ABSTRACT
We propose a novel method for velocity estimation and detection of a moving object in an image sequence. This purpose is achieved by using discrete wavelet transform and parallel bank of extended complex Kalman filter in the transform/spatio-temporal mixed-domain. In the mixed-domain, image sequence processing is replaced by 1-dimensional(1-D) complex signal processing. Then, trajectory signal with approximately constant speed and orientation is considered as a bunch of 1-D complex sinusoidal waves. By applying extended complex Kalman filter to each 1-D complex signal, object’s velocity can be estimated. In addition, these estimates can be combined with depending on the energy of the subband images. Through some simulation results, it is shown that moving object’s velocity is accurately estimated and the object is detected effectively.

Key Words: Wavelet Transform, Extended Complex Kalman Filter, Mixed-Domain, Extraction of Moving Objects

1 INTRODUCTION

Under the present condition that computer vision is playing a very important role in many fields, the study for velocity vector estimation and enhancement of moving objects in an image sequence is very interesting.

An object moving with a constant orientation at a certain velocity is considered as a linear trajectory signal(LTS). In addition, moving object along arbitrary trajectory is, if its velocity vector changes smoothly, approximately considered as LTS within a short time interval. The detection problem is solved by designing the frequency selective filter using the fact that the spectral composition of LTS is limited on a specified plane in the 3-dimensional(3-D) frequency domain[1]. But in this 3-D approach, there are some difficulties owing to the multidimensionality, such as system’s instability and number of coefficients.

L.T.Bruton has proposed mixed-domain filtering method that 2-D spatial variables are transformed into 3-D frequency domain, and 1-D temporal variable is processed by linear difference equation[2]. In this manner, it is reported its arithmetic complexity and storage requirements are effectively reduced in comparison with another approach. In addition, another method has been proposed by decomposing the original images into subbands using Discrete Wavelet Transform(DWT), and neglecting some of low power subbands, to save total filtering operations[3][4].

Another approach to extraction of moving objects is to use extended Kalman filter(EKF). In Ref.[5], the application of parallel bank of EKF(PEKF) to the reduction of noise in sequential images containing a moving object and to the estimation of the object’s velocity is carried out. The interest of using PEKF is that 2-D spatial variables are transformed into frequency domain and 1-D temporal variable is processed by EKF, that is, EKF operates on data at a single spatial frequency and estimates both the Fourier coefficients of the image and velocity.

In this paper, we propose a novel method for moving object’s velocity estimation and object’s detection. In our approach, by applying the extended complex Kalman filter to each 1-D complex signal in the mixed-domain, moving object’s velocity can be correctly and rapidly estimated and the object can be effectively detected. In addition, DWT is applied to the original image sequence, and depending on the energy of the subband images, it is selectively processed with the subbands.

2 THE FILTER STRUCTURE

2.1 Analysis of Trajectory Signal in Mixed-Domain

In the mixed-domain, 2-D spatial variables are transformed into frequency domain, and 1-D temporal variable is processed by linear difference equation. In this manner, the trajectory signal is considered as a bunch of 1-D complex sinusoidal waves whose frequency are described as follows,

\[ \omega_{ij} = -(v_x(k)\omega_{x1} + v_y(k)\omega_{y1}) \quad (i, j = 1, 2, \ldots, N) \]
and it is shown in Fig.1, where the spatio-temporal domain index is given by \((i,j,k)\). The extended complex Kalman filter (ECKF) is applied to these waves, to estimate object's velocity and detect moving object.

![Figure 1: Trajectory Signal in the Mixed-Domain](image)

### 2.2 Trajectory Signal Model in Mixed-domain

We assume that the 2-D Fourier transform of the image, \(U_{ij}(k)\), is "measured." Then the measurements, \(U_{ij}(k)\), consist of noisy versions of each of the sequential images in the mixed-domain:

\[
U_{ij}(k) = Z_{ij}(k) + N_{ij}(k) \quad (k = 1, 2, \cdots),
\]

where \(Z_{ij}(k)\) is the 2-D Fourier transform of trajectory signal and \(N_{ij}(k)\) is zero mean, Gaussian white noise (variance \(\sigma^2_n\)). Since \(Z_{ij}(k)\) is a complex sinusoidal wave, it can be written by

\[
Z_{ij}(k) = a \exp(j\omega_{ij}k)
\]

For simplicity of description, spatial location 'ij' is omitted.

Let us consider that ECKF is employed to estimate the state of this system. The state consists of both the velocity of the moving object, \(\omega_t\), and the image in the spatial frequency domain, \(Z(k)\). Then, measurements \(U(k)\) can be written by the following system:

\[
\begin{align*}
X(k + 1) &= g_k(X(k)) \\
U(k) &= h_k(X(k)) + N(k)
\end{align*}
\]

where

\[
X(k) = [X_1(k) \quad X_2(k)]^T = [e^{j\omega_t}Z(k - 1)]^T
\]

\[
Z(k) = h_k(X(k))
\]

\[
g_k(X(k)) = \begin{bmatrix} 1 & 0 \\ 0 & e^{j\omega_t} \end{bmatrix} \begin{bmatrix} e^{j\omega_t}Z(k - 1) \\ X_1(k)X_2(k) \end{bmatrix}
\]

\[
h_k(X(k)) = [Z(k - 1) \quad 0] \begin{bmatrix} e^{j\omega_t} \\ Z(k - 1) \end{bmatrix}
\]

and \(T\) indicates transposition.

### 2.3 Mixed-Domain Filtering with Extended Complex Kalman Filter

In our approach, complex Kalman filter instead of real Kalman filter is applied to 1-D complex signals in the mixed-domain. The use of complex Kalman filter has the following advantages in comparison of real Kalman filter: (a) Procedure of computing complex signal is simple. (b) Amount of multiplication to obtain the desirable estimate is reduced to about 1/2[6]. And the extended complex Kalman filter is nonlinear, recursive filter that can be employed to generate estimates of the state of a system whose evolution is governed by a nonlinear state equation.

The extended complex Kalman filter equations for a system with a state equation (4), additive input noise, and a measurement equation (5) are:

\[
\begin{align*}
\hat{P}(k + 1|k) &= G(k)\hat{P}(k|k)G^*T(k) \\
K(k) &= \hat{P}(k|k - 1)H^*T(k) \\
\cdot [H(k)\hat{P}(k|k - 1)H^*T(k) + 1]^{-1} \\
\hat{X}(k|k) &= \hat{X}(k|k - 1) + K(k)[U(k) - h_k(\hat{X}(k|k - 1))] \\
\hat{X}(k + 1|k) &= \begin{bmatrix} 1 & 0 \\ 0 & \hat{X}_1(k|k) \end{bmatrix}^\dagger \hat{X}(k|k) \\
\hat{P}(k|k) &= \hat{P}(k|k - 1) - K(k)H(k)\hat{P}(k|k - 1),
\end{align*}
\]

where

\[
G(k) = \left. \frac{\partial g_k(X(k))}{\partial X(k)} \right|_{X(k)=\hat{X}(k|k)}
\]

\[
H(k) = \left. \frac{\partial h_k(X(k))}{\partial X(k)} \right|_{X(k)=\hat{X}(k|k - 1)}
\]

\[
\hat{P}(k|k) = \hat{\Sigma}(k|k)/\sigma^2_n, \quad \hat{P}(k + 1|k) = \hat{\Sigma}(k + 1|k)/\sigma^2_n
\]

\[
\hat{\Sigma}(k|k) = \mathbf{E}[(X(k) - \hat{X}(k|k))(X(k) - \hat{X}(k|k))^\dagger]
\]

\[
\hat{\Sigma}(k + 1|k) = \mathbf{E}[(X(k + 1) - \hat{X}(k + 1|k))(X(k + 1) - \hat{X}(k + 1|k))^\dagger]
\]

\(\mathbf{E}[\cdot]\) is the expected value operator and ‘\(^*\)' denotes complex conjugation.

The mixed-domain filtering with ECKF is processed as follows,

(Proposed Algorithm)

1. Trajectory signal \(z_{ij}(k)\) is transformed into the mixed-domain by applying 2-D DFT with respect to \((i,j)\)

\[
Z_{ij}(k) = DFT[z_{ij}(k)]
\]

2. At each 2-tuple \((\omega_{xi}, \omega_{yj})\), ECKF operates on measured Fourier coefficients of the image, to estimate actual Fourier coefficients and temporal frequency \(\omega_{tij}\) of complex sinusoidal wave.
3. From eq.(1), estimates $\hat{\omega}_{tij}$ at each $(\omega_{xi}, \omega_{yj})$ are combined using least-squares algorithm, to yield the velocity estimate $\hat{v} = (\hat{v}_x, \hat{v}_y)$.

4. This velocity estimate is fed back to the individual ECKFs for the next iteration.

5. The obtained output $\hat{F}_{ij}(k)$ is transformed into the image by applying the 2-D inverse DFT with respect to $(\omega_{xi}, \omega_{yj})$,

$$\hat{f}_{ij}(k) = IDFT[\hat{F}_{ij}(k)] \quad (18)$$

Figure 2: Filter Structure

In this process, 3-D filtering for $N \times N$ size of image is accomplished by $N \times N$ ECKFs and the filter structure is shown in Fig.2. The filter consists of a parallel bank of ECKF(PECKF) to yield the final velocity estimate. The individual filters in this bank are just ECKFs for the individual Fourier coefficients, and defined by eqs.(10)-(16).

### 2.4 Combined Discrete Wavelet Transform/Parallel Extended Complex Kalman Filter

In the mixed-domain, velocity estimation must be taken care of the following aspect: (a) At the lower frequency region $(\omega_{xi}, \omega_{yj})$, it is forced to estimate $\omega_t$ more accurately. (b) On the other hand, at the higher frequency $(\omega_{xi}, \omega_{yj})$, desired spectrum tends to be contaminated with noise. Thus, it is reasonable to select some high-energy subbands for effective processing.

The original image is decomposed into subbands by DWT and selectively processed with the subbands. The filter structure with DWT is illustrated in Fig.3.

Using the DWT in this manner provides a controlled mechanism to partition the spectrum of the input signal into subband signals, which then may be selectively filtered during PECKF step of the mixed-domain filtering. The final estimates $\hat{v}$ is given by

$$\hat{v} = c_{LL}\hat{v}_{LL} + c_{LH}\hat{v}_{LH} + c_{HL}\hat{v}_{HL} + c_{HH}\hat{v}_{HH}, \quad (19)$$

where weighting factors $c$ are determined so that the following equation is satisfied, depending upon the energy of the subband images.

$$c_{LL} + c_{LH} + c_{HL} + c_{HH} = 1, \quad (0 \leq c_{LL}, c_{LH}, c_{HL}, c_{HH} \leq 1) \quad (20)$$

### 3 SIMULATION RESULTS

**Example 1**

Fig.4 illustrates the application of the proposed method to a sample sequence of 128 x 128 pixel images. The images in the sequence were corrupted by additive Gaussian white noise(variance $\sigma^2 = 0.3$) as shown in Fig.4(a). We set the motion vector of target LTS at (1.0ppf, -2.0ppf) and $c_{LL} = 1.0, c_{LH} = c_{HL} = c_{HH} = 0.0$. The filtered image is given in Fig.4(b).

![Input image](a)  ![Output image](b)

**Figure 4: Filtering Result 1 (40th frame)**

**Example 2**

In this example, we show the application of the proposed
method to image sequence (128×128 pixel) with noise moving object. We set the motion vector of target object at (2.0ppf,2.0ppf), that of noise object at (-2.0ppf,-2.0ppf), and $c_{LL} = 0.0, c_{LH} = c_{HL} = 0.3, c_{HH} = 0.4$. Input and output images are shown in Fig.5.

![Input image](image1.png) ![Output image](image2.png)

Figure 5: Filtering Result 2 (30th frame)

The time history of velocity estimates is given in Fig.6, for these examples. The velocity estimates converge to the actual velocity vector.

These results show that desired object’s velocity is correctly estimated and target object is effectively enhanced.

4 CONCLUSION

In the proposed method, moving object’s velocity vector is well estimated and desired objects are effectively detected. The main contributions are summarized as follows: 1) by applying ECKF to each 1-D complex signal, both the Fourier coefficients of the image and object’s velocity can be estimated, 2) original image is separated into subbands by DWT, and depending upon the containing energy, these estimates can be combined.

Acknowledgment

This work was supported by a Scientific Grant-in-Aid from the Ministry of Education, Science, Sports and Culture of Japan.

References


