Wigner Distribution for 2D Motion Estimation from Noisy Images

Vinay G. Vaidya

The Boeing Company, P.O. Box 3999, Seattle, Washington 98124

AND

Robert M. Haralick

Intelligent Systems Laboratory, Department of Electrical Engineering, University of Washington, Seattle, Washington 98195

The Wigner distribution has been extensively used for speech processing. However, its use for image processing is new. Past attempts at using the Wigner distribution for motion detection have led to complex and often impractical algorithms. This paper takes a new look at the use of the Wigner distribution, develops necessary theory for its use, presents a new technique for reducing the undesirable cross-term effects, and shows practical applications. Our results are based on 125,000 simulations in a 64 × 64 size frame. The results show that even with a signal-to-noise ratio (SNR) of 5 dB, one can detect rectangular objects of at least 6 pixels in length and width, to within 0.5 pixel location accuracy. The misdetection rate is near zero and the average false detection rate is one false object per frame in a 5-dB SNR environment. This false detection rate is reduced to zero by using the pseudopeak elimination technique developed here. The method is then applied to an image sequence obtained from a real dynamic scene of a 747 airplane taking off. Using the new method, the 747 take-off speed was was predicted to be 142 knots, which is well within the typical range of 140 to 150 knots. Another sequence consisting of cars in motion also gives good results. The results presented here clearly show the successful application of the method to real life situations. © 1993 Academic Press, Inc.

1. INTRODUCTION

Motion estimation has several applications in different disciplines. In order to solve the motion estimation problem, a number of methods have been proposed. Most of these methods are inadequate for handling multiple objects and/or noise.

In this paper we develop a new method using the Wigner distribution to estimate motion parameters. Consider, for example, the image shown in Fig. 1. The problem at hand is to identify where the objects are located. This task becomes relatively simple if the Wigner distribution based method, described in this paper, is used. After the Wigner distribution is computed, the original image is transformed into the image shown in Fig. 2.

From the transformed image one can easily identify object locations by identifying the peaks.

The results presented here clearly show that the method can be successfully applied to real life situations. In this paper we present a performance characterization of the new algorithm. This performance characterization is based on more than 125,000 simulated images. The simulated images are 64 × 64 in size and have three objects in motion. Then we present performance results of real dynamic scenes. These scenes constitute an airplane taking off and cars on a highway.

Outline of the Paper

The problem statement is given as a formal mathematical statement in Section 2. In Section 3, we discuss some of the methodologies prevalent in this field and more directly related to the problem stated here. We begin with the Fourier transform approach and its shortcomings. Then we discuss one of the most widely used methods, namely, optic flow.

In Section 4, we discuss the Wigner distribution. We begin with a definition, followed by its properties. The Wigner distribution is then modified so that it can be used for image processing. One of the main complications about using the Wigner distribution is the cross terms resulting from two or more nearby signals. This in our application causes pseudopeaks. Section 5 is devoted to the discussion of these pseudopeaks and how to overcome their effects.

In Section 6, experimental protocol is discussed. The algorithm for simulating images, type of noise model used, and an algorithm for generating an image sequence are outlined in this section. Then the details of how real dynamic scenes were obtained is discussed.

When any new algorithm is evaluated, it is essential to generate performance characterization curves. Such curves are given in Sections 7 and 8. In Section 7 we also
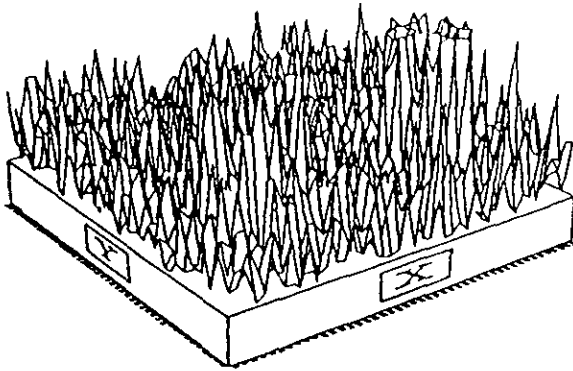
FIG. 1.   Simulated noisy image with SNR = −1 dB.

discuss the algorithm for motion estimation. Results obtained from real dynamic scenes are also analyzed in Section 8. Finally, we conclude this work in Section 9.

2. ASSUMPTIONS AND PROBLEM STATEMENT

The applicability of a method for solving a class of motion estimation problems depends on certain assumptions regarding the images and the world itself. In this section we present such assumptions in detail and give a formalized problem statement.

2.1. Assumptions

Our end goal is to develop a method that can estimate motion parameters from a sequence of images without using any a priori knowledge of the object or the world. We do not use point correspondence, CAD models, or object features. Since we do not consider range data or stereoscopic images, we limit the discussions to the 2D world. The objects are assumed rigid and are free to
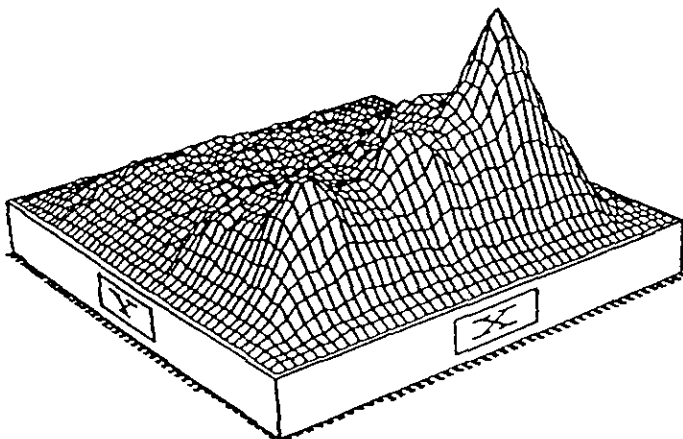

FIG. 2.   Pseudo-Wigner distribution.

translate and/or rotate in a 2D plane. We assume that the images from the 3D world have been projected onto the 2D plane by perspective projection. Motion between successive frames is assumed to be relatively small for perspective projections. This restriction is removed as one takes orthographic projections.

The images are assumed to be grey scale images with gray-scale values greater than 0. The gray-scale values are assumed to be greater than 1 throughout the image. The experimental results presented here use gray-scale ranging from 1 to 255. However, any positive range is acceptable. In the Wigner distribution method one takes the sum of symmetric correlations. If a gray value of 0 is encountered then the product would also result in a zero value. Thus, no clue can be obtained about the value of other nonzero pixel. This reason becomes more apparent in Section 4, where it is presented in detail.

During computation, if the image coordinates go beyond the frame size, it is assumed that the gray-scale values are 0. In other words, the gray-scale values do not repeat outside the frame boundary.

It is equally important to give consideration to image contrast. Thus, we assume that the image contrast is good. The adjective "good," however, needs to be quantitatively stated. When one thinks about detecting features or objects from a given scene, then two questions need to be answered. The first one is how large is the object that you want to detect and the other is how good is the image contrast. Answers to these questions are interdependent. If the object of interest is very small, then the required contrast has to be high. On the other hand, if we are interested in a large object detection, then a relatively poor contrast might suffice. A relationship between the object size and image contrast is derived for a specific case in Section 4. It is further assumed that the objects of interest are brighter than the background. In cases where this condition cannot be met, the images can be inverted in gray scale. Such an example is given in Section 9.

In a multiple object scene, it is assumed that the objects are separated from each other by at least a 2-pixel-wide boundary. In other words, we assume that the objects are not occluded. In a given dynamic scene, one may have some objects only translating, while others may only be rotating, and still more objects may be undergoing rotation and translation simultaneously. It is assumed that the object rotation is around the axis passing through the origin of the image and that the objects do not rotate around the centroid. These rotations and translations can be small or large.

The method presented here is based on three frames. Thus, it is assumed that the rotational and translational velocities of the objects remain constant between the first pair and last pair of frames.

2.2. Problem Statement

The problem can be formalized in a mathematical statement as follows. Given a sequence of 2D perspective projection images, $I_1, I_2, \ldots, I_n$, with m rotating and translating rigid objects, each separated at least by a 2-pixel boundary, to find the translation vector t_m and the rotation vector R_m for all m's where θ_m is the counterclockwise rotation angle related to the rotation matrix by

$$R_m = \begin{pmatrix} \cos \theta_m & -\sin \theta_m \\ \sin \theta_m & \cos \theta_m \end{pmatrix} \tag{1}$$

and the translation vector t_m is given by

$$t_m = \begin{pmatrix} x_m \\ y_m \end{pmatrix}. \tag{2}$$

Let p_{mi} be any point on an object O_m. The point p_{mi} is given by

$$p_{mi} = \begin{pmatrix} x_{mi} \\ y_{mi} \end{pmatrix} \tag{3}$$

Further, let p_{mj} be a point on the object O_m before the motion, and let p_{mk} be a corresponding point on the object after the motion. These two points appear in two images, namely, I_j and I_k. The relationship between the points is given by

$$p_{mk} = R_m * p_{mj} + t_m. \tag{4}$$

3. LITERATURE REVIEW

A number of methods have been proposed for estimating 2D motion when CAD models of the objects are not available. Some of these may be classified as matching, the method of differentials, and the Fourier method (Huang and Tsai [9]).

The method of differentials is a simple method to determine motion. This method, although simple to use, does not always work. The main reason for its failure is the presence of noise in the images and/or any variations in the gray-scale intensities from one image to the other.

The Fourier method involves taking the Fourier transform of two images. The phases of the two images are subtracted, and this result is obtained for two separate frequencies. This leads to a set of two algebraic equations with two unknowns.

Haralick and Shapiro [7] gave a good review of the current work in motion estimation. They also gave an extensive bibliography of the work done.

In this paper the main focus is on estimating motion without point correspondence. There are two methods that are most commonly used for solving this type of problem, namely, the Fourier method and the optical flow method.

3.1. Fourier Transform Approach

The Fourier transform method has been successfully used by Jain [11] for object classification. Rajala et al. [14] used the Fourier transforms for tracking moving objects in noisy environments by averaging the power spectra over the frequencies. Lin et al. [12] made use of the Fourier transform approach for determining motion from a set of points on an object of interest. They made use of the range data from two time instances to determine 3D motion parameters. The method assumed noise-free data and no missing points from one set to the other set. This method is more suitable where range data is available or points on an object from a 2D scene are available.

The Fourier approach, for motion estimation, consists of taking the Fourier transform of the two images and utilizing the phase difference relationship to compute translation. Similarly, there is a simple relationship between space-domain rotation and frequency-domain rotation. Although this approach is simple, it has certain limitations.

Shortcomings of the Fourier Transform Approach

The Fourier transform approach works well for estimating motion parameters from images with a single object. If the image has multiple objects, this method cannot distinguish the motion of one object from that of the other. Thus, it cannot estimate the motion of each object individually. In addition to this problem, the Fourier method does not give correct results for images with a nonuniform background (Huang and Tsai [9]).

With the Fourier approach, once space-domain information is converted into frequency domain, the new function has only frequency variables. The information about space-domain variables is lost in the frequency domain. This is a major drawback.

The Wigner distribution based method, discussed in Section 4, overcomes these problems of the Fourier approach.

3.2. Optic Flow Method

A vast amount of literature is available on optic flow. Recently, Willick and Yang [18] evaluated motion constraint equations. They evaluate Horn and Schunk's [8], Schunk's [15], and Nagel's [13] equations. In the following discussion we elaborate on a method for computing optic flow using the Wigner distribution.

Wigner Distribution for Optical Flow Computation

Jacobson and Wechsler [10] were the first ones to apply the Wigner distribution to image processing problems. They use the Wigner distribution to compute optic flow. Their approach is a modification to Gafni and Zeevi's [6] approach to optical flow computation. Gafni and Zeevi used Fourier transform, which gives a single velocity for the entire image. This limitation was realized by Jacobson and Wechsler. Thus, they proposed replacing the Fourier transform with the Wigner distribution.

The following steps are required to compute optical flow using Jacobson and Wechsler's method.

1. Given a time-varying sequence of images, $f(x, y, t)$, compute its 3D Wigner distribution for a selected value of (x, y, t):

$$W_f(\omega_x, \omega_y, \omega_t; x, y, t)$$

$$= \int_\infty^\infty \int_\infty^\infty \int_\infty^\infty \left[f\left(x + \frac{\alpha}{2}, y + \frac{\beta}{2}, t + \frac{\tau}{2}\right) \right. \quad (5)$$

$$\left. f\left(x - \frac{\alpha}{2}, y - \frac{\beta}{2}, t - \frac{\tau}{2}\right) e^{-j(\alpha\omega_x + \beta\omega_y \tau\omega_t)} \right] d\alpha \, d\omega \, d\beta.$$

2. Compute the modulus of W_f and raise it to the power n. In their paper, Jacobson and Wechsler use 2 as the value of n:

$$|W_f(\omega_x, \omega_y, \omega_t; x, y, t)|^n. \quad (6)$$

3. Select a set of integer values for velocities V_x and V_y.

4. Compute the velocity polling function C_f for the selected values of V_x and V_y. The function C_f is given by

$$C_f(V_x, V_y; x, y, t)$$

$$= \int_\infty^\infty \int_\infty^\infty \int_\infty^\infty |W_f(\omega_x, \omega_y, \omega_t; x, y, t)|^n$$

$$\delta(V_x\omega_x + V_y\omega_y + \omega_t)d\omega_x \, d\omega_y \, d\omega_t. \quad (7)$$

5. Repeat steps 3 and 4 until all values of V_x and V_y have been tried. If the range of possible values of V_x and V_y is from -5 to $+5$ then step 4 needs to be computed 121 times.

6. Find the velocity coordinate of peak in C_f. This gives optical flow

$$V_f(x, y, t) = \left(\begin{bmatrix} V_x(x, y, t) \\ V_y(x, y, t) \end{bmatrix} \right). \quad (8)$$

7. Repeat steps 1–6 for all (x, y, t).

Within the main loop there is another loop for discrete values of V_x and V_y which needs to be repeated 121 times

for a range of -5 to $+5$. If the user-guessed range of velocity values is incorrect, the correct answer will not be obtained. This is a serious shortcoming of this method. Jacobson and Wechsler [10] reported results based on simulated grating images and only considered uniform translation. The effect of noise was not investigated. They did not use images with a distinct foreground and background. The error obtained in calculating velocities is nonzero, although there is no noise present. The paper does not mention the problem of cross terms which is inherent to the Wigner distribution. When these factors are taken into consideration, it becomes clear that the method is not well suited for general purpose use. Their approach, nevertheless, makes a significant step in making use of the *spatiotemporal–frequency representation* for image processing.

4. WIGNER DISTRIBUTION

The Wigner distribution is a generalized time–frequency representation. In some sense it is a true representation of the phenomenon that takes place everyday in nature. When we watch the sunset, we are unaware of the fact that what we watch is nothing but the frequency variation as a function of time. The characteristic of the Wigner distribution, to be a function of both time and frequency, is remarkable. The Fourier transform, on the other hand, is strictly a function of frequency. Thus, once we convert a function using the Fourier transform, we strictly deal in the frequency domain. If we take the inverse Fourier transform, we end up strictly in the space domain. The Wigner distribution, on the other hand, allows one to simultaneously use the time- and frequency-domain information.

Wigner [17], in 1932, proposed this function for the study of quantum mechanics. Ville [16] proposed it again in 1948. However, researchers did not pay much attention to the method until the 1980s. Researchers, in the speech processing area, extensively used this concept in the 1980s.

Although the concept of mixed representation has been applied extensively to the signal processing area, research papers applying the Wigner distribution to the image processing area are limited. Jacobson and Wechsler [10] modify Gafni and Zeevi's [6] approach for calculating optic flow using the Wigner distribution.

Claasen and Mecklenbrauker [4, 5] published a series of papers on the Wigner distribution in order to facilitate and promote the use of this technique in other areas of research. These papers put emphasis on the properties of the distribution so that this concept of a generalized space–frequency representation could be applied in several disciplines. The following sections, 4.2 and 4.3, closely follow Claasen and Mecklenbrauker's papers [4, 5].

4.1. Definition of the Wigner Distribution

The Wigner distribution of two signals, $f(t)$ and $g(t)$, is defined as

$$WD_{f,g}(t, \omega) = \int_{-\infty}^{\infty} e^{-j\omega k} f(t + k/2) g^*(t - k/2) \, dk, \quad (9)$$

where ω is the frequency, t is time, and g^* is the complex conjugate of the function $g(t)$. The auto-Wigner distribution is given by

$$WD_f(t, \omega) = \int_{-\infty}^{\infty} e^{-j\omega k} f(t + k/2) f^*(t - k/2) \, dk. \quad (10)$$

The auto-Wigner distribution for a real function $f(t)$ is given by

$$WD_f(t, \omega) = \int_{-\infty}^{\infty} e^{-j\omega k} f(t + k/2) f(t - k/2) \, dk. \quad (11)$$

One can also define the WD for the spectra F and G by

$$WD_{F,G}(\omega, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\eta t} F(\omega + \eta/2) G^*(\omega - \eta/2) \, d\eta. \quad (12)$$

In the discrete domain the Wigner function is defined as

$$WD_{f,g}(t, \omega) = 2 \sum_{k=-\infty}^{\infty} e^{-2j\omega k} f(t + k) g^*(t - k). \quad (13)$$

The auto-Wigner distribution is defined as

$$WD_f(t, \omega) = 2 \sum_{k=-\infty}^{\infty} e^{-2j\omega k} f(t + k) f^*(t - k). \quad (14)$$

If the function f is real then the auto-Wigner distribution is defined as

$$WD_f(t, \omega) = 2 \sum_{k=-\infty}^{\infty} e^{-2j\omega k} f(t + k) f(t - k). \quad (15)$$

The above equation would be useful in the development of the Wigner distribution for image processing.

Properties of the Wigner Distribution

Claasen and Mecklenbrauker [4, 5] gave a number of properties of the Wigner distribution. Here, we discuss only a few properties related to the present work.

First, for any two signals $f(t)$ and $g(t)$,

$$WD_{f,g}(t, \omega) = WD^*_{g,f}(t, \omega). \quad (16)$$

Hence, the WD of any function, real or complex, is always real.

Another way of arriving at the same conclusion is as follows. Since

$$\sin(-k\theta) = -\sin(k\theta) \quad (17)$$

$$\sum_{k=-\infty}^{\infty} \sin(-k\theta)[f(x + k) * f(x - k)] = 0. \quad (18)$$

Therefore, the Wigner distribution is always a real function.

The WD of a real function is an even function of the frequency:

$$WD_f(t, \omega) = WD^*_f(t, -\omega). \quad (19)$$

The WD is a spectrum of the symmetric correlation function $f(t + k/2) g^*(t - k/2)$ which is a function of k for a fixed value of t. Thus, the WD can be computed by using the FFT technique.

Since the Wigner distribution is always a real function, Eq. (15) can be written as

$$WD(t, \omega) = 2 \sum_{k=-\infty}^{\infty} \cos(2\omega k) f(t + k) f(t - k). \quad (20)$$

4.2. Pseudo-Wigner Distribution

The concept of windows is widely used in signal processing. The same concept can be applied to the WD. This was first proposed by Claasen and Mecklenbrauker [4]. As we see later, this concept is very useful for the image processing application presented here. The windowing in the time domain involves applying a sliding window to the functions $f(t)$ and $g(t)$. Let w_f and w_g be such windows. Then at time t,

$$f_t(\tau) = f(\tau) w_f(t - \tau) \quad (21)$$

and

$$g_t(\tau) = g(\tau) w_g(t - \tau). \quad (22)$$

The Wigner distribution WD_{f_t, g_t} is the convolution of $WD_{f,g}$ and WD_{w_f, w_g} and is given by

$$WD_{f_t, g_t}(\tau \, \omega)$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} WD_{f,g}(\tau, \eta) WD_{w_f, w_g}(\tau - t, \omega - \eta) d\eta, \quad (23)$$

where t is a parameter indicating the position of the sliding window. The above equation also makes it clear that

it gives the whole family of WDs. To get an individual WD at time $\tau = t$, one would use

$$WD_{f_t,g_t}(t\ \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} WD_{f,g}(t, \eta) WD_{w_f, w_g}(0, \omega - \eta) d\eta$$

(24)

This family of WD's is called the pseudo-Wigner distribution.

In our application the function $f(t)$ is real and discrete. If we apply a window w of duration $2d + 1$, then

$$PWD_f(t, \omega)$$

$$= 2 \sum_{k=-\infty}^{\infty} \cos(2\omega k) w(k) f(t + k) w(-k) f(t - k).$$ (25)

Since the duration of the window is $2d + 1$,

$$PWD_f(t, \omega)$$

$$= 2 \sum_{k=-d}^{d} \cos(2\omega k) w(k) f(t + k) w(-k) f(t - k).$$ (26)

4.3. The Wigner Distribution for Image Processing

To use the Wigner distribution function for image processing, one needs to extend it to two-dimensional space. Such an extension results in a four-dimensional Wigner distribution function. The function has two space-domain variables x and y, and two frequency-domain variables u and v. The extension to 2D space is then

$$WD(x, y, u, v)$$

$$= \frac{4}{MN} \sum_{l=-N/2}^{N/2} \sum_{k=-M/2}^{M/2} e^{-j4\pi[uk/M+vl/N]} f(x + k, y - l)$$

$$* f(x - k, y - l)$$ (27)

$$= \frac{4}{MN} \sum_{l=-N/2}^{N/2} \sum_{k=-M/2}^{M/2} \cos(\theta) f(x + k, y + l)$$

$$* f(x - k, y - l),$$

where M is the number of columns, N is the number of rows, $\theta = 4\pi[uk/M + vl/N]$, and f is the gray-scale function.

4.4. Information from the Peaks in the Wigner Distribution

The Wigner distribution, as defined from 2D space, becomes a function of four variables. To plot a function one needs a five-dimensional space. To overcome this difficulty one can keep three variables constant and plot the function in 2D space.
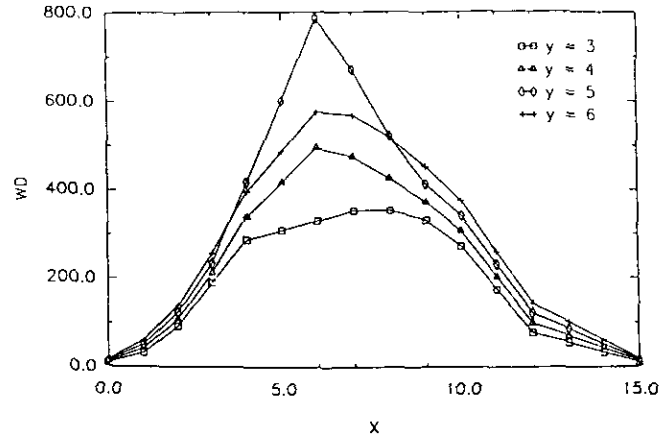

FIG. 3. Wigner distribution curves for a constant y.

Let us set the frequency variables, u and v, to 0. This gives stationary contributions of the signal. Such a contribution can be plotted as a 2D graph where x or y is held constant. Figure 3 shows such a graph for y equal to a constant. Thus, one obtains a series of curves for a constant y. These curves reach a peak and then start going down. The peaks represent points on the object. For symmetric objects with uniform gray-scale values the highest peak occurs at the centroid of the object. The pixels on the boundary may not have sharp peaks.

Let us review some of the assumptions made. We assumed that the gray scale is greater than 0, and the gray scale for the object is greater than the background. This means that when one travels from left to right along a constant-y line, the value of the WD is the highest at the centroid of the object.

4.5. Analytical Development

In this section we develop a relationship between the object size and the gray-scale contrast. This relationship is required to effectively use the Wigner distribution method.

Let G be the background gray-scale value and F be the foreground gray-scale intensity. Let us assume that the background and foreground are uniform throughout the image. The difference between the foreground and background is denoted by d. It is assumed that the foreground intensity is greater than the background intensity. Thus,

$$F = d + G.$$ (28)

Let γ be the contrast ratio defined as
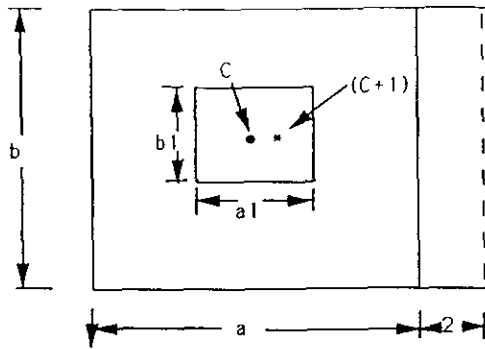
$$\gamma = \frac{d}{G}.$$ (29)

FIG. 4.   General case. Object size $a1 \times b1$. Frame size $a \times b$. Object centroid is C.

Let u and v in the Wigner distribution calculations be 0. For simplicity in the equations let W be defined as follows:

$$W = \frac{WD}{4/MN}. \tag{30}$$

Let M be the number of columns and N be the number of rows.

General Case for the Wigner Distribution

Figure 4 shows an object of size $a_1 \times b_1$ located in a frame of size $a \times b$. Let C be the point at which the Wigner distribution is calculated first. This point is a center of a region $a_1 \times b_1$. The next point at which the WD is calculated is $(C + 1)$. Since we traveled 1 pixel along the x axis, we have added $2b$ pixels in the calculation of the WD. We want to calculate the slope of a WD curve at point $(C + 1)$. Inside the object we have foreground gray-scale values and outside the object we have background gray-scale values. The WD at point C is given by

$$W_C = (ab - a_1b_1)G^2 + a_1b_1(G + d)^2 \tag{31}$$

and the WD at $(C + 1)$ is

$$W_{C+1} = [(a_1 - 2)b1(1 + 2\gamma + \gamma^2) + 4b_1(1 + \gamma)$$
$$+ (a + 2)b - (a_1 + 2)b_1]G^2. \tag{32}$$

The slope at point $(C + 1)$ is given by

$$\text{slope} = (2b - 2b_1\gamma^2) * G^2. \tag{33}$$

When using the WD curves, we look for peaks. These peaks should occur at the centroid C. Thus, for a peak to exist at point C, the slope at point $(C + 1)$ should be less

than or equal to zero. For slope to be less than or equal to 0,

$$2b - 2b_1\gamma^2 \le 0. \tag{34}$$

This gives the following condition:

$$\gamma^2 \ge \frac{b}{b_1}. \tag{35}$$

The highest γ can be obtained by

$$\gamma^2 \ge \frac{\text{maximum (rows, columns)}}{\substack{\text{minimum (length or width of the} \\ \text{smallest object to be detected})}}. \tag{36}$$

General Case for the Pseudo-Wigner Distribution

For the pseudo-Wigner distribution let x_w and y_w be the window size in the x and y directions. This window is assumed to be larger than the object size. The equations for this case become

$$W_C = (x_wy_w - a_1b_1)G^2 + a_1b_1(G + d)^2 \tag{37}$$

and the WD at $(C + 1)$ is

$$W_{C+1} = [(a_1 - 2)b1(1 + 2\gamma + \gamma^2) + 4b_1(1 + \gamma)$$
$$+ x_wy_w - (a_1 + 2)b_1]G^2. \tag{38}$$

The slope at point $(C + 1)$ is given by

$$\text{slope} = (-2b_1\gamma^2) * G^2. \tag{39}$$

As before, for the slope to be less than 0,

$$-2b_1\gamma^2 < 0. \tag{40}$$

This gives a new condition for the PWD. The condition is

$$\gamma^2 > 0. \tag{41}$$

As can be seen, this condition is far less restrictive than the condition for the regular Wigner distribution.

The Selection of Frequencies and the Use of Windows

The WD for image processing is a function of four variables, namely, x, y, u, and v. Our end goal is to develop a technique whereby one can observe the peaks and find out object locations. Due to physical limitations, we can visualize only 3D information. Thus, we need to keep two variables constant. One way is to keep u and v, the frequency variables, constant. Another way, of course, is to keep x and y constant. If we keep x and y constant,

then for each pair of (x, y) we need to compute the WD for every discrete value of u and v. Then we have to plot the WD as a function of u and v for every single pixel. This means analyzing as many plots as the number of pixels in an image. This certainly makes the other option of keeping frequencies constant more appealing. Having decided to view the WD as a function of x and y by keeping the u and v constant, the next question is how does one go about selecting these frequencies.

Windows can be looked at from a statistical estimation point of view. Quite often in statistical estimation theory one uses the concept of weights. Windows or the pseudo-Wigner distribution can be looked upon as a means to apply weights to each pixel value. The simplest window is a 2D rectangular window. This means that the pixels within the window have a complete influence on the WD computation. However, the pixels outside of the window have zero influence. Considering the kernel values as weights applied to individual pixels, we have come up with different possible 2D windows. Each of these cases are evaluated as stated in the Experimental Protocol section. The results of the experimental evaluation are reported in the Results section.

5. PSEUDOPEAKS

In the last section we saw how to use the Wigner distribution to represent each object by a peak. The method works well so long as there is only one object in a given frame. However, if there are multiple objects in a frame then this results in a number of peaks larger than the number of objects. The additional peaks, which do not correspond to any objects, are termed pseudopeaks. In this section we see the reasons for the occurrence of pseudopeaks and how to eliminate them.

5.1. Pseudopeaks in the Wigner Distribution

Consider, for example, two symmetrically located objects such as these shown in Fig. 5. In Fig. 5 we see two rectangular objects of size 3×3 simulated in a 16×16 frame. The object gray scale is 50 and the background gray scale is 10. These objects are symmetrically located around the point $(7, 7)$. Figure 6 shows curves for this simulation. In the curves, there are two peaks located at $(4, 4)$ and $(10, 10)$. The peaks correctly estimate the object centroids. In addition, there is also a peak located at the point $(7, 7)$. In reality, however, no object is located at this point.

If a Wigner distribution is calculated at the point $(7, 7)$, due to the symmetric location of both objects, the WD peaks at this point. Although there is no object at the point $(7, 7)$ the result here indicates that there is an object located at the point $(7, 7)$. Thus, the peak at the point $(7, 7)$ is termed a pseudopeak.
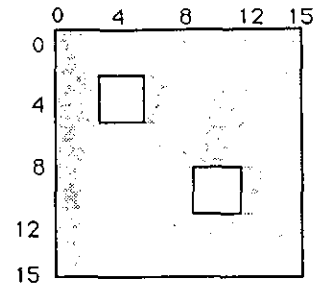

FIG. 5. Two symmetrically located objects in a 16×16 frame.

Pseudopeaks can also occur at the center of the frame. The WD is maximum at the frame center since the maximum number of pixels are involved in the calculation. For a nonzero positive gray scale, the sum of the products would be maximal if the sum were carried out over a maximum possible range. This makes the center a peak point in the WD curves. If there is no object present at the frame center, then such a peak is also a pseudopeak.

5.2. Mathematical Explanation for Pseudopeaks

From Section 4 we know that

$$WD_f(x, \omega) = \int_{-\infty}^{\infty} e^{-j\omega k} f\left(x + \frac{k}{2}\right) f\left(x - \frac{k}{2}\right) dk. \quad (42)$$

Let $p(x)$ and $q(x)$ be two functions such that

$$f(x) = p(x) + q(x)$$

$$WD_{p+q}(x, \omega) = \int_{-\infty}^{\infty} e^{-j\omega k} \left[p\left(x + \frac{k}{2}\right) q\left(x + \frac{k}{2}\right) \right] \quad (43)$$

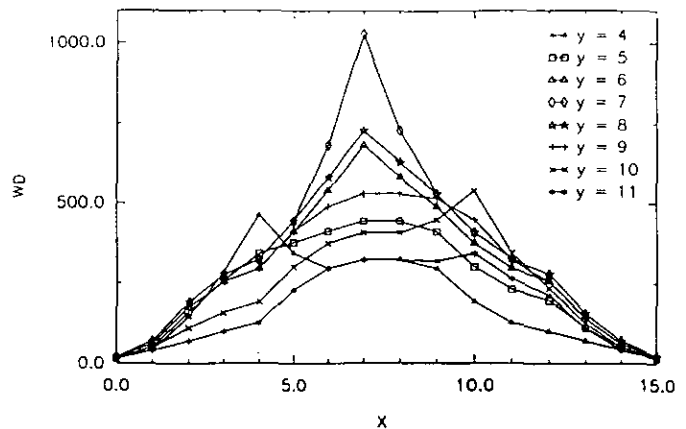$$* \left[p\left(x - \frac{k}{2}\right) + q\left(x - \frac{k}{2}\right) \right] dk.$$

FIG. 6. Wigner distribution curves for two objects. A pseudopeak occurs at $(7, 7)$.

Thus,

$$WD_{p+q}(x, \omega) = WD_p(x, \omega) + WD_q(x, \omega)$$
$$+ WD_{p,q}(x, \omega) + WD_{q,p}(x, \omega). \quad (44)$$

In the above equation, the last two terms, $WD_{p,q}(x, \omega)$ and $WD_{q,p}(x, \omega)$, show the interaction of the two signals p and q. This interaction is the cause of pseudopeaks. It is commonly referred to as cross terms in the signal processing literature.

The next question is, how can one eliminate these pseudopeaks. This is the topic of the next section.

5.3. Elimination of Pseudopeaks

To eliminate the peak we need to shape the kernel in such a way that this contribution to the WD is nullified.

Figure 7 shows the projection of objects shown in Fig. 5. This projection is along the $y = 7$ line. When the WD is calculated at the point $(7, 7)$, the contribution from the objects would occur for $k = 2, 3,$ and 4. For this range of k, $(x - k)$ ranges from 3 to 5. This is the object location for the first object. Also, for the same values of k, $(x + k)$ ranges from 9 to 11. This is the object location for the second object.

Let the kernel values be C_1, C_2, and C_3 for $k = 2, 3,$ and 4, respectively. Further, let us assume that the kernel values for negative k's are same as C_1, C_2, and C_3. Let α be the sum of the contributions to the WD when k is 2, 3, and 4. This is given by

$$\alpha = C_1 * f_9 * f_5 + C_2 * f_{10} * f_4 + C_3 * f_{11} * f_3. \quad (45)$$

Let the contribution to the WD be β when k is $-2, -3,$ and -4. β is given by

$$\beta = C_1 * f_5 * f_9 + C_2 * f_4 * f_{10} + C_3 * f_3 * f_{11}. \quad (46)$$

Now if we add α and β, the net contribution should be zero. This would eliminate pseudopeaks. If the grayscale intensity of the two objects is uniform, then
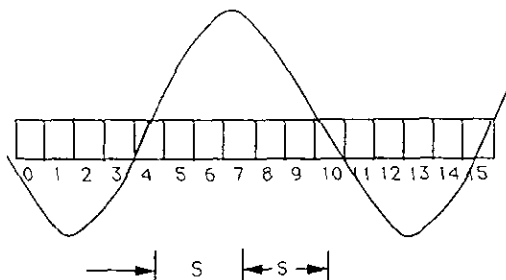

FIG. 7. Kernel shaping. A cosine curve is passed through the pixels.

$$f_3 = f_4 = f_5 \quad (47)$$

and

$$f_9 = f_{10} = f_{11}. \quad (48)$$

Therefore, in the above equations for α and β, we have

$$f_5 * f_9 = f_{11} * f_3. \quad (49)$$

Thus, to make α equal to zero we need to meet the following kernel values:

$$C_1 = -C_3 \quad (50)$$

$$C_2 = 0. \quad (51)$$

Such a kernel would eliminate pseudopeaks. The next question is how can one shape the kernel to achieve the above described result. This is described in the next section.

5.4. Kernel Shaping for Pseudopeak Elimination

In Fig. 7, one object spans from pixel 3 to pixel 5. The centroid of this object is at pixel 4 and is referred to as point A in the following discussion. The second object spans from pixel 9 to pixel 11 and its centroid is at pixel 10, which is referred as point B. A and B are symmetrically located with respect to point C i.e., pixel 7. After computing the Wigner distribution, one would see peaks at points A, B, and C. The distance between A and C is S. Due to symmetry, the distance between points B and C is also S. We want the kernel to have positive values in half of the region and negative values in the other half. From the Wigner distribution equation we know that the kernel is a cosine wave form. The desired results may therefore be achieved by passing the cosine through the center of the object. The cosine would have a zero value at the center of the object. In one-half of the object region, the cosine would be positive and in the other half it would be negative. Thus, let

$$\theta_A = -\frac{\pi}{2} \quad (52)$$

$$\theta_B = \frac{\pi}{2} \quad (53)$$

and

$$\theta_C = 0. \quad (54)$$

If there is indeed an object located at C, we do not want its contribution to be nullified. This is achieved by setting $\theta_C = 0$. At point B we have

$$\theta_B = \frac{\pi}{2}. \tag{55}$$

Therefore,

$$\frac{\pi}{2} = 4\pi \left[\frac{uk}{M} + \frac{vl}{N} \right]. \tag{56}$$

For $v = 0$, and $k = S$ we get

$$\frac{\pi}{2} = 4\pi \frac{uS}{M}. \tag{57}$$

Therefore,

$$u = \frac{1}{8} \frac{M}{S} \tag{58}$$

and

$$v = 0. \tag{59}$$

The above equations give frequency values for shaping the kernel to eliminate pseudopeaks.

5.5. Kernel Modification to Eliminate Pseudopeaks at the Frame Center

The idea behind the kernel modification to eliminate pseudopeaks at the frame center is similar to the one outlined in the previous section. In this case the kernel should look like the one shown in Fig. 8. At the center, the angle θ should be 0 and at the edges it should be π. Thus, for $v = 0$ and $k = M/2$,
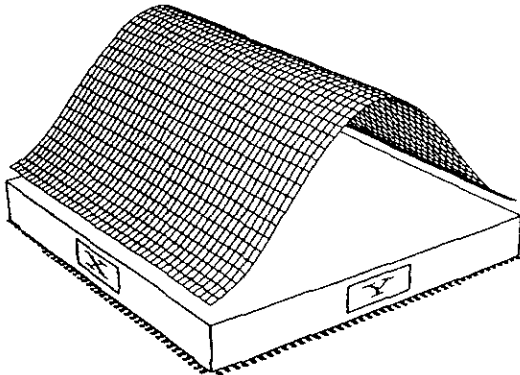
$$\theta = \pi = 4\pi \frac{u[M/2]}{M}. \tag{60}$$

FIG. 8. Kernel shaping to eliminate the pseudocenter peak.

Therefore,

$$u = 0.5 \tag{61}$$

and

$$v = 0. \tag{62}$$

5.6. Procedure Summary

The procedure used here for eliminating pseudopeaks may be summarized as follows. Find all peaks appearing in the WD plot. Using this set of peaks, find out groups of three peaks such that one peak out of these three is exactly in between the other two. Using S, the distance between the farthest peaks in a group, calculate the frequency u by using the formula

$$u = \frac{1}{8} \frac{M}{S} \tag{63}$$

and set v to 0. Use these frequencies and calculate the WD. Check the location where the peak in doubt occurred. If a peak appears again, then the peak is a true peak. Otherwise it was a pseudopeak. Repeat this procedure for every peak in question.

6. EXPERIMENTAL PROTOCOL

To test the new Wigner distribution based method, we carried out experiments on simulated as well as real images. The experimental data may be divided into three categories. The first set of data consists of simulated images used for characterizing performance error in detecting centroids of rectangular objects. In this set we have only single-frame images and an image sequence is not simulated.

In the second set of data we have simulated an image sequence consisting of three frames. The number of objects per frame is 3. In this set we simulate two rectangular objects that are rotating and translating. This data set is used to characterize the translational and rotational error.

The third data set has a real dynamic image sequence. The images include traffic scenes and the airplane take-off scenes.

Simulated Image Generation

The simulated images, used for experiments presented here, are 64 × 64 in size. The objects generated in these images translate and rotate within the 64 × 64 size, and all the objects remain in view at all times. The basic parameters of an image are the background gray-scale value, the contrast ratio, and the foreground gray-scale

value. The generated objects are rectangles rotated between 0° and 180°.

Real Images

The real image sequences used here constitute images from traffic scenes and airplane take-off scenes. These images were taken from a building on Highway 99 near Seattle–Tacoma International Airport in Seattle, Washington. The camera used for capturing the images is "palmcorder" made by Panasonic. The recorded images were transferred from a VHS cassette to a real-time data disc made by Applied Memory Technology.

Camera Calibration

For the images described above, calibrating the camera is the most difficult task. However, we have tried to calibrate it with the available information. In case of some traffic scenes we found some markers used for dividing lanes. These are circular in shape and are separated by about 3 ft. We measured these marks on Highway 99 near Seattle–Tacoma International Airport, Seattle, Washington. The same marks were measured by displaying the
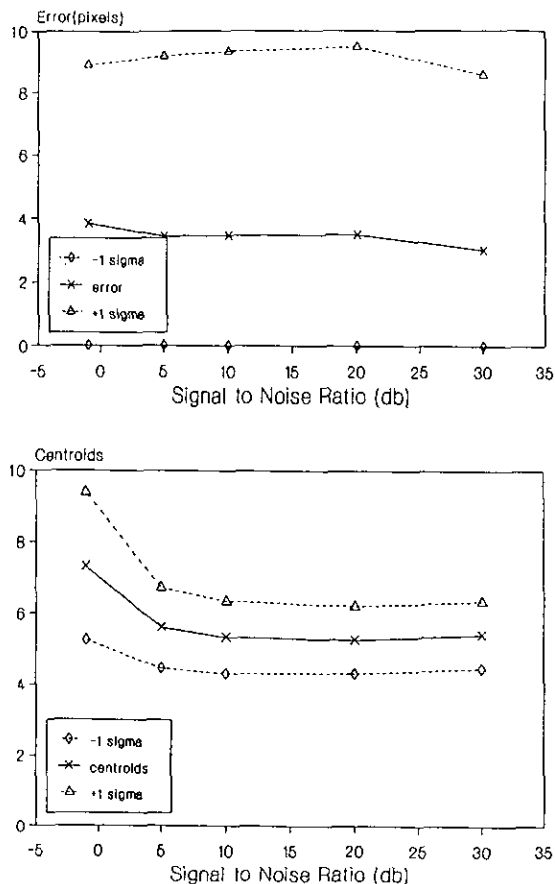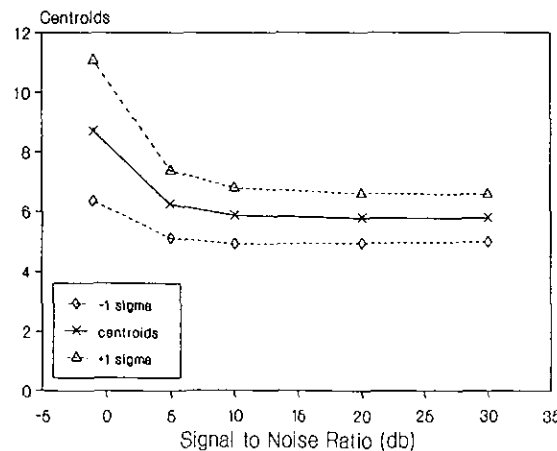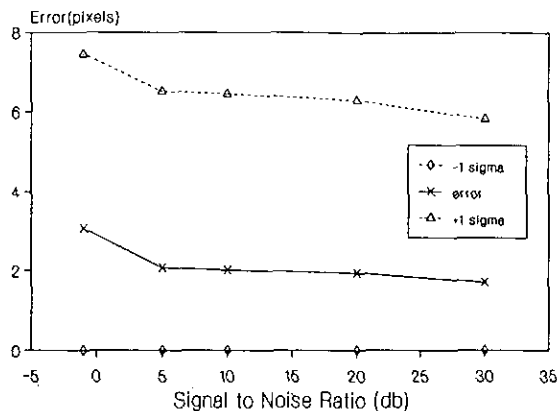

FIG. 10. Case B: $\gamma = 2.8$, $u = 0.25$, $v = 0$, *WD*.

images on a screen. This gave us the pixels-per-inch ratio.

In the case of airplanes, we identified the type of airplane and obtained its structural data for calibrating the images.

Performance Measure

The first set of data consists of simulated images. The objective of these, the error in estimating object centroids, averaged over 1000 experiments, is plotted as a function of the signal-to-noise ratio. The standard deviation, σ_{error}, of the error is also calculated. Two more curves representing (error + σ_{error}) and (error − σ_{error}) are also plotted.

Another important criterion for algorithm evaluation is the number of computed centroids. This is a measure of the elimination of pseudopeaks. We know that there were only three objects. Thus, any number higher than 3 comes from pseudopeaks. The number of centroids is plotted as a function of the signal-to-noise ratio. The standard deviation of the number of centroids, σ_{cent} is also calculated. Two more curves indicating (centroids + σ_{cent}) and (centroids − σ_{cent}) are also plotted.

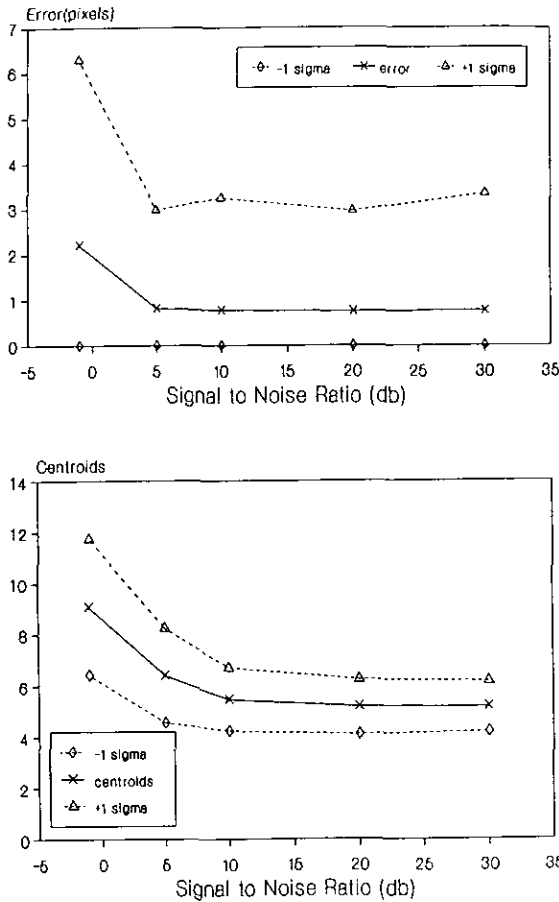**FIG. 9.** Case A: $\gamma = 2.8$, $u = 0$, $v = 0$, *WD*.

FIG. 11.   Case C: $\gamma = 2.8$, $u = 0.5$, $v = 0$, *WD*.

7.   PERFORMANCE CHARACTERIZATION AND THE ESTIMATION ALGORITHM

Performance Characterization of the Object Location Algorithm

A brief discussion of each case is given below. The contrast ratio, γ, is defined as the ratio of foreground to background gray scale. For all the cases the contrast ratio, γ, is kept constant at 2.8. The top graph in every figure shows absolute error, averaged over 1000 trials, versus signal-to-noise ratio (SNR). In addition, it shows two curves, one for $+\sigma$ and the other for $-\sigma$. These curves indicate deviation of the error from the average. The minimum value of the $-\sigma$ curve was restricted to 0, since negative absolute error is not meaningful. The bottom graph shows average number of centroids computed by the algorithm. This graph also has two additional curves showing deviation from the average. In all these cases the number of centroids should be 3, since there were three objects in each image. However, due to pseudopeaks and noise we see more peaks than the actual number of objects. This indicates a false detection

rate. In all of the following cases, the misdetection rate is 0.

Case A.   In this case the frequencies u and v are set to 0 and the Wigner distribution is computed. This case evaluates the performance of the symmetric correlation. Figure 9 shows that the average object location error is within 3 to 4 pixels and remains relatively constant for all values of SNR. The curve showing $+\sigma$ deviation indicates that the error could go as high as 9 pixels. The number of centroids detected by this algorithm is more than 7 for -1 dB SNR. This further goes down to about 6. The actual number of objects was 3. Thus, the false detection rate is about 3. This is indeed high. In our application this means that we will have three additional false objects in motion.

Case B.   This case is used to evaluate a kernel with $u = 0.25$ and $v = 0$. Figure 10 shows that the average object location error and the number of centroids. This case performs better in terms of location error since the error is now reduced to 2 pixels compared to about 3 to 4 in Case A. However, its performance in terms of false
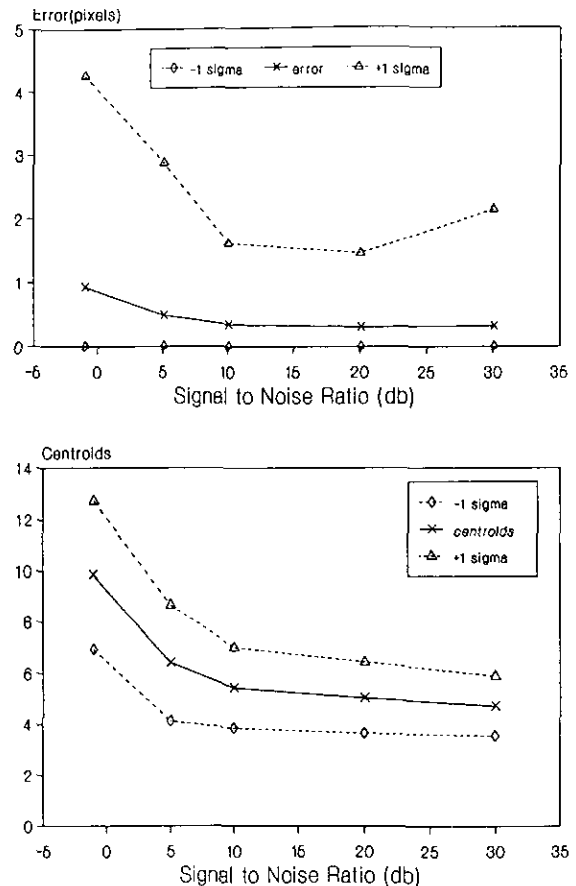

FIG. 12.   Case D: $\gamma = 2.8$, $u = 0.125$, $v = 0.125$, window size = 10, *PWD*.
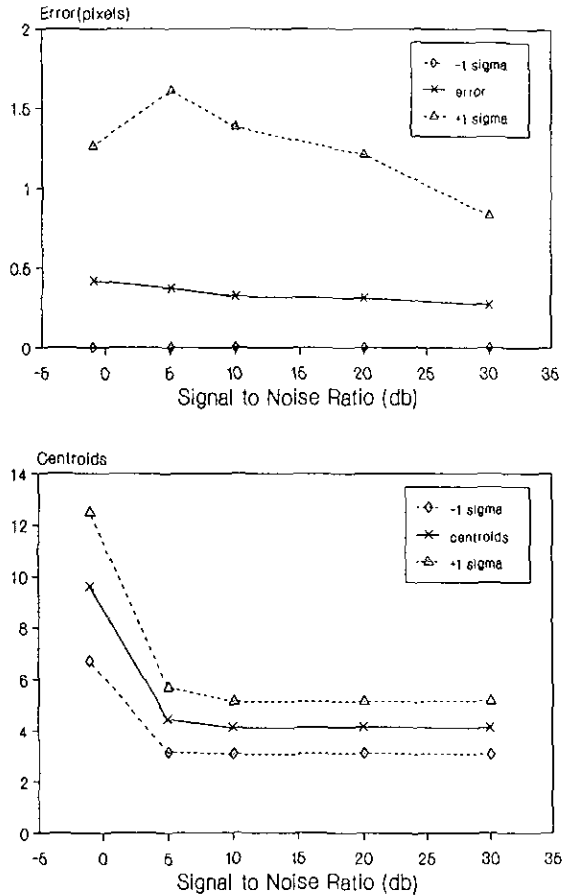
FIG. 13. Case E: $\gamma = 2.8$, $u = 0.4$, $v = 0.4$, window size $= 10$, PWD.

detection rate is the same as in Case A. We still have about three false objects per frame.

Case C. In this case u is changed to 0.5. All other parameters are same as in Case B. In Fig. 11 we see that the error is further reduced to 1 pixel. However, the false detection rate remains around 3.

Case D. In this case we evaluate the use of the pseudo-Wigner distribution (PWD). We use a window size of 10 pixels and u and v are 0.125 each. From Fig. 12 we see that the location error is less than 0.5 pixels for a SNR of 5 dB and higher. However, the false detection rate continues to be 3 or higher. The location error is approximately 1 for a SNR of -1 dB.

Case E. Here we evaluate $u = 0.4$, and $v = 0.4$ with the PWD. The window size is selected to be 10 pixels. This kernel gives better results than the one used in Case D.

The location error is less than half that for a SNR of -1 dB and higher. This case is the best case if we want to have minimum location error. The false detection rate is about 1 for a SNR of 5 dB and higher. For a SNR of -1

dB we get an excessively high false detection rate, which is 7. The curves for this case are shown in Fig. 13.

The Effect of Contrast Ratio on Performance

Figure 14 shows performance curves as a function of contrast ratio. It is seen that as the contrast ratio is increased, the error in object location goes down. In this case the pseudo-Wigner distribution with a window size of 10 × 10 is used.

Performance curves for contrast ratios 3.3, 4.3, and 5.3 were also analyzed and are similar to the ones given here.

ALGORITHM. From the above set of five cases one was selected for evaluating the motion estimation algorithm. This case uses the PWD with a window size equal to 10 pixels. The frequencies u and v are given by

$$u = \frac{1}{8}\left(\frac{M}{2S}\right) \tag{64}$$

and

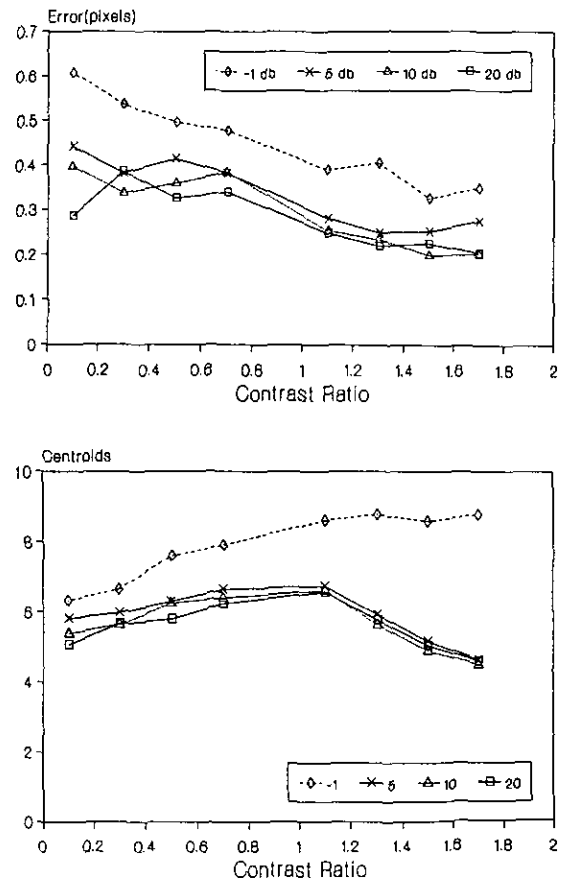$$v = \frac{1}{8}\left(\frac{N}{2S}\right), \tag{65}$$
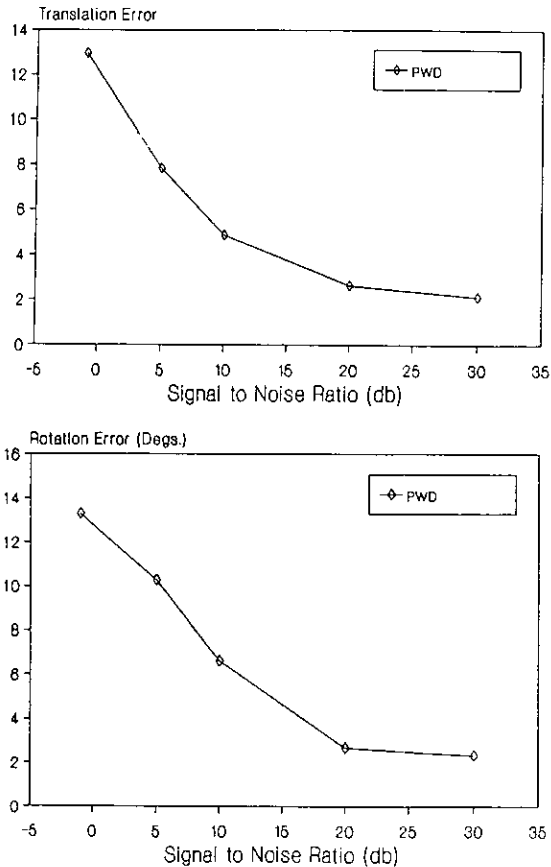

FIG. 14. The effect of contrast ratio on performance.

FIG. 15.   Performance curves for pure translation.

lation velocities ranged from 1 to 10 pixels per frame in either direction. The algorithm uses three frames to solve the problem. The figure shows that for the *PWD* the translational error rapidly goes down as the SNR is increased. It should be noted that if we were to know beforehand that there is no rotation then the results would be very accurate. This conclusion is based on the performance characterization of object location error.

Translation and Rotation

Figure 16 shows the results for the same objects as before but in this simulation the objects are translating and rotating. The translational velocities are within 1 to 10 pixels in either direction. The rotational velocity is randomly selected between 0° and 30° per frame. It is seen that the translational error is higher than before. This is partly due to the fact that we are simultaneously solving the equations for translation and rotation. Thus, overall error gets distributed between translation and rotation.

where M and N are the number of columns and rows, respectively. The algorithm is as follows: First, we compute the *WD* for each frame. From the *WD* data we determine the locations of the peaks. In the case of a noisy environment, it is difficult to detect peaks. In such cases we first eliminate peaks with values of 20% or lower than the highest peak observed in an image. If within a small neighborhood the number of peaks is greater than 1 then a centroid of all such peaks is computed to the give location value of the peak. This neighborhood is 5 × 5 for the *WD*. Then a check is made for any pseudopeaks. If pseudopeaks are present then they are eliminated by the method discussed earlier in this paper. After locating true peaks, we find point correspondence. The point correspondence is found from matching the peaks from one frame to the closest peak in the other frame. The weighted least-squares method is then used on these matched data points.

8.   RESULTS AND DISCUSSION

Pure Translation

Figure 15 shows the results of two objects translating in different directions with different velocities. The trans-
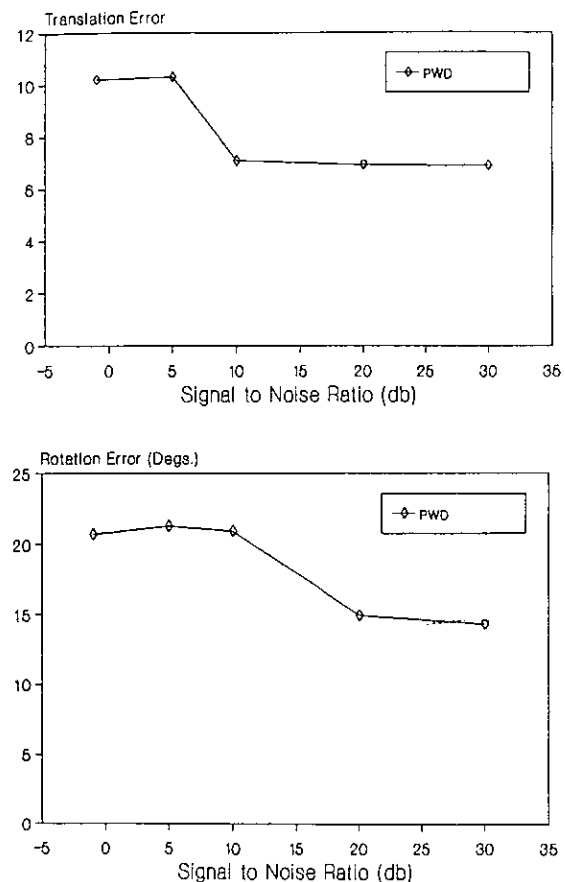

FIG. 16.   Performance curves for translation and rotation.

Real Image Sequences

We have two image sequences with which the *PWD* method is evaluated. The first sequence is of a car moving near a busy intersection on a highway in Seattle, Washington. The first frame of this sequence is shown in Fig. 17. The car in motion is the white car. The white circular marks seen in the figure are line dividers. These were for calibrating the images. Using the *PWD* method the speed of the car was determined to be 26.54 mph. If one were to actually count the number of pixels moved and compute the speed of the car, it would be 28.06 mph. Thus, we see that the speed from using the *PWD* is in close agreement with the estimated speed.

The second sequence consists of a 747 airplane take-off. The first frame is shown in Fig. 18. It is seen that the object in motion is darker than the background. Since one of our assumptions is that the object is brighter than the background, we cannot directly use this sequence. The original images are first inverted in gray scale. After inversion, the object is motion is brighter than the background. Thus, we meet our assumption. For calibration

we use the airplane dimensions at various locations. Based on the new method, the speed of the airplane is estimated to be 141.13 knots. This is well within the typical range of 140 to 150 knots.

Discussion

In this section we presented results for simulated as well as real image sequences. In simulated sequences when there was pure translation, the new method gives perfect results for a SNR of 5 dB or higher. This in itself is an accomplishment. This result shows that the method can be applied to noisy image sequences and we can expect to get near-perfect answers.

When a sequence was generated with rotation, the method did not work well. This could be due to various reasons. The data obtained after rotation has errors in it due to quantization. Such errors do not exist with pure translation. Another problem is that we have only three data points. This seriously limits our ability to correct for any errors. Besides, all our information is based upon one point from an object. So if the object rotates by, say,

FIG. 17.   First frame of the car sequence.

FIG. 18.   First frame of the airplane sequence.

10°, the point representing the peaks may or may not rotate much. In terms of pixels this movement could be zero due to quantization. Additionally, since the rotation is around the origin of the frame, in a discrete case a point is more likely to show motion if it is far away from the origin than if it is near the origin.

The real sequences involve pure translation and, as expected, we obtain excellent results. We have also presented a case where the background is brighter than the foreground and shown how to meet the required conditions for the algorithm to work. Also, it is worth noting that although the images for the airplane sequence have poor contrast, we still get good results.

9. CONCLUSION

Our performance characterization results are based on over 125,000 simulations in which the object size, object location, object orientation, and gray-scale intensity were all selected at random. We characterize performance for noise levels as high as -1 dB SNR. The results presented here show that the misdetection rate is near 0 and the false detection rate is one object per image of three objects each. In addition, this false detection rate can be reduced to 0 by shaping the kernel as outlined in Section 5.

Since our primary objective is to develop a method that can be used in real applications, it is necessary to find out how the method would work with real images. Thus, we use real image sequences to validate the method. It is seen that our algorithm predicted the speed of the car to be about 25 mph, which was quite reasonable for that particular intersection. With the help of the next image sequence we predicted the speed of the 747 as 142 knots, which is well within the range of 140 to 150 knots.

In conclusion, in this paper we have developed a strong analytical basis for using the Wigner distribution for image processing. We have also developed a method to eliminate the cross terms. Our simulation results as well as results from real life scenes clearly give the required confidence to machine vision researchers and practitioners to use the method in noisy environments.

As with any other research, we have continued our investigation for pseudopeak elimination methods. Cohen [3] formulated a generalized time–frequency distribution of which the Wigner distribution is a special case. His work led to many more kernels such as the exponential kernel of Choi and Williams [2] and the cone kernel of Zhao et al. [19]. As expected by one of the referees of this paper, the results based on the exponential kernel and the cone kernel give better results than the pseudo-Wigner distribution. Our new research findings on the

generalized space–frequency distribution and the two kernels mentioned above have been recently sent to a journal for publication.

REFERENCES

1. L. Capodiferro, R. Cusani, G. Jacovitti, and M. Vascotto, A correlation based technique for shift, scale, and rotation independent object identification, in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, 1987*, Vol. 1, pp. 221–224.

2. H. Choi and W. Williams, Improved time frequency representation of multicomponent signals using exponential kernels, *IEEE Trans. Acoust. Speech Signal Process.* **37**(6), 1989, 862–871.

3. L. Cohen, Generalized phase-space distributions functions, *J. Math. Phys.* **7**, 1966, 781–786.

4. T. A. C. M. Claasen and W. F. G. Mecklenbrauker, The Wigner Distribution—A tool for time-frequency signal analysis, part I: Continuous time signals, *Phillips J. Res.* **35**, 1980, 217–250.

5. T. A. C. M. Claasen and W. F. G. Mecklenbrauker, The Wigner Distribution—A tool for time-frequency signal analysis, part II: Discrete time signals, *Phillips J. Res.* **35**, 1980, 276–300.

6. H. Gafni and Y. Y. Zeevi, A model for separation of spatial and temporal information in the visual system, *Biol. Cybernet.* **28**, 1977, 73–82.

7. R. M. Haralick and L. G. Shapiro, *Robot and Computer Vision*, Addison–Wesley, Reading, MA, 1992.

8. B. K. P. Horn and B. G. Schunk, Determining optical flow, *Artif. Intell.* **17**, 1981, 185–203.

9. T. S. Huang and R. Y. Tsai, Image sequence analysis: Motion estimation, in *Image Sequence Analysis*, pp. 1–18, Springer-Verlag, Berlin/New York, 1981.

10. L. Jacobson and H. Wechsler, Derivation of optical flow using a spatiotemporal-frequency approach, *Comput. Vision Graphics Image Process.* **38**, 1987, 29–65.

11. A. K. Jain, Object classification and registration by radon transform based invariants, in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, 1987*, Vol. 1, pp. 225–228.

12. Z. Lin, T. Huang, and S. D. Blostein, Motion estimation from 3D point sets with and without correspondences, in *IEEE Workshop on Motion Estimation, 1986*, pp. 194–201.

13. H. Nagel, On a constraint equation for the estimation of displacement rates in image sequences, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-11**, 1989, 13–30.

14. S. Rajala, A. N. Riddle, and W. E. Snyder, Application of the one dimensional Fourier transform for tracking moving objects in noisy environments, *Comput. Vision Graphics Image Process.* **21**, 1983, 280–293.

15. B. G. Schunk, Image flow continuity equations for motion and density, in *Proceedings of Workshop on Motion: Representation and Analysis, Charleston, SC, May 1986*, pp. 89–94.

16. J. Ville, Theorie et applications de la notion de signal analytique, *Cables et Transmission 2eA* **1**, 1948, 61–74.

17. E. Wigner, On the quantum correction of thermodynamic equilibrium, *Phys. Rev.* **40**, 1932, 749–759.

18. D. Willick and Y. Yang, Experimental evaluation of motion constraint equations, *CVGIP: Image Understand.* **54**, 1991, 206–214.

19. Y. Zhao, L. E. Atlas, and R. J. Marks II, The use of cone-shaped kernels for generalized time frequency representations of nonstationary signals, *IEEE Trans. Acoust. Speech Signal Process.* **38**(7), 1990, 1084–1091.

VINAY G. VAIDYA was born on May 25, 1954, in Bombay, India. He received a Bachelor of Engineering (Electrical) degree from the University of Bombay and an M.S.E.E. and a Ph.D. in electrical engineering from the University of Washington in Seattle. In 1977 he worked as a research fellow at the national Institute for Training in Industrial Engineering in Bombay. He joined Boeing, Seattle, in 1981. There he worked on the design of automatic throttle system for the 747 and the design of windshear control law on the 767. Currently at Boeing he is with the Advanced Technology group where his job responsibilities include developing new methods for electronics manufacturing. His research interests include image and signal processing, and mathematical modeling for tumor growth. He has served as associate editor for *Modeling and Simulation,* Instrumentation Society of America, in 1987 and 1989.

ROBERT M. HARALICK was born in Brooklyn, New York, on September 30, 1943. He received a B.A. degree in mathematics from the University of Kansas in 1964, a B.S. degree in electrical engineering in 1966, and an M.S. degree in electrical engineering in 1967. In 1969, after completing his Ph.D. at the University of Kansas, he joined the faculty of the Electrical Engineering Department there where he last served as Professor from 1975 to 1978. In 1979 Dr. Haralick joined the Electrical Engineering Department at Virginia Polytechnic Institute and State University where he was a professor and Director of the Spatial Data Analysis Laboratory. From 1984 to 1986 Dr. Haralick served as Vice President of Research at Machine Vision International, Ann Arbor, MI. Dr. Haralick now occupies the Boeing Clairmont Egtvedt Professorship in the Department of Electrical Engineering at the University of Washington. Professor Haralick has made a series of contributions in computer vision. In the high-level vision area, he has identified a variety of vision problems which are special cases of the consistent labeling problem. In the low-level and mid-level areas, Professor Haralick has worked in image texture analysis using spatial gray tone co-occurrence texture features. These features have been used with success on biological cell images, X-ray images, satellite images, aerial images, and many other kinds of images taken at small and large scales. Professor Haralick's work in shape analysis and extraction uses the techniques of mathematical morphology. He has developed the morphological sampling theorem, which establishes a sound shape/size basis for the focus of attention mechanisms which can process image data in a multiresolution mode, thereby making some of the image feature extraction processes execute more efficiently. Professor Haralick's most recent work is concentrating on the performance characterization of vision algorithms. Here he is concentrating on determining the appropriate random perturbation models for low-, mid-, and high-level vision data structures and the propagation of these kinds of random perturbation through different computer vision algorithms. This work is opening up a new area in the field of computer vision. Professor Haralick is a Fellow of IEEE for his contributions in computer vision and image processing. He is listed in the 46th edition of *Who's Who in America* and in the 1992 edition of *Who's Who in the World.* He serves on the editorial board of *IEEE Transactions on Pattern Analysis and Machine Intelligence.* He is the computer vision area editor for *Communications of the ACM.* He also serves as associate editor for *Computer Vision, Graphics, and Image Processing* and *Pattern Recognition.*