Video Restoration with Motion Prediction Based on the Multiresolution Wavelet Analysis

Kei Akiyama^{1,2,5}, Zhi-wei Luo^{3,2}, Masaki Onishi^{4,2}, Shigeyuki Hosoe², Kouichi Taji¹, and Yoji Uno¹

¹ Nagoya University, Graduate School of Engineering, Furo-cho, Chikusa-ku Nagoya, 464-8603 Japan {k akiyama, taji, uno}@nuem.nagoya-u.ac.jp ² Bio-mimetic Control Research Center, RIKEN 2271-130, Anagahora, Shimoshidami, Moriyama-ku Nagoya, 463-0003 Japan, hosoe@bmc.riken.jp ³ Kobe University, Faculty of Engineering, 1-1 Rokkohdai-cho, Nada-ku, Kobe, 657-8501, Japan, luo@gold.kobe-u.ac.jp ⁴ Information Technology Research Institute, AIST 1-1-1 Umezono, Tsukuba, 305-8568, Japan onishi@ni.aist.go.jp ⁵ Hitachi Global Storage Technologies Japan, Ltd. 1 Kirihara-cho, Fujisawa, 252-8588, Japan

Abstract. We propose a novel method for image sequence restorations. It is based on the wavelet domain image restoration method proposed by Belge et al. for static images [1]. In this paper, by combining the iteration procedure in the Belge's method with the renewing process for sequentially given images and by employing Kalman filer for predicting the foreground movement of the images in the wavelet domain, considerable reduction of the computational cost is shown to be achievable. This is verified by computer simulations on artificially degraded images.

Keywords: Multiresolution wavelet analysis, Video restoration, Motion dynamics, Nonlinear optimization.

1 Introduction

A video sequence acquired by a camera often contains blur and/or disturbance by various causes. In many applications like image surveillance or broadcasting, these degradation factors need to be automatically removed in order to facilitate higher level recognitions. So far many restoration methods have been proposed especially for static images [1,2]. In recent years the number of restoration method for video sequences is gradually increasing. Pizurica et al. proposed a method [3] which combines spatially adaptive noise filtering in the wavelet domain and temporal filtering in the signal domainD Rares et al. presented an algorithm dealing with degradation related to severe artifacts [4]. In [5,6], Kornprobst et al. proposed some restoration methods utilizing motion compensation.

M. Ishikawa et al. (Eds.): ICONIP 2007, Part II, LNCS 4985, pp. 77–86, 2008.

⁻c Springer-Verlag Berlin Heidelberg 2008

Fig. 1. Block diagrams for the multiresolution wavelet decomposition of an image

For further references, please see [7,8]. With these developments, however, to handle the degradation like optical blur and disturbance which happens more frequently in real environment, more work is needed for restoration.

In this paper, we propose a restoration method for video sequences which are degraded by optical blur and noise. In our previous work [9], we proposed a video restoration method by extending Belge et al.'s restoration method for static images to video case. This method could reduce the computational cost compare to the method of [1] by modeling a class of image sequences by a state equation and predicting future frames based on it. However, there was yet a room for further improvement since we had to execute a multiresolution wavelet reconstruction (MWR, to be described) for making a prediction of the future image in state space. In the present paper, we propose an improved method. Here the prediction of the future image is directly realized in the wavelet domain. This contributes not only reducing the above redundancy but also making the prediction more efficient by utilizing the property of multiresolution wavelet decomposition (MWD) images. We verify our method by computer simulation of an artificially degraded image sequence.

2 Image Restoration Using Multiresolution Wavelet Decomposition

2.1 Multiresolution Wavelet Decomposition

Let q be a lexicographically ordered static image. The block diagram of the MWD of *g* is shown in Fig. 1. In the figure $p(\cdot)$ and $q(\cdot)$ generally represent an 1-D low-pass and high-pass filter, respectively. From the input image, four downsampled images are obtained [1]. Furthermore, by repeating the decomposition, we can get multiresolution images $[10]$. An MWD image \hat{g} calculated by L level MWD is presented as

$$
\hat{\boldsymbol{g}} := \left(\hat{g}_{(L,0)}^T, \cdots, \hat{g}_{(L,3)}^T, \hat{g}_{(L-1,1)}^T, \cdots, \hat{g}_{(1,3)}^T\right)^T.
$$
\n(1)

Fig. 2. An example of the MWD images

An MWD result for the image 'Cameraman' is shown in Fig. 2 as an example. In Fig. 2(b), component $\hat{g}_{(3,0)}$ represents a scaled down image of the original one and other components $\hat{g}_{(l,i)}$ correspond to extracting the 1-D features (horizontal, vertical and diagonal) of the original image. Note that the number of the total pixels is unchanged during the decomposition. Since the decomposition operation has the orthogonal property, we can get the reconstruction operation and it can completely recapture the original image from the decomposed one. We call it the wavelet multiresolution reconstruction (MWR).

2.2 Degradation Process and Its Restoration for MWD Images

Given the low-resolution image sequence $g = \{g^{[1]}, g^{[2]}, \cdots, g^{[K]}\}\$ of the original image sequence $f = \{f^{[1]}, f^{[2]}, \cdots, f^{[K]}\}$ of length K. The sequences $\hat{g} =$ $\{\hat{g}^{[1]}, \hat{g}^{[2]}, \cdots, \hat{g}^{[K]}\}\$ and $\hat{f} = \{\hat{f}^{[1]}, \hat{f}^{[2]}, \cdots, \hat{f}^{[K]}\}\$ denote respectively the MWD of *g* and *f*. In this paper, we consider a restoration problem for a given degraded MWD image sequence $\{\hat{\bm{g}}^{[1]}, \cdots, \hat{\bm{g}}^{[K]}\}$ which is degraded from its original image sequence $\{\hat{\pmb{f}}^{[1]},\cdots,\hat{\pmb{f}}^{[K]}\},$ where the superscripts denote the frame number. First, we formulate a degradation process for MWD images [1] by

$$
\hat{\boldsymbol{g}}^{[k]} = \hat{\boldsymbol{H}} \hat{\boldsymbol{f}}^{[k]} + \hat{\boldsymbol{u}}^{[k]}.
$$
\n
$$
(2)
$$

In equation(2), the vector $\hat{u}^{[k]}$ is an additive noise and the matrix \hat{H} represents an linear distortion or optical blur, which can be assumed to be constant with respect to frames since the change is sufficiently small.

When considering a restoration for the degradation process of equation (2) , one natural way would be to apply some of the known restoration procedures to each frames one by one, regarding them as static images, and then make necessary modifications to make the computation more efficient and improve

Fig. 3. Block diagram of the image sequence restoration [9]

restoration by considering relationships existing among the frames. In this paper, we follow this way. As a basic restoration method for static images, the one proposed by Berge et al. will be used. The method has been derived by considering minimization of the cost function given by

$$
J_k\left(\hat{f}^{[k]},\lambda\right) = \left\|\hat{g}^{[k]} - \hat{H}\hat{f}^{[k]}\right\|_2^2 + \lambda_{(L,0)}\left\|\hat{f}^{[k]}_{(L,0)}\right\|_p^p + \sum_{l=1}^L \sum_{j=1}^3 \lambda_{(l,j)}\left\|\hat{f}^{[k]}_{(l,j)}\right\|_p^p. (3)
$$

The first term of equation (3) represents the closeness of the original image. The second and the third terms express the closeness to the statistical prior knowledge of the natural image in the wavelet domain. Lambdas are the regularization parameters. The optimal restored images can be calculated by a numerical optimization [1]. The algorithm can realize the edge preserving restoration by assigning different parameters to each decomposed image.

However, generally the calculation will become very huge since we have to repeat iterative computations with big size matrices for every frames. To cope with this problem, in [9] by combining the iterative procedure in the Berge's method with the renewing process for sequentially given images and employing Kalman filter, we showed that considerable reduction of the calculation cost can be realized. The block diagram is shown in Fig. 3. In the 'Restoration Calculation' box, the optimal restored image (denoted by $\hat{f}_{rst}^{[k]}$) is calculated by using the following modified equation of the optimization method in [1]:

$$
\left(\hat{\boldsymbol{H}}^T \hat{\boldsymbol{H}} + \frac{p}{2} \boldsymbol{D}_{prd}^{[k]}\right) \hat{\boldsymbol{f}}_{rst}^{[k]} = \hat{\boldsymbol{H}}^T \hat{\boldsymbol{g}}^{[k]} \tag{4}
$$

$$
\mathbf{D}_{prd}^{[k]} = \text{diag}\left[\frac{\lambda(i)}{(|\hat{\mathbf{f}}_{prd}^{[k]}(i)|^{2} + \beta)^{1-p/2}}\right]_{i=1}^{N^{2}}.
$$
 (5)

Notice that instead of restored image $f_{rst}^{[k]}$ in [1] its predicted image $f_{prd}^{[k]}$ is used here. Predicted image $f_{prd}^{[k]}$ is computed in the 'Predictor' box. To carry out the computations in Predictor box we had made the following assumptions.

A1 An original image sequence consists of a foreground and a background.

- **A2** The change of the background is small enough to be set as a static image.
- **A3** The change of the foreground can be formulated or approximated by a known dynamic equation.
- **A4** The foreground is assumed to be a single rigid body and maintain its orientation.

With the assumptions **A1** and **A2**, we can utilize the restoration result of previous frame directly as an initial estimation of the background for each frame. On the other hand, we can predict a new position of the foreground from the previous restoration result and the information about motion dynamics (**A3**) by using Kalman filter.

This algorithm can reduce the calculation cost for an image sequence restoration compared to the frame by frame optimization based on Belge et al.'s method, while the qualities of the restoration results being almost unaffected. However, yet some redundant calculations are included because it needs MWR calculations of restored images for the sake of making prediction for next frames, and again calculate MWD after a predicted image is obtained. If we could get a predicted image directly in the wavelet domain, the redundancy of this algorithm can be reduced. We state this modified image sequence restoration method in the next section.

3 Video Restoration Algorithm in Wavelet Domain

At first, we show the overall sketch to our new video restoration algorithm in wavelet domain (Fig. 4). In the following, the restoration image and the predicted image of \hat{f} will be represented as $\hat{f}_{rst} = \{ \hat{f}_{rst}^{[1]}, \hat{f}_{rst}^{[2]}, \cdots, \hat{f}_{rst}^{[K]} \}$ and $\hat{f}_{prd} = \{ \hat{f}_{prd}^{[1]}, \hat{f}_{prd}^{[2]}, \cdots, \hat{f}_{prd}^{[K]} \}$ respectively. According to the structure of MWD, $\hat{f}^{[k]}$ (similarly for $\hat{f}^{[k]}_{rst}$ or $\hat{f}^{[k]}_{prd}$) will be represented as

$$
\hat{\pmb{f}}^{[k]} := \left(\hat{f}_{(L,0)}^{[k]^T},\cdots,\hat{f}_{(L,3)}^{[k]^T},\hat{f}_{(L-1,1)}^{[k]^T},\cdots,\hat{f}_{(1,3)}^{[k]^T}\right)^T.
$$

Step 1 Initialization. Let

$$
\hat{f}_{prd}^{[1]} = \hat{g}^{[1]},
$$

\n
$$
\hat{f}_{rst}^{[1]}
$$
 is given by (4) and (5),
\n
$$
\hat{f}_{prd}^{[2]} = \hat{f}_{rst}^{[1]},
$$

\n
$$
\hat{f}_{rst}^{[2]}
$$
 is given by (4) and (5).

Fig. 4. Block diagram of the proposed method for kth frame

Step 2 With using $\hat{f}^{[k-1]}_{rst}$ and $\hat{f}^{[k]}_{rst}$, for all

$$
(l, j) = (L, 0), (L, 1), (L, 2), (L, 3), (L-1, 1), \cdots, (1, 3),
$$

compute the followings in turn

$$
\hat{\mathbf{f}}_{fg(l,j)}^{[k]} : \text{ foreground (Sec. 3.2)},
$$

$$
\hat{\mathbf{f}}_{bg(l,j)}^{[k]} : \text{background (Sec. 3.2)},
$$

$$
\hat{\mathbf{f}}_{bd(l,j)}^{[k]} : \text{background domain (Sec. 3.2)},
$$

$$
\mathbf{v}_{(L,0)}^{[k]}, \text{ and } \mathbf{v}_l^{[k]} \ (l = L, L-1, \dots, 1).
$$

Step 3 Compute the prediction of $v_l^{[k+1]}$ using Kalman filter for (11). **Step 4** Compute

$$
\hat{f}_{bd(l,j)}^{[k+1]} \text{ and } \hat{f}_{fg(l,j)}^{[k+1]} \text{ (by (8)),}
$$
\n
$$
\hat{f}_{prd(l,j)}^{[k+1]} \text{ by inserting estimated } \hat{f}_{fg(l,j)}^{[k+1]} \text{ into } \hat{f}_{bg(l,j)}^{[k]}.
$$

Step 5 Compute $\hat{f}_{rst}^{[k]}$ by (4) and (5). **Step 6** Compute $f_{rst}^{[k]}$ by MWR. **Step 7** If $k = K$, stop. Othewise $k = k+1$ and go to **Step 2**.

In executing the algorithm if it happens that we can not continue computation by the frame out of the moving object or by a sudden change of the background we have to cancel the prediction till the next movement is observed.

3.1 Definition of the Dynamics for the MWD Image Sequence

Based on the above assumptions made in **2.2**, we model the dynamics of an original MWD image sequence as follows. First, we define the variables as in

Table 1. Definition of the variables for kth frame

Item name	Definition
$\hat{\bm{f}}[k]$	Original image
$\hat{f}^{[k]}_{bd} \ \hat{f}_{bg} \ \hat{f}^{[k]}_{\hat{f}}$	Original background domain (0: foreground, 1: others)
	Original background image
	Original foreground image
$\hat{\bm{a}}^{[k]}$	Degraded image

Table 1. By these definitions, each component of the original MWD image sequence is represented as:

$$
\hat{f}_{(l,j)}^{[k]} = \left\{ I_{(N/2^l)^2} - \text{diag}\left[\hat{f}_{bd(l,j)}(i)\right]_{i=1}^{(N/2^l)^2} \right\} \cdot \hat{f}_{bg(l,j)}^{[k]} + \hat{f}_{fg(l,j)}^{[k]}.
$$
\n
$$
((l,j) = (L,0) \text{ and } l = 1, \dots, L, j = 1, 2, 3)
$$
\n(6)

We introduce a transition of a foreground between k th and $k+1$ th frames. For this, a motion of a foreground object is described by

$$
\begin{bmatrix} \boldsymbol{v}_l^{[k+1]} \\ \boldsymbol{a}_l^{[k+1]} \end{bmatrix} = \begin{bmatrix} I_2 & I_2 \\ 0_{2 \times 2} & I_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{v}_l^{[k]} \\ \boldsymbol{a}_l^{[k]} \end{bmatrix}, \tag{7}
$$

where $v_l^{[k]}$ and $a_l^{[k]}$ are velocity and acceleration per a frame of a characteristic point for each decomposed image of kth frame, respectively. I_2 denotes a 2×2 identity matrix. Equation (7) represents an uniform accelerated motion on a 2-D plane. In correspondence with the difference of the initial condition, the various movement (straight lines or parabola-shaped motions in the 2-D plane, for example) can be described in this way. Since from assumption **A4**, the distance between each element of $\hat{f}_{bd(l,j)}^{[k]}$ and its corresponding element of $\hat{f}_{bd(l,j)}^{[k+1]}$ remains the same, the relationship between $\hat{f}_{bd(l,j)}^{[k]}$ and $\hat{f}_{bd(l,j)}^{[k+1]}$ is written by

$$
\hat{f}_{bd(l,j)}^{[k+1]}((n-1)N+m)
$$
\n
$$
= \hat{f}_{bd(l,j)}^{[k]} \left(\left[\left\{ (n-v_{lv}^{[k+1]}) \mod N/2^l \right\} - 1 \right] N/2^l + \left\{ (m-v_{lh}^{[k+1]}) \mod N/2^l \right\} \right). \tag{8}
$$
\n
$$
(n = 1, \cdots, N/2^l, \quad m = 1, \cdots, N/2^l)
$$

Equation (8) can be expressed using a matrix $T_l(v_l^{[k]})$ as follows:

$$
\hat{f}_{bd(l,j)}^{[k+1]} = T_l \left(\boldsymbol{v}_l^{[k+1]} \right) \hat{f}_{bd(l,j)}^{[k]} \tag{9}
$$

$$
T_l\left(v_l^{[k+1]}\right) = \text{diag}\left(C_{lh}^{v_{lh}^{[k+1]}}, \dots, C_{lh}^{v_{lh}^{[k+1]}}\right) \cdot C_{lv}^{v_{lv}^{[k+1]}}.
$$
 (10)

We call $T_l\left(v_l^{[k+1]}\right)$ the transition matrix of level *l*. C_{lv} and C_{lh} in equation (10) are an $(N/2^l)$ ² × $(N/2^l)$ ² dimension block circulant matrix and $N/2^l \times N/2^l$

Fig. 5. Motion estimation for MWD images

dimension circulant matrix respectively, which are the same as in [9]. The transition of foreground image $\hat{f}_{fg(l,j)}^{[k+1]}$ can also be described exactly in the same way as in (9).

3.2 Foreground Extraction and Motion Prediction

First, we estimate optical flow for $(L, 0)$ image by taking squared error between local areas of k−1th and kth frames. To avoid mismatching, squared errors over a certain threshold would not be recognized as a motion. Second, we extract the foreground object domain, in which the optical flows are similar each other. To detect a motion vector $v_{(L,0)}^{[k]}$ of the foreground object, we take an average of the flows within the foreground object domain.

Since from the properties of MWD, $(L, 1)$, $(L, 2)$ and $(L, 3)$ images are reflecting the vertical, horizontal and diagonal characteristics of an original image more strongly [11], we will use them to detect motions along one direction only. We detect motion of these three images for each one dimension $((L, 1)$ for horizontal direction, $(L, 2)$ for vertical direction and $(L, 3)$ for diagonal direction) (Fig. $5(a)$) within the corresponding domain detected by $(L, 0)$ image, and extract foreground objects of each image. Then, we take averages of the motion within the foreground object domains for $(L, 1)$ and $(L, 2)$ images and assign these values to the motion vector $v_L^{[k]}$. Repeat the above process from level L–1 to level 1, and detect the motion vector for each level. Motion search in each level is done within the foreground object domain detected in the upper level (Fig. $5(b)$). By this hierarchical searching method, calculation cost can be smaller than the full searching method. The motion vector for level *l* is denoted as $v_l^{[k]} := (v_{lv}^{[k]}, v_{lh}^{[k]}),$ of which $v_{lv}^{[k]}$ and $v_{lh}^{[k]}$ are results of the 1-D (vertical and horizontal) motion estimations. $(l, 3)$ images are not used for motion estimations, since they may be strongly affected by noise in the original image [11].

Now, with the detected motion vector $v_l^{[k]}$ as above and with the assumed model (7) for the movement of the foreground object, we can get a prediction

(c) Degraded image of 10th frame (d) Restored image of 10th frame

Fig. 6. Simulation result of the proposed method (3rd and 10th frames)

concerning the foreground location of the next frame image, by constructing a Kalman filter for

$$
\begin{cases}\n\begin{bmatrix}\n\boldsymbol{v}_l^{[k+1]}\n\end{bmatrix} = \begin{bmatrix}\nI_2 & I_2\n\end{bmatrix} \begin{bmatrix}\n\boldsymbol{v}_l^{[k]}\n\end{bmatrix} + \boldsymbol{w}_{(l,j)}^{[k]}\n\end{cases}
$$
\n
$$
\begin{bmatrix}\nv_{lv}^{[k]}\n\end{bmatrix} = \begin{bmatrix}\nI_2 & 0_{2\times 2}\n\end{bmatrix} \begin{bmatrix}\n\boldsymbol{v}_{l}^{[k]}\n\end{bmatrix} + \boldsymbol{n}_{(l,j)}^{[k]}.
$$
\n(11)

4 Simulation

We performed a simulation of the proposed method with known degradation parameters and we verified the performance of the proposed method. We generated an artificial image sequence in 64×64 pixels and 10 frames. We used a test image 'Text' for the background and a triangle object with changing pixel value for the foreground. The foreground was supposed to move with constant velocity. We made the original image sequence $f^{[k]}$ by equation (6) and calculted its degraded image sequence $g^{[k]}$ by equation (1). We considered an optical blur for **H** in equation (1) and used a Gauss function of the variance $\sigma^2 = 1.2$ with the 7×7 discretized elements. The disturbance $u^{[k]}$ was assumed to be a Gaussian noise of average zero and the SN ratio of 30dB independently for each frames. In the restoration calculation, the level of the wavelet multiresolution decomposition (L) was assumed to be three and we selected the three tap wavelet [11].

The degraded and the restored images for some frames are shown in Fig. 6. In both frames, the background and the foreground of the restored images could be much clearly recognized than the original degraded images by the proposed method. The total calculation time for 10 frames was 90 sec. and the prediction time (Steps 2–5) was about10 sec. The calculation time for our prediction algorithm is sufficiently short.

5 Conclusion

We proposed an effective restoration method for degraded video sequence in this paper. The dynamics of the MWD image sequence is modeled and a novel calculation algorithm is proposed. Computer simulation for an artificial image sequence shows favorable result qualitatively. More quantitative verifications such as calculation time or restoration quality are remained for future works.

Since this formulation is based on several restrictive assumptions, further extension is needed such as for multiple moving objects, more complex movement other than parabolic translation or shape change in an image sequence.

References

- 1. Belge, M., Kilmer, M.E., Miller, E.L.: Wavelet domain image restoration with adaptive edge-preserving regularization. IEEE Trans. on IP 9(4), 597–608 (2000)
- 2. Osher, S., Ruden, L.I., Fatemi, E.: Nonlinear total variation based noise removal algorithms. Phisical Riview D 60, 259–268 (1992)
- 3. Pizurica, A., Zlokolika, V., Philips, W.: Combined wavelet domain and temporal video denoising. In: Proc. IEEE Intl. Conf. on Advanced Video and Signal based Surveillance (AVSS) (2003)
- 4. Rares, A., Reinders, M.J.T., Biemond, J., Lagendijk, R.L.: A spatiotemporal image sequence restoration algorithm. In: Proc. IEEE Intl. Conf. on IP (2002)
- 5. Kornprobst, P., Deriche, R., Aubert, G.: Image sequence restoration: A PDE based coupled method for image restoration and motion segmentation. In: Burkhardt, H., Neumann, B. (eds.) ECCV 1998. LNCS, vol. 1407, p. 548. Springer, Heidelberg (1998)
- 6. Gangal, A., Kayikçiouglu, T., Dizdariglu, B.: An improved motion-compensated restoration method for damaged color motion picture films. Signal Processing: Image Communication 19(4), 353–368 (2004)
- 7. Gee, T.F., Karnowski, T.P., Tobin Jr., K.W.: Multiframe combination and blur deconvolution of video data. In: Proc. SPIE Image and Video Communications and Processing, vol. 3974, pp. 788–795 (2000)
- 8. Selesnick, I.W., Li, K.Y.: Video denoising using 2d and 3d dual-tree complex wavelet transforms. In: Proc. SPIE Wavelets: Appl. Signal Image Processing X, vol. 5207, pp. 607–618 (2003)
- 9. Akiyama, K., Luo, Z.W., Onishi, M., Hosoe, S.: Restoration of degraded moving image for predicting a moving object. IEE J. Trans. EIS 127(7) (2007) (in Japanese)
- 10. Mallat, S.: A theory for multiresolution signal decomposition: the wavelet representation. IEEE Trans. on PAMI 11(7), 674–693 (1989)
- 11. Daubechies, I.: Ten Lectures on Wavelets. SIAM, Philadelphia (1992)