Automatic Shadow Removal Algorithm for VOP, DWT Based Watermarking Algorithm for VOP and Generation of Super Resolved VOP

Alwyn R. Pais[1], John D'Souza[2], Ram Mohana Reddy[3], P.Hari Krishna[4]
[1] *Dept. of CS&E, alwyn.pais@gmail.com*
[2] *Dept. of E&C, john_krec@yahoo.com*
[3] *Dept. of IT, profgrmreddy@gmail.com*
[4] *Dept. of CS&E, harikrishnarudra@gmail.com*
National Institute of Technology Karnataka, Surathkal, INDIA

Abstract

Removal of shadow from Video Object Planes (VOPs) will assist in surveillance applications for comprehensive detection of activities. We have proposed a method for removal of shadows from the VOP. Also noise removal is done using existing methods from the VOP. To authenticate the surveillance VOP, digital watermarking is used. We have proposed digital watermarking using localized Biorthogonal wavelets for VOP. Super-resolved VOP is generated using multi-frame method. Edge model based super resolution method is used to get the better results. Also the effect of digital watermarking is studied for the super-resolved VOP. A number of test cases have been proposed and found out a best method for video surveillance application. Our proposed super resolution (SR) method gives better results than bilinear and bi-cubic methods.

Keywords: *Noise, Shadow removal, Super resolution, VOP, Watermarking.*

1. Introduction

Computer Vision is concerned with the theory of building artificial systems that obtain information from images. The image data can have many forms such as video sequence, views from multiple cameras or multi-dimensional data from a medical scanner. There are many characteristics of computer vision of which are considered to be relevant are image processing/image analysis which focuses on 2D images and their transformations, imaging which focuses on producing images, focus on 3D images, machine vision focuses on manufacturing applications and pattern recognition. Typical tasks of computer vision are recognition, motion, scene reconstruction, image restoration, image acquisition, pre-processing, feature extraction, detection and high-level processing. Organization of computer vision systems is application dependent. Video Surveillance is an important application of Computer Vision for an organization, from the security point of view. This is one application where an automated system can replace human beings, as well as have inputs which are not possible through human surveillance alone.

A shadow occurs when an object partially or totally occludes direct light from a source of illumination. In general, shadows can be divided into two major classes: self shadows and cast shadows. A self shadow occurs in the portion of an object which is not illuminated by direct light. A cast shadow is the area projected by the object in the direction of direct light.

Cast shadows can be further classified into umbra and penumbra region, which is a result of multi-lighting and self shadows also have many sub-regions such as shading and inter reflection. Usually, the self shadows are vague shadows and do not have clear boundaries. On the other hand, cast shadows are hard shadows and always have a violent contrast to background. Because of these different properties, algorithms to handle these two kinds of shadows are different.

Detection and tracking is vital to many applications dealing with image sequences such as video surveillance. However, an unexpected shadow causes difficulties such as object merging, shape distortion and even loss of object [1]. The difficulties in shadow detection arise because they share two important visual features. First, since shadows typically differ from background, shadow points are detected as foreground objects. Second, shadows move in same motion as their objects casting them [1]. For these reasons, shadow detection and removal in images/videos is serious problem and has become an active research area.

Any foreground object segmented out from a video which have a semantic meaning is called Video Object plane (VOP). An innate problem of VOP generation is that objects of interest are not homogeneous with respect to low level features such as color, intensity, texture and optical flow. Hence, conventional low-level segmentation algorithms will fail to extract required meaningful partitions. Hence automatic segmentation of VOPs is active research topic which has lot of challenging problems like shape changing, rotations, intra object motion, rigid parts, cluttered background, moving background and eliminating shadows etc., [2].

The unwanted information penetrated into image which contaminates the image details is referred as noise. Most well known impulsive noise is “Salt” and “Pepper” noise which is caused due to bit errors induced in transmission, dead pixels, and analog to digital conversion [3].

Digital media is transmitted through unstructured networks which makes vulnerable to many attacks. Data hiding is required to protect the digital media content from attacks as the copyright owners compensated every time their work is used [4]. Watermarking is a technique of data hiding which embeds data into a digital signal.

With the development of modern technology, more attention has been shifted from tracking and detection toward spatial resolution for object recognition, mapping and image capabilities [5]. Super resolution (SR) technique refers to generation of high resolution (HR) image from low resolution (LR) image by adding some additional details of the image [6]. In tracking, the objects segmented out needs to observe comprehensively for which, mere zooming causes the object to blur and loose its features. Hence, super resolution is required to enhance the image details.

In this paper we have presented an algorithm for shadow removal in VOP based on [7] and a digital watermarking technique for VOP based on localized Biorthogonal wavelets. In our work, we have implemented a system for super resolution of the watermarked VOP sequences after object is tracked(without shadow and noise). Further sections of the paper are organized as follows. Related work is explained in section 2. Section 3 discuss about the proposed system. Experimental results and qualitative analysis of different test cases are given in section 4. We conclude the paper in Section 5.

2. Related Works

2.1 Shadow Removal

A large number of shadow detection and removal algorithms have been proposed in past literature. The different approaches proposed can be classified as blob tracking, active contour tracking, Markov random field tracking, color and pattern based tracking, luminance information, gradient density information, color information, texture and intensity reduction, texture attributes, image block statistics [8] and RGB chroma model [9]. A number of approaches and taxonomy are also presented for shadow detection and removal [1].The background suppression [10] approach requires a computationally expensive background update, but is more general and thus we focus on it in the following. Our paper presents a gradient based subtraction method focused to eliminate shadows from VOP.

2.2 VOP

Many segmentation algorithms exist in the literature. Depending upon the processing, they are classified into 3 types as spatial, temporal and spatiotemporal segmentation. We have used our own segmentation algorithm presented in [2] to generate VOP. In this algorithm, first the background model is extracted using initial frames. Frame differences between two consecutive frames are calculated by block matching method to track the objects in motion. The blocks which changed in current frame compared with preceding and succeeding frames are considered as changed blocks and are tracked. Similarly, rigid parts are also tracked by using extracted background model and using average threshold method between blocks.

Quantization and boundary detection of the frames is the final step. For selection of boundaries of VOPs, a quantization technique is applied on the blocks based on peer group filtering [11]. A pixel in each block is used to quantize by computing Euclidean distances among all pixels of the block with respect to selected pixel. By combining all the edges tracked after the quantization process, boundaries of the object can be established. By this method, shadows are not detected as the algorithm works on motion of foreground objects with respect to background since shadows have same motion as their objects casting them. We have used this method to generate VOPs.

2.3 Filters

There are many filters proposed to remove noise from images in the past literature. In [12], the authors used fourth order partial differential equations for image enhancement. To remove salt and pepper noise, a method proposed based on histogram distributions and soft decision to distinguish impulsive noise from legitimate pixels [13]. In [14], the authors compared different types of smoothing filters for impulsive noises incurred through transmission of a digital image such as mean filter, simple median filter, component median filter, vector median filter, spatial median filter and modified spatial median filter. We have used component median filter (CMF) here to remove salt and pepper noise.

2.4 Watermarking

Watermarking is necessary for authentication. Currently digital watermarking is divided into two categories by embedding position – spatial domain and transform domain [15]. The various techniques of spatial domain category are least significant bit (LSB) modification

[16], patchwork, texture block coding [15] etc. Watermarking technology shifted to transform domain in recent years since transform domain is more robust and compatible to popular image compression standards. Possible methods of frequency transformation include Discrete Fourier transform (DFT), discrete cosine transforms (DCT) and discrete wavelet transform (DWT). We have proposed a DWT watermarking technique as it provides high robustness and can easily applicable to true color images.

2.5 Super Resolution

Generation of high resolution images from low resolution images can be achieved by reconstruction and learning based methods [6]. Multiple low resolution images of the same scene are used to extract certain information from each image and combine them resulting in a single high resolution image. It is necessary that each image has some extra information and achieved through sub-pixel shifting. Frequency domain approach is one of the methods used for reconstruction [17]. On the other hand, learning based approaches maintains a relation between LR and HR image. They rely on learning of characteristics of a specific image set to gather extra information required for HR generation. There are many super-resolution techniques and algorithms. Some uses kernel based interpolation techniques like bi-cubic, bilinear and Sub-pixel edge localization. In [6], we proposed a novel reconstruction approach of super resolution based on multiple images. In [2], we proposed another method for HR image generation based on edge modeling and laplacian sub band. We have used both methods in this system to create high resolution shadow removed VOP.

3. Proposed Methodology

The proposed system consists of four modules viz., denoising of VOP, Shadow removal from VOP, Watermarking and Super-resolution by Multiframe method as shown in Fig. 1. We will discuss these modules in subsequent subsections.

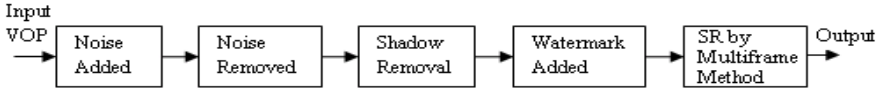


Fig. 1. Proposed System

3.1 Noise Removal

To remove "Salt and Pepper" noise from image sequences, we have used component median filter (CMF) specified in [14] as it preserves image details accurately. A filter mask (usually of size 3X3 or 5X5) is applied over a point to be studied. The component median filter finds medians for each component of RGB color space under the mask and combined to form a new point, used to replace the pixel studied.

$$CMF(X_1,.,X_N) = [MED(X_{1R},.,X_{NR})\ MED(X_{1G},..,X_{NG})\ MED(X_{1B},..,X_{NB})\ ]^T \quad (1)$$

Where $X_1,\ldots,X_N$ are the pixels which fall under filter mask of size NxN. $MED(X_{1i},..,X_{Ni})$ indicates median of pixels and 'i' denotes R, G and B color components. It is better than all other filters explained in [14] as it works on each component of RGB color space. We have imparted "salt and pepper" noise level of 5 percent to the image sequences.

Fig. 2(a) and fig. 2(b) shows original video frame and VOP generated respectively. Fig. 2(c) and fig. 2(d) shows the VOP frames after noise added and removed.

3.2 Shadow Removal

In this section, we present Shadow detection and removal algorithm for VOP based on gradient based background subtraction (GBS) and contours adopted using [7]. The algorithm is presented in Table.1.

Table.1. Algorithm: Shadow Removal of VOP

1. 1. *Contour Generation*: The contours of the foreground objects of the VOP without shadow are generated by using Gradient Based Background Subtraction. 2. 2. *Closure of Contours*: The contours generated are needed to be closed and this can be done by a closure operation. 3. 3. *Detection of boundaries of blobs*: Subtract the closed contours from foreground objects to result in a number of blobs which consist of shadow regions. The boundaries of each blob are determined. 4. 4. *Shadow blobs detection and removal*: The shadow regions in the blobs are detected using neighborhood ratio calculation and are removed from foreground objects of VOP.

3.2.1 Contour Generation: Separation of shadow blobs and foreground objects is a very difficult task as they are well connected and no other image details are available. For detection of contours, we pass VOP through Gaussian mixture model (GMM) based subtraction and gradient based subtraction (GBS) [18]. We have used Gaussian mixture model presented by Stauffer and Grimson [19] to generate foreground pixels as shown in fig. 2 (e). In GMM model, the recent history of each pixel is maintained using 'k' (usually 3 to 5) Gaussian distributions which are associated with attributes like weight, mean and variance. According to [18], GBS depends on GMM based background subtraction attributes. In the GBS, number of Gaussians models each pixel to find whether it belongs to contour or foreground. Gradient magnitude (D_m) and gradient direction (D_d) are the features of each pixel considered to be a vector Δ [D_m, D_d]. A distribution is needed for Δ to model the gradient of pixel intensities. The gradients are calculated by the gray values of the pixels and the distribution $F(D_m, D_d)$ is generated [18]. If for a certain gradient vector, the probability of being generated from background gradient distribution is less than some threshold, the pixel belongs to foreground which acts as the contour. Fig. 2(e) and fig. 2(f) shows the results of foreground objects of VOP and gradient based subtraction for VOP frame.

3.2.2 Closure of Generated Contours: On application of gradient based subtraction to VOP, it can't detect accurate contours as there are no background details available. The ineffective parameters and thresholding makes some part of the object contour eradication. So the output will be partial contours of the foreground object. To fill up the contours, a close operation can be performed with appropriate structuring element [20] as shown in fig. 2(g).

3.2.3 Detection of Boundaries of Blobs: The contour filled objects obtained in the above section are subtracted from the foreground object VOP to result in number of blobs which are identified using connected component analysis [20] (fig. 2(h)). The boundaries around each blob are detected with canny edge detection method [20] as shown in fig. 2(i).

3.2.4 Shadow Blobs Detection: For each blob generated from the above step, a neighborhood ratio [7] is calculated. The eight neighbors of each boundary pixel are checked if they overlap with the closed contour generated in section 3.2.2. If either of the pixels exists, a match count (M_c) is incremented. After all the boundary pixels of a particular blob are covered, a neighborhood ratio (N_r) is calculated by taking ratio of match count and total number of boundary pixels of that blob (Q).

$$N_r = M_c/Q \quad (2)$$

If N_r of a certain blob is less than some threshold T_H, then that blob is considered to be shadow region and discarded from foreground VOP. The process is continued for all blobs and are discarded which fall under threshold from the foreground VOP (fig 2(e)). A mask operation is applied on it with the true color VOP to get shadow removed VOP as shown in fig. 2(l).

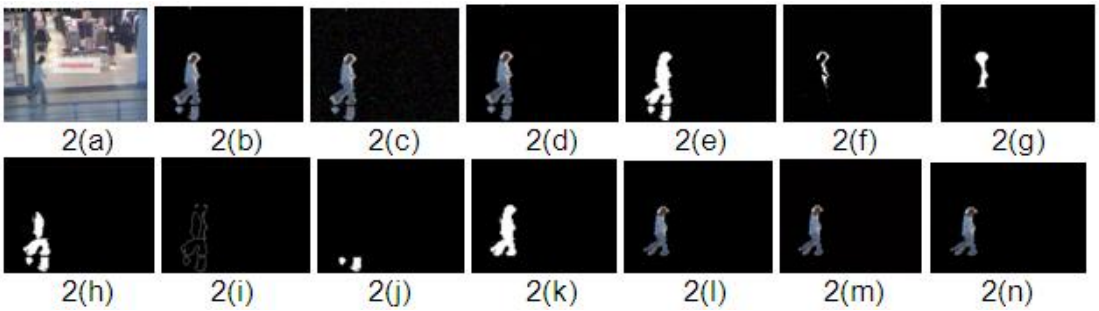


Fig. 2. Sequence of frames. 2(a) Original frame 74. 2(b) VOP. 2(c) Noise added. 2 (d) Noise removed. 2(e) Foreground pixels. 2(f) GBS result. 2(g) Contour closure. 2(h) Blobs identified. 2(i) Boundary detection. 2(j) Shadow detected. 2(k) Result image. 2(l) Masked output. 2(m) Watermark added. 2(n) HR image with multiframe method.

3.3 Digital Watermarking

A DWT based watermarking is proposed for VOP authentication based on [21 & 22]. As wavelet transformation is robust against simple image processing operations like low pass filtering, blurring and contrast enhancement and can also endure cropping operations. DWT transformation decomposes the image/frame into different frequency coefficients such as approximation coefficients (LL_1) and detail coefficients. The approximation coefficients are low frequency coefficients which contain largely image details. The other detail coefficients are given as vertical detail coefficients (LH_1), horizontal detail coefficients (HL_1) and diagonal detail coefficients (HH_1) as shown in fig. 3. The conventional wavelet watermarking techniques embeds the watermark in the components of the first level DWT. As VOP contains no other details except extracted foreground objects, we have used one level DWT decomposition to embed a watermark in the proposed technique.

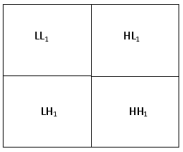


Fig. 3. First Level DWT Decomposition

In this paper, Digital Image Watermarking using Localized Biorthogonal wavelets for VOP is used. Biorthogonal wavelets have the properties of perfect reconstruction, symmetry and have higher embedding capacity if they are decomposed into different channels. The watermark embedded in the VOP and the extracted watermarks are shown in fig. 4(a) and 4(b) respectively. The watermark embedded VOP is shown in fig. 2(m).

Fig. 4(a). Watermark Inserted in VOP

Fig. 4(b).Watermark Extracted from Frame 74

3.4 Super-Resolution

The VOP is super resolved using multiframe edge model based superresolution technique.

3.4.1 SR Based on Multiple Frames: Super resolution in this method is based on multiple frames of sharpened edges. For this, meaningful segments are divided using Histogram based segmentation which does the task of dividing manageable frames by calculating the number of pixels of each gray level of consecutive frames. The difference of number of pixels D is calculated as

$$\sum_{k=1}^{n} | H_p(k) - H_c(k) | \tag{3}$$

Where $H_p(k)$ and H_c (k) denotes number of k^{th} gray level pixels of previous frame and current frame respectively. Rotation angle between current frame f_c and previous frame f_p is calculated by Planner Motion Estimation (angle Φ_m) and shift estimation i.e., horizontal (X) and vertical (Y) shifts are calculated [6] by registration process. The laplacian pyramid (L_1) images are generated by using symmetric residue pyramid (SRP) technique which in turns uses Gaussian pyramid images (Fig.5). Let G_0 be the original image and G_1 be the 1st level Gaussian pyramid image. G_1 is a result of applying a low-pass filter to G_0. L_0 is the error given by $L_0 = G_0 - G_1$.

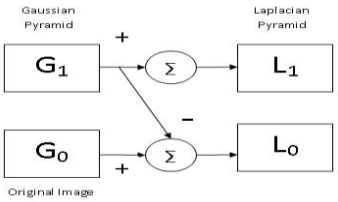


Fig. 5. Laplacian Generation Method

An iterative symmetric residue pyramid (SRP) algorithm is used to generate edge representation. By applying unsharp mask (UM) filter technique, sharpness of the laplacian images is enhanced. Converting RGB color space of images into YC_bC_r color space, the sharpened edge details are injected into Y color space of original frames and reconvert them to RGB color space. The HR image generated using this method is shown in fig. 2(n).

3.4.2 Algorithm:

1. Histogram based segmentation is used to group the similar frames together i.e, segmentation in temporal domain. Here similar frames mean frame related to same scene.
2. The above video segments are processed one by one. Two frames from a video segment are taken at a time - previous (reference) frame f_p and current frame f_c.
3. Rotation angle between current frame f_c and previous frame f_p is calculated by Planner Motion Estimation. Let's say angle is ϕ_m.
4. Rotate the image f_c by - ϕ_m to cancel the rotation.
5. Estimate shifts i.e, horizontal (X) and vertical (Y) shifts between current frame f_c and previous frame f_p.
6. Reconstruct the high resolution image (f_h) using current frame f_c ,previous frame f_p ,ϕ_m, X and Y.
   (a) Map the co-ordinates of each pixel of the current frame f_c to the previous frame fp using ϕ_m, X and Y.
   (b) Using these samples, interpolate the pixel values on High resolution grid.
7. Compute the edge representation (L-1) of the reference image fp by SR method.
8. Apply sharpness enhancement on the edge representation (L-1) by Unsharp method.
9. Inject the sharpened edge representation in the high resolution image (f_h) to get final HR image.
10. Repeat Steps 2-9 with fp = fc and fc = next frame for a video segment.
11. Repeat Steps 2-10 for all video segments generated by histogram based segmentation.

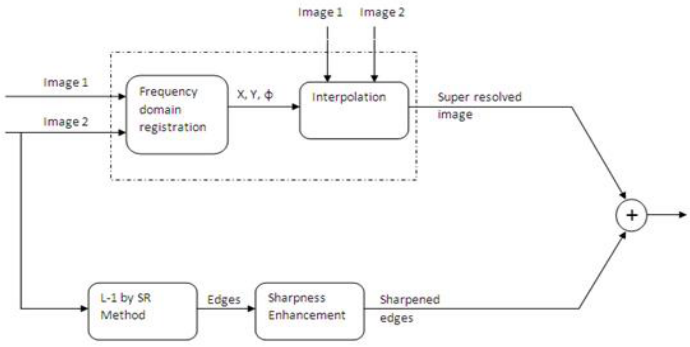


Fig. 6. Block Diagram of Process of Generating Super-resolved Image and Injecting Sharpened Edges.

3.4.3 Temporal Video Segmentation: It is the first step towards Super resolution of digital video sequences [9]. Its goal is to divide the video stream into a set of meaningful and manageable segments that are used as basic elements for SR algorithm. When two images are sufficiently dissimilar, there may be a cut. Based 19 on the metrics used to detect the difference between successive frames, the algorithms can be divided broadly into three categories: pixel, block based and histogram comparisons [9]. So to divide given input video to meaningful segments we can choose histogram segmentation which is having the advantage of reducing sensitivity to camera and object movements by comparing the histograms of successive images [9].

3.4.3.1 Histogram Based Video Segmentation: In the histogram based video segmentation, differences of the no of pixels of each gray level of consecutive frames are calculated. Let's

say histogram of a frame f_p is an n-dimensional vector $H_p(k)$, k=1,2,.. n, where n is the number of gray levels, $H_p(k)$ is the number of pixels from the frame f_p with gray level j and $H_c(k)$ is the number of pixels from the frame f_c with gray level k. If the absolute sum of histogram differences (D) between two successive frames is greater than a threshold T as given in [2], a cut is declared.

$$D = \sum_{k=1}^{n} | H_p(k) - H_c(k) |$$

3.4.4 Registration: As discussed earlier, if the image have sub-pixel shifts then each image cannot be obtained from the others, assuming each image has different shifts. New information is therefore contained in each low-resolution image, and it can be exploited to obtain a high-resolution image. To use of that unique information, first we need to apply a process called registration. Image registration is the process of transforming the different sets of data into one coordinate system. Registration is necessary in order to be able to integrate the data obtained from different samples. Relation between images is calculated during Image registration giving the motion vector information i.e., shifts and rotation as output. Once the motion vector information is obtained, all pixels from available frames are mapped back onto reference frame based on the motion vector information, to obtain an high-resolution frame. This process of mapping the pixels from different samples is called as reconstruction of an image. The motion can be described as a function of three parameters - horizontal shifts X, vertical shifts Y, and a planar rotation angle φ. A frequency domain approaches is used for estimating 20 these motion parameters between the reference image and the other image. A frequency domain approach allows estimation of the horizontal and vertical shift and the (planar) rotation separately. In this section, a planar motion estimation method is described for the registration images taken in sequence from a single obtained video segment from histogram based segmentation. As the Fourier transforms explicitly used, our method is limited to band limited signals described in the Fourier basis. Although this limits the applicability of such a method, it is a very reasonable assumption in practice. The optical system of a digital camera typically acts as a low-pass filter, and attenuates or blocks all high frequencies. The captured image will therefore be an essentially band-limited signal. This method computes the planar shift and rotation parameters between a pair of images[10].

3.4.4.1 Planar Motion Estimation: We use a frequency domain algorithm to estimate the motion parameters between the reference image and each of the other images. Only planar motion parallel to the image plane is allowed. Let's say the motion is described as a function of three parameters: horizontal and vertical shifts, Δx_1 and Δx_2, and a planar rotation angle φ. Fourier based image registration methods only allow global motion in a plane parallel to the image plane. A frequency domain approach allows us to estimate the horizontal and vertical shift and the (planar) rotation separately. Assume we have a continuous two-dimensional reference signal $f_1(x)$ and its shifted and rotated version $f_2(x)$:

$$f_2(\boldsymbol{x}) = f_1(\boldsymbol{R}(\boldsymbol{x} + \Delta\boldsymbol{x})),$$

$$\text{with } \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \Delta\boldsymbol{x} = \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \end{bmatrix}, \quad \boldsymbol{R} = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}$$

This can be expressed in Fourier domain as

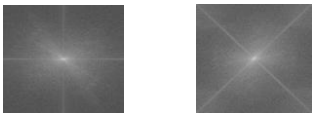
$$\begin{aligned} F_2(\boldsymbol{u}) &= \iint_{\boldsymbol{x}} f_2(\boldsymbol{x}) e^{-j2\pi \boldsymbol{u}^T \boldsymbol{x}} d\boldsymbol{x} \\ &= \iint_{\boldsymbol{x}} f_1(\boldsymbol{R}(\boldsymbol{x} + \Delta\boldsymbol{x})) e^{-j2\pi \boldsymbol{u}^T \boldsymbol{x}} d\boldsymbol{x} \\ &= e^{j2\pi \boldsymbol{u}^T \Delta\boldsymbol{x}} \iint_{\boldsymbol{x}'} f_1(\boldsymbol{R}\boldsymbol{x}') e^{-j2\pi \boldsymbol{u}^T \boldsymbol{x}'} d\boldsymbol{x}' \end{aligned}$$

with $F_2(u)$ the two-dimensional Fourier transform of $f_2(x)$ and the coordinate transformation $x' = x + \Delta x$. After another transformation $x'' = Rx'$, the relation between the amplitudes of the Fourier transforms can be computed as

$$\begin{aligned} |F_2(\boldsymbol{u})| &= \left| e^{j2\pi \boldsymbol{u}^T \Delta\boldsymbol{x}} \iint_{\boldsymbol{x}'} f_1(\boldsymbol{R}\boldsymbol{x}') e^{-j2\pi \boldsymbol{u}^T \boldsymbol{x}'} d\boldsymbol{x}' \right| \\ &= \left| \iint_{\boldsymbol{x}'} f_1(\boldsymbol{R}\boldsymbol{x}') e^{-j2\pi \boldsymbol{u}^T \boldsymbol{x}'} d\boldsymbol{x}' \right| \\ &= \left| \iint_{\boldsymbol{x}'} f_1(\boldsymbol{x}'') e^{-j2\pi \boldsymbol{u}^T (\boldsymbol{R}^T \boldsymbol{x}'')} d\boldsymbol{x}'' \right| \\ &= \left| \iint_{\boldsymbol{x}''} f_1(\boldsymbol{x}'') e^{-j2\pi (\boldsymbol{R}\boldsymbol{u})^T \boldsymbol{x}''} d\boldsymbol{x}'' \right| \\ &= |F_1(\boldsymbol{R}\boldsymbol{u})| \end{aligned}$$

We can see that $F_2(u)$ is a rotated version of $F_1(u)$ over the same angle φ as the spatial domain rotation (see Fig. 3). $F_1(u)$ and $F_2(u)$ do not depend on the shift values Δx, because the spatial domain shifts only affect the phase values of the Fourier transforms. Therefore we can first estimate the rotation angle φ from the amplitudes of the Fourier transforms $F_1(u)$ and $F_2(u)$. After compensation for the rotation, the shift 4 can be computed from the phase difference between $F_1(u)$ and $F_2(u)$.

3.4.4.2 Rotation Estimation: The rotation angle between $F_1(u)$ and $F_2(u)$ can be computed as the angle φ for which the Fourier transform of the reference image $F_1(u)$ and the rotated Fourier transform of the image to be registered $F_2(u)$ have maximum correlation. Fig.7 shows the effect of image rotation in Fourier domain.



(a) (b)

Fig.7. Shows the effect of image rotation on Fourier transform representation.(a) Fourier transform representation of original lena image.(b) Fourier transform representation of lena rotated image (φ = 45 degrees)

3.4.4.3 Shift Estimation: A shift of the image parallel to the image plane can be expressed in Fourier domain as a linear phase shift:

$$\begin{aligned} F_2(\boldsymbol{u}) &= \iint_{\boldsymbol{x}} f_2(\boldsymbol{x}) e^{-j2\pi \boldsymbol{u}^T \boldsymbol{x}} d\boldsymbol{x} = \iint_{\boldsymbol{x}} f_1(\boldsymbol{x} + \Delta\boldsymbol{x}) e^{-j2\pi \boldsymbol{u}^T \boldsymbol{x}} d\boldsymbol{x} \\ &= e^{j2\pi \boldsymbol{u}^T \Delta\boldsymbol{x}} \iint_{\boldsymbol{x}'} f_1(\boldsymbol{x}') e^{-j2\pi \boldsymbol{u}^T \boldsymbol{x}'} d\boldsymbol{x}' = e^{j2\pi \boldsymbol{u}^T \Delta\boldsymbol{x}} F_1(\boldsymbol{u}) \end{aligned}$$

It is well known that the shift parameters Δ can thus be computed as the slope of the phase difference ($F_2(u)/F_1(u)$). To make the solution less sensitive to noise, we fit a plane through the phase differences using a least squares method.

3.4.5 Super-Resolved Image Reconstruction: It involves generation of high resolution image by using multiple low resolution image and relation between them. Here relation means Rotation and Shift estimation between low resolution images. In this reconstruction algorithm, the samples of the different low-resolution images are first expressed in the coordinate frame of the reference image. Then, based on these known samples, the image values are interpolated on a regular high-resolution grid. We chose bi-cubic interpolation because of its low computational complexity and good results.

What is the optimal number of images to use when reconstructing a high-resolution image? The exact answer to this question depends on many parameters, such as the registration accuracy, imaging model, total frequency content, and so forth. Intuitively, two effects need to be balanced. On one hand, the more images there are, the better the reconstruction should be. On the other hand, there is a limit to the improvements that can be obtained: even from a very large number of very low-resolution images of a scene, it will not be possible to reconstruct a sharp, high-resolution image. Blur, noise, and inaccuracies in the signal model limit the increase in resolving power that can be obtained. In our case, the registration accuracy will limit the selection of number of input images. The optimal number of images used at a time to reconstruct the high resolution image is two in our case. High resolution image is reconstructed from two consecutive low resolution images in video sequence at a time. Coordinates of the pixels in the second frame are estimated using the registration parameters of the first frame. The high resolution image is obtained by interpolating these known samples to the high resolution grid.

3.4.6 Gaussian Pyramid: The Gaussian Pyramid is a hierarchy of low-pass filtered versions of the original image, such that successive levels correspond to lower frequencies. The low-pass filtering is done using convolution with a Gaussian filter kernel. The Gaussian pyramid on image I is defined as:

$G_0(x; y) = I$
$G_{i+1}(x; y) = REDUCE(G_i(x, y))$

The REDUCE operation is carried out by convolving the image with a Gaussian low pass filter.

Fig.8 shows the gaussian pyramid. The original image is repeatedly filtered and subsampled to generate the sequence of reduced resolution images G_1, G_2 etc [15].

Fig.8. The Gaussian Pyramid

Fig.9. The Laplacian Pyramid

3.4.7 Laplacian Pyramid: The Laplacian is then computed as the difference between the original image and the low pass filtered image.

As shown in the Fig.10, the Laplacian image is the result of the low pass filtered image subtracted from the original image. Let G_0 be the original image and G_1 be the 1st level Gaussian pyramid image. G_1 is a result of applying a low-pass filter to G_0.

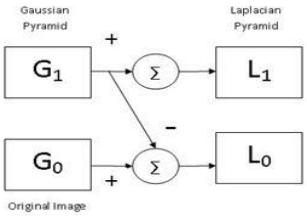


Fig. 10. Shows Laplacian Generation Method

The error L_0 is then given by $L_0 = G_0 - G_1$

The first subband L_0, captures all the sharp edges in the image. The Laplacian pyramid will give the edge representation as shown in Fig.9.

For generating an HR video frame, the sharp edges must be identified in the LR and their sharpness must be restored in the HR image. The current work makes use of recent results related to Laplacian pyramid to identify edges. The Laplacian pyramid, like all other multiresolution representations, creates a hierarchy of subbands encoding edges of decreasing sharpness [13][15].

3.4.8 Symmetric Residue Pyramid: Another method - Symmetric Residue Pyramids [4] was proposed as an extension to Burt Laplacian pyramids. This extension of the Laplacian pyramid exploits the redundancy present in the Laplacian pyramid to achieve better signal compaction and non redundant edge representation.

The process of generating a L_i given a G_i starts with an initial guess, which may even be a blank (zero) image. An iterative process is deployed to get one of the acceptable L_i: (exp is the expand/interpolation operation, ss is subsampling and lpf is low-pass filtering).

- $L_i[0]$ = Initial guess (may even be 0 image)
- $G_i[k] = \exp(G_i+1) + L_i[k]$ (usual pyramid reconstruction)
- $L_i[k + 1] = G_i[k] - \exp(ss(lpf\ (G_i[k])))$

The output $L_i[k + 1]$ will contain the detected edges. Here L_{-1} is nothing but a zoomed edge representation.

3.4.9 L_{-1} Generation: The first subband L_0 of Laplacian Pyramid, captures all the sharp edges in the image. But its resolution is same as that of input image. But for super-resolution, up-sampled edges are required. So according to Laplacian pyramid, we need to find the sub-band in reverse order which is L_{-1}. So the algorithm mentioned previous section for Symmetric residue pyramid becomes as follows -

- $L_{-1}[0]$ = Initial guess (may even be 0 image)
- G-1 [k] = exp(G0) + L-1[k] (usual pyramid reconstruction)
- L-1[k + 1] = G-1[k] - exp(ss(lpf (G-1[k])))

3.4.10 Sharpness Enhancement: The visual appearance of an image may be significantly improved by emphasizing its high frequency contents to enhance the edge and detail information in it. The classic unsharp masking (UM) technique is often employed for this purpose. From a signal-processing standpoint, an unsharp mask is generally a filter that amplifies high-frequency components.

An unsharp mask cannot create additional detail, but it can greatly enhance the appearance of detail by increasing small-scale acutance. Acutance describes how quickly image information transitions at an edge, and so high acutance results in sharp transitions and detail with clearly defined borders.

In image processing, deconvolution is the process of approximately inverting the process that caused an image to be blurred. While unsharp masking increases the apparent sharpness of an image in ignorance of the manner in which the image was acquired. For deconvolution to be effective, all variables in the image scene and capturing device need to be modelled, including aperture, focal length, distance to subject, lens and media refractive indices and geometries.

Therefore applying deconvolution successfully to general-purpose camera images is usually not feasible, since the geometries of the scene are not set. But for general purpose use, methods such as unsharp mask are usually sufficient, and generally cost-effective.

In our proposed method, the edges L_{-1} calculated in previous sections are sharpened by using unsharp masking (UM) technique. Let's say output of unsharp masking is $L`_{-1}$.

3.4.11 Injecting Edge Details in Image: Color image consists of 3 channels R,G and B. Any one of these channel may not describe the structure of the scene e.g. objects or edges. So R,G and B channels are not convenient for injecting the edges. So image is converted to another color space called YC_bC_r consisting of 3 channels Y, C_b and C_r. Y channel contains the structure of the scene e.g. objects and edges in gray scale. It is a weighted average of the all R, G and B color. But our eyes are more sensitive to green color so weight for green color is given more. C_b and C_r contains the blue and red color information. To modify the edges in an image, we can use only Y channel. Now sharpened edges $L`_{-1}$ needs to be added to the zoomed image I`. As I` can be a color image, sharpened edges $L`_{-1}$ can't be added directly to image. So first I_0 is converted to YC_bC_r color space from RGB color space. Y channel of YC_bC_r color space is used to add edges. After adding edges, image is converted back to RGB color space.

4. Experimental Results

The VOP sequences are generated using the VOP generation algorithm given by [2]. Shadows of the VOP were removed using the algorithm explained in section 3.2. Digital watermarking of the VOP is done using Discrete Wavelet Transform. The VOPs were super-resolved using the algorithms proposed in [2, 6]. We have tested the shadow removal algorithm for VOP sequences for different scenarios. The threshold value TN specified for neighborhood calculation is set to 0.8 which means that less than 80 percent of the shadow boundaries are connected to object contour. We have calculated Root Mean Square Error (RMSE), Peak Signal to Noise Ratio (PSNR) and blur index for different frames after applying super resolution. The algorithms have been implemented in Matlab and tested on 2.99 GHz Core2Duo processor with RAM of size 2GB.

For the qualitative analysis of the generated SR VOP sequences, we have used MSU Video Quality Measurement Tool from Graphics & Media Lab, Moscow State University, Russia [24]. We used two metrics for this comparison. They are the MSU Blurring Metric and DCT-based video quality metric (VQM). For two values in blurring metric, whichever greater, is sharper than the other. VQM result gives the perception difference between two methods. For two values in VQM result whichever greater, is more perceptible than the other.

The proposed algorithms for shadow removal and digital watermarking of VOP and other modules, noise removal and super resolution by two methods are considered in different ways to have different test cases. The qualitative results are found out for each test case and found out a best method for video surveillance application. The quality checking metrics such as blurring metric and VQM perception quality difference are also found out.

4.1 Test Case 1

Initially 5 percent noise level is added to the input VOP to obtain noisy VOP. The added noise is removed by applying CMF filter. Shadow has been removed from the VOP with the proposed algorithm. Watermark is added to the resultant VOP using the above mentioned technique and then super resolved by the multiframe method. The watermark is extracted after super resolving the VOP. To know the effect, the quality checking metrics have been used. The blurring metric and VQM perception quality difference had also been found out. The setup is shown in fig. 11 and the results are tabulated in table 2. Fig.12 gives the super resolved VOPs using this test case experimental setup and watermarks extracted. Fig.13 provides blur index and VQM measurement graphs for both VOPs.

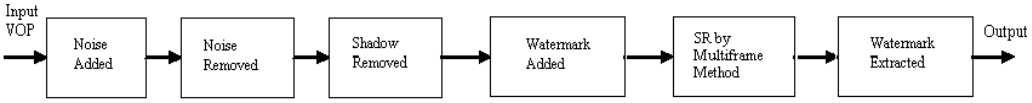


Fig. 11. Proposed System for Test case 1

Table. 2. Quality Checking Metrics for Test case 1

Frame No	Method	MSE	RMSE	PSNR	Blur Index
VOP 1: 75	Bicubic	0.9735	0.9866	48.2477	1.6195
	Multiframe	0.9462	0.9727	48.3711	2.3432
	Bilinear	0.9278	0.9863	48.2507	1.4964
81	Bicubic	0.9648	0.9822	48.2864	2.1510
	Multiframe	0.9446	0.9719	48.3783	2.9299
	Bilinear	0.9650	0.9823	48.2857	1.9699
VOP 2: 76	Bicubic	0.9568	0.9782	48.3227	1.2684
	Multiframe	0.9338	0.9663	48.4282	1.6243
	Bilinear	0.9556	0.9776	48.3279	1.1918
100	Bicubic	0.9572	0.9784	48.3207	1.6510
	Multiframe	0.9379	0.9684	48.4094	1.9070
	Bilinear	0.9567	0.9781	48.3229	1.5284

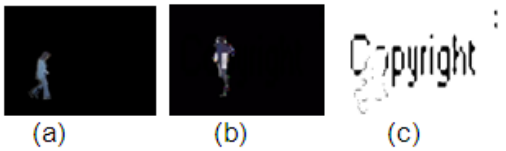
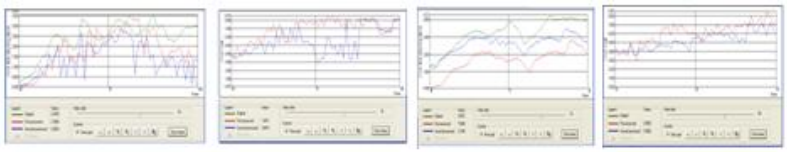


Fig.12. (a) Super resolved frame 75 of 1st VOP by Multiframe method (b) Super resolved frame 100 of 2nd VOP by Multiframe method (c) extracted watermark from frame 75 of 1st VOP which is super resolved by Multiframe method.



(a) (b) (c) (d)

Fig. 13. (a) & (b) Blur index and VQM measurement of 1st VOP respectively (c) & (d) Blur index and VQM measurement of 2nd VOP respectively.

4.2 Test Case 2

In the second test case, noise is added to input VOP and it is removed prior to shadow elimination. After elimination of shadow, it is super resolved by multiframe method and authenticated by DWT watermarking. The experimental setup of test case 2 is shown in fig. 14 and results are tabulated in table 3. Fig. 15 gives the super resolved VOPs and watermarks extracted using the test case 2 setup. Fig. 16 gives the blur index and VQM measurement graphs of two VOP sequences.

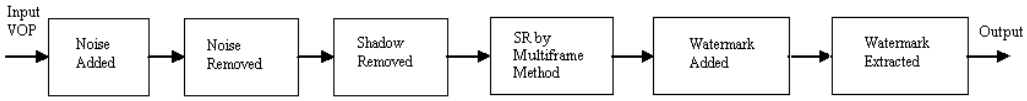


Fig.14. Proposed System for Test case 2

Table. 3. Quality Checking Metrics for Test case 2

Frame No	Method	MSE	RMSE	PSNR	Blur Index
VOP 1: 75	Bicubic	0.9740	0.9869	48.2454	2.2988
	Multiframe	0.9638	0.9817	48.2909	5.6360
	Bilinear	0.9732	0.9865	48.2488	2.1751
81	Bicubic	0.9663	0.9830	48.2798	2.7960
	Multiframe	0.9597	0.9796	48.3096	7.4777
	Bilinear	0.9656	0.9827	48.2827	2.6254
VOP 2: 76	Bicubic	0.9563	0.9779	48.3249	1.9491
	Multiframe	0.9355	0.9779	48.4205	2.6146
	Bilinear	0.9556	0.9672	48.3280	1.8796
100	Bicubic	0.9572	0.9784	48.3206	2.3279
	Multiframe	0.9390	0.9690	48.4043	2.8052
	Bilinear	0.9574	0.9785	48.3200	2.2102

(a) (b) (c)

Fig.15. (a) 75 frame of 1st Super resolved VOP by Multiframe method. (b) Super resolved frame 100 of 2nd VOP by Multiframe method. (c) extracted watermark from frame 75 of 1st SR VOP by Multiframe method.

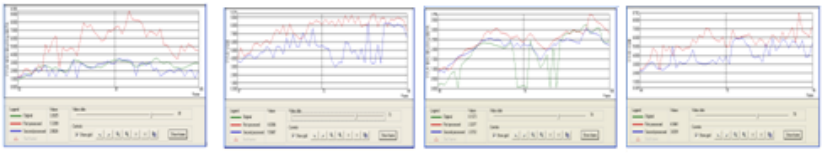


Fig. 16. (a) & (b) Blur index and VQM measurement of 1st VOP respectively (c) & (d) Blur index and VQM measurement of 2nd VOP respectively.

4.3 Test Case 3

In this test case, initially shadow has been eliminated from the input VOP. Addition of noise and reduction of noise is performed in subsequent steps. A watermark is added to the resultant VOP which then is super resolved by multiframe method to provide two super resolved VOPs. Watermark embedded is extracted to check the distortion made by super resolution.The experimental setup is shown in fig.17. Results are tabulated in table 4. Fig.18 shows super resolved VOPs and watermarks extracted. Fig.19 provides with blur index and VQM measurement graphs.

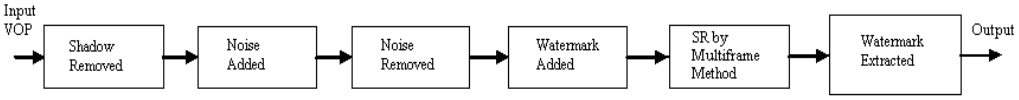


Fig.17. Proposed System for Test case 3

Table. 4. Quality Checking Metrics for Test Case 3

Frame No	Method	MSE	RMSE	PSNR	Blur Index
VOP 1: 75	Bicubic	0.9717	0.9857	48.2555	1.6218
	Multiframe	0.9467	0.9730	48.3687	2.1849
	Bilinear	0.9715	0.9856	48.2564	1.4867
81	Bicubic	0.9627	0.9812	48.2958	2.0865
	Multiframe	0.9447	0.9720	48.3777	2.7508
	Bilinear	0.9627	0.9812	48.2957	1.9209
VOP 2: 76	Bicubic	0.9478	0.9735	48.3637	1.7847
	Multiframe	0.9282	0.9634	48.4543	2.5307
	Bilinear	0.9479	0.9736	48.3633	1.6439
100	Bicubic	0.9634	0.9815	48.2926	2.0728
	Multiframe	0.9438	0.9715	48.3821	2.5281
	Bilinear	0.9633	0.9815	48.2932	1.8805

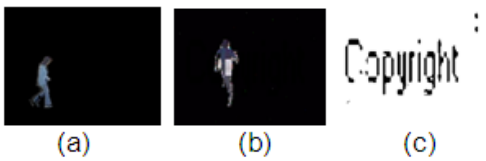


Fig.18. (a) Super resolved frame 75 of 1st VOP by Multiframe method (b) Super resolved frame 100 of 2nd VOP by Multiframe method (c) extracted watermark from frame 75 of 1st VOP which is super resolved by Multiframe method.

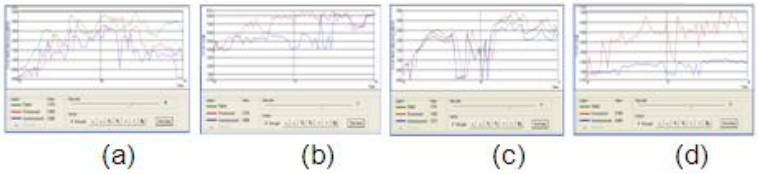


Fig.19. (a) & (b) Blur index and VQM measurement of 1st VOP respectively (c) & (d) Blur index and VQM measurement of 2nd VOP respectively

4.4 Test Case 4

The test case 4 procedure is similar to test case 3 but addition and extraction of watermark is performed after super resolving the VOP. Fig.20 provides the system setup for test 4. The results are tabulated in table 5. Fig.21 and fig.22 provides with super resolved VOPs, extracted watermarks and blur index, VQM measurement graphs of super resolved VOP for two different VOP sequences respectively.

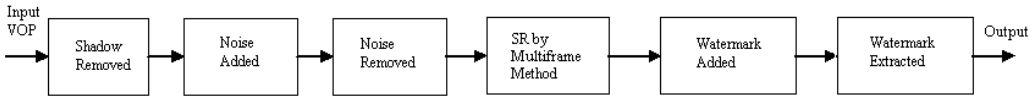


Fig.20. Proposed System for Test Case 4

Table. 5. Quality Checking Metrics for Test Case 4

Frame No	Method	MSE	RMSE	PSNR	Blur Index
VOP 1: 75	Bicubic	0.9721	0.9859	48.2539	2.2663
	Multiframe	0.9467	0.9730	48.3688	2.1503
	Bilinear	0.9723	0.9860	48.2530	2.1585
81	Bicubic	0.9632	0.9814	48.2936	2.7265
	Multiframe	0.9452	0.9722	48.3757	2.6975
	Bilinear	0.9626	0.9811	48.2963	2.5752
VOP 2: 76	Bicubic	0.9476	0.9735	48.3644	2.4208
	Multiframe	0.9279	0.9633	48.4559	2.6863
	Bilinear	0.9476	0.9735	48.3645	2.2937
100	Bicubic	0.9632	0.9814	48.2935	2.7095
	Multiframe	0.9475	0.9734	48.3649	2.6355
	Bilinear	0.9637	0.9817	48.2914	2.5331

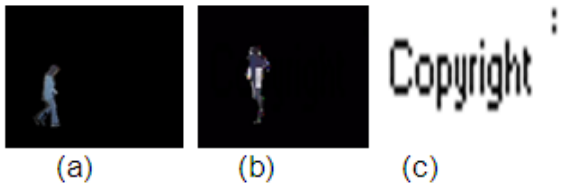


Fig. 21. (a) Super resolved frame 75 of 1st VOP by Multiframe method (b) Super resolved frame 100 of 2nd VOP by Multiframe method (c) extracted watermark from frame 75 of 1st VOP which is super resolved by Multiframe method.

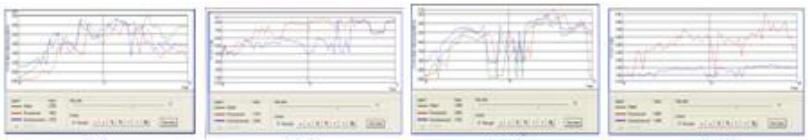


Fig.22. (a) & (b) Blur index and VQM measurement of 1st VOP respectively (c) & (d) Blur index and VQM measurement of 2nd VOP respectively

4.5 Test Case 5

Initially, shadow is eliminated from the input VOP. The resultant shadow eliminated VOP is super resolved by multiframe method. To study the effect of noise and watermarking on super resolved VOP, 5 percent level of noise is added which is then reduced. For authentication, a watermark is embedded into the resultant VOP. The qualitative results have been found out and are tabulated in table 6. Fig.23 gives the test case experimental setup. Fig.24 gives the super resolved VOPs by multiframe method and of 2 VOP sequences. Fig. 25 provides with blur index and VQM measurement graphs of both VOP sequences.

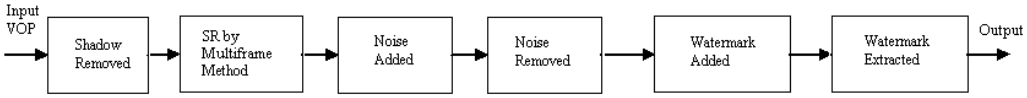


Fig.23. Proposed System for Test Case 5

Table. 6. Quality Checking Metrics for Test Case 5

Frame No	Method	MSE	RMSE	PSNR	Blur Index
VOP 1: 75	Bicubic	0.9623	0.9800	48.2978	1.9574
	Multiframe	0.9493	0.9810	48.3568	2.1094
	Bilinear	0.9627	0.9743	48.2960	1.8313
81	Bicubic	0.9551	0.9773	48.3305	2.4016
	Multiframe	0.9470	0.9732	48.3671	2.3902
	Bilinear	0.9549	0.9772	48.3310	2.3067
VOP 2: 76	Bicubic	0.9377	0.9684	48.4100	2.0943
	Multiframe	0.9284	0.9635	48.4536	2.5858
	Bilinear	0.9377	0.9684	48.4104	2.0026
100	Bicubic	0.9556	0.9776	48.3279	2.3901
	Multiframe	0.9454	0.9723	48.3746	2.8007
	Bilinear	0.9557	0.9776	48.3276	2.2252

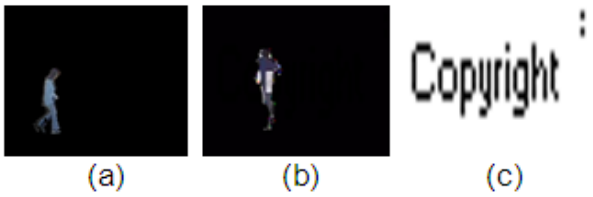


Fig.24. (a) Super resolved frame 75 of 1st VOP by Multiframe method (b) Super resolved frame 100 of 2nd VOP by Multiframe method (c) extracted watermark from frame 75 of 1st VOP which is super resolved by Multiframe method.

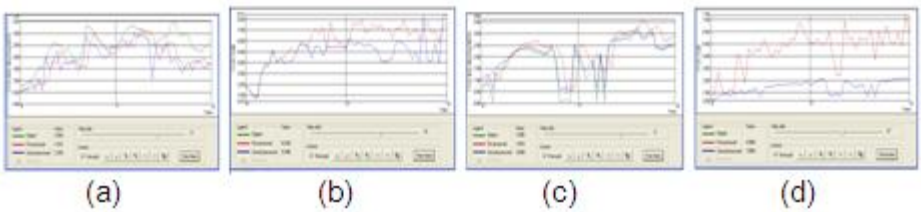


Fig.25. (a) & (b) Blur index and VQM measurement of 1st VOP respectively (c) & (d) Blur index and VQM measurement of 2nd VOP respectively

4.6 Test Case 6

The experimental setup of test case 6 is similar to setup of test case5. Some modules have been interchanged to have complete view of results of different tests. Initially, shadow is eliminated from input VOP and is super resolved by multiframe method. A watermark is embedded by using DWT watermarking for authentication. To study noise effect, some noise is inserted which is then tried to reduce by applying noise removal smoothing filter on the resultant VOP. The watermarks are extracted then to discover the distortion occurred to them. Fig. 26 provides with the setup. The results are tabulated in table 7. Fig. 27 provides with the VOP frames which have been undergone through all the setup process. Fig. 28 provides the blur index and VQM measurement graphs.

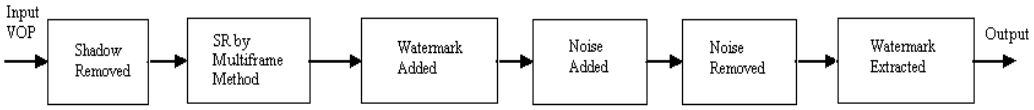


Fig.26. Proposed System for Test Case 6

Table. 7. Quality Checking Metrics for Test Case 6

Frame No	Method	MSE	RMSE	PSNR	Blur Index
VOP 1: 75	Bicubic	0.9630	0.9813	48.2946	1.8754
	Multiframe	0.9485	0.9739	48.3603	2.1160
	Bilinear	0.9632	0.9814	48.2937	1.7946
81	Bicubic	0.9560	0.9778	48.3262	2.3437
	Multiframe	0.9464	0.9728	48.3703	2.3810
	Bilinear	0.9558	0.9777	48.3271	2.2219
VOP 2: 76	Bicubic	0.9387	0.9689	48.4055	2.1094
	Multiframe	0.9274	0.9630	48.4579	2.6600
	Bilinear	0.9386	0.9688	48.4039	2.0645
100	Bicubic	0.9565	0.9780	48.3230	2.4120
	Multiframe	0.9449	0.9721	48.3770	2.8381
	Bilinear	0.9560	0.9778	48.3260	2.2379

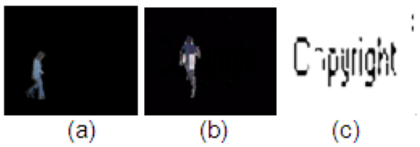


Fig. 27. (a) Super resolved frame 75 of 1st VOP by Multiframe method (b) Super resolved frame 100 of 2nd VOP by Multiframe method (c) extracted watermark from frame 75 of 1st VOP which is super resolved by Multiframe method.

Fig. 28. (a) & (b) Blur index and VQM measurement of 1st VOP respectively (c) & (d) Blur index and VQM measurement of 2nd VOP respectively.

4.7 Test Case 7

To study only super resolution and watermark effects, another experimental procedure is arranged as shown in fig. 29. The input VOP is authenticated by embedding a watermark which is then super resolved by the multiframe method. The watermarks are extracted from sample frames of either super resolved VOPs to know about the distortion occurred under super resolution. The qualitative results are tabulated in table 8. The super resolved frames and the watermarks extracted from original two different VOPs are given in fig.30. Fig. 31 provides with graphs of blur index and VQM measurements which measures the sharpness and perception difference of VOP sequences.

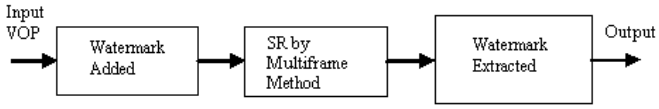


Fig.29. Super Resolution of Watermarked Original VOP

Table. 8. Quality Checking Metrics for Test Case 7

Frame No	Method	MSE	RMSE	PSNR	Blur Index
VOP 1:75	Bicubic	0.9832	0.9915	48.2046	2.2990
	Multiframe	0.9725	0.9861	48.2521	3.1842
	Bilinear	0.9806	0.9903	48.2157	2.1351
81	Bicubic	0.9783	0.9891	48.2260	3.1605
	Multiframe	0.9644	0.9820	48.2884	3.5166
	Bilinear	0.9757	0.9878	48.2375	2.9024
VOP 2: 76	Bicubic	0.9771	0.9885	48.2314	2.8469
	Multiframe	0.9658	0.9827	48.2819	3.8906
	Bilinear	0.9747	0.9873	48.2422	2.6130
100	Bicubic	0.9782	0.9890	48.2267	3.1952
	Multiframe	0.9669	0.9833	48.2772	3.9122
	Bilinear	0.9781	0.9889	48.2269	2.8800

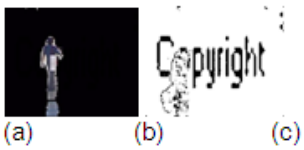
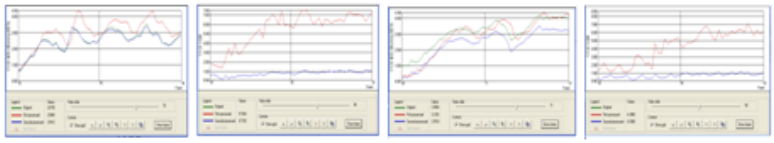


Fig. 30. (a) Super resolved frame 75 of 1st VOP by Multiframe method (b) Super resolved frame 100 of 2nd VOP by Multiframe method (c) extracted watermark from frame 75 of 1st VOP which is super resolved by Multiframe method.



(a) (b) (c) (d)

Fig. 31. (a) & (b) Blur index and VQM measurement of 1st VOP respectively (c) & (d) Blur index and VQM measurement of 2nd VOP respectively

In all the tables presented above for test cases, (i.e. table 2 – table 8), MSE indicates Mean square Error value, RMSE indicates Root mean square value, PSNR indicates Peak signal to noise ratio. Blur index values denotes the sharpness of that particular method.

From the tables, it is evident that SR using multiple frames is better than bicubic and bilinear methods. The Blur index and PSNR values of this method are higher and MSE, RMSE values are lower which formulates that multiframe super resolution is better. As blur index indicates sharpness, SR using multiframes is sharper than other methods specified.

For the entire Blur index and VQM graphs shown above, the measurements have taken for 2 VOP sequences. In all the respective fig.s of each procedure, (a) & (b) indicates blur index and VQM for 1st VOP and (c) & (d) indicates blur index and VQM for 2nd VOP. Red line indicates the multiframe super resolved VOP. Green line indicates the original VOP sequence. From all the graphs presented above, it is evident that super resolution using multiple frames is more perceptible than other methods.

The initial six test cases are considered to compare their performance and to find out the best method for video surveillance application. The test case 7 is not considered as the shadow elimination is not included. Out of 6 test cases, test case 1 is better by comparing all the qualitative results. Taking blur index into account, test case 2 provides better results than test case 1. Considering VQM measurements which measures perception quality, test case 1 provides better results as the values ranges from 4.000 to 9.000 and most frames of it have higher VQM values.

By considering the above scenarios, it shows that test case 1 is considered best for the surveillance applications. Moreover, the qualitative results of all test cases show that SR by multiframe method is better than other SR methods. Besides, the watermarks extracted from super resolved by multiframe VOPs are accurate than extracted watermarks of SR by other methods . Therefore, it can be concluded that test case 1 with SR by multiframe method is best suitable for video surveillance applications.

5. Conclusion

A new method for shadow removal from VOP has been proposed. This algorithm gives good results for the VOP and also computation efficient. A new DWT based watermarking algorithm using localized Biorthogonal wavelets for VOP has been proposed. The generated VOP has been super-resolved using multi-frame super resolution video generation algorithms as proposed by us in [2 & 6]. The super resolved VOP using multi-frame super resolution algorithm gives better results than bilinear, bicubic super resolution methods. Six test cases have been formulated with all the modules to find out the best scenario for video surveillance application. It has been concluded that the test case 1 with super resolution by multi frame method is best suitable for video surveillance applications.

References

[1] Andrea Prati, Ivana Mikic, Mohan M. Trivedi and Rita Cucchiara. Detecting Moving Shadows: Algorithms and Evaluation. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.25, no. 7, July 2003.

[2] Alwyn R Pais, Naveen Pasupuleti, Rajashekhar Gidnavar, John D'Souza. Automatic Video Object Plane Segmentation and Super resolution: A Novel Approach. Indian International Conference on Artificial Intelligence, 16-18 Dec, 2009.

[3] Jaakko Astola, Petri Haavisto and Yrjo Neuvo. Vector Median Filters. Proceedings of the IEEE, Vol. 78, No. 4, pp. 678-689, 1990.

[4] N Nikolaidis and I Pitas. Digital Image Watermarking: an Overview. IEEE International Conference on Multimedia Computing and Systems, vol. 1, pp.1 – 6, 7-11 June, 1999.

[5] Nanzhi Jiang, Renbiao Wu and Jian Li. Super Resolution Feature Extraction of Moving Targets. IEEE Transactions on Aerospace and Electronic Systems, vol. 37, no.3, pp. 781-793, 2001.

[6] Alwyn Roshan Pais, John D'Souza, Ram Mohana Reddy and Sandip Patil. Super Resolution of Video with Sharpened Edges Using Multiple Frames - A Novel Approach. International Conference on Advances in Recent Technologies in Communication and Computing, pp. 98-102, 27-28 Oct, 2009.

[7] Shoaib M, Dragon R, Ostermann J. Shadow Detection for Moving Humans using Gradient-Based Background Subtraction. IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP, pp. 773-776, 2009.

[8] Liang Xu, Liang Zhang, Ping Han, Ren-biao Wu. Adaptive Threshold Shadow Detection Based On Image Block Statistics. ICSP Proceedings, 2008.

[9] Song Xuehua, Ding Yan, Gen Jianfeng, Chen Yu. Shadow Removal of Vehicles in a Video System Based on RGB Chroma Model. International Conference on Computer Science and Software Engineering, vol.1, pp. 977-980, 2008.

[10] R Cucchiara, C Grana, M Piccardi and A Prati. Detecting objects, shadows and ghosts in video streams by exploiting color and motion information. 11th International Conference on Image Analysis and Proceedings, pp. 360-365, 2001.

[11] Y Deng, S Kenney, M S Moore and B S Manjunath. Peer group filtering and perceptual color image quantization. Proc. IEEE International Symposium on Circuits and Systems VLSI , (ISCAS'99), Orlando, FL, no. 4, pp.21-24, June 1999.

[12] Yu-Li You and M.Kaveh. Image Enhancement using fourth order partial differential equations. IEEE Transactions on Image Processing, 2000.

[13] Soheil Feizi and Sina Zahedpour. Salt and pepper noise removal for image signals. International conference on telecommunications, pp.1-5, 2008.

[14] James C Church, Yixin Chen, and Stephen V Rice. A Spatial Median Filter for Noise Removal in Digital Images. IEEE southeastcon, pp.618-623, 2008.

[15] Potdar, V M, Han, S, Chang E. A Survey of Digital Image Watermarking Techniques. 3rd IEEE International Conference on Industrial Informatics (INDIN), pp.709 – 716, Aug. 2005.

[16] M D Swanson, B Zhu and A H Tewfik. Robust Data Hiding for Images. IEEE Digital Signal Processing Workshop, pp. 37-40, 1996.

[17] Malay Kumar Nema, Subrata Rakshit, and Subhasis Chaudhuri. Edge Model Based High Resolution Image Generation. ICVGIP, Springer, 2006.

[18] O Javed, K Shafique and M Shah. A hierarchical approach to robust background subtraction using color and gradient information. Proc. of IEEE Workshop on Motion and Video Computing, 2002.

[19] C Staufer, W E L Grimson. Adaptive background mixture models for real-time tracking. Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, Vol. 2, pp. 246-252, June 1999.

[20] Rafael C Gonzalez, Richard E Woods and Steven L Eddins. Digital Image Processing using Matlab. Pearson Education, Inc. 2004.

[21] Guangmin Sun and Yao Yu. DWT based Watermarking Algorithm of Color Images. IEEE Conference on Industrial Electronics and Applications, ICIEA, pp.1823-1826, 2007.

[22] Suhad Hajjara, Moussa Abdallah, Amjad Hudaib, Digital image watermarking using localized Biorthogonal wavelets. European Journal of Scientific Research. Vol.26 No.4 pp.594-608, 2009.

[23] Xin Li, Michael T Orchard. New Edge-Directed Interpolation. IEEE transactions on image Processing, vol. 10, no. 10, Oct 2001.

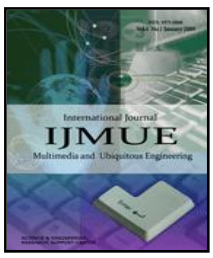
[24] MSU Video Quality Measurement Tool. http://compression.ru/video/qualitymeasure/indexen.html. Mar 2010.

Authors

Alwyn R. Pais is currently the Assistant Professor,Department of Computer Science and Engineering, National Institute of Technology Karnataka, Surathkal,INDIA. He was born in Karnataka,India. He completed his Bachelor of Engineering from the Mangalore University, Karnataka and Master of Technology from IIT Bombay. He is currently pursuing his PhD from NITK, Surathkal. He is also the coordinator of the ISEA project at NITK. His areas of interest are Algorithms, Cryptography and Computer Vision. His email is alwyn.pais@gmail.com and has personal webpage at http://isea.nitk.ac.in/faculty/alwyn/

John D'Souza is currently the Associate Professor, Department of Electronics and communication Engineering, National Institute of Technology Karnataka, Surathkal, INDIA. He holds M.Tech and PhD from IIT, Kharagpur. His areas of interest are Digital communication, Information Theory and Cryptography.

Ram Mohana Reddy is currently the Professor, Department of Information Technology, National Institute of Technology Karnataka, Surathkal, INDIA. He holds a PhD from Edinburgh University UK.



Hari Krishna is a post Graduate Student at NITK, Surathkal