Automatic target recognition during sensor motion and vibration

Ofer Hadar, MEMBER SPIE
A. Kuntsevitsky
M. Wasserblat
Norman S. Kopeika, MEMBER SPIE
Stanley R. Rotman, MEMBER SPIE
Ben-Gurion University of the Negev
Department of Electrical and Computer Engineering
P.O. Box 653
84105 Beer-Sheva, Israel
E-mail: kopeika@bgu.ac.il

Abstract. The resolution capability of imaging systems is affected by blur resulting from vibration and motion during the exposure. This blur is often more severe than electronic and optical resolution limitations inherent in the system. Such image quality degradation must be considered when dealing with the development and analysis of automatic target recognition (ATR) systems. This research analyzes the influence of image vibrations and motion on the probability of acquiring a target with an ATR system. The analysis includes accepted metrics that characterize the relationship existing between the target and its background. A high level of correlation is expected between these factors and the probability of target detection permitting efficient performance in the prediction and evaluation of any ATR system. Such correlations are considered here in the presence of sensor motion and vibration and situations are considered in which the probability of recognition is improved by the motion, despite the blur. The results of this research can be implemented in military applications as well as in developing image restoration procedures for image-blur conditions.

Subject terms: optical transfer function; automatic target recognition; motion; vibration; segmentation.

Optical Engineering 34(10), 3062—3068 (October 1995).

1 Introduction

Image motion and vibrations are dominant factors that affect the resolution of the received image. In many high-resolution vehicular or airborne imaging systems, despite the use of high-quality sensors, resolution is limited by image motion and, as a result, the high-resolution capability of the sensor may be wasted.1 This reduction of resolution is one of the major factors in the reduction of the ability to acquire a target by an automatic target recognition (ATR) system.

The degradation of image quality as a result of motion in the image plane can take several forms, such as linear motion, acceleration, or sinusoidal vibrations. Each of these motion forms degrades the image in a different way. The image degradation is described by the optical transfer function (OTF). This function can be derived from a priori knowledge of the relative displacement between the object and the camera during the exposure time.3,3 The mathematical model for extracting the OTF from the image motion is described elsewhere3; here only a brief review is presented. In Sec. 2, we discuss the OTF for three kinds of motion that are considered in this paper: linear motion and high- and low-frequency vibrations. In Sec. 3, three image quality measures are considered: target-to-background contrast (TBC), target versus background entropy (ETB), and signal-to-clutter ratio (SCR). These parameters have been obtained as a function of blur radius \(d\) for images degraded by the three types of image motion. The original image includes a target in the middle and a fractal background around it. The measure TBC is found to be reciprocal to \(d\), whereas SCR is directly related to \(d\). However, ETB was found to be reciprocal to \(d\) only for large blur radius. Section 4 presents the theory of extracting a target from its background by using a thresholding operation.7,9 Results of the segmentation process are presented for several blur radii and several types of image motion. The degree of success of the segmentation is obtained by two quantitative measures: fraction of extracted target (FET) and fraction of extracted background (FEB). Results of the segmentation process on restored images are also presented to establish the worthwhileness of such restoration. One surprising result from this analysis, as suggested in Ref. 10, is that in some cases image motion can improve the ability to extract the target from its background.
2 Representation of Image Motion with a Spatial Filter

This section describes a new approach to the transformation process of image motion to a spatial filter. The bases of this method are the numerical calculations of the OTF in real time appropriate for any kind of image motion. Here, only the results for linear motion and high- and low-frequency vibrations are reviewed. This method is based on a priori knowledge of the relative displacement $x(t)$ between the camera and the object, which is easily measurable with a microelectronics accelerometer or other sensor. The line spread function (LSF) that represents the motion degradation process in the spatial domain is the probability density function (pdf) or the histogram of that relative displacement. In this paper, the LSF was obtained from a simulation of image motion in the image plane. The degradation blur radius $d$ is measured in pixels along the horizontal axis and is limited by the image size ($256 \times 256$ pixels).

The advantage of this method is that it can be applied to any kind of motion. The practical application of it is to evaluate the histogram of $x(t)$ using a numerical calculation appropriate to real-time imaging systems. The transfer function in the spatial frequency domain can be obtained by a 1-D Fourier transform of the LSF,

$$
\text{LSF}(x) = \frac{1}{V_t} = \frac{1}{d}, \ (0 < x < d).
$$

The smearing caused by linear motion is uniform from $x = 0$ to $d$, where $d$ is the spatial extent of the blur in pixels and is equal to $V_t$.

Therefore, from Eq. (1)

$$
\text{MTF}(f) = \text{sinc}(\pi f d), \ (f < f_{\text{max}} = \frac{1}{d}),
$$

where $f_{\text{max}}$ is the maximum spatial frequency, MTF is the modulation transfer function, and PTF is the phase transfer function. The degradation process was obtained only from the MTF, which is more affected by image quality than the PTF. The optical system cannot resolve at frequencies higher than $1/d$ because such detail is smaller than the actual blur radius $d$. Hence, the MTF is limited to spatial frequencies less than $f_{\text{max}} = 1/d$. The mathematical MTF at higher spatial frequencies is false resolution. Therefore the MTF that is used here is set to zero from $f_{\text{max}}$ for the maximum image resolution, which is 1 cycle/(2 pixels) = 128 line pairs in the spatial frequency domain.

A second type of motion that is considered here is sinusoidal image motion. The effect of this type of image motion, which often results from mechanical vibrations, can be divided into high- and low-vibration frequencies. In the first case, the time exposure $t_e$ is longer than the vibrational period $T_0$, and the image blur is therefore the entire peak-to-peak translation $2D$ of the image. The image motion is given by

$$
x(t) = D \cos \frac{2\pi t}{T_0}.
$$

The MTF for the high-vibration frequency case is given by

$$
\text{MTF}(f) = J_0(2\pi f D) .
$$

The low-vibration frequency situation involves time exposures shorter than the vibration period. In this case, the blur radius $d$ and the MTF are random processes that depend on the exact time the exposure takes place. This type of blur is often more degrading than the high-vibration frequency blur because in real-life situations for low-vibration frequencies is in many cases much greater than $2D$ for high-vibration frequencies. The statistical nature of the MTF is described elsewhere in detail. The LSF for this kind of motion can be derived from the LSF of high-frequency vibration, and this method is described elsewhere. The important parameter for this kind of motion is the relative exposure time $t_e/T_0$, as $t_e/T_0$ increases, the mean blur radius also increases. Three different values of $t_e/T_0$ are considered here, 0.05, 0.1 and 0.15.

3 Target-to-Local Background Measures

3.1 Introduction

Image motion and vibrations have a significant impact on the imagery generated by an imaging sensor. To develop metrics to describe the target detectability, image measures that quantitatively describe ATR system inputs are needed. The motivation for using image measures is based on the anticipated properties of the image. The important property is that fewer image quality measures than scenario domain parameters are required to characterize ATR performance. The selected image measures must adequately represent all possible input parameters (atmosphere, vibrations, etc.) affecting ATR systems. This section presents three target-to-local-background measures that are required to characterize ATR performance: TBC, ETB, and SCR. The measures TBC and ETC were found to be parameters highly correlated to target probability of detection. The SCR measure was used to characterize the influence of clutter background on target detection by human observers. The purpose of the analysis here is to investigate how these measures change under different kinds of image motion and different blur radii. This kind of analysis can be very useful for estimating the capability of target extraction from backgrounds under image motion blur conditions. It can also be useful for estimating the probability of detection for a target in scenarios that involve image motion.

3.2 Target and Background Area

Figure 1 indicates the regions of image pixels that were used to compute the image measures. The regions consist of two...
3.3 Description of the Image

The image that is used in this experiment contains a target in the middle and a fractal background around it. The target is a tank with a large range of gray levels, as shown in Fig. 2. The size of the inner rectangle around the tank is 25 × 49 pixels. Acquisition performance probabilities are derived here for targets brighter than the background. These results, however, can be extended for the case of targets darker than the background for the same image motion conditions. The background in this experiment is a fractal image with high clutter. The fractal dimension is 2.5. For other types of backgrounds, further work is needed.

3.4 Description of Measures

3.4.1 Target-to-background contrast

A gray-level domain measure is computed as the ratio of the average target T pixel intensity minus the average local background B pixel intensities to the standard deviation of the local background pixel intensities. Symbolically,

\[ TBC = \frac{\mu_T - \mu_B}{\sigma_B}, \quad (6) \]

where \( \mu_X \) denotes the mean variable \( X \), and \( \sigma_X \) denotes the standard deviation of variable \( X \).

3.4.2 Target versus background entropy

This gray-level domain measure characterizes the difference in entropy for the target-segmented region versus the background region. Symbolically the measure is

\[ ETB = \left| E_T(m) - E_B(m) \right|, \quad (7) \]

where \( E(X) \) denotes the entropy of variable \( X \) and \( m \) represents the gray level of a given image pixel. The target and the background entropy are given by Eqs. (8a) and (8b), respectively:

\[ E_T(m) = \sum_{i=0}^{255} P^T_i \log_{10}(P^T_i) \quad (8a) \]

\[ E_B(m) = \sum_{i=0}^{255} P^B_i \log_{10}(P^B_i) \quad (8b) \]

It is evident that the entropy of an image is a function of the gray-level probability \( P_x \). It is interesting to find the image histogram that yields the maximum entropy. Because the entropy is a measure of uncertainty, the probability distribution that generates the maximum uncertainty will have the maximum entropy. Hence, a large entropy number, approaching 1.0, indicates that many gray levels are represented in the selected region. As the ETB measure decreases, it is an indication that the target and the background region have very similar entropy levels; thus it is more difficult to extract the target from its local background.

3.4.3 Signal-to-clutter ratio

The SCR is a function of the degree of ‘‘clutter’’ in the image. The clutter in an image can greatly affect the probability of detection and time required for detection of a target in a scene. It is evident that the clutter is a crucial parameter for determining the probability of detection of military targets in genuine battlefield conditions both by the human observer and by ATR systems. There are many definitions for clutter. One of the most common ones was suggested by Schmieder and Weathersby. They proposed to divide the scene into \( N \) blocks (where the length of the block in each dimension is twice the size of the target), measure the variance within the block, and take the square root of the average of all the blocks in the picture. Hence,

\[ \text{clutter} = \left( \frac{1}{N} \sum_{i=1}^{N} \sigma_i^2 \right)^{1/2}, \quad (9) \]

where \( \sigma_i^2 \) is the variance within the \( i \)th block. The SCR of the process is given by the contrast of the target divided by the clutter,

\[ \text{SCR} = \frac{\text{target}_\max - \mu_B}{\text{clutter}}. \quad (10) \]
3.5 Results

The results presented here are for three kinds of motion: linear motion and high- and low-frequency vibrations. Figure 2 describes one sample comparison between the original image and the image blurred by linear motion. Figures 2(a) and 2(b) present the original image and the target and background histograms, respectively. Figures 2(c) and 2(d) present the same results but for an image blurred by linear motion where \( d = 15 \) pixels. In Figs. 3(a), 3(b), and 3(c), the three measures, TBC, ETB, and SCR are plotted for linear motion, respectively. The blur radius parameter \( d \) varies from 1 to 30 pixels. In Figs. 4(a), 4(b), and 4(c), the same measures are plotted for high-frequency vibrations. The variable parameter is the sinusoidal amplitude \( D \), where \( d \) equals \( 2D \) and varies from 2 to 40 pixels. Figures 5(a), 5(b), and 5(c) present results for low-frequency vibrations. For this kind of motion, each of the figures contains three graphs for \( t_e/T_0 \) equal to 0.05, 0.1, and 0.15. The sinusoidal amplitude for this simulation was determined to be \( D = 100 \) pixels. The blur radius for low-frequency vibration varies from the minimum value \( d_{\text{min}} \), which occurs near the peaks of the sine wave, to the maximum value \( d_{\text{max}} \), which occurs near the zero-crossing points of the sine wave. The range of each of these graphs changes depending on \( t_e/T_0 \) (Ref. 1 to 3). The blur radius parameter \( d \) is reciprocal to \( d \). This behavior is appropriate for all three types of motion. However, our second metric ETB begins to decrease only after a certain blur radius. For high-frequency vibrations, ETB increases with \( d \) until \( d = 23 \) pixels; from this point the reduction is sharp until \( d = 40 \) pixels. For low-frequency vibrations ETB is almost constant for the case \( t_e/T_0 = 0.05 \), whereas for the other two cases, \( t_e/T_0 = 0.1 \) and 0.15, ETB decreases with increasing values of \( d \). From the point where ETB begins to decrease, this reduction can be approximated by a linear function for the three types of motion. This linear function enables us to predict TBC and ETB according to the blur radius in the image. If these parameters are correlated to the probability of detection, it should be possible to predict the probability of detection \( P_D \) for a given type of image motion and blur radius.

In contrast, our third metric, SCR, is directly related to \( d \) for the three image motions. The reason for it is surprising and derives from the fact that the background without image motion is fractal with a large value of clutter. Image motion MTF causes the background to be smoothed so that the image is no longer fractal and therefore the clutter is decreased. The reduction of clutter, which is the denominator of Eq. (10), causes an increase in SCR. This increase can be approximated to a linear function for low-frequency vibrations. This approximation is appropriate for low-frequency vibrations only for \( t_e/T_0 = 0.05 \). For \( t_e/T_0 = 0.1 \) and 0.15, the linear approximation can be applied only from specific points along the blur radius axis, \( d = 18 \) pixels and \( d = 35 \) pixels, respectively. The reason for such behavior is that the motion of low-frequency vibrations is close to linear motion for short relative exposure time \( t_e/T_0 \) (Ref. 1 to 3). For large values of \( t_e/T_0 \), linear motion applies only to exposure times near the zero-crossing points of the sine wave. The value of SCR as a function of the blur radius and image motion type can be helpful for the prediction of the probability of detection. The analysis in Ref. 5 presents the influence of degradation by a Gaussian MTF on the probability of detection. This analysis relates the probability of detection to the resolution of the image for different SCR values for human observers. The results in Ref. 5 can be applicable here while the blur radius can be related to the resolution of the target.

From these results two measures, TBC and ETC, are degraded while SCR is improved by image motion or vibration. Specific ATR algorithms must be tested to determine if they are predictable from the TBC and ETB metrics or from the SCR metric.
4 Target Segmentation as a Function of Image Motion

4.1 Introduction

This section examines the ability to discriminate target from natural clutter in scenes during image motion and vibration. The output of the extraction process must be a specific decision as to which pixels belong to the target, and which to background. One acceptable, albeit simplistic, way to extract the target from its background is by using a thresholding operation. Thresholding is an operation in which the value of each pixel in the output image depends on the value of the corresponding input pixel relative to a single value known as the threshold. The proper choice of the threshold is very important if one wants to extract the target correctly. A commonly used method of threshold selection is based on examining the histogram of the image gray levels. Threshold selection is relatively easy when the histogram in the image is strongly bimodal with two peaks comparable in size and separated by a deep valley. In a typical histogram, such as in this case, there is no prominent peak corresponding to the target and, in addition, there is no well-defined valley between the two peaks. It becomes extremely difficult to determine an optimum threshold level just on the basis of such an intensity distribution. In Ref. 8, a way is suggested to find the optimal threshold according to this intensity-gradient of the image. The image that is used in this simulation contains a target in the middle and a fractal background around it, as in Fig. 2. The results here are for a target that is brighter than the background; however, the results for darker targets exhibit the same behavior under image motion conditions.

4.2 Segmentation Results

The results of the extracting process are presented in Fig. 6. This figure includes the original image [Fig. 6(a)], the segmentation result [Fig. 6(b)], and results for images degraded by linear motion [Figs. 6(c) and 6(d)]; the nonlinear component of low-frequency vibration [Figs. 6(e) and 6(f)] and high-frequency vibration [Figs. 6(g) and 6(h)] for blur radii equal to 20 pixels. In this figure, there are several pixels that belong to the background and pass the threshold; these are false alarms. For the purpose of quantifying the segmentation process, a different approach is suggested for investigating the threshold selection as a function of image motion and vibration. This approach is similar to the theory of signal detection (TSD). Instead of dealing with probability of detection and probability of false alarm, here the two investigated parameters are FET and FEB. FET is obtained from the number of pixels in the target above the threshold divided by all the pixels in the target, and FEB is the number of pixels in background above the threshold divided by all the pixels in the background. Both FET and FEB are limited between 0 and 1. These two quantitative parameters are an indication of the segmentation process success. The decision process is illustrated in Fig. 7, where it is assumed that the target is brighter than the background. The image that was used in this analysis is the same as before, where positions of the target and the background are known. As shown in Fig. 7, FET and FEB are both functions of the threshold; for each selected threshold, there is a different pair of {FET, FEB}. The selected threshold varies from the maximum gray level of the background to the minimum gray level of the target. The first value of the threshold (threshold1) creates the first point in the graph {FET = 0, FEBmin} and the last value of the threshold (thresholdN) creates the last point in the graph {FEBmax, FET = 1}. The same procedure was applied for different image motion degradations that were previously described. The results for FET as function of FEB are presented in Figs. 8(a), 8(b), and 8(c) for linear motion and high- and low-frequency vibrations, respectively. Each of these figures
contains four different blur radii, the segmentation operation curve (SOC). For the cases of linear motion and high-frequency vibrations, \( d = 20, 40, \) and 60 pixels. For the low-frequency vibrations, the results are for \( t_e/T_0 = 0.1 \) where \( d = d_{\text{min}} = 4.89 \) pixels, \( d_{\text{med}} = 43.7 \) pixels, and \( d_{\text{max}} = 61.8 \) pixels. The case of 'no motion' was plotted as a reference in each of these graphs.

To evaluate effects of restoration on target segmentation, these images have been restored by the constrained least-squares (CLS) filter defined in Ref. 13. The degree to which such real-time restoration is worthwhile is considered. The segmentation process was applied to the restored images according to the optimal threshold described elsewhere. The results of segmentation after restoration are presented in Fig. 9 for the case of linear image motion and \( d = 15 \) pixels.

5 Discussion

The motion MTF for all types of motions can be considered as a low-pass filter (LPF) in the spatial frequency domain. The passage of an image through the image motion MTF causes reduction in contrast and moreover a reduction in the dynamic range of the image gray levels. This reduction of dynamic range is more noticeable for the target histogram than for the background histogram. This means that the gray level of the target histogram is increased and seems to be much brighter. This occurs because the lower gray levels of the target disappear while the higher gray levels do not change. The reason for this phenomenon is that the original target contains small areas of low gray levels. Because of the motion MTF, these areas are assimilated in the background. These changes in the histograms cause the target to be more prominent against the background.

The practical implementation of this behavior is that there are some cases in which image motion blur can contribute to extracting the target from its local background. This phenomenon usually appears when the target and the background have similar gray-level histograms. The image motion gives rise to a new form of the original image in which the histograms of the target and the background are more separate. The same result was indicated in Ref. 11, where it was suggested to blur the image so as to obtain more success in the segmentation process.

Consider now the results of the SOC curves. By comparing the segmentation curves, it can be shown that as the SOC crowds into the upper left-hand corner of the graph, it indicates that the target is easier to extract from the background. The demand for easy extraction of the target from its background is that, for a given FEB, it is desirable to obtain a maximum value of FET. An interesting phenomenon can be seen in Fig. 8. The image motion MTF reduces the fraction of extracted target for a given fraction of extracted background from \( FEB = 0 \) until a critical point of \( FEB = -0.2 \) for all kinds of motion. From this point, image motion improves SOC such that FET becomes larger for a given FEB than for the same FEB in the case of 'no motion' or 'no vibrations.'
The blur radius $d$ is also an important parameter; as $d$ increases, the SOC is improved from the mentioned critical point. This improvement is limited for certain values of $d$; above it, the SOC begins to decrease. The meaning of this result is that image motion increases the ability to extract the target from its background in most of the cases considered. This phenomenon was also indicated in Ref. 10 for detecting a moving target with ATR systems (or detecting a target with moving sensors).

Consider now the results of segmentation after image restoration. In this case, image restoration causes an increase in false alarm targets. On the other hand, however, the target itself after segmentation in the binary image is more precise, i.e., the edges are sharper, and it is easier to identify more fine details in the target. These two contrary conclusions can lead to the following solution. For the purpose of locating the target in the image, it is preferable to apply the segmentation process to the blurred image. After the target location has been found, it is worthwhile to restore the image to identify the target type (for example, tank or track) and to obtain more fine details of the target. It is also important to note that the blur effect is more prominent for a noise-limited situation. As the background becomes noisier, as in the case of fractal images, the blur reduces the noise and makes the background smoother so that it becomes easier to detect the target.

6 Conclusions

The influence of image motion and vibrations on an ATR has been investigated for various blur radii. The analysis includes the influence of motion on three image quality measures that are correlated with the probability of detection. The segmentation process has been presented in a manner similar to the TSD. It was found that image motion and vibrations can, in some cases, improve the ability to extract the target from its local background.

Acknowledgments

This research was supported partially by the Ministry of Science and the Arts in Jerusalem and by the Paul Ivanier Center for Robotics and Production Management at the Ben-Gurion University of the Negev.

References


Ofer Hadar received the BSc degree in electrical and computer engineering from Ben-Gurion University of the Negev, Israel, in 1990 and is now in a combined MS-PhD program at the same university. He is a research assistant in the electro-optics program. His current research interest is in the effects of vibrations and image motion on target acquisition. He has also worked on the influence of motion on the human visual system. Hadar is a member of IEEE and SPIE.

Norman S. Kopeika received the BS, MS, and PhD degrees in electrical engineering from the University of Pennsylvania, Philadelphia, in 1966, 1968, and 1972, respectively. In 1973 he joined the Department of Electrical Engineering, Ben-Gurion University of the Negev, Beer-Sheva, Israel, where he is currently a professor and incumbent of the Reuben and Francis Feinberg Chair in Electro-Optics. From 1978 to 1979 he was a visiting associate professor in the Department of Electrical Engineering, University of Delaware, Newark. His current research interests include atmospheric optics, effects of surface phenomena on optoelectronic device properties, optical communication, electronic properties of plasmas, laser breakdown of gases, the optogalvanic effect, electromagnetic (EM) wave-plasma interaction in various portions of the EM spectrum, and utilization of such phenomena in EM wave detectors and photopreionization lasers. He has published over 100 reviewed journal papers in these areas. His textbook Image System Engineering Principles is to be published by SPIE Press in 1996. Dr. Kopeika is a member of the IEEE, SPIE, the Optical Society of America, and the Laser and Electrooptics Society of Israel.

Stanley R. Rotman received the BS, MS, and PhD degrees from the Electrical Engineering and Computer Science Department of the Massachusetts Institute of Technology in 1979, 1980, and 1985, respectively. His doctoral thesis was entitled “Defect structure of luminescent garnets.” In 1985 he joined the staff of the Science and Technology Division of the Institute for Defense Analyses, where he did research on solid-state lasers, IR target acquisition, and electronic materials characterization. In 1987 he was appointed senior lecturer in the Department of Electrical and Computer Engineering at Ben-Gurion University of the Negev, where he is now a associate professor. His present research interest are the characterization of optical materials for solid-state lasers, modeling IR target acquisition, and digital image processing. He is a member of OSA, IEEE, and SPIE.

Biographies and photographs of other authors not available.