig. 1. The first operation is filling in the missing regions of LR image using Criminisi [24] based inpainting algorithm. The second operation super resolves the inpainted LR image to get a high resolution (HR) image. The block diagram is discussed below.

2.1 Frame extraction and down-sampling

The frames are extracted from a video, and these frames are down-sampled by factor of two to get LR images. Inpainting a LR image is not dependent on noise [26], and hence computational time is low when compared to inpainting a HR image. After getting the LR image, subtract the current image from the previous image. If the residual is zero or less than the threshold value of pixels, the inpainting process is skipped as it indicates that the current frame is the same as the previous frame. This step avoids the inpainting of redundant frames. If the residual is non-zero, the inpainting process is carried out, as discussed in the below sections.

2.2 Dual-tree complex wavelet transform

The author Kingsbury [27] proposed that DTCWT can be used to overcome the disadvantages of the traditional wavelet transform. The complex wavelet transform (CWT) is the complex-valued

extension to standard DWT. CWT uses complex value filtering that decomposes the real/complex signal into real and imaginary parts in the transform domain. The real and imaginary coefficients are used to compute amplitude and phase information. DTCWT has separated sub-bands for positive and negative orientations. DTCWT calculates the complex transform of the signal using two separate discrete wavelet transform (DWT) decomposition.

The DTCWT decomposes the input image into its 16 subbands. Out of 16 subbands, 4 are the average subbands which have the average image contents. The other 12 subbands contain high-frequency information. Each subband contains the edge information, which is oriented at various angles. The subbands in the DTCWT are oriented at the angles $\{\pm 15^\circ, \pm 45^\circ, \pm 75^\circ\}$. So, we get the oriented subbands A1 to A12. The subbands A13 to A16 are discarded as they are the average subbands [28], and they do not contain any special information.

2.3 Inpainting method

Criminisi's inpainting algorithm is applied to four sub-bands of the image formed after applying DTDWT. The inpainting approach is as below (Fig. 2):

- For each pixel k on boundary $\partial\Omega$
- Determine the priority

$$P(k) = C(k)D(k) \quad (1)$$

where $C(k)$ is confidence measure. It is measure of information in the neighbourhood pixel and $D(k)$ data term.

In this

$$C(k) = \frac{\sum_{q \in (\psi_k(R-\Omega))} C(q)}{\|\psi_k\|} \quad (2)$$

where $\|\psi_k\|$ is the area of patch ψ_k of pixel and Ω is the target region of image R .

This confidence term has high values near the border of the initial mask, and decreases near the centre of Ω . It tends then to in-paint first pixels having the most valid neighbours.

The initial conditions of confidence term $C(q)$ are calculated as below:

$$C(q) = \begin{cases} 0, & q \in R - \Omega \\ 1, & q \in \Omega \end{cases}$$

The confidence term is updated as below after the patch has been filled with the pixels

$$C(q) = C(k)$$

The term $D(k)$ takes care of structures in ψ_k . In this method, structure tensors [29] approach is used to calculate data term

$$D(k) = \|W_p n_k\| \quad (3)$$

where $W_p = \sum_{q \in (\psi_k \cap (I-\Omega))} w \nabla I_k \nabla I_k^T$

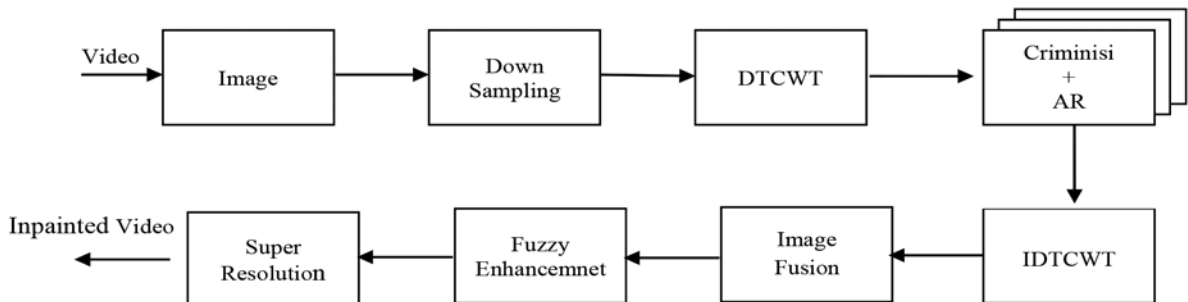


Fig. 1 Proposed block diagram for video inpainting

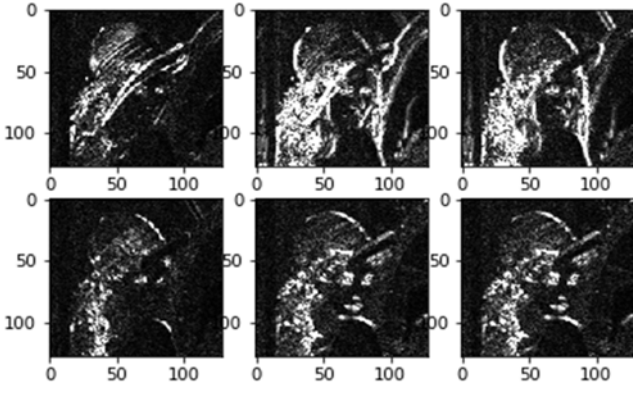


Fig. 2 Sub-band of image after taking DTCWT

The colour gradient ∇I_k at k , n_k is unit vector orthogonal to boundary $\partial\Omega$ at k and w is a normalised 2D Gaussian function centred at p .

(iii) Once these priorities are calculated, the patch $\psi_{k'}$ around the pixel k' that has maximum priority is considered for filling. Since the pixels on the boundary of the region $\partial\Omega$ to be in-painted get more priority, the selected patch $\psi_{k'}$ will always consist of both the known and missing pixels. Hence, we need to use contents or similar patches called exemplar. The patch $\psi_{k'}$ is compared with a patch ψ_l around every pixel l to find an exemplar in an image. As the image size increases, the execution time required to find an exemplar also increases. However, the similar patches may be found in nearby region. Instead of searching the whole image, the search is restricted for matching the patches to large-sized search window $W_{k'}$ of size 90×90 (set empirically) around the patch to be filled up. This step reduces the number of computations required. Find the similarity between two patches using the sum of squared differences (SSD) measure. The patch ψ_l which gives the minimum SSD is considered as exemplar $E_{k'}$. Due to variation in intensity or contrast, the SSD calculated for $E_{k'}$ will be high. In such cases, patches will be visible in in-painted image. To achieve better inpainting, estimate

the neighbourhood pixels using AR model. The AR parameters suggest the values of neighbouring pixels towards the centre pixel of every 3×3 region in that set.

(iv) Estimate the missing pixels in the $\psi_{k'}$ with the help of the exemplar $E_{k'}$ and AR parameters. Di Zenzo [29] and Perez *et al.* [30] methods are used to copy the pixels from the source region in an image into the missing region in a same or different image. Once a patch is processed, its pixels are excluded from the region to be inpainted. The updated missing region is then used in the next iteration and after each iteration, the missing region shrinks. The algorithm terminates when all the missing pixels are filled.

(v) Repeat above steps for all target pixels in each sub-band in the current level until all the missing pixels are filled.

2.4 Inverse dual-tree complex wavelet transform (IDTCWT)

To make the proposed method robust, the wavelet images are inpainted by patch size having 7×7 , 9×9 and 11×11 and also 11×11 rotated by 180° to allow the filling order. By this, four inpainted LR images are obtained. Then IDTCWT is applied to reconstruct the original image.

2.5 Image fusion

The obtained LR images are combined to get one LR image using the fusion technique based on variance calculated in DCT domain [31]. The DCT coefficients are calculated for image blocks of size 8×8 pixels.

2.6 Fuzzy enhancement

The fuzzy logic-based histogram equalisation [32] is used to enhance the image. This method preserves image brightness and improves the local contrast of the original image.

2.7 Super-resolution

SR is a technique to reconstruct the HR image using LR images. The high-frequency content lost during the image acquisition process has to be recovered in the SR techniques. From under-sampled LR observations, the primary concern of the SR algorithm is to reconstruct HR images, and it produces high-quality images from blurred, noisy and degraded images.

The proposed iterated back-projection (IBP) based resolution enhancement algorithm is primarily inspired by the recent work of Deshpande *et al.* [33, 34]. The modified IBP algorithm is as below.

(i) LR fused image is divided into patches of $7 \times 7 L_{lr}$. Initialise iteration.

(ii) The LR image patch is up-sampled by the up-sampling factor f to obtain U_r using Bicubic approach.

$$U_r = \uparrow_u L_{lr} \quad (4)$$

Table 1 Input video sequences

Category	Video	Frames	Frames size	Description
baseline	Highway	400	320×240	videos without camera motion
	Pets	800	720×576	
	Office	200	360×240	
	Pedestrians	500	320×240	
bad weather	Blizzard	800	720×576	videos captured in bad weather
	Snowfall	400	360×240	
dynamic background	Fountain02	300	320×240	dynamic background motion with complex textures (e.g. water)
	Canoe	300	320×240	
	Fall	200	360×240	
	Overpass	200	320×240	
general	Akiyo	300	960×720	Lena and Akiyo videos are manually degraded by removing a few pixels from each frame
	Lena	300	960×720	



Fig. 3 Input videos (first row) and output videos (second row)

(iii) Initialise iteration index $i = 1$.

$$P^g = \frac{1}{d} (U_r^{i*} \text{Point Spread Function}) \quad (5)$$

Table 2 Influence of various wavelet transforms on the proposed framework

Video	Algorithm	PSNR	SSIM	VIFP	Time, s
Lena	DWT	36.83	0.863	0.861	42
	NSCT	38.26	0.901	0.869	56
	NSST	37.91	0.894	0.856	63
	DTCWT	38.88	0.921	0.883	47
Akiyo	DWT	36.74	0.879	0.863	49
	NSCT	38.1	0.884	0.879	61
	NSST	37.81	0.882	0.871	67
	DTCWT	38.56	0.898	0.894	52
Snowfall	DWT	33.41	0.763	0.755	66
	NSCT	34.67	0.814	0.795	79
	NSST	36.19	0.856	0.822	84
	DTCWT	37.55	0.901	0.867	52
Blizzard	DWT	31.63	0.746	0.719	52
	NSCT	34.11	0.791	0.766	63
	NSST	35.29	0.825	0.811	76
	DTCWT	36.44	0.894	0.862	41

Bold indicates the DTCWT gives better results compared to other transforms.

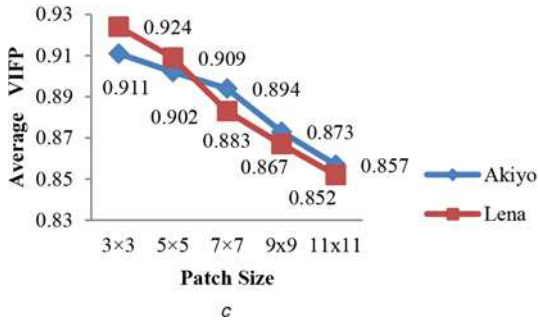
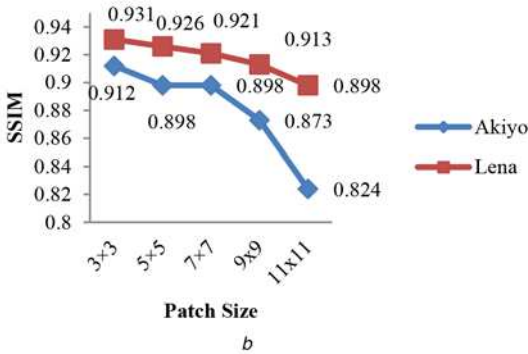
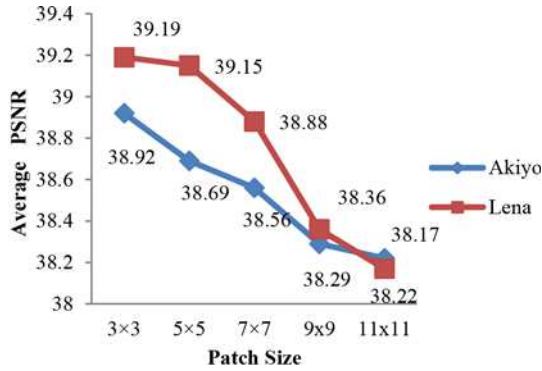


Fig. 4 Performance analysis of patch size

a PSNR versus patch size
b SSIM versus patch size
c VIFP versus patch size

Apply Lanczos3 low-pass filtering to remove high-frequency components to avoid the aliasing effect.

Point spread function at position (x, y) is given by

$$\text{PSF}_{x,y,\sigma} = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}. \quad (6)$$

where $\sigma = 0.5$
(iv)

$$E_{CF}^g = \|L_{lr} - P^g\| \quad (7)$$

where error E_{CF}^g

If $E_{CF} < \text{maximum error } \|E_{CF}^g\|_1$ then

$$U_r^{g+1} = U_r^g + \text{GP}^* \uparrow (E_{CF}^g) \quad (8)$$

where GP is the Laplacian of Gaussian back-projection matrix

$$\text{BM} = \begin{bmatrix} 1 & -2 & 1 \\ 2 & 8 & 2 \\ 1 & -2 & 1 \end{bmatrix}$$

else- U_r^g .

Set maximum error = E_{CF} .

(v) Repeat steps (iii)–(v) for all video frames.

(vi) At the last iteration, the output is the enhanced image.

(vii) Enhance the patch using a high-pass filter.

The steps (i)–(vii) are applied for all the patches of LR images. The super-resolved patches are stitched to get the super-resolved image.

3 Results

The proposed work is implemented using Matlab2012 on Intel Core i3 machine with 1.8 GHz processor speed and 4 GB RAM. The performance of the proposed framework is evaluated for two applications:

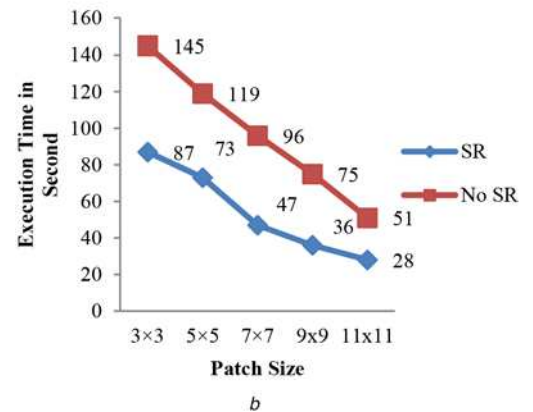
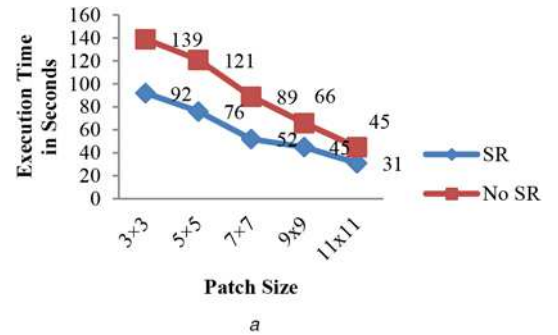


Fig. 5 Average execution time analysis for

a Akiyo video
b Lena video

object removal and error concealment. In object removal, the missing area is in arbitrary shape and in loss concealment; the missing area is square or rectangle shape and arbitrary shape.

3.1 Object removal

The performance of the proposed algorithm is tested on 12 video sequences having 4700 frames. These video sequences are available in the dataset [35] composed of 6 groups of 31 real-world videos having more than 70,000 frames. The reference image, in the form of binary maps, is provided for all these videos to indicate where the changes occur. The input videos which are used to evaluate the performance of the proposed framework are as shown in Table 1. Due to space limitation, five samples of the input frame and output frame are shown in Fig. 3.

For quality analysis, a new object, as a mask, is inserted into original frames and these frames are inpainted by using the proposed framework. Then the quality of the inpainted frames is analysed using [36–39] for peak-signal-to-noise ratio (PSNR), structural similarity index matrix (SSIM) and visual information fidelity in pixel domain (VIFP). The analysis of the influence of DWT, non-sub-sampled contourlet transform (NSCT), non-subsampled Shearlet transform (NSST) and DTCWT on the proposed framework is shown in Table 2. From the analysis, as shown in Table 2, DCTWT performs well compared to other algorithms. So, DTCWT is used in inpainting framework.

3.2 Super-resolution

The SR plays a significant role in reducing the execution time of the framework for inpainting the video sequences. The LR images are divided into small patches of size 7×7 . This is decided by analysing the performance of various patch sizes on two video sequences Lena and Akiyo. The quality analysis for various patch sizes is as shown in Fig. 4. The execution time analysis is carried out on super-resolved and not super-images, as shown in Fig. 5. From Figs. 4 and 5, it can be seen that the quality of the inpainted image is better in small path sizes (3×3 and 5×5) than other patches. In this experiment, 7×7 patch size is selected as more execution time is required for small size patches, and there is a very minimal amount of change in image quality. Fig. 5 shows that, the use of SR approach reduces the time required to inpaint the frames.

3.3 Error concealment

The performance of the proposed framework is further analysed in a context of error concealment. In this test, four videos are considered. The state-of-the-art algorithms of Giryes and Elad [40], Zarif *et al.* [41], Newson *et al.* [20], Janardhana Rao *et al.* [21] and Yang Li *et al.* [22] are used as the comparison baseline, as shown in Table 3. As mentioned in Table 3, three loss rates are considered, viz., 10, 25 and 45% for each video sequence. This loss of pixels is manually done.

Table 3 Analysis of error concealment application

Loss of pixel, %	Video	Quality	Giryes and Elad [40]	Zarif <i>et al.</i> [41]	Newson <i>et al.</i> [20]	Janardhana Rao <i>et al.</i> [21]	Yang Li <i>et al.</i> [22]	Proposed
10	Akiyo	PSNR (dB)	36.17	36.34	36.81	37.25	38.23	38.44
		SSIM	0.849	0.863	0.877	0.883	0.887	0.896
		VIFP	0.795	0.819	0.842	0.859	0.871	0.892
		time (s)	91	74	89	67	53	44
		PSNR (dB)	34.29	35.83	36.05	36.62	37.28	37.71
25		SSIM	0.754	0.784	0.803	0.811	0.822	0.829
		VIFP	0.673	0.718	0.731	0.788	0.797	0.804
		time (s)	156	128	91	88	74	62
		PSNR (dB)	28.066	30.339	30.59	30.66	31.37	31.6
		SSIM	0.615	0.657	0.688	0.693	0.712	0.726
45		VIFP	0.594	0.616	0.638	0.671	0.688	0.705
		time (s)	229	196	133	117	97	78
10	Lena	PSNR (dB)	35.26	36.83	36.92	37.55	38.43	38.79
		SSIM	0.855	0.872	0.884	0.896	0.904	0.913
		VIFP	0.846	0.858	0.859	0.870	0.876	0.879
		time (s)	82	65	62	55	46	41
		PSNR (dB)	34.51	35.29	35.88	36.19	36.37	36.64
25		SSIM	0.716	0.781	0.793	0.805	0.813	0.829
		VIFP	0.753	0.769	0.788	0.792	0.805	0.816
		time (s)	167	143	89	78	63	58
		PSNR (dB)	29.13	32.06	32.51	32.76	33.02	33.19
		SSIM	0.624	0.693	0.708	0.712	0.718	0.725
45		VIFP	0.517	0.619	0.639	0.697	0.715	0.734
		time (s)	264	218	189	115	97	89
10	Snowfall	PSNR (dB)	35.82	36.3	36.59	36.87	37.14	37.55
		SSIM	0.812	0.83	0.851	0.876	0.892	0.901
		VIFP	0.794	0.81	0.826	0.830	0.849	0.867
		time (s)	91	87	81	68	56	52
		PSNR (dB)	33.88	34.7	34.88	35.05	35.11	35.19
25		SSIM	0.658	0.72	0.753	0.781	0.795	0.811
		VIFP	0.711	0.73	0.755	0.762	0.769	0.766
		time (s)	168	145	115	89	72	66
		PSNR (dB)	27.98	31.6	31.73	31.95	32.01	32.08
		SSIM	0.607	0.63	0.672	0.694	0.701	0.706
40		VIFP	0.535	0.59	0.607	0.613	0.619	0.627
		time (s)	244	205	181	152	134	118
10	Blizzard	PSNR (dB)	34.18	35.7	35.76	35.93	36.21	36.44
		SSIM	0.829	0.87	0.881	0.886	0.892	0.894
		VIFP	0.767	0.82	0.839	0.854	0.859	0.862
		time (s)	85	73	65	59	48	41
		PSNR (dB)	32.49	33.7	33.91	34.16	34.52	34.98
25		SSIM	0.764	0.79	0.804	0.811	0.817	0.825
		VIFP	0.709	0.74	0.749	0.751	0.759	0.755
		time (s)	155	131	115	92	79	72
		PSNR (dB)	28.33	30.2	30.57	30.93	31.29	31.55
		SSIM	0.589	0.61	0.654	0.672	0.681	0.688
45		VIFP	0.511	0.54	0.598	0.601	0.603	0.604
		time (s)	221	189	152	113	101	95

Table 4 Average execution time per frame

Video	Frames	Frames size	Loss of pixels, %	Time, s
Highway	400	320 × 240	15	32
Pets	800	720 × 576	24	68
Office	200	360 × 240	16	37
Pedestrians	500	320 × 240	14	29
Blizzard	800	720 × 576	17	54
Snowfall	400	360 × 240	14	56
Fountain02	300	320 × 240	15	55
Canoe	300	320 × 240	12	47
Fall	200	360 × 240	16	56
Overpass	200	320 × 240	17	53
Akiyo	300	960 × 720	11	44
Lena	300	960 × 720	09	39

From Table 3, it can be seen that as the loss of pixels increases, the quality of the inpainted frame decreases. However, compared to state-of-the-art algorithm, the proposed framework performs well for the increase in loss of pixels. The average execution time required to inpaint the sample video sequences is given in Table 4.

From Table 4, it can be seen that the execution time depends on frame resolution and number of missing pixels. The execution time is more for more missing pixels. For Lena and Akiyo videos missing pixels are very less compared to Pets and Blizzard videos and hence, the time required to inpaint the Lena and Akiyo videos is less compared to other video sequences.

4 Conclusion

A novel video inpainting framework is proposed to inpaint the video sequences. The framework consists of inpainting, AR and SR methods. In the proposed novel work, inpainting algorithm is applied on sub-bands of images, which are decomposed using DTCWT. This framework uses the AR model to avoid the discontinuity which arises due to variation in brightness or contrast of patches. This framework removes seam in the target region that arises due to variation in brightness or contrast of patches. The fuzzy logic-based histogram equalisation is used to further enhance the image by preserving the image brightness and improve the local contrast. The performance of the proposed framework is evaluated for object removal and error concealment applications. The framework is compared with existing state-of-the-art algorithms. It can be concluded that by using the proposed approach, the video frames are inpainted without much loss of information and at a minimal amount of execution time.

5 References

- [1] Guillemot, C., Le Meur, O.: 'Image inpainting: overview and recent advances', *IEEE Signal Process. Mag.*, 2014, **31**, (1), pp. 127–144
- [2] Barnes, C., Shechtman, E., Finkelstein, A., et al.: 'Patch match: a randomized correspondence algorithm for structural image editing', *ACM Trans. Graphics (Proc. SIGGRAPH)*, 2009, **28**, (3), pp. 24:1–24:11
- [3] Glasner, D., Bagon, S., Irani, M.: 'Super-resolution from a single image'. IEEE 12th Int. Conf. on Computer Vision, Kyoto, 2009, pp. 349–356
- [4] Bertalmio, M., Sapiro, G., Caselles, V., et al.: 'Image inpainting'. Proc. of the ACM SIGGRAPH Conf. on Computer Graphics (SIGGRAPH 2000), New Orleans, LA, USA, July 2000, pp. 417–424
- [5] Chan, T., Shen, J.: 'Mathematical models of local non-texture inpaintings', *SIAM J. Appl. Math.*, 2002, **62**, (3), pp. 1019–1043
- [6] Chan, T., Shen, J.: 'Non-texture inpainting by curvature-driven diffusions (CDD)', *J. Vis. Commun. Image Represent.*, 2001, **12**, (4), pp. 436–449
- [7] Granados, M., Kim, K.I., Tompkin, J., et al.: 'Background inpainting for videos with dynamic objects and a free-moving camera'. Proc. of the 12th European Conf. on Computer Vision (ECCV '12), Florence, Italy, 2012, pp. 682–695
- [8] Liu, Y., Caselles, V.: 'Exemplar-based image inpainting using multiscale graph cuts', *IEEE Trans. Image Process.*, 2013, **22**, (5), pp. 1699–1711
- [9] Xia, A., Gui, Y., Yao, L., et al.: 'Exemplar-based object removal in video using GMM'. Int. Conf. on Multimedia and Signal Processing, Guilin, Guangxi, 2011, pp. 366–370

- [10] Ghanbari, A., Soryani, M.: 'Contour-based video inpainting'. 7th Iranian Conf. on Machine Vision and Image Processing, Tehran, 2011, pp. 1–5
- [11] Joshi, M.V., Chaudhuri, S., Panuganti, R.: 'Super-resolution imaging: use of zoom as a cue', *Elsevier Image Vis. Comput.*, 2004, **22**, (14), pp. 1185–1196
- [12] Deshpande, A., Patavardhan, P.: 'Super resolution of long range captured multiframe iris polar images', *IET Biometrics*, 2017, **6**, (5), pp. 360–368
- [13] Deshpande, A., Patavardhan, P.: 'Multiframe super-resolution for long range captured iris polar image', *IET Biometrics*, 2017, **6**, (2), pp. 108–116
- [14] Rudin, L.I., Osher, S., Fatemi, E.: 'Nonlinear total variation based noise removal algorithms', *Phys. D. Nonlinear Phenom.*, 1992, **60**, pp. 259–268
- [15] Sun, L., Jeon, B., Zheng, Y., et al.: 'A novel weighted cross total variation method for hyperspectral image mixed denoising', *IEEE Access*, 2017, **5**, pp. 27172–27188
- [16] Amrani, N., Serra-Sagrà, J., Peter, P., et al.: 'Diffusion-based inpainting for coding remote-sensing data', *IEEE Geosci. Remote Sens. Lett.*, 2017, **14**, (8), pp. 1203–1207
- [17] Newson, A., Almansa, A., Fradet, M., et al.: 'Towards fast, generic video inpainting'. Proc. of the 10th European Conf. on Visual Media Production, London, 2013, pp. 7:1–7:8
- [18] Newson, A., Almansa, A., Fradet, M., et al.: 'Video inpainting of complex scenes', *J. Imaging Sci.*, 2014, **7**, (4), pp. 1993–2019
- [19] Huang, J., Tang, X.: 'A fast video inpainting algorithm based on state matching'. 9th Int. Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), Datong, 2016, pp. 114–118
- [20] Newson, A., Almansa, A., Gousseau, Y., et al.: 'Non-local patch-based image inpainting', *Image Process. Line*, 2017, **7**, pp. 373–385
- [21] Janardhana Rao, B., Chakrapani, Y., Srinivas Kumar, S.: 'Image inpainting method with improved patch priority and patch selection', *IETE J. Educ.*, 2018, **59**, (1), pp. 26–34
- [22] Li, Y., Jiang, B., Lu, Y., et al.: 'Fine-grained adversarial image inpainting with super resolution'. Int. Joint Conf. on Neural Networks (IJCNN), Budapest, Hungary, 2019, pp. 1–8
- [23] Jun Kim, Y., Hazra, D., Byun, Y., et al.: 'Old document restoration using super resolution GAN and semantic image inpainting'. Proc. of the Int. Workshop on Artificial Intelligence and Education (WAIE 2019), Singapore, Singapore, 2019, pp. 34–38
- [24] Hui, Z., Li, J., Wang, X., et al.: 'Image fine-grained inpainting', arXiv preprint arXiv:2002.02609, 2020
- [25] Criminisi, A., Pérez, P., Toyama, K.: 'Region filling and object removal by exemplar-based image inpainting', *IEEE Trans. Image Process.*, 2004, **13**, (9), pp. 1200–1212
- [26] Le Meur, O., Ebdelli, M., Guillemot, C.: 'Hierarchical super-resolution-based inpainting', *IEEE Trans. Image Process. Instit. Electr. Electron. Eng.*, 2013, **22**, (10), pp. 3779–3790
- [27] Kingsbury, N.: 'A dual-tree complex wavelet transform with improved orthogonality and symmetry properties'. Proc. 2000 Int. Conf. on Image Processing (Cat.No.00CH37101), Vancouver, September 2000, pp. 375–378
- [28] Patil, S., Talbar, S.: 'Multiresolution analysis using complex wavelet and curvelet features for CBIR', *Int. J. Comput. Appl.*, 2012, **47**, (17), pp. 6–10
- [29] Di Zenzo, S.: 'A note on the gradient of a multi-image', *Comput. Vis. Graphics Image Process.*, 1986, **33**, (1), pp. 116–125
- [30] Pérez, P., Gangnet, M., Blake, A.: 'Poisson image editing', *ACM Trans. Graph.*, 2003, **22**, (3), pp. 313–318
- [31] Haghighat, M., Aghagholzadeh, A., Seyedarabi, H.: 'Multi-focus image fusion for visual sensor networks in DCT domain', *Comput. Electr. Eng.*, 2011, **37**, pp. 789–797
- [32] Magudeeswaran, V., Ravichandran, C.G.: 'Fuzzy logic-based histogram equalization for image contrast enhancement', *Math. Probl. Eng.*, 2013, **2013**, pp. 1–10
- [33] Deshpande, A., Patavardhan, P.: 'Single frame super resolution of non-cooperative iris images', *ICTACT J. Image Video Process.*, 2016, **7**, (2), pp. 1362–1365
- [34] Deshpande, A., Patavardhan, P., Rao, D.H.: 'Iterated back projection based super-resolution for iris feature extraction', *Elsevier Procedia Comput. Sci.*, 2015, **48**, pp. 269–275
- [35] Available at <http://jacarini.dinf.usherbrooke.ca/dataset2014/> (Last accessed on October 2018)
- [36] Wang, Z., Bovik, A.C., Sheikh, H.R., et al.: 'Image quality assessment: from error visibility to structural similarity', *IEEE Trans. Image Process.*, 2004, **13**, (4), pp. 600–612
- [37] Sheikh, H.R., Bovik, A.C., de Veciana, G.: 'An information fidelity criterion for image quality assessment using natural scene statistics', *IEEE Trans. Image Process.*, 2005, **14**, pp. 2117–2128
- [38] Lukes, T., Fliegel, K., Klima, M., et al.: 'Performance evaluation of image quality metrics with respect to their use for super-resolution enhancement'. IEEE Fifth Int. Workshop on Quality of Multimedia Experience (QoMEX), Klagenfurt, July 2013, pp. 42–43
- [39] Zhou, X., Bhanu, B.: 'Evaluating the quality of super-resolved images for face recognition'. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition Workshops, Anchorage, AK, June 2008, pp. 1–8
- [40] Giryes, R., Elad, M.: 'Sparsity based Poisson inpainting'. IEEE Int. Conf. on Image Processing (ICIP), Paris, October 2014, pp. 2839–2843
- [41] Zarif, S., Faye, I., Rambli, D.R.A.: 'Video completion method for static object removal based on GLCM'. IEEE 11th Int. Colloquium on Signal Processing & Its Applications (CSPA), Kuala Lumpur, March 2015, pp. 57–62