SPIE Aerosense 2001 Conference on Signal Processing, Sensor Fusion, and Target Recognition X, April 16-20, Orlando FL. (Minor errors in published version corrected.)

A Signal-Level Fusion Model for Image-Based Change Detection in DARPA's Dynamic Database System

Mark J. Carlotto
Veridian Corporation, 1400 Key Blvd., Suite 100, Arlington VA 22209

Abstract

Improving change detection performance (probability of detection/false alarm rate) is an important goal of DARPA's Dynamic Database (DDB) program. We describe a novel approach based on fusing the outputs from two complementary image-based change detection algorithms. Both use historical imagery over the region of interest to construct normalcy models for detecting change. Image level change detection (ILCD) segments the set of images into temporally co-varying pixel sets that are spatially distributed throughout the image, and uses spatial normalcy models constructed over these pixel sets to detect change in a new image. Object level change detection (OLCD) segments each image into a set of spatially compact objects, and uses temporal normalcy models constructed over objects associated over time to detect change in the new image. Because of the orthogonal manner in which ILCD and OLCD operate in space-time, false alarms tend to decorrelate. We develop signal-level statistical models to predict the performance gain (output/input signal to noise ratio) of each algorithm individually, and combined using 'and' fusion. Experimental results using synthetic aperture radar (SAR) images are presented. Fusion gains ranging from slightly greater than unity in low clutter backgrounds (e.g., open areas) to more than 20db in complex backgrounds containing man-made objects such as vehicles and buildings have been achieved and are discussed.

Key Words: Change detection, algorithm fusion , image-level fusion, performance modeling, synthetic aperture radar (SAR).

1. Introduction

Multi-sensor fusion has been applied in a variety of remote sensing and reconnaissance applications including land use/land cover classification, terrain and feature extraction, change detection, and object recognition. In DARPA's Dynamic Database (DDB) program (Kessler 1990) fusion techniques are used to improve change detection and object classification performance. This paper describes a new way to improve change detection performance by fusing the outputs from two *different* change detection algorithms operating on data from the *same* sensor. We show that because of the orthogonal nature of the change detection algorithms used it is possible to achieve significant fusion gains by this approach.

After briefly summarizing the two change detection algorithms in Section 2, signal-level statistical models are derived in order to predict the performance of each algorithm individually, and combined using 'and' fusion (Section 3). Results from several experiments over different image backgrounds are presented in Section 4. Areas for future work are discussed in Section 5.

2. Detecting Change in Space and Time

DDB employs two different algorithms for detecting change. The first, known as object-level change detection (OLCD), was originally developed under DARPA's Semi-Automated IMINT Processing (SAIP) effort (Welby 1999). OLCD detects and maintains a database of target-like regions in SAR (Tom et al 2000) and electro-optical (EO) imagery (Hoogs and Mundy 2000). Regions which appear, disappear or change state are labeled as possible changes. Conceptually, OLCD spatially segments images into compact target-like regions, models the behavior of these regions, and detects changes in time.

The second algorithm, image level change detection (ILCD) was developed under the DDB program for detecting changes in complex backgrounds (Carlotto 1999, 2000b). ILCD, which operates on EO as well as SAR, segments a set of reference images (i.e., images that do not contain targets of interest) into spatially distributed, temporally co-varying pixel sets. Each pixel set corresponds to a unique background type (Carlotto 2000a). Changes in a new image are detected spatially by comparing the values within each pixel set to the average over the pixel set.

Changes detected by ILCD and OLCD are fused by the object level change fusion (OLCF) component in DDB (Berlin et al 2000). Because of the orthogonal manner in which ILCD and OLCD operate in space-time (ILCD is temporal segmentation followed by spatial normalcy modeling and detection while OLCD is spatial segmentation followed by temporal normalcy modeling and detection), false alarms tend to decorrelate. OLCF exploits this property to enhance change detection performance by reducing the false alarm rate.

3. Performance Modeling

To simplify our analysis we assume that OLCD operates on pixels instead of objects, and approximate the spatial and temporal normalcy models used by ILCD and OLCD as simple averages in space and time (Figure 1).

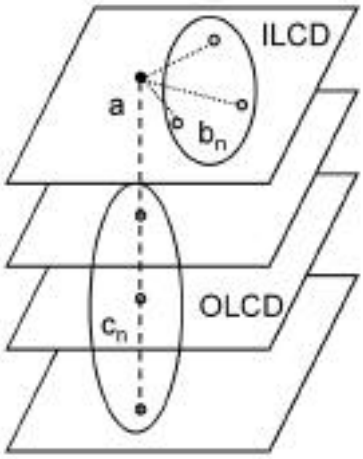

Figure 1 ILCD and OLCD normalcy models

Consider the pixel a in Figure 1. Let b_n be the other pixels in the same co-varying pixel set as a. The error in predicting a from the spatial average of the b_n (ILCD prediction error) is

$$e_b = a - \frac{1}{N}\sum b_n \tag{1}$$

where the total number of pixels in the set is $N+1$. Let c_n be the regions (assumed here to be a single pixel in extent) associated with a over time. The OLCD prediction error is the error in predicting a from the temporal average of the c_n

$$e_c = a - \frac{1}{M}\sum c_n \,. \tag{2}$$

where M is the number of reference images.

The mean squared ILCD prediction error is:

$$\begin{aligned} E[e_b^2] &= E\left[\left(Na - \sum b_n\right)^2\right] \\ &= \frac{1}{N^2}\left(N^2 E[a^2] - 2N^2 E[ab_n] + NE[b_n^2] + (N^2 - N)E[b_n b_m]\right) \end{aligned} \tag{3}$$

Assuming jointly normal random variables we can parameterize the statistics in terms of the means and covariances as follows:

$$E[a^2] = \sigma_a^2 + \mu_a^2 \quad E[ab_n] = \sigma_{ab}^2 + \mu_a\mu_b \quad E[b_n^2] = \sigma_b^2 + \mu_b^2 \quad E[b_n b_m] = \sigma_{bb'}^2 + \mu_b^2 \tag{4}$$

from which we obtain

$$E\left[e_b^2\right] = \frac{1}{N^2}\left(N^2(\sigma_a^2 + \mu_a^2) - 2N^2(\sigma_{ab}^2 + \mu_a\mu_b) + N(\sigma_b^2 + \mu_b^2) + (N^2 - N)(\sigma_{bb'}^2 + \mu_b^2)\right). \tag{5}$$

The signal to noise ratio (SNR) is the signal divided by the noise power:

$$SNR_b = \frac{\mu_a^2 - 2\mu_a\mu_b + \mu_b^2}{\sigma_a^2 - 2\sigma_{ab}^2 + \frac{1}{N}\sigma_b^2 + (1-\frac{1}{N})\sigma_{bb'}^2} = \frac{(\mu_a - \mu_b)^2}{\sigma_a^2 - 2\sigma_{ab}^2 + \frac{1}{N}\sigma_b^2 + (1-\frac{1}{N})\sigma_{bb'}^2}. \tag{6}$$

If we assume:

$$\sigma_a^2 = \sigma_b^2 = \sigma^2 \quad \sigma_{ab}^2 = \sigma_{bb'}^2 = \beta\sigma^2 \tag{7}$$

where $0 \le \beta \le 1$, when N is large

$$SNR_b = \frac{(\mu_a - \mu_b)^2}{\left[1 - 2\beta + \frac{1}{N} + (1-\frac{1}{N})\beta\right]\sigma^2} = \frac{(\mu_a - \mu_b)^2}{\left[(1-\beta)(1+\frac{1}{N})\right]\sigma^2} \approx \frac{(\mu_a - \mu_b)^2}{\sigma^2(1-\beta)}. \tag{8}$$

As the covariance between random variables in the ILCD normalcy model increases (approaches unity) the SNR increases because we can do a better and better job of estimating a spatially from the b_n.

The OLCD SNR is derived in the same fashion and is

$$SNR_c = \frac{(\mu_a - \mu_c)^2}{\sigma_a^2 - 2\sigma_{ac}^2 + \frac{1}{M}\sigma_c^2 + (1-\frac{1}{M})\sigma_{cc'}^2} = \frac{(\mu_a - \mu_c)^2}{\left[1 - 2\gamma + \frac{1}{M} + (1-\frac{1}{M})\gamma\right]\sigma^2} \tag{9}$$

where as above

$$\sigma_a^2 = \sigma_c^2 = \sigma^2 \quad \sigma_{ac}^2 = \sigma_{cc'}^2 = \gamma\sigma^2. \tag{10}$$

As the number of images M increases

$$SNR_c = \frac{(\mu_a - \mu_c)^2}{\left[(1-\gamma)(1+\frac{1}{M})\right]\sigma^2} \approx \frac{(\mu_a - \mu_c)^2}{\sigma^2(1-\gamma)}. \tag{11}$$

As the covariance between random variables in the OLCD normalcy model increases (approaches unity) its SNR also increases because it can do a better and better job of estimating a temporally from the c_n.

In DDB OLCF fuses ILCD and OLCD outputs at the object level. ILCD objects are connected regions where the squared prediction error exceeds a specified constant false alarm rate (CFAR) threshold. OLCD objects are compact regions that have changed based on a similar CFAR criterion. Here we approximate the performance of OLCF by modeling it, as we have ILCD and OLCD, at the signal (prediction error) level.

The expected value of the product of the ILCD and OLCD prediction error images is

$$\begin{aligned} E[e_b e_c] &= E\left[\left(Na - \sum_N b_n\right)\left(Ma - \sum_M c_n\right)\right] \\ &= E[a^2] - E[ab_n] - E[ac_n] + E[b_n c_m] \\ &= (\sigma_a^2 + \mu_a^2) - (\sigma_{ab}^2 + \mu_a\mu_b) - (\sigma_{ac}^2 + \mu_a\mu_c) + (\sigma_{bc}^2 + \mu_b\mu_c) \end{aligned} \tag{12}$$

from which we derive the OLCF SNR

$$SNR_{bc} = \frac{\mu_a^2 - \mu_a\mu_b - \mu_a\mu_c + \mu_b\mu_c}{\sigma_a^2 - \sigma_{ab}^2 - \sigma_{ac}^2 + \sigma_{bc}^2} = \frac{\mu_a^2 - \mu_a\mu_b - \mu_a\mu_c + \mu_b\mu_c}{\sigma^2(1-\beta-\gamma+\alpha)} \tag{13}$$

where $\sigma_{bc}^2 = \alpha\sigma^2$. This is analogous to an 'and' fusion rule.

Table 1 summarizes the algorithm correlations: α is the correlation of random variables between ILCD and OLCD normalcy models, and β and γ are the correlations within the ILCD and OLCD normalcy models respectively.

Table 1 ILCD, OLCD, and OLCF parameters

	a	b	b'	c	c'
a	1	β	β	γ	γ
b	β	1	β	α	α
b'	β	β	1	α	α
c	γ	α	α	1	γ
c'	γ	α	α	γ	1

To show the benefit of fusion we define the OLCF to ILCD, and OLCF to OLCD processing gains:

$$Gain_{bc/b} = \frac{1-\beta}{1-\beta-\gamma+\alpha} \qquad Gain_{bc/c} = \frac{1-\gamma}{1-\beta-\gamma+\alpha} \tag{14}$$

assuming $\mu_b = \mu_c$. In general, ILCD performs well when β is large, OLCD performs well when γ is large, and OLCF performs well when α is small. Figure 2 plots theoretical fusion gains for $\gamma = \beta$. Appendix A describes statistical methods for estimating α, β, and γ from images.

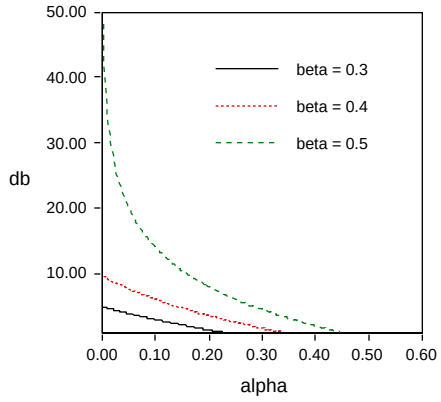

Figure 2 Fusion gains predicted by signal level performance model

Table 2 Targets in open

SNR	Measured SNR (db)	Measured Processing Gain (db)	Predicted Processing Gain (db)	Estimated Component Parameters
Input	43.1			
ILCD	42.3	-0.8	2.8	$\hat{\beta} = 0.34$
OLCD	23.5	-19.6	0	$\hat{\gamma} = 0.25$
OLCF	44.3	1.2	6.5	$\hat{\alpha} = 0.06$

Table 3 Targets in trees

SNR	Measured SNR (db)	Measured Processing Gain (db)	Predicted Processing Gain (db)	Estimated Component Parameters
Input	35.2			
ILCD	38.7	3.5	4.6	$\hat{\beta} = 0.47$
OLCD	24.7	-10.5	0.7	$\hat{\gamma} = 0.31$
OLCF	49.8	14.6	8.7	$\hat{\alpha} = 0.14$

Table 4 Targets near facility

SNR	Measured SNR (db)	Measured Processing Gain (db)	Predicted Processing Gain (db)	Estimated Component Parameters
Input	26.2			
ILCD	38.6	12.4	9.5	$\hat{\beta} = 0.69$
OLCD	17.5	-8.7	2.2	$\hat{\gamma} = 0.41$
OLCF	48.4	22.2	16.1	$\hat{\alpha} = 0.27$

Table 5 Targets embedded in manmade clutter

SNR	Measured SNR (db)	Measured Processing Gain (db)	Predicted Processing Gain (db)	Estimated Component Parameters
Input	27.2			
ILCD	46.9	19.7	8.5	$\hat{\beta} = 0.66$
OLCD	14.2	-13	4.1	$\hat{\gamma} = 0.53$
OLCF	56.2	29	16.4	$\hat{\alpha} = 0.34$

4. Experimental Results

The performance of a detection algorithm is determined by its output signal and noise probability distributions $p(y|signal)$ and $p(y|noise)$. The probabilities of detection (Pd) and false alarm (Pfa) depend on the operating point t

$$P_d(t) = p(y > t|signal) \qquad P_{fa}(t) = p(y > t|noise). \tag{15}$$

Algorithm performance can also be expressed in terms of the signal to noise ratio (SNR)

$$SNR = (\mu_s - \mu_n)^2 / (\sigma_s^2 + \sigma_n^2). \tag{16}$$

Computed over the new image (input SNR), the SNR is indicative of the performance of single image Neyman-Pearson (CFAR) detector. Computed over the prediction error image (output SNR), the SNR is indicative of the CFAR performance of ILCD or OLCD at the signal (pre-detection) level. The difference between input and output SNR (in db) is a measure of the performance gain of the change detector.To evaluate the performance models we selected four SAR images collected by Veridian's DCS sensor over the DDB Eglin Site 12 study area. These images represent the deployment of targets in four different clutter environments: in the open, near trees, near a facility, and embedded in man-made clutter. Tables 2-5 summarizes the experimental results. The first column in each table shows the input SNR and the output SNR from ILCD, OLCD, and OLCF (pre-detection level). The measured processing gain (second column) is the output minus input SNR. The predicted processing gains (third column) were obtained from Eqs. 8, 11, and 14 using component parameters (fourth column) estimated from the images as discussed in the appendix. As the complexity of the scene increases the fusion gain increases with the highest gains achieved in manmade backgrounds.

Although the predicted and measured performance gains agree in terms of their overall trend there are significant differences between them. In heterogeneous scenes, using a single global statistic to represent the covariances within different background types is one potential source of error. Another is attempting to compute statistics over background typescontaining a limited number of samples (i.e., when M and/or N is small).

5. Future Work

Improving the accuracy of our performance model is one important area for future work. In particular we shall explore the possibility of developing local estimates of performance over each background type.

Because OLCD fuses ILCD and OLCD outputs at the object level, in order to obtain a more meaningful measure of performance it is necessary to model the performance of the components following CFAR detection. This is a second area for future work.

Detecting change in high clutter backgrounds with significant target obscuration is an important operational challenge. Combining detections across looks at the signal level using 'or' fusion (Carlotto 2000b) or using adaptive fusion strategies (Liggins and Nebrich 2000) are two possible alternatives. Extending our performance model to other fusion rules is a third area for future work.

Appendix - Parameter Estimation

β characterizes the performance of ILCD. To estimate β we first compute the correlation coefficient between the new image y and its average over S_k

$$\rho_i = E\left[\left(\frac{1}{N}\sum_n y_n\right)\left(\frac{1}{N}\sum_{n'} y_{n'}\right)\right] \Big/ \sqrt{E\left[\left(\frac{1}{N}\sum_n y_n\right)^2\right] E\left[y_{n'}^2\right]} \quad n,n' \quad S_k$$
$$= \sigma_{ab}^2 \Big/ \sqrt{\frac{1}{N^2}\left[N\sigma_b^2 + \left(N^2 - N\right)\sigma_{bb'}^2\right]\sigma_a^2} = \frac{\beta}{\sqrt{\left(\frac{1}{N}\right) + \left[1 - \left(\frac{1}{N}\right)\right]\beta}} \approx \sqrt{\beta} \tag{A-1}$$

when N is large. We use the average of the estimates over all pixel sets as the estimated value of β for the image.

γ characterizes the performance of OLCD. To estimate γ we measure the correlation coefficient between a pixel in the new image y and the temporal average of the associated reference image pixels x_n, and average the results to obtain an estimate of γ for the image:

$$\rho_o = E\left[\left(\frac{1}{M}\sum_m x_m\right) y_{m'}\right] \Big/ \sqrt{E\left[\left(\frac{1}{M}\sum_m x_m\right)^2\right] E\left[y_{m'}^2\right]}$$
$$= \sigma_{ac}^2 \Big/ \sqrt{\frac{1}{M^2}\left[M\sigma_c^2 + \left(M^2 - M\right)\sigma_{cc'}^2\right]\sigma_a^2} = \frac{\gamma}{\sqrt{\left(\frac{1}{M}\right) + \left[1 - \left(\frac{1}{M}\right)\right]\gamma}} \approx \sqrt{\gamma} \tag{A-2}$$

when M is large.

Finally to obtain an estimate of α, which characterizes the performance gain of OLCF, we measure the correlation coefficient between the ILCD and OLCD predicted error images:

$$\rho_f = \frac{Var\left[e_b e_c\right]}{\sqrt{Var\left[e_b^2\right] Var\left[e_c^2\right]}} \tag{A-3}$$

where

$$Var\left[e_b e_c\right] = \sigma_a^2 - \sigma_{ab}^2 - \sigma_{ac}^2 + \sigma_{bc}^2$$
$$Var\left[e_b\right] = \sigma_a^2 - 2\sigma_{ab}^2 + \frac{1}{N}\sigma_b^2 + \left(1 - \frac{1}{N}\right)\sigma_{bb'}^2 \tag{A-4}$$
$$Var\left[e_c\right] = \sigma_a^2 - 2\sigma_{ac}^2 + \frac{1}{M}\sigma_c^2 + \left(1 - \frac{1}{M}\right)\sigma_{cc'}^2$$

Expanding Eq. 7 we have

$$\rho_f = \frac{\sigma_a^2 - \sigma_{ab}^2 - \sigma_{ac}^2 + \sigma_{bc}^2}{\sqrt{\sigma_a^2 - 2\sigma_{ab}^2 + \frac{1}{N}\sigma_b^2 + (1-\frac{1}{N})\sigma_{bb'}^2}\sqrt{\sigma_a^2 - 2\sigma_{ac}^2 + \frac{1}{M}\sigma_c^2 + (1-\frac{1}{M})\sigma_{cc'}^2}} = \frac{1-\beta-\gamma+\alpha}{\sqrt{\left(1+\frac{1}{N}\right)(1-\beta)}\sqrt{\left(1+\frac{1}{M}\right)(1-\gamma)}} \quad \text{(A-5)}$$

from which the following expression is obtained

$$\hat{\alpha} = \rho_f \sqrt{\left(1+\frac{1}{N}\right)(1-\beta)}\sqrt{\left(1+\frac{1}{M}\right)(1-\gamma)} - (1-\beta-\gamma). \quad \text{(A-6)}$$

References

Mark Berlin, Mark Nebrich, and Martin Liggins, "Wide-Area SAR-EO Object Level Change Fusion Process for DDB," *2000 Meeting of the MSS National Symposium on Sensor and Data Fusion*, Kelly AFB, San Antonio TX, June 2000.

Mark Carlotto, "Image-level change detection," *1999 Meeting of the MSS National Symposium on Sensor and Data Fusion*, Applied Physics Laboratory, Johns Hopkins University, Laurel MD, May 1999.

Mark Carlotto, "Detecting partially-obscured objects using multi-look SAR," *2000 Meeting of the MSS National Symposium on Sensor and Data Fusion*, Kelly AFB, San Antonio TX, June 2000.

Mark Carlotto, "Nonlinear background estimation and change detection for wide area search," *Optical Engineering*, Vol. 39, No. 5, May 2000.

Anthony Hoogs and Joseph L. Mundy, " Information Fusion for EO Object Detection and Delineation," *2000 Meeting of the MSS National Symposium on Sensor and Data Fusion*, Kelly AFB, San Antonio TX, June 2000.

Martin E. Liggins II and Mark A. Nebrich, "Adaptive multi-image decision fusion," *SPIE Aerosense*, April 2000, Orlando FL.

Otto Kessler, http://web-ext2.darpa.mil/tto/Programs/ddb.html, July 2000.

Victor Tom, Helen Webb, and David Lefebvre, " Wide-Area SAR Object Level Change Detection," *2000 Meeting of the MSS National Symposium on Sensor and Data Fusion*, Kelly AFB, San Antonio TX, June 2000.

Steven Welby, http://www.darpa.mil/iso/sAIP_Briefing/index.htm, 1999.