OPTIMAL WEIGHTING FUNCTIONS FOR CROSS CORRELATION
OF DISTORTED IMAGES†

by

Hassan Mostafavi
Systems Control, Inc.
Palo Alto, California

ABSTRACT

The detection performance of the two-dimensional cross correlation algorithm applied to two images with relative geometric distortion is presented. The performance sensitivity to changes in distortion and window function parameters is demonstrated. The expression for peak-to-sidelobe ratio is shown to be maximized by the window function

$$h(x) = R_g((I-A)x)$$

where $x$ is the two-dimensional variable in the image plane, $R_g(x)$ is the autocorrelation function of the image random pattern and the $2 \times 2$ matrix, $A$, represents the geometric distortion coordinate transformation. The maximum achievable peak-to-sidelobe ratio is shown to be $1/|\det(I-A)|^{1/2}$. For the special case of Gaussian shaped image autocorrelation and window functions, the performance sensitivity to changes in distortion and window function parameters is demonstrated.

1. INTRODUCTION

In many applications it is desired to estimate the offset between two images of a scene, produced by two sensors with different positions and orientations. A particular method that is widely used consists of computing the two-dimensional cross correlation function of the two images and using the location of the peak of this function as the desired offset.

The cross correlation algorithm computes the function

$$C(\mathbf{z}) = \int h(x) R_g(x - z) I_1(x) dx$$

(1)

where $I_1(x)$ and $I_2(x)$ are the two images of the scene, called reference and sensor images respectively, and $x$ and $z$ denote the two-dimensional variables in the plane. The window function, $h(x)$, specifies the spatial weighting applied to the reference image before cross correlation with the sensor image. It, in effect, specifies the area of integration around the reference point.

Almost always, the difference in position and orientation of the sensors produces relative geometric distortions that cause the two images to be different by more than a simple translation. This means that after aligning the images of a prespecified reference point by translating one image with respect to the other, the images of most other points do not exactly coincide. This causes a reduction in the value of and a widening of the cross correlation peak. Thus, the performance

†This work was supported by DARPA Strategic Technology Office, Contract DAAK40-77-C-0113.

in terms of both the probability of false peak detection and the accuracy of the offset estimation is degraded.

The effect of geometric distortion on the performance of the correlation algorithm has been investigated in the past. [1-4]. Through numerical examples it has been observed [3,4] that for a fixed geometric distortion and a square window function there is an optimum size that results in the best performance of the algorithm. It also appears from these results that in many cases the choice of the shape of the window function can be just as important as the size.

In this paper the optimum window function for a fixed geometric distortion is derived. In this study, the detection performance rather than the estimation accuracy is emphasized. The performance measure which is maximized is the ratio of the average peak value of the cross correlation function to the standard deviation of its sidelobes. For a fixed probability of correct peak detection, the probability of false detection is a monotonic decreasing function of the peak-to-sidelobe ratio.

The mathematical model used for obtaining the average detection performance is identical to the one used in references [3] and [4]. The reference images are modeled as sample functions of a homogeneous Gaussian random pattern defined in terms of its autocorrelation function, $R_g(x)$. The additive noise is modeled by a homogeneous random pattern that is independent of the image.

The geometric distortion is represented by an affine transformation of the image coordinates. This is a good model for the cases where the scene is approximately planar, the differences in sensor position and attitude are relatively small, and the field of view of the sensor on the scene plane is small compared with the sensor distance from the reference point in the scene. An affine transformation of the image coordinates is represented by

$$\mathbf{z}' = A\mathbf{z} + \mathbf{z}_0$$

(2)

where $\mathbf{z}_0$ is the offset vector to be estimated and the $2 \times 2$ matrix $A$ represents the geometric distortion. For many scene and sensor geometries the matrix $A$ can be computed. The derivation of the optimum window function uses the assumption that the geometric distortion is relatively small and the difference between $A$ and the identity matrix, $I$, is approximately bounded by

$$|\det(I-A)|^{1/2} < .2$$

(3)

This corresponds to ±20 percent magnification or rotation by ±11.5 degrees.
In Section 2 the general expression for the peak-to-sidelobe ratio is derived. It is shown that, for a fixed distortion, this expression is maximized by the window function
\[ h_A(x) = R_p[(I-A)x] \]
which depends only on the image autocorrelation function and the distortion. The maximum achievable peak-to-sidelobe ratio is shown to be
\[ r_{\text{max}} = \frac{1}{\det(I-A)^{1/2}} \left( \frac{1}{1 + \frac{1}{\text{SNR}_I}} \right)^{1/2} \]
where SNR is the input signal-to-noise ratio. Assuming complete freedom in the choice of the window function, this gives the limit on the detection performance in terms of the distortion and the additive noise.

Section 3 applies these results to the important special case where the image autocorrelation function is assumed to have a low-pass Gaussian shape. For this special case the sensitivity of the performance to the window function parameters is studied. Also the effect of using a rectangular window function rather than the optimum is examined. In Section 4 the sensitivity of performance to variations of geometric distortion is considered. Also comments are given about the optimum window function when the geometric distortion is random and only its statistics are known.

2. MAXIMUM PEAK-TO-SIDELobe RATIO

The differences between the reference and sensor images are divided into two categories. One is due to additive noise which is assumed to be independent of the images, and the other is due to geometric distortion which is modeled as an affine transformation of image coordinates. Therefore, the reference and sensor images can be defined by
\[ I_r(x) = P(x) \]
and
\[ I_s(x) = P(Ax + t_0) + N(x) \]
respectively.* Without loss of generality the reference point in the reference image is assumed to be the origin of the image plane. Thus, \( t_0 \) is its image in the sensor image plane. The objective of the cross correlation algorithm is to estimate \( I_r \). From (1), (6) and (7) it follows that a correct estimation takes place if \( C(t_o) \) is the peak of the cross correlation function, \( C(I_o) \).

The image and additive noise are modeled by two zero mean, Gaussian and homogeneous random patterns, \( P(x) \) and \( N(x) \) respectively, which are assumed to be uncorrelated. Their autocorrelation functions are denoted by \( R_p(x) \) and \( R_n(x) \) respectively.

The cross correlation function \( C(t) \) is a random pattern whose statistics determine the performance of the correlation algorithm. The mean function of the random pattern \( C(t) \) is given by [3]

\[ E[C(t)] = \int h(x) R_p[(I-A)x + t - t_0] \, dx \]
which for \( t = t_0 \) gives the average peak value,
\[ E[C(t_o)] = \int h(x) R_p[(I-A)x] \, dx \] (9)

Since the function \( R_p(x) \) tends to zero for large values of \( x \), from (8) it follows that for any finite window the mean function goes to zero for large values of \( || x - t_0 || \). In such regions of the cross correlation surface, \( C(t) \) is defined as a sidelobe.

The peak-to-sidelobe ratio is defined as
\[ r = \frac{E[C(t_o)]}{\text{Var}[C(t_o)]}^{1/2} \]
where \( t_o \) is any offset in the region of sidelobes. The variance of the sidelobe \( C(t_o) \) is given by [3]
\[ \text{Var}[C(t_o)] = \int g(x) [R_p(Ax + t_0) + R_p(x)] R_n(x) \, dx \] (11)
where the function \( g(x) \) is defined by
\[ g(x) = \int h(x) h(x + x) \, dx \] (12)

For the small geometric distortions considered in this study, as far as the value of the integral in (11) is concerned the effect of \( A \) is very small. Thus, \( A \) can be replaced by the identity matrix, i.e.,
\[ \text{Var}[C(t_o)] = \int g(x) [R_p(x) + R_p(x)] R_n(x) \, dx \] (13)

This means that the effect of the geometric distortion on the sidelobe variance is much smaller than its effect on the average peak value given by (9).

For window functions that are approximately five or more times wider than \( R_p(x) \), using (12) and (13), the sidelobe variance can be well approximated by
\[ \text{Var}[C(t_o)] = [\int h^2(x) \, dx] [\int R_p^2(x) + R_p(x) R_n(x)] \, dx \]
\[ = [\int h^2(x) \, dx] [\int R_p^2(x) \, dx] (1 + \frac{1}{\text{SNR}_I}) \] (14)

where \( \text{SNR}_I \) denotes the input signal-to-noise ratio defined by
\[ \text{SNR}_I = \frac{\int R_p^2(x) \, dx}{\int R_p^2(x) \, dx} \] (15)

This is the ratio of the average image power to that portion of the additive noise power which is within the image spectral band.

Combining (9), (10) and (13) the peak-to-sidelobe ratio in the general case becomes
\[ r = \frac{\int h(x) R_p[(I-A)x + t - t_0] \, dx}{\int g(x) [R_p^2(x) + R_p(x) R_n(x)] \, dx}^{1/2} \]
and using (14), for wide enough window functions this ratio is given by
where $r_p$ is the peak-to-sidelobe ratio in the absence of additive noise defined by

$$ r_p = \frac{1}{1 + \frac{1}{\text{SNR}_1}} $$

In order to find the window function which maximizes $r$, the Schwartz inequality can be applied to the numerator of (16) to obtain

$$ r^2 \leq \frac{\int h^2(q)^2 dq}{\int R^2_p(I-A)x^2 dq} \cdot \frac{1}{1 + \frac{1}{\text{SNR}_1}}. $$

This bound is achieved by choosing $h(q)$ as

$$ h_A(q) = R_p(I-A)x. $$

For geometric distortions limited by the bound in (3), $h_A(q)$ is five or more times wider than $R_p(q)$. Therefore for this window function substituting (14) in (19) yields

$$ r^2 \leq \left( \frac{\int R^2_p[I-A)x^2 dq}{\int R^2_p(q) dq} \right)^{1/2} \cdot \left( \frac{1}{1 + \frac{1}{\text{SNR}_1}} \right)^{1/2} $$

Since this bound is independent of the window function it is the maximum of $r^2$ and is achieved by the choice of window function given in (20), i.e.,

$$ r_{\text{max}} = \left( \frac{\int R^2_p[I-A)x^2 dq}{\int R^2_p(q) dq} \right)^{1/2} \cdot \left( \frac{1}{1 + \frac{1}{\text{SNR}_1}} \right)^{1/2} $$

$$ = \frac{1}{|\text{det}(I-A)|^{1/2}} \cdot \left( \frac{1}{1 + \frac{1}{\text{SNR}_1}} \right)^{1/2}. $$

### 3. GAUSSIAN AUTOCORRELATION FUNCTIONS

Almost all images encountered in practice can be modeled as two-dimensional, low-pass, random patterns. A simple model whose parameters can be estimated rather easily for these images involves the assumption of a Gaussian autocorrelation function. Using such a model the only parameters that need to be estimated are the second central moments of the autocorrelation function.

A general Gaussian autocorrelation function is given by

$$ R_p(q) = \text{E} \exp \left( -\frac{1}{2} \chi^T D^{-1} \chi \right) $$

where the matrix $D$, specifies the mean square widths and orientation of the surface defined by $R_p(q)$.

Assuming a Gaussian autocorrelation function, from (2) it follows that the optimum window function is also Gaussian. A general Gaussian window function is defined by

$$ h(q) = k \exp \left( -\frac{1}{2} \chi^T C^{-1} \chi \right) $$

where the positive definite matrix $C$ specifies the mean square widths and orientation of the window function. From (20) it follows that the optimum window, $h_A(q)$, is obtained by choosing matrix $C$ as

$$ C_A = (I-A)^T D^{-1} (I-A) $$

in (24). The maximum achievable peak-to-sidelobe ratio is given by (22).

For example suppose the matrix $D$ in (23) is

$$ D = \begin{bmatrix} \Delta^2 & 0 \\ 0 & \Delta^2 \end{bmatrix} $$

and the geometric distortion matrix is

$$ A = \begin{bmatrix} 1+e_0 & -\alpha_0 \\ \alpha_0 & 1+e_0 \end{bmatrix} $$

where $e_0$ is the small relative magnification factor and $\alpha_0$ is the small relative rotation (in radians) of one image with respect to the other. Then from (25) it follows that the optimum window is specified by the matrix

$$ C_0 = \begin{bmatrix} \frac{\Delta^2}{e_0^2 + \alpha_0^2} & 0 \\ 0 & \frac{\Delta^2}{e_0^2 + \alpha_0^2} \end{bmatrix} $$

and the maximum peak-to-sidelobe ratio is

$$ r_{\text{max}} = \frac{1}{(e_0^2 + \alpha_0^2)^{1/2}} \cdot \left( \frac{1}{1 + \frac{1}{\text{SNR}_1}} \right)^{1/2}. $$

Another case of interest is that in which a Gaussian window function (not necessarily the optimum) is used and the system performance is to be estimated for a given distortion. Substituting (23) and (24) in (18) for wide enough window functions gives

$$ r = \frac{2(\text{det} C^{-1} \cdot \text{det} D^{-1})^{1/4}}{\text{det}[C^{-1} + (I-A)^T D^{-1} (I-A)]^{1/2}} \cdot \left( \frac{1}{1 + \frac{1}{\text{SNR}_1}} \right)^{1/4}. $$

The sensitivity of the peak-to-sidelobe ratio to the window function parameters can be demonstrated by considering a Gaussian window function with the same orientation as the optimum but with different rms width and length. The expression for peak-to-sidelobe ratio in this case reduces to

$$ r = r_{\text{max}} \cdot \left( \frac{1}{w_0^2 + \frac{\Delta^2}{e_0^2 + \alpha_0^2}} \right)^{1/2} \cdot \left( \frac{1}{1 + \frac{1}{\text{SNR}_1}} \right)^{1/2}. $$
where \( w_0 \) and \( \ell_0 \) denote the rms width and length of the optimum window function, \( w \) and \( \ell \) denote the rms width and length of the window function used and \( r_{\text{max}} \) is given by (22).

It can be seen that \( r \) as a function of the width and length has a separable form and therefore each dimension of the window function can be optimized independently. Figure 1 shows \( r \) as a function of \( \frac{w}{w_0} \) for a fixed value of \( \frac{\ell}{\ell_0} \). From (31) it follows

\[
Q(T) = \int_{-\infty}^{+\infty} \exp\left(-\frac{1}{2} u^2\right) du
\]

and \( \ell_0 \) and \( w_0 \) are again the rms length and width of the optimum window function given by (20).

The dashed curve of Figure 1 shows \( r \) as a function of \( \frac{w}{w_0} \) for a fixed \( \frac{\ell}{\ell_0} \) as given by (32). It can be seen that the maximum \( r_0 \) is achieved by \( \frac{w}{w_0} \approx 0.82 \).

The peak value of this plot is equal to

\[
0.943 \cdot r_{\text{max}} \cdot \left(\frac{\ell}{\ell_0}\right)^{1/4} \left(\frac{w}{w_0}\right)^{1/2}
\]

If both length and width of the rectangular window are chosen optimally, the value of \( r \) is \((0.943)^2 \approx 0.89\) times the \( r \) corresponding to the optimum Gaussian window function. The comparison of the two curves in Figure 1 shows that for window sizes in the order of the optimum and smaller, the analytic expressions for the Gaussian window function can be used to estimate the performance of the rectangular window. Furthermore, the size of the optimum rectangular window can be obtained from the dashed curve in Figure 1.

4. VARIATION OF GEOMETRIC DISTORTION

In most practical situations the geometric distortion is not known exactly. Therefore it is useful to study the sensitivity of the peak-to-sidelobe ratio to changes of geometric distortion from the value for which the window function has been optimized.

The relation between \( r \) and the geometric distortion for a fixed window function is given, in general by (17) and (18) and for the special Gaussian shaped functions, by (30). As an example consider the example in Section 3, and suppose the window function is optimized for magnification, \( \varepsilon \), and rotation, \( \alpha \). The peak-to-sidelobe ratio, when the actual magnification and rotation are \( \varepsilon' \) and \( \alpha' \) respectively, can be written as

\[
r = r_{\text{max}} \frac{2}{1 + \frac{\varepsilon^2 + \alpha^2}{\varepsilon'^2 + \alpha'^2}}
\]

where \( r_{\text{max}} \) is given by (30).

From (34) and also from (28) it follows that for this example the optimum window function and the value of \( r \) depend on \( \varepsilon^2 + \alpha^2 \) and \( \varepsilon'^2 + \alpha'^2 \) only and in particular they do not depend on the direction of rotation or the sign of the relative magnification factor. This is true for many geometric distortions and autocorrelation functions encountered in practice.

The above result also shows that the peak value of \( r \) resulting from zero distortion is twice the value of \( r_{\text{max}} \) corresponding to the actual distortions \( \varepsilon = \varepsilon_0 \) and \( \alpha = \alpha_0 \), and thus decreases with increasing \( \varepsilon_0 \) and \( \alpha_0 \). This latter statement is in general true for any distortion matrix, \( A \), and Gaussian shaped functions and can be deduced from (30). Figure 2 shows for different window functions the plots of \( r \) as a function of geometric distortion, normalized to the reduction factor due to the additive noise. Each window function is optimized for a different geometric distortion.
The choice of the window function width also depends on both the range of uncertainty in the geometric distortion and the amount of the additive noise. For example, window functions that are optimized for larger geometric distortions result in values of $r$ that are lower and at the same time less sensitive to variations in geometric distortion. However, the input signal-to-noise ratio may cause $r$ to fall so low that detection performance is unacceptable. On the other hand, an increase in the value $r$ beyond, for example 10, causes an insignificant improvement in the detection performance.

Figure 2. Peak-to-sidelobe ratio as a function of the geometric distortion for various window functions, each optimized for a different distortion.

5. REFERENCES


