

Object Alignment for Military Optical Surveillance

Oscar J.G. Somsen, and Fok Bolderheij

Abstract—Electro-optical devices are increasingly used for military sea-, land- and air applications to detect, recognize and track objects. Typically, these devices produce video information that is presented to an operator. However, with increasing availability of electro-optical devices the data volume is becoming very large, creating a rising need for automated analysis. In a military setting, this typically involves detecting and recognizing objects at a large distance, i.e. when they are difficult to distinguish from background and noise. One may consider combining multiple images from a video stream into a single enhanced image that provides more information for the operator. In this paper we investigate a simple algorithm to enhance simulated images from a military context and investigate how the enhancement is affected by various types of disturbance.

Keywords—Electro-Optics, Automated Image alignment

I. INTRODUCTION

ELECTRO-OPTICAL devices are increasingly used in military applications to supplement information obtained by more common sensors such as radar, but also to provide information that is hard to obtain by other means. For example, the air defense and command frigates of the Royal Netherlands navy employ far infrared vision to detect sea-skimming missiles. The Joint Strike Fighter will be equipped distributed aperture system for 360° vision. Well known are visual and infrared systems for night vision and infrared systems for missile guidance.

Some electro-optical devices are directly aimed at providing the operator with enhanced vision. Others have as primary function, early detection and recognition of objects. Typically, such objects occupy a small fraction of the image and are obscured by noise, inhomogeneous background, etc. Such objects are easy to miss for an operator that inspects the image by eye, especially when a large volume of data is available from multiple devices. The type of disturbance depends on the applications. Airborne objects such as planes, helicopters and missiles are often observed against a relatively homogeneous background, but require very early detection due to their high velocity. At sea and land horizontal detection of ships and vehicles involves dealing with a much more structured background that often includes the horizon and foreground features such as bushes and waves. Perhaps most easy is vertical detection of ships and vehicles from air or space.

Combining a number of images from an object can help to distinguish it from the background, remove obscurity by foreground features and reduce noise. Also, it can help to

enhance contrast and resolution. If the object is the same in a number of images, a simple averaging procedure will be sufficient to create a sharp object while smoothing out the anomalies. However, the object will not typically be steady within the image due to its own movement or that of the detector. Also, the orientation may change or even the object itself, e.g. when a missile is launched or even when lights are switched on and off. Thus, an enhancement algorithm should at least include an approach to correctly overlay the images before they are averaged or otherwise combined.

In this study, we aim to investigate how multiple images of an object can be aligned to enable enhancement. We consider the military contexts described above. While many techniques for image enhancement are available [1-4], we will first consider simple averaging and correlation algorithms and characterize how these are affected by disturbances such as noise in background variability. Once these disturbances have been sufficiently characterized we may design or select suitable techniques to correct for them. We use an approach of increasing complexity. We start our study with a simple case of artificial images and white noise. When the enhancement of such images is sufficiently clear, we can include variability of background and object before we attempt to analyze experimental data.

In the remainder of this paper we discuss the method of aligning images by correlation, demonstrate it on a simple set of simulate images and characterize how the alignment is affected by variation of contrast and margin around the object.

II. METHODS

The essence of our approach is that we align and average multiple images of the same object with the aim of obtaining a better image. The alignment is done with the help of correlation. Briefly, the correlation of two grey-scale images is obtained by displacing one image, multiplying the values of corresponding pixels of the two images and summing the result over all pixels. For grey-scale images of $N \times M$ pixels the correlation $C(dx, dy)$ can be obtained for any displacement (dx, dy) where dx and dy are integers:

$$C(dx, dy) = \text{corr}(A, B)(dx, dy) \\ = \sum_{x=1}^N \sum_{y=1}^M A(x, y) \cdot B(x + dx, y + dy)$$

Where $A(x, y)$ is the gray-level of image A at pixel (x, y) . The same holds for image B. For computational efficiency we assume the images to be periodic. In that case the correlation is calculated with little computational effort using fast Fourier

O.J.G. Somsen and F. Bolderheij are with the Netherlands Defense Academy, Den Helder, Netherlands. (e-mail: o.jg.somsen@nlda.nl, f.bolderheij@nlda.nl)

transforms. If the object is entirely within the image and the background is sufficiently homogeneous this assumed periodicity will not affect the results.

If the two images are identical, their correlation is at its maximum value when the displacement is zero. Therefore, we define the optimal alignment of one image to another as the displacement for which the correlation is at its maximum value. In this work we consider only translational displacement. When the object is rotated or even deformed between images a more advanced “alignment” is required. A procedure that involves only translation may be applicable in some cases. But, more importantly, our results may help to develop more advanced approaches.

Our research involves simulated images with artificial noise. The noise can be introduced in various ways. Perhaps simplest is, to add a random number to the value of each pixel. The width and shape of the distribution that is used define the type and magnitude of the noise. However, preliminary results showed that a very large noise amplitude (compared to the range of gray-levels in the image) was often needed to see any effect on the alignment procedure. This led to unrealistic images. Therefore, we adopt a different approach and introduce the noise by randomly resetting a fraction of the pixels. Unless stated otherwise the value of the reset pixels is randomly set to 0 (black) or 255 (white) with equal probability. The probability at which each pixel is reset, i.e. the fraction between 0 and 1, determines the amount of noise. This type of noise is harder to describe analytically but it produces more realistic images.

For the actual enhancement procedure, various approaches are possible. When more than two images are to be aligned, it is possible to select one image and align the other images to that “master” image. However, better results are to be expected if the master image is already enhanced, i.e. if the effects of noise and background-heterogeneity have been cancelled (as well as possible). Such a master image may have been obtained from a previous run or from a different type of initialization procedure. We will not consider that here. In general we assume that a master image of sufficient quality is available.

III. RESULTS

The results of a typical result are shown in figure 1. In this case the object was a simple gray-scale top view drawing of a ship (50×200 pixels) on top of a homogeneous background of 150×300 pixels. We created 24 test images by placing the object at a random position in the image and introducing noise as described in the methods section. In this case the noise was considerable. As can be seen in the enlarged image, this obscures most of the object’s-features. Only the contour and some of the largest features can still be distinguished.

The 24 images were aligned with the original (noise free) image. One observes that two sides of the aligned images are black. The figure does not show how well the alignment actually was. It does however show the average of aligned

images. If the alignment is good enough one should expect a reduction of the noise amplitude with a factor $\sqrt{24} \approx 5$. Indeed, a number of features of the original object that were obscured in the test images have been recovered in the final result. Even fine lines, such as the rotor blades of the helicopter, which were completely lost in the test images, can be observed in the final results. This indicates that the alignment of the test images is almost perfect, notwithstanding the strong noise in this example.

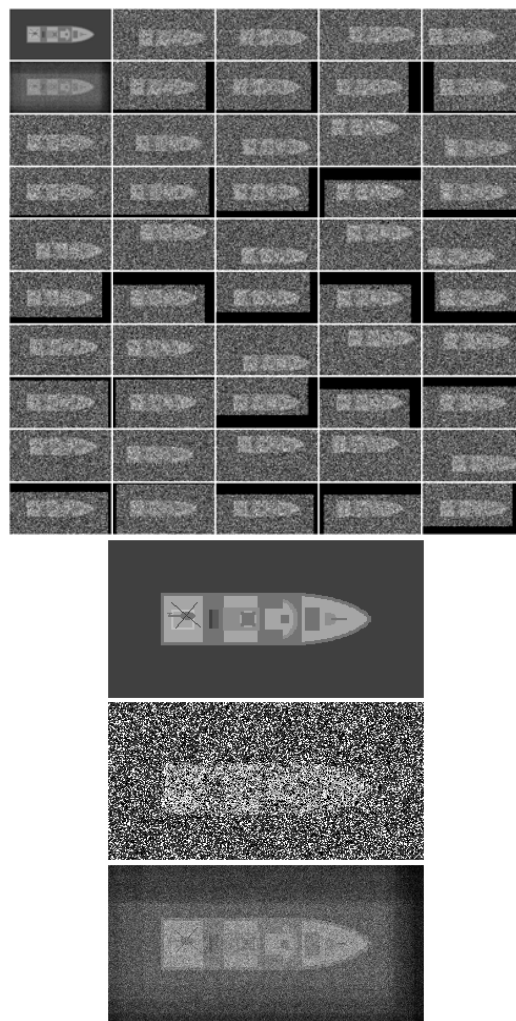


Fig. 1 Top window: Typical results of an “align and average”-procedure. Odd rows: original image (top left) and 24 test images obtained by shifting the object and adding noise (see text). Even rows: Aligned copies of each image and the final result (below the original image). Bottom window: enlarged copies of (top to bottom): original picture, one of the test images and the final result

Of course this simulated case is far from any real data since we have used identical images, a perfectly homogeneous background and perfectly random noise. Also, a perfect reference was available. Such a situation could only be approached with long observation against a very calm sea or sky. Nevertheless, our results show that excellent correlation

and image enhancement can be obtained even in the presence of a large amount of noise at least in theory. This provides a basis to investigate how well our approach works in less ideal circumstances.

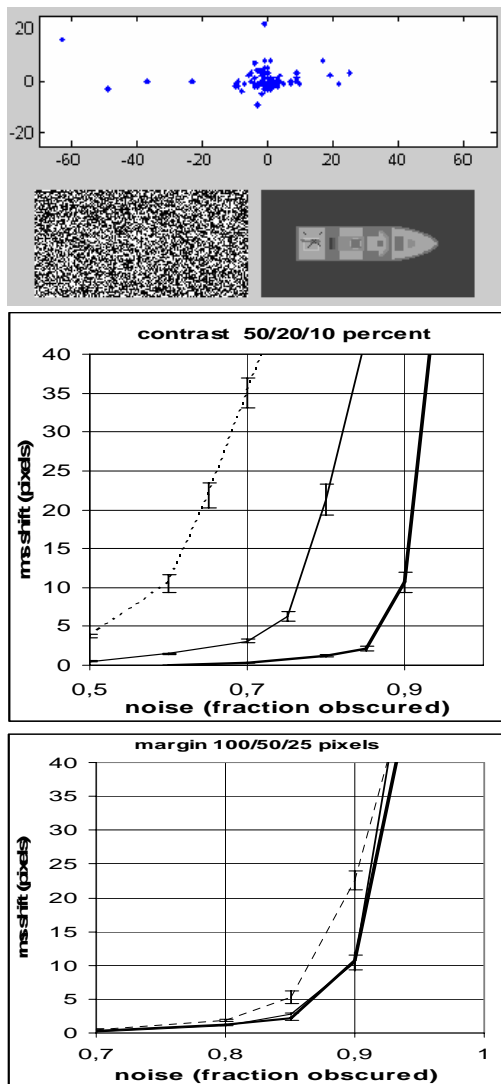


Fig. 2 Obscuring by noise. Top panel: Distribution of shifts (measured in pixels) induced by adding noise. Left: One of the experimental images. 90% of the pixels has been obscured. Right: Original image (200x50 pixel pattern with 50 pixels margin). Middle Panel: contrast at 50% (thick), 20% (thin) or 10% (dotted) of the full possible contrast (256 gray levels). Bottom panel: margin at 100 (thick), 50 (thin) or 25 (dotted) pixels around the object (200x50 pixels). The error bars represent the fluctuation in 3-10 runs each with 100 simulated images

The central part of our approach involves the alignment procedure. We shall now consider this more quantitatively. First, the top panel of figure 2 shows a distribution of alignment errors for a typical run. The object is not shifted within the image but only obscured by noise. Any shift found by the alignment algorithm then corresponds to error. In this

example the rms-error is 7.5 pixels horizontally and 1.4 pixels vertically. This is relatively small compared to the drawing which measures 50x200 pixels, especially since the noise is very high. In fact, 90% of the pixels is obscured and the object no longer distinguishable to the naked eye (top middle image). In some cases the alignment error is much larger than the rms-value. Possibly, the distribution of the alignment errors is not Gaussian. Instead it appears that a "lock" between the images is lost when the noise is unfavorable. The procedure may be improved by identifying such images in the data and discarding them.

To further characterize the performance of the alignment procedure we varied the contrast of the object. The object within the image includes gray levels from 0 to 255. Default, this is multiplied by 0.5 and 64 is added, so that the background is gray instead of black. In this series we reduced the contrast by decreasing the multiplication factor to 0.2 and 0.1. As may be expected the results in figure 2 show increased alignment errors. The threshold at which the shift reaches 10 pixels decreases from 0.9 to 0.6 when the contrast is reduced from 50 to 10%. Let us compare this to the areas within the image that are covered by the object and by noise. The latter increases fourfold (from 10% to 40% of the image) while the former decrease with a factor 1.5 (from 90% to 60% of the image). The ratio of the two areas increases with a factor $4 \cdot 1.5 = 6$ which agrees nicely with the fivefold reduction of contrast. This indicates that the required contrast may be proportional to the ratio of noise- and object-covered areas within the image

In addition, we varied the margin around the object. This size of the margin is critical when selecting small sub-images. When a small margin is used, there is a risk that the object is not completely within the sub-image. But, when a large margin is used the sub-image is largely composed of background and noise and it may be very hard for a computer algorithm to distinguish, and in our case align, the object. In figure 2 good results were obtained with margins from 25 pixels (half the width of the object within the image) to 100 pixels. In the latter case the object occupies less than 10% of the image. The alignment was good in all cases: more than 80% of the pixels had to be obscured by noise to cause a significant misalignment. Contrary to the aforementioned expectation, the results are best when the margin is large. The increased amount of amount of noise-covered background within the image does not appear to affect alignment procedure, at least in this case. Of course, the background in this case was homogeneous and composed only of noise. Problems may arise when the background is not homogeneous. But, under the conditions presented here, it may be useful to select a relatively large margin since it improves the probability that the object is within the image while the alignment performance does not necessarily decrease.

IV. DISCUSSION AND CONCLUSION

We presented a preliminary study on the performance of a simple technique that produces a single enhanced image from multiple noisy images of an object with reduced noise. Such techniques may be used to assist human operators to identify objects in surveillance video. We used a straightforward correlation procedure. While more advanced techniques are available [1-4], this approach will help us to characterize the effect of various types of disturbance. We considered objects at a large distance (i.e. covering a small part of the image) against a relatively homogeneous background which is relevant for some military and especially naval application.

We considered the enhancement results under various circumstances. We used a very simple scenario of an artificial object that was translated against a homogeneous background and obscured by noise. In ideal cases we found almost perfect alignment even when the images were largely obscured by noise and difficult to view by naked eye. When the noise increases an increased contrast is needed which may be estimated from the fraction of the image that is obscured by it. To our surprise we found that a relatively large margin around the object appears to be beneficial for the alignment, which may make the detection easier.

Our results are not yet sufficient to design a useful procedure. We will need to consider other effects such as non-homogeneous backgrounds and rotational alignment. We hope to study this in the near future. For more complicated cases it will be necessary to also consider possible changes of the morphology of the object [4]. Also, detection against the horizon or other structured backgrounds is difficult and requires advanced techniques. Nevertheless, with some basic aspects included the procedure may be considered in some cases of actual surveillance and our results could help to develop better surveillance techniques.

REFERENCES

- [1] Sylman D, Zalevsky Z, Mico V, et al. 2009. Two-dimensional temporal coherence coding for super resolved imaging. *OPTICS COMMUNICATIONS* 282: 4057-4062.
- [2] Temizel A, Vlachos T. 2006. Wavelet domain image resolution enhancement. *IEEE PROCEEDINGS-VISION IMAGE AND SIGNAL PROCESSING*: 153: 25-30.
- [3] Lee J, Kim YH, Nam JH. 2008. Adaptive noise reduction algorithms based on statistical hypotheses tests. *IEEE TRANSACTIONS ON CONSUMER ELECTRONICS*: 54: 1406-1414.
- [4] Mattiucci N, D'Aguanno G, Scalora M, et al. 2009. Transmission function properties for multi-layered structures: Application to super-resolution. *OPTICS EXPRESS* 17: 17517-17529.