

Local Co-Registration Adjustment for Anomalous Change Detection

James Theiler and Brendt Wohlberg

EDICS: GEO-MULT

Abstract—We describe an approach for improving the robustness to misregistration of pixel-wise anomalous change detection (ACD) algorithms. The aim of ACD is to distinguish actual anomalous changes from the irrelevant incidental differences that occur throughout the scene. For such change detection to be effective, it is important that corresponding pixels in the two images of interest correspond to the same location in the scene. Indeed, one of the most confounding sources of incidental differences is the inevitable imprecision in the co-registration of the two images. We address this with small local adjustments to the co-registration which leads to a modified misregistration-insensitive measure of anomalousness. Several variants are considered, and the resulting performance improvements are evaluated using both real and simulated changes, and real and simulated misregistration.

Index Terms—Change detection, Anomaly detection, Anomalous change detection, Co-registration, Multispectral imagery, Hyperspectral imagery

I. INTRODUCTION

Anomalous change detection (ACD) seeks the small, rare, and/or unusual changes that occur in a scene, based on two (or more) images taken of the same scene at different times [1]. The emphasis of ACD is on distinguishing these anomalous changes from the pervasive differences (due, for instance, to differences in calibration, illumination, atmospheric conditions, *etc.*) that occur throughout a scene.

While there are many cases in which pervasive differences are also of interest – *e.g.* flood damage, crop failure, urbanization – the motivation for ACD is that the small but unusual changes could easily escape the notice of a human analyst (who ultimately decides whether a given change is actually *interesting* or *meaningful*). What automated ACD offers is a way to cull through a collection of imagery, and to narrow down the changes that the analyst may need to examine.

One of the most confounding sources of “pervasive difference” is misregistration of the images. Thus it is important to align the images as precisely as possible in the first place, so that corresponding pixels in the two images correspond to the same position in the scene [2], [3], [4], [5], [6], [7], [8]. In some cases this alignment can be performed to subpixel accuracy, but there is always some residual misregistration in the aligned images. Since the effects of this misregistration are pervasive over the whole scene, ACD already provides

some robustness to misregistration [9], [10], [11]. In this paper, we develop a more active misregistration compensation strategy, Local Co-Registration Adjustment (LCRA), based on preliminary work reported earlier [12], [13]. We emphasize that LCRA is *not* a co-registration algorithm in its own right, but a scheme that is applied after the co-registration process. Its purpose is not to improve that co-registration, but to improve the anomalous change detection by making it more robust to the residual misregistration that inevitably remains. As such, its efficacy diminishes with increasing accuracy of the initial co-registration, but we have observed performance improvements even when the misregistration errors are smaller than a pixel.

This paper is organized as follows. Section II describes the direct pixel-wise ACD algorithms that underlie the local adjustment schemes that we are advocating here. These pixel-wise algorithms define an “anomalousness” which our local adjustments attempt to minimize. Before diving into the details of how these local adjustments work, however, we step back in Section III to briefly describe how ACD algorithms can be evaluated. While this evaluation is carried out on real data, quantitative results may require simulated misregistration and/or simulated anomalous pixels; in an effort to keep this section brief, the simulation details are described in the Appendix. With the evaluation criteria well-defined, we describe the details of local adjustment algorithms in Section IV, and provide computational experiments demonstrating the performance. Finally, in Section V, we conclude.

II. PIXEL-WISE ANOMALOUS CHANGE DETECTION

Let χ and γ denote the two images of interest. At the position indexed by (k, l) , we have vector-valued pixels $\mathbf{x} = \chi_{k,l} \in \mathbb{R}^{d_x}$ and $\mathbf{y} = \gamma_{k,l} \in \mathbb{R}^{d_y}$, with d_x and d_y the number of spectral channels in χ and γ , respectively. Our goal is to produce an “anomalousness” image A , in which each scalar-valued pixel $A_{k,l}$ represents how anomalous the change is at the position (k, l) . For pixel-wise ACD algorithms, we can write the anomalousness at (k, l) in terms of a function $\mathcal{A}(\mathbf{x}, \mathbf{y})$ that depends only on the pixel values at (k, l) , *i.e.*

$$A_{k,l} = \mathcal{A}(\chi_{k,l}, \gamma_{k,l}). \quad (1)$$

Algorithms that have been proposed for pixel-wise ACD include the chronochrome [14], neural net prediction [15], covariance equalization [16], multivariate alteration detection [17], and a machine learning framework [18] that is related to mutual information, and which has led to a number of variations that optimize for different situations, such as

Both authors are with Los Alamos National Laboratory, Los Alamos, NM 87545, USA. Email: jt@lanl.gov and brendt@lanl.gov

This work was carried out under the auspices of the National Nuclear Security Administration of the U.S. Department of Energy at Los Alamos National Laboratory under Contract No. DE-AC52-06NA25396 and was supported by the NNSA’s Laboratory Directed Research and Development Program.

sub-pixel anomalies [19] or fat-tailed elliptically contoured data distributions [20]. A more direct information theoretic change detection has also been proposed [21]. In all of these pixel-based algorithms, a scalar “anomalousness” value is assigned to every pixel in the image, and those pixels with the highest anomalousness value are the top candidates for the locations of anomalous change. Our approach for misregistration compensation can be applied to any of these pixel-based ACD algorithms, but in this paper we concentrate on a family of detectors constructed from the covariance of the data [22]; these detectors employ both the auto-covariances of the individual images and the cross-covariance of the two aligned images. Specific members of this family used here are chronochrome (CC) [14], the hyperbolic anomalous change detector (HACD) [18], and the elliptically-contoured change detector (ECACD) [20].

III. EVALUATING ACD ALGORITHMS

A difficulty that arises in any kind of rare target detection problem (such as anomaly or anomalous change detection) is that the targets, being rare, provide few examples for evaluating algorithm performance. While results on real data, with real anomalous changes, are certainly important, these evaluations are inevitably anecdotal, and should be augmented with more controlled experiments with a statistically useful number of anomalous changes.

In this paper, we concentrate on two datasets for evaluating the relative performance¹ of different approaches for ACD. This performance is quantified in terms of receiver-operator characteristic (ROC) curves that plot detection rate of anomalous changes against false alarm rate. Because we care about low false alarm rates in particular, we plot the false alarm rate on a logarithmic axis.

1) *Airborne remote sensing*: We used a 126-channel HyMap image pair with 280×800 pixels from the RIT Target Detection Blind Test [23]. The radiance images were used, and anomalous changes were simulated using the framework described in the Appendix. (The “self test” image was treated as the *normal* image and the “blind” image as the *base* image.) The two images are based on separate overflights, and have been georegistered, using ground control points [24], as part of the blind test experiment. An advantage of this dataset is that the residual misregistration is realistic. A disadvantage is that the detailed structure of the misregistration is unknown.

2) *Desktop clutter*: A pair of RGB images of desktop clutter is displayed in Fig. 2. Unlike the aerial imagery, the misregistration for the desktop clutter is effectively zero, which means that simulated misregistration (using the methodology described in the Appendix) of precisely known detail can be imposed. Also, the anomalous changes in this image pair are real changes in the scene; while this is arguably more realistic than simulated anomalous changes, it provides only a very limited range (and a very limited number) of anomalous changes.

¹We caution against reading too much into *absolute* performance numbers; while trends in relative performance of algorithms can provide useful comparative information, the absolute performance will vary dramatically from problem to problem.

IV. LOCAL CO-REGISTRATION ADJUSTMENT

Pixel-based ACD algorithms effectively assume the two images are perfectly co-registered, and one way to improve ACD performance is to provide a more accurate co-registration. But even the best co-registration algorithms will exhibit residual misregistration, and our suggestion is to make local co-registration adjustments to compensate for this misregistration.

We have previously noted [13, Sec. 2.1] the similarity of the approach proposed here to the use of mutual information in co-registration algorithms [2], [3], [4], [7], as well as to the use of mutual information for simultaneous co-registration and change detection [9]. In particular, our local adjustments are similar to the “local rigid displacements” described by Inglada and Giros [5] (see their Fig. 1). Both methods look for small adjustments to optimize a similarity measure (or, in our case, the *dissimilarity* measure given by “anomalousness”). But while the rigid displacements are deliberate attempts to actually improve the image co-registration on local patches, our adjustments are at the single pixel level, and are not explicitly treated as improvements to the co-registration itself. As an operational issue, the local rigid displacements might be performed to produce a better pair of co-registered images, and that would be followed by our local adjustment scheme. The ultimate aim of our scheme is not to produce more precisely registered images; it is to provide more resilience to residual misregistration for pixel-wise change detection algorithms. This distinction is motivated by the informal observation that the “min” operator is generally more stable than the “argmin.”

The following subsections describe the local co-registration adjustment algorithm in more detail. In IV-A, we describe the basic algorithm, and note its inherent asymmetry. This is followed in IV-B with the description of a symmetric variant. Both of these algorithms consider integer pixel offset adjustments, but in IV-C, the extension to non-integer offsets is considered. An additional extension that uses the local adjustments to tweak up the underlying pixel-wise detector by recomputing the cross-covariance matrix is investigated in IV-D.

A. Asymmetric Algorithm

The following scenario provides a motivation for the simplest variant of the local co-registration adjustment algorithm. We emphasize that the asymmetric nature of the *algorithm* corresponds to a fundamental asymmetry in the *problem*: we are explicitly interested in finding anomalous changes with respect to γ that occur in χ (for example, the appearance in χ of an object not present in γ).

If a pixel \mathbf{x} in image χ *really is* an anomalous change, then we expect \mathbf{x} not only to be different from its corresponding pixel \mathbf{y} in γ , but to differ with all of \mathbf{y} ’s neighbors as well. That is, $\mathcal{A}(\mathbf{x}, \mathbf{y})$ is large and so is $\mathcal{A}(\mathbf{x}, \mathbf{y}_m)$ for all the \mathbf{y}_m that are neighbors of \mathbf{y} . But if \mathbf{x} *merely seems like* an anomalous change, because of misregistration, then although \mathbf{x} and \mathbf{y} may be different, we expect \mathbf{x} to be similar to some of \mathbf{y} ’s neighbors. Thus, even if $\mathcal{A}(\mathbf{x}, \mathbf{y})$ is large, at least one of the neighbors \mathbf{y}_m of \mathbf{y} will have a small value of $\mathcal{A}(\mathbf{x}, \mathbf{y}_m)$.



Fig. 1. RGB image made from one of the two 126-channel HyMap images taken as part of the RIT Target Detection Blind Test [23], [24] experiment.

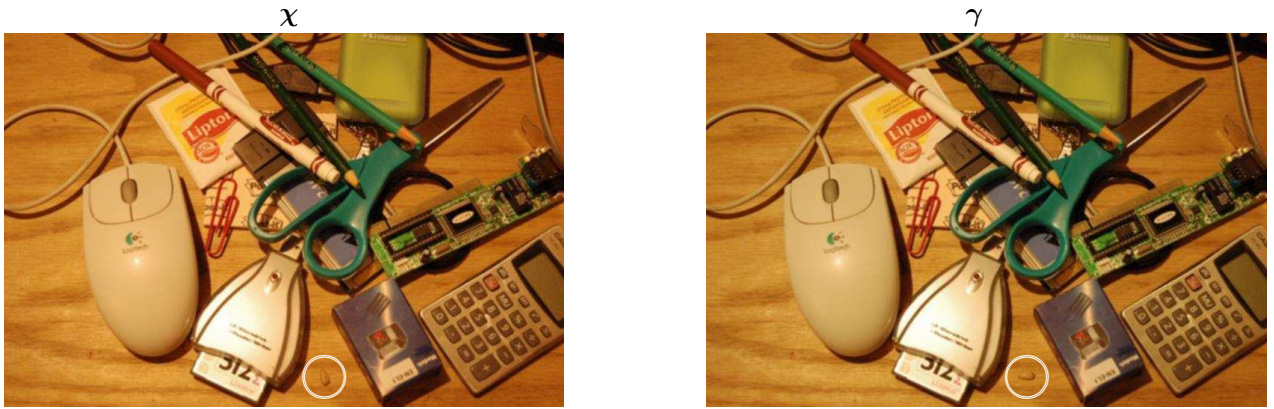


Fig. 2. Two images of desktop clutter were taken under (very) slightly different lighting conditions with a camera on a stable tripod. The actual misregistration between the images is much less than a pixel, and so we use these images as a starting point for simulating the effects of larger misregistration. The anomalous change is small (a rotated sunflower seed indicated by the white circle) and does not have a well-defined direction of asymmetry.

This argument immediately suggests the following misregistration compensation scheme (see Fig. 3): For each pixel in χ , consider a window about the corresponding pixel in γ , and choose the pixel within this window that gives the lowest anomalousness. Formally, we write

$$A_{k,l} = \min_{(m,n) \in W} \mathcal{A}(\chi_{k,l}, \gamma_{k+m,l+n}) \quad (2)$$

for the LCRA anomalousness at pixel (k,l) . Here, \mathcal{A} is the anomalousness function provided by the underlying pixel-wise ACD algorithm, and W is a set of integer pairs defining the optimization window. One simple choice for W is the 3×3 window about the central pixel:

$$W = \{(-1, 1), (-1, 0), (-1, -1), (0, 1), (0, 0), (0, -1), (1, 1), (1, 0), (1, -1)\}. \quad (3)$$

In general, we will consider square and circular optimization windows of radius r , defined respectively as

$$W_{S,r} = \{(n, m) \in \mathbb{Z}^2 \mid |n| \leq r, |m| \leq r\}, \quad (4)$$

$$W_{C,r} = \{(n, m) \in \mathbb{Z}^2 \mid n^2 + m^2 \leq r^2\}. \quad (5)$$

The computational cost of LCRA is a factor of $|W|$ times the cost of the corresponding pixel-wise ACD algorithm, where $|W|$ is the number of pixels in the window. For a square

window of radius r , we have $|W_{S,r}| = (2r + 1)^2$; for a circle, we have $|W_{C,r}| \approx \pi r^2$. (Specifically, $|W_{C,1}| = 5$, $|W_{C,2}| = 13$, and $|W_{C,3}| = 29$.)

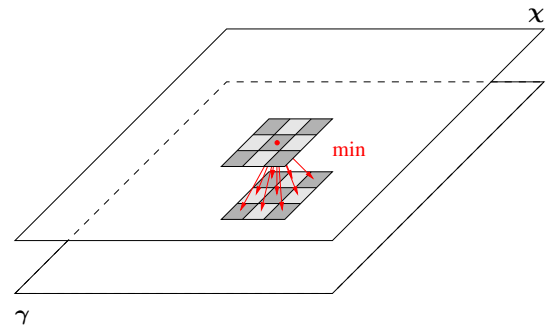


Fig. 3. Construction of a set of pixel pairs from a single pixel in image χ and all pixels within a window about the corresponding pixel in image γ .

The offset $(m, n) \in W$ that minimizes $\mathcal{A}(\chi_{k,l}, \gamma_{k+m,l+n})$ is naturally interpreted as the misregistration at point (k, l) in the image. We do not, however, treat it as an accurate estimator of misregistration *per se*, but interpret it more loosely as a way to compensate for the effect of misregistration on ACD performance.

Fig. 4 shows that LCRA can usefully be applied to different

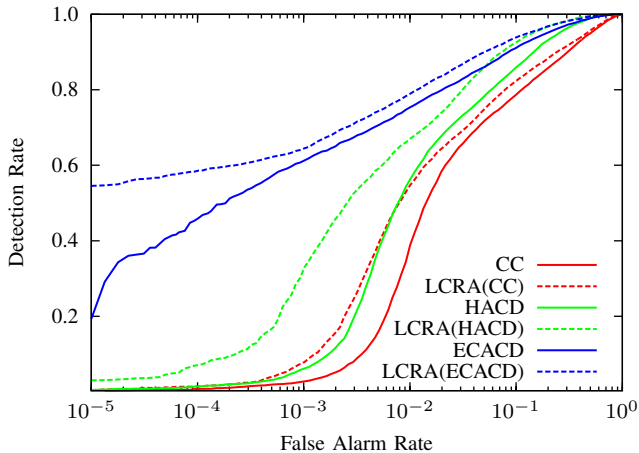


Fig. 4. ROC curves for three different pixel-wise ACD algorithms – chronochrome (CC), hyperbolic anomalous change detection (HACD), and elliptically-contoured anomalous change detection (ECACD) – along with the performance of the LCRA (using $r = 1$) approach, using each of these algorithms as the underlying pixel-wise detector. This is for the RIT blindtest data, using the natural misregistration between the “blind” and “self” images, but employing artificial anomalous changes as described in Appendix subsection B.

pixel-wise anomalous change detectors, and for an image such as the RIT blind test data in which there exists residual misregistration, can improve the performance of those pixel-wise detectors.

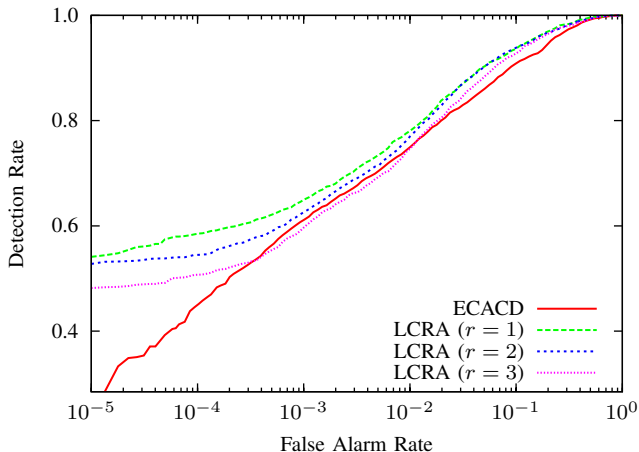


Fig. 5. ROC curves for LCRA using three different window sizes: $r \in \{1, 2, 3\}$. Although improvement in the low false alarm rate regime is observed for all choices of r , the best performance is seen here for $r = 1$. Note that the ECACD and LCRA ($r = 1$) curves are the same as those seen in Fig. 4.

Choosing the parameter r for the size of the window W trades off several factors. In general, r should be large enough to cover the expected magnitude of the misregistration between the two images, but we note that substantial performance improvement is observed when r is smaller than that. Larger r is both more expensive and more prone to missing real targets because of serendipitous matching of the anomaly with something in the surrounding background. And any changes that involve targets moving by fewer than r pixels between

images will fail to be detected. Fig. 5 shows that $r = 1$ gives the best performance.

B. Symmetric Algorithm

The motivation for the asymmetric LCRA algorithm is based on a scenario in which the anomalous change occurs in a particular image of the pair (*i.e.*, χ) and this asymmetry carries through to the resulting algorithm. But if the anomalous change is a pixel in the other image (*i.e.*, γ), it may elude detection. Note, however, that while a true anomalous change will give either a high or low anomalousness depending on the direction of minimization (*i.e.* χ to γ , or *vice versa*), a spurious anomalous change due to misregistration is symmetric, in that similar minimum anomalousness should be obtained independent of the direction of minimization. This observation suggests the Symmetric LCRA algorithm [13]

$$A_{k,l} = \max \left\{ \begin{array}{l} \min_{(m,n) \in W} \mathcal{A}(\chi_{k,l}, \gamma_{k+m,l+n}), \\ \min_{(m,n) \in W} \mathcal{A}(\chi_{k+m,l+n}, \gamma_{k,l}) \end{array} \right\} \quad (6)$$

in which minimization is performed in both directions, and the maximum of the resulting minima is selected as the final result, as illustrated in Fig. 6.

Fig. 7 shows that the performance of the asymmetric LCRA dramatically depends on which direction the LCRA is applied. If the anomalous change is in χ , then Eq. (2) is the correct choice and LCRA(χ, γ) substantially improves the performance of the pixel-wise ACD. But the “backward” LCRA(γ, χ) – which reverses the role of χ and γ in Eq. (2) – is terrible! The symmetric algorithm defined in Eq. (6), SLCRA, does not require that the user correctly guess which image has the anomalous change, and although it does not do quite as well as LCRA(χ, γ), it still outperforms the pixel-wise ECACD algorithm.

The experiment in Fig. 8 compares SLCRA with the pixel-wise ECACD in two situations. In the “aligned” situation, there is essentially no misregistration between the two images, and in that case one pays a small penalty for using SLCRA instead of the pixel-wise algorithm. In the second part of the experiment, the images are misregistered (using the scheme described in Appendix subsection A3). In the misregistered situation, SLCRA outperforms the pixel-wise detector.

Fig. 9 compares the performance of SLCRA with different values of r . Although the simulated misregistration (see Appendix subsection A3) is up to three pixels, we see substantial improvement using only $r = 1$. Larger values of r provide better performance in the low false alarm rate regime, though they are poorer in the regime of higher false alarms. Since larger values of r are also more computationally expensive, we recommend being conservative in the choice of r .

Figs. 10 and 11 further emphasize the distinction between using LCRA for better co-registration and using LCRA for better change detection. That the latter works even though the former does not is one of the main observations of this paper. This is illustrated by taking the $r = 3$ curve in Fig. 9, and looking at the intermediate results in more detail. A map of

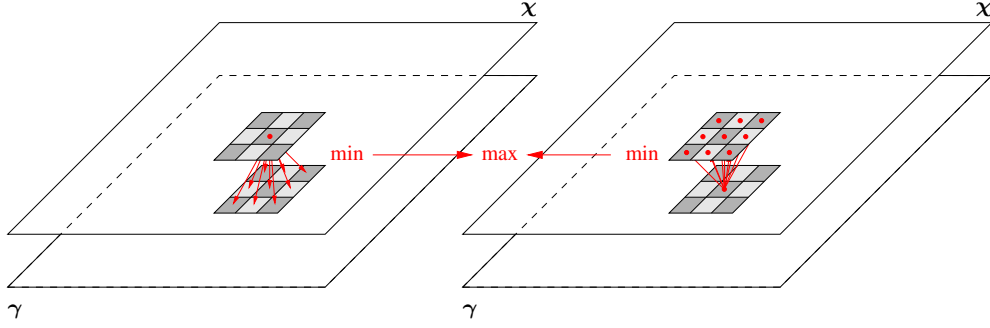


Fig. 6. Symmetric version of the LCRA algorithm illustrated in Fig. 3. At each pixel, the minimum anomalousness pair is selected in each direction (*i.e.* image χ to γ , and image γ to χ) and the final anomalousness value is the maximum of these two minima.

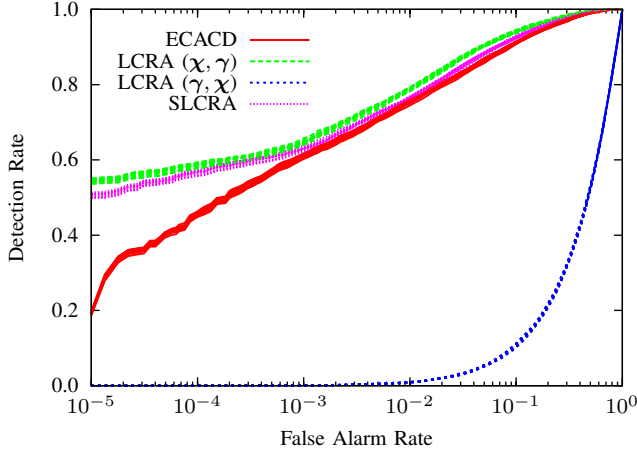


Fig. 7. ROC curves comparing the asymmetric LCRA to the symmetric SLCRA algorithms. We observe that SLCRA improves the performance of the pixel-wise ECACD algorithm, and it does so without having to know which image includes the anomalous changes. In this case, the SLCRA is nearly as good as the “correct” LCRA, shown as $\text{LCRA}(\chi, \gamma)$, but avoids the disastrous performance seen in $\text{LCRA}(\gamma, \chi)$, which is obtained when the user incorrectly guesses which image has the anomalous change. Again, the ECACD and $\text{LCRA}(\chi, \gamma)$ curves are the same as those that appear in Fig. 4 and Fig. 5. Multiple runs are shown (with different randomly chosen simulated anomalous changes) to characterize to statistical variability in the ROC curves.

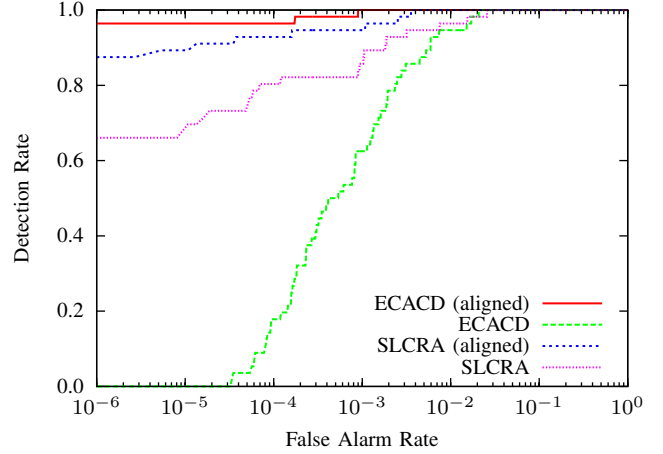


Fig. 8. ROC curves for desktop clutter images with no misregistration (aligned) and with simulated misregistration (smooth random translation with $p = 2$). The comparison of ECACD (aligned) with SLCRA (aligned) shows that when the images truly are aligned, there is a small performance penalty associated with SLCRA. Those two curves also provide an upper bound on the expected performance of ECACD and SLCRA when the images are misregistered. Comparing the unaligned ECACD to the other unaligned curve shows that the underlying pixel-wise ACD is again improved by SLCRA. Both SLCRA results are computed using $r = 1$.

how the vertical component of the simulated misregistration varies across the image is shown in Fig. 10(a). A similar map, of the estimated LCRA offsets, is shown in Fig. 10(b). A visual comparison of those two panels indicates that the adjustment offsets are correlated with the actual misregistrations, but the correlation is not strong (the correlation coefficient is 0.27 and 0.33 in the horizontal and vertical directions respectively). Indeed, the weakness of that correlation is evident from the histograms shown in Fig. 11. We see that error made by the estimated offsets is on average (in the root mean square sense) larger than the offsets themselves. (In other words, if we estimated the offsets with zero instead of the argmin in LCRA, we would have a smaller RMSE).

But in spite of these poor estimates, we still see improvements in the ROC curves for LCRA-based ACD, as we have seen in Fig. 9. What is happening (we speculate) is that larger offset errors are observed in parts of the image where misregistration is less important (for instance, where the image

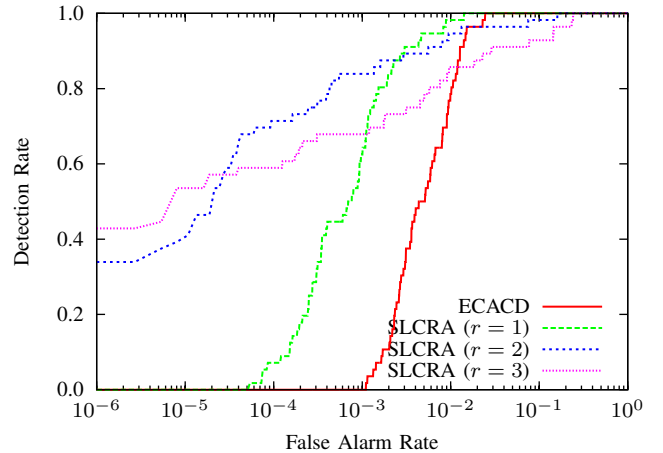


Fig. 9. ROC curves for desktop clutter images with simulated misregistration (smooth random translation with $p = 3$). This comparison shows that $r = 1$ gives some performance gain even when the maximum misregistration is of the order of 3 pixels, but performance is improved at low false alarm rates for values of r closer to the actual maximum misregistration.

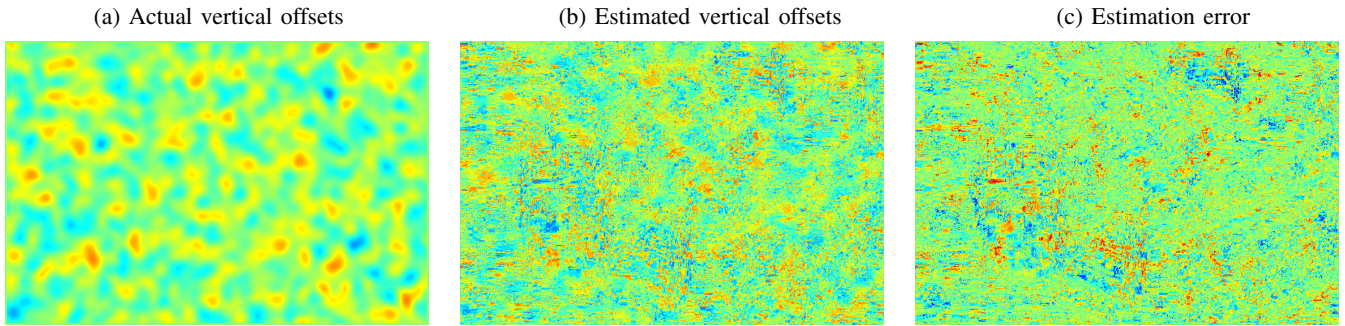


Fig. 10. Maps of actual and estimated vertical offsets applied at each position by the smooth random translation with $p = 3$ (corresponding to the SLCRA $r = 3$ results in Fig. 9). The estimated local adjustments correspond roughly to the actual local translations, but with a lot of noise. The (S)LCRA scheme is not, *per se*, a scheme for making local adjustments to improve the co-registration; it is focused instead on improving the anomalous change detection performance. Note that some spatial clusters of large estimation error correspond to relatively uniform regions in the test images (dark blue and dark red correspond to the largest negative and positive values respectively).

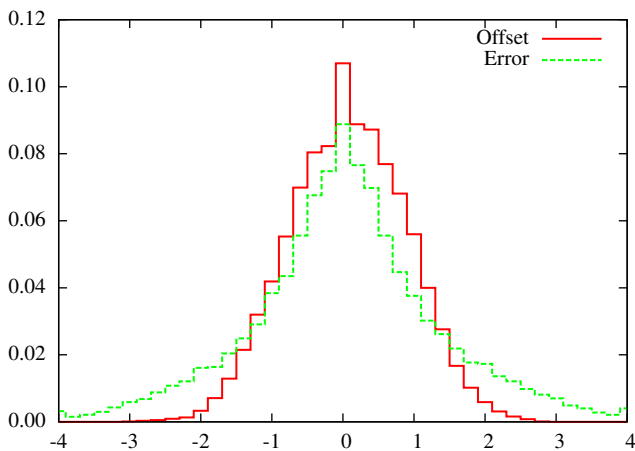


Fig. 11. Histograms of actual vertical offset and error in estimated vertical offset in Fig. 10. (Similar behavior is observed in the horizontal offsets.) The Root Mean Square (RMS) error of the estimated vertical offset is 1.33 pixels, which is larger than the 0.82 pixels RMS of the offsets themselves. This is an additional indication that the estimated offsets are not good estimates of misregistration *per se*.

is smooth, so that $\mathcal{A}(x, y_m)$ doesn't vary much with m). Those errors don't have much effect on anomalous change detection. But in parts of the image where there is higher contrast and/or finer spatial structure, *i.e.*, where misregistration is more of an issue, the local adjustments are better.

C. Sub-pixel Co-Registration Adjustment

We have seen that integer-based (S)LCRA can provide a significant benefit even when the actual misregistration between a pair of images does not consist of integer pixel offsets (which in practice of course it never does). This benefit is diminished when the misregistration is small, but by extending LCRA to sub-pixel adjustments, we can improve the performance of ACD even when the residual misregistration is smaller than a pixel.

1) *Interpolated LCRA*: A simple approach that does not require a substantially modified implementation is to redefine the optimization windows to include fractional pixel offsets. These are applied using interpolation to resample χ and γ

at higher resolution. Specifically, for $s \in \mathbb{Z}^+$, we consider subsampling to $1/s$ of a pixel. The algorithm follows the approach in Eq. (2) (or Eq. (6)), but the windows

$$W_{S,r,s} = \left\{ \left(\frac{n}{s}, \frac{m}{s} \right) \mid (n, m) \in \mathbb{Z}^2, |n| \leq sr, |m| \leq sr \right\}$$

$$W_{C,r,s} = \left\{ \left(\frac{n}{s}, \frac{m}{s} \right) \mid (n, m) \in \mathbb{Z}^2, n^2 + m^2 \leq (sr)^2 \right\}$$

will be larger by a factor of s^2 , and the computational expense will scale similarly.² Fig. 12 shows two cases for which the misregistration is known precisely (because it is simulated using the constant offset approach in Appendix subsection A2); in one case it is very small ($1/3$ pixel), and in the other it is more moderate ($4/3$ pixels), but in both cases it is well away from being an integer pixel offset. For the small offset, we see that the integer-based LCRA is actually worse than the straight pixel-wise detector, but the subpixel LCRA provides a small but significant improvement. For the larger offset, integer-based LCRA gives substantial improvement over the pixel-wise detector, but the subpixel approach provides only marginal improvement over that.

In general, larger s gives better performance, but unlike the case of larger r , there is less concern with overfitting, since anomalousness varies smoothly across a pixel.

2) *Quadratic Fit LCRA*: In the case that the pixel-wise detector is quadratic – *e.g.*, for CC and HACD, but not for ECACD – linear interpolation of subpixel image values leads to anomalousness measures that are quadratic functions of subpixel position. This suggests a way to take what amounts to the $s \rightarrow \infty$ limit of the subpixel interpolation scheme described above.

We effectively construct a continuous version of the LCRA function $\mathcal{A}(\chi_{k,l}, \gamma_{k+m+\alpha, l+n+\beta})$ with fractions $-0.5 \leq \alpha \leq 0.5$ and $-0.5 \leq \beta \leq 0.5$ real-valued instead of integer-valued.

²The computation requires $O(r^2 s^2)$ pixel-wise anomalousness evaluations per pixel. A more economical variant only considers the sub-pixel locations in the vicinity of the offset (n, m) with minimum anomalousness; this requires only $O(r^2) + O(s^2)$ anomalousness evaluations per pixel, but it also requires that the computation be done in two steps that are not as readily vectorizable.

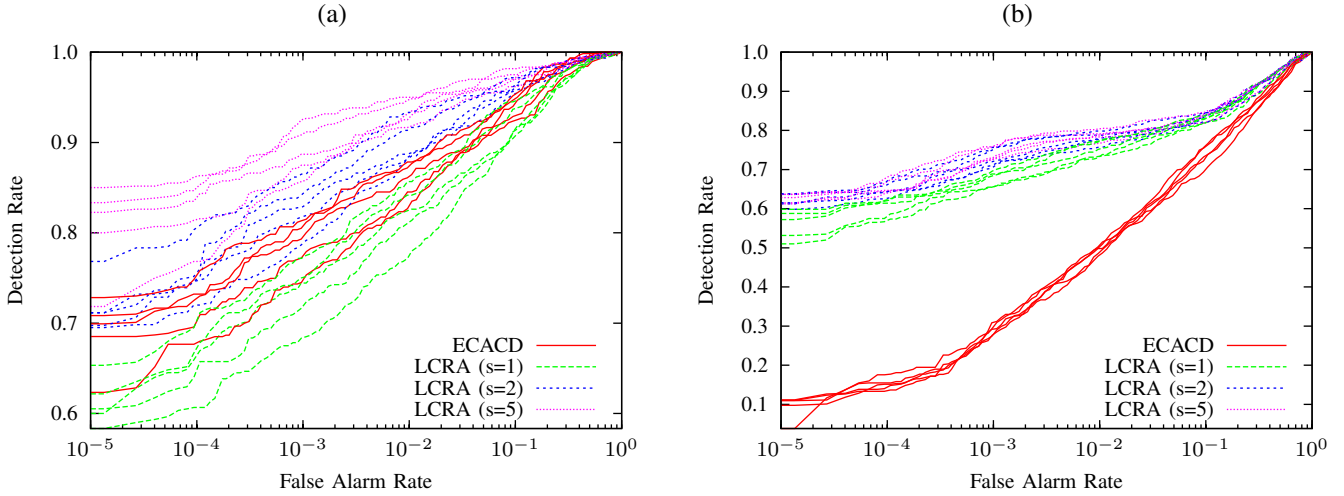


Fig. 12. Two images derived from a single image in the RIT blind test dataset have a simulated misregistration (using the process described in Appendix subsection A2) that is fixed over the full image. In (a) the shift is $1/3$ pixel, and in (b) the shift is $4/3$ pixels. The first three principal components of the two images are used. Anomalous changes are simulated in one of the images, using the scheme described in Appendix subsection B3. Shown here are ROC curves that result from the pixel-wise ECACD detector, and from integer-pixel ($s = 1$) and subpixel ($s > 1$) LCRA. The relative performance of ECACD and LCRA ($s = 1$) in (a) shows that integer LCRA does not give any improvement (and can in fact degrade performance) when the offset is significantly smaller than one pixel. The relative performance of these two methods in (b), on the other hand, shows that when the offset is on the order of a pixel, the integer-based LCRA is nearly as good as the sub-pixel LCRA. Multiple runs are shown (with different randomly chosen simulated anomalous changes) to characterize to statistical variability in the ROC curves.

The goal is to estimate a value

$$A_{k,l} = \min_{\substack{(m,n) \in W \\ -0.5 \leq \alpha \leq 0.5 \\ -0.5 \leq \beta \leq 0.5}} \mathcal{A}(\chi_{k,l}, \gamma_{k+m+\alpha, l+n+\beta}) \quad (7)$$

and the way this is achieved in practice is in two steps. First, the integer values (m, n) are identified that minimize the whole-pixel LCRA anomalousness in Eq. (2). Then, linear interpolation is employed to estimate the half-integer values

$$\gamma_{k+m\pm 0.5, l+n} = (\gamma_{k+m, l+n} + \gamma_{k+m\pm 1, l+n})/2, \quad (8)$$

$$\gamma_{k+m, l+n\pm 0.5} = (\gamma_{k+m, l+n} + \gamma_{k+m, l+n\pm 1})/2 \quad (9)$$

and from these, we compute $\mathcal{A}(\chi_{k,l}, \gamma_{k+m\pm 0.5, l+n})$ and $\mathcal{A}(\chi_{k,l}, \gamma_{k+m, l+n\pm 0.5})$. Combined with the already known $\mathcal{A}(\chi_{k,l}, \gamma_{k+m, l+n})$, and the previously computed $\mathcal{A}(\chi_{k,l}, \gamma_{k+m\pm 1, l+n})$ and $\mathcal{A}(\chi_{k,l}, \gamma_{k+m, l+n\pm 1})$, we can fit quadratic functions to each of the four quadrants around the point (m, n) . Specifically, we can write

$$\widehat{\mathcal{A}}(\chi_{k,l}, \gamma_{k+m+\alpha, l+n+\beta}) = a + b\alpha + c\beta + d\alpha^2 + e\beta^2 \quad (10)$$

where the five parameters a, b, c, d, e can be fit based on the five points³ $(\alpha, \beta) \in \{(0, 0), (0, 0.5), (0, 1), (0.5, 0), (1, 0)\}$. We compute the minimum of the quadratic in Eq. (10), constrained to the bound $[0, 0.5] \times [0, 0.5]$. This is done for each of the quadrants, and the minimum of those four minima is used for $A_{k,l}$ in Eq. (7). As an implementation issue, because Eq. (10) is separable in α and β , what would be a set of four two-dimensional quadratic-programming problems can be

³This quadratic function can be fit with only five points because we have constrained the cross-term (proportional to the product $\alpha\beta$) to be zero. We also considered the more general case of a nonzero cross-product, which uses $(\alpha, \beta) = (0.5, 0.5)$ as a sixth point, but the performance was not substantially altered and the extra computation is considerable.

decomposed into four one-dimensional quadratic problems.

We refer to this method as Quadratic Fit LCRA (QFLCRA). Following the discussion in Section IV-B, we can symmetrize the algorithm by taking the maximum of the two QFLCRA anomalous measures, and the result is the Quadratic Fit SLCRA (QFSLCRA).

We remark that this linear interpolation has a smoothing effect, and one result of that can be a bias toward half-integer offsets, as was pointed out by Inglada *et al.* [25]. But because we are more interested in the “min” than the “argmin” we are less concerned about these artifacts than we would be if we were trying to infer explicit sub-pixel misregistration offsets.

Fig. 13 shows that this quadratic fit LCRA can substantially improve the performance of the quadratic HACD detector, and even outperforms the subpixel LCRA detector with $s = 2$, which employs a comparable number of anomalousness evaluations.⁴

D. Covariance Re-Estimation

For LCRA, we can write $m(k, l)$ and $n(k, l)$ as the offsets that minimize Eq. (2) as a function of pixel position (k, l) . If we make a new image, γ' , with $\gamma'_{k,l} = \gamma_{k+m(k,l), l+n(k,l)}$, then we can treat γ' as a “corrected” image, and another way to express the LCRA algorithm is to write $A_{k,l} = \mathcal{A}(\chi_{k,l}, \gamma'_{k,l})$ where \mathcal{A} is the underlying pixel-wise change detector. In practice, we do not actually use this corrected image, partly for reasons of computational efficiency, but also because we don’t expect the offsets $m(k, l)$ and $n(k, l)$ to very accurately reflect the actual misregistration.

In computing $\mathcal{A}(\chi_{k,l}, \gamma'_{k,l})$, we use a pixel-wise detector \mathcal{A} that is based on covariance matrices that are computed from

⁴Actually QFLCRA uses four fewer evaluations, corresponding to the offsets $(\alpha, \beta) = (\pm 0.5, \pm 0.5)$

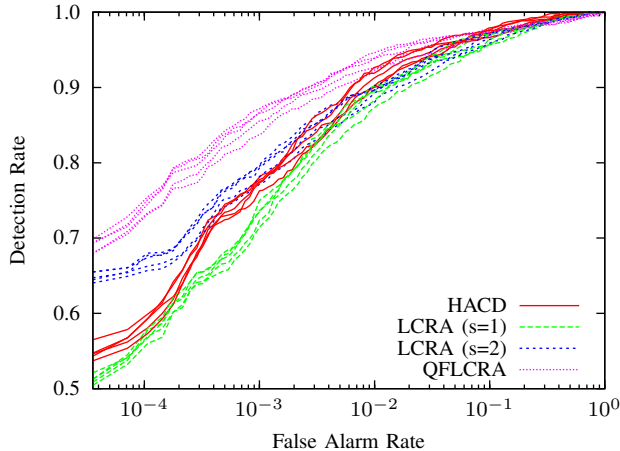


Fig. 13. We use the same blind test dataset image we used in Fig. 12, but with an offset of 1/8 pixel. We observe that QFLCRA substantially improves the performance of the HACD detector even for this very small misregistration. The QFLCRA outperforms the subpixel LCRA with $s = 2$, even though QFLCRA involves fewer evaluations of the anomalousness function. There were a total of 25 runs with different randomly simulated anomalous changes; each curve is the average of five such runs.

the original, misregistered, images. The idea of covariance re-estimation is to recompute the covariances using χ and γ' rather than χ and γ (it is the cross-covariance, in particular, that will change). In principle, the process is iterative: with new covariances, we can obtain new offsets $m(k, l)$ and $n(k, l)$, and thereby a new γ' . Note that only the covariances are modified in each iteration of this algorithm, and each iteration is applied to the original images χ and γ rather than χ and updated γ' . As might be expected based on the poor accuracy of the estimated offsets, using the updated images was found experimentally to give significantly worse detection performance. But in spite of the inaccuracy of the individual estimated offsets, we nonetheless expect the covariance estimate to improve. Partly this is because the covariance matrix is an average over all pixels, and partly because the covariance depends on the values of γ' and not the values of the offsets m and n .

In practice, we find diminishing returns after the first covariance re-estimation, and therefore we recommend only one or two iterations. The computation time for each iteration is essentially the same as the computation time for regular LCRA.

For SLCRA, the idea is basically the same. In this case, we get offsets for both images – $m_\chi(k, l)$ and $n_\chi(k, l)$, as well as $m_\gamma(k, l)$ and $n_\gamma(k, l)$ – though for any given position (k, l) only one of the images will have a nonzero offset. With corrected images χ' and γ' , we can, as in the case for LCRA, produce re-estimated covariances, and use these to improve the SLCRA performance.

Fig. 14 compares the performance of SLCRA and SLCRA with a single step of covariance re-estimation for the RIT data, and indicates a small improvement. Fig. 15(a) provides a similar comparison for the desktop clutter images. In this case the detection rate variance across multiple trials is high (due to the small number of anomalous pixels in this example), but

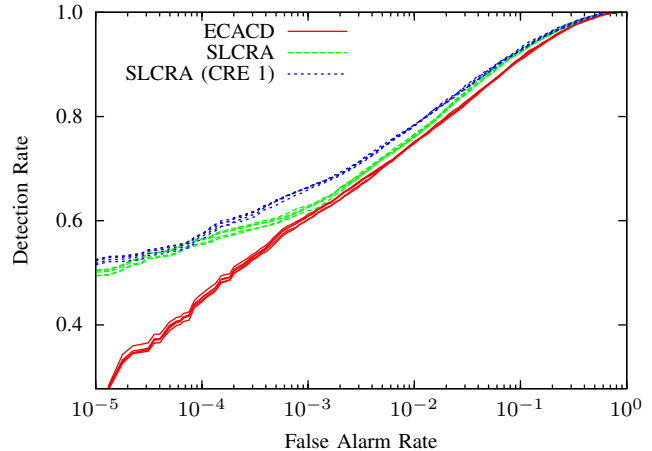


Fig. 14. ROC curves comparing ECACD, SLCRA, and SLCRA with covariance re-estimation for RIT data. SLCRA is computed using $r = 1$.

Fig. 15(b) shows that the improvement is positive in all five runs.

V. CONCLUSIONS

In general, the more accurately two images can be co-registered, the more effective change detection will be. Some residual misregistration is inevitable, however, and we have shown that local co-registration adjustments (LCRA) – even if they are “too local” to improve the co-registration itself – can lead to improved performance from anomalous change detectors. The integer-pixel adjustments provided by LCRA and its symmetric variant SLCRA achieve this improvement even though the actual misregistrations are not limited to integer pixel displacements. We have also shown that some further improvement can be obtained with sub-pixel adjustments, and have shown how local quadratic fits to the anomalousness function enables an efficient implementation of those sub-pixel adjustments. We have also seen incremental improvement using an iterative covariance re-estimation scheme. In making these performance comparisons, we have relied upon a hybrid framework of real and simulated data.

APPENDIX

SIMULATION FRAMEWORK FOR ACD

In the framework proposed in [22], there are two distinct simulations, either of which can be performed independently of the other. One can simulate *pervasive differences* by applying some operator (brightening, smoothing, misregistering, etc.) to every pixel in a *base* image \mathcal{B} in order to produce a *normal* change image γ . In the second kind of simulation, *anomalous changes* are generated by applying some operator to a single pixel in the *base* image, and thereby obtaining an *anomalous* change image χ . This is illustrated in Fig. 16.

This simulated image pair – the normal change γ and the anomalous change χ – can be used as input to an ACD algorithm. The challenge is to find the pixel where the anomaly was introduced. Because this is a simulation, it can be repeated many times, and a ROC curve can be produced that shows the trade-off of false alarm rate and detection rate.

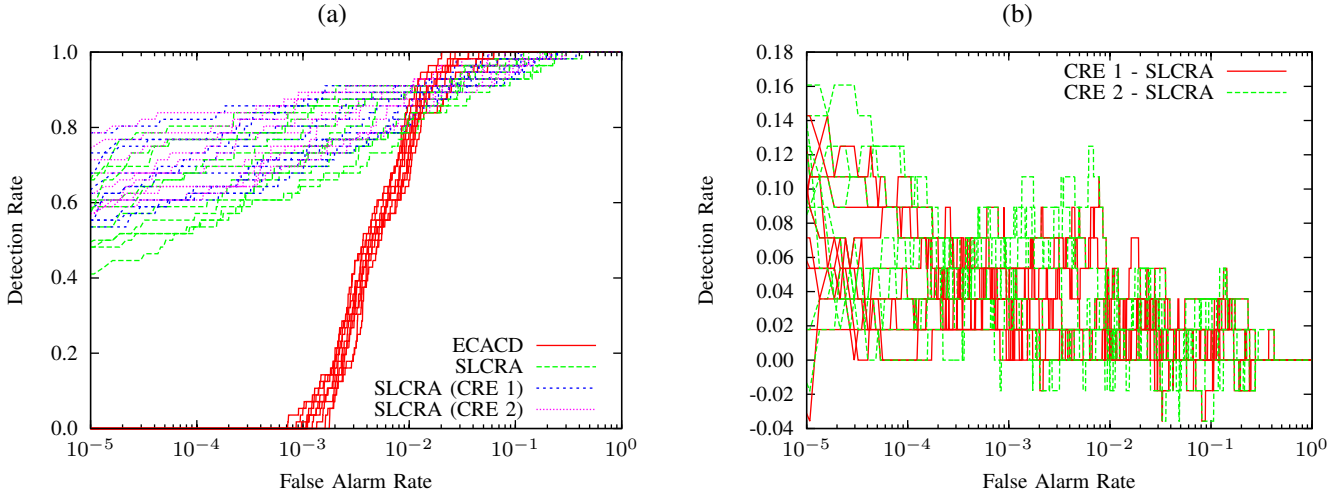


Fig. 15. ROC curves comparing ECACD, SLCRA, and SCLRA with covariance re-estimation for the desktop clutter images with synthetic radius 3 smooth misregistration. (Each of the 10 trials includes generation of a new misregistration.) Since it is not possible to identify the corresponding curves for each method in a single trial in (a), differences for each trial are plotted in (b), which shows a small improvement (note that most differences are positive) for one step of covariance re-estimation, and less convincingly suggests a further improvement for an additional step.

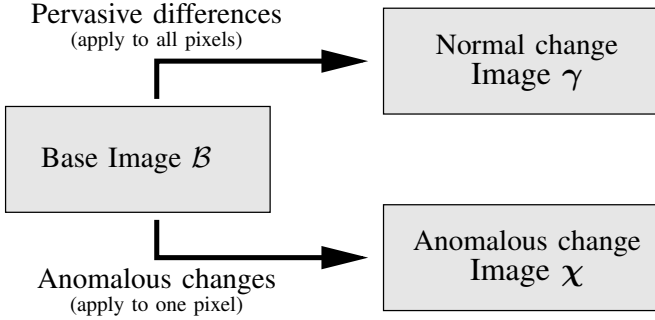


Fig. 16. Construction of normal and anomalous images in simulation framework. The *normal* change image χ is generated from the base image by applying some operator over the whole image. The *anomalous* change image γ is generated from the base image by altering one of the pixels. The resulting images, χ and γ , are input to the ACD algorithm being evaluated. A good algorithm will find the one altered pixel.

A. Simulating pervasive differences: misregistration

Since our concern in this paper is with the effects of misregistration, it is this pervasive difference that we emphasize in our simulations. We employ three levels of sophistication in our misregistration simulations.

In all cases, we trim an appropriate number of pixels from the edges of the image to avoid the artifacts that the various shifting procedures introduce there.

1) *Simple translation*: The simplest way to simulate misregistration is to translate one of the images by a fixed integer number of pixels in the horizontal and/or the vertical directions. That is, $\gamma_{k,l} = \mathcal{B}_{k+m,l+n}$ for some choice of (m,n) that is fixed over the whole image.

2) *Downsampled translation*: To simulate translation by a rational (*i.e.* fractional) number of pixels, one can begin by misregistering by an integer number n of pixels, smoothing the images with a smoothing kernel of size s pixels, and then downsampling the images by an integer factor of s . The resulting misregistration is n/s . The resulting images have

fewer pixels (by a factor of s) than the original images, and therefore less statistical power, but the simulation provides a direct model for sub-pixel misregistration.

3) *Smoothed random translation*: In both of the previous schemes, the translation is the same over the whole scene. In actual use, however, we can generally expect that any overall translation will have been taken out, and that we are left with a *residual* misregistration that varies across the image. That is, the offsets m and n are themselves functions of the pixel location (k,l) . This often arises, for instance, when the elevation varies across a scene; rotation and perspective changes can also be modeled as local translational motions that vary across the image.

We simulate this situation with a two-step procedure. The first step applies a different randomly-chosen misregistration to each pixel; that is: for each position (k,l) , integers m and n are randomly chosen from a uniform distribution over a fixed range: $-p \leq m \leq p$ and $-p \leq n \leq p$. The second step is a low-pass filtering of the offsets as a function of (k,l) ; *e.g.*, see Fig. 10(a). In our experiments, we used a Gaussian smoothing kernel with a standard deviation of ten pixels. For the non-integer (*i.e.*, sub-pixel) offsets produced by this smoothing, we employ bilinear interpolation to generate the *normal* change image γ .

B. Simulating anomalous changes

Whereas the pervasive differences are simulated across the full image, the anomalous changes are simulated at single pixels (or small local patches of pixels [26]). A variety of choices are possible for these anomalous changes [22], but in this paper we simulate anomalous changes with randomly sampled pixels from somewhere else in the image.

1) *Conceptual scheme*: The idea is to produce pixels for which the change is unusual, without employing pixels that are individually unusual. Put another way, the change $\gamma_{k,l} \rightarrow \chi_{k,l}$ is unusual compared to the other changes in $\gamma \rightarrow \chi$, but $\chi_{k,l}$ is not unusual compared to the rest of the pixels in χ .

The way we do this by choosing two positions at random in the image: (k, l) and (k', l') . We let $\chi = \mathcal{B}$ for all pixels except the anomalous change; for that pixel: $\chi_{k,l} = \mathcal{B}_{k',l'}$.

2) *Purely spectral scheme*: When purely spectral ACD algorithms are employed, one can take a shortcut and produce an *anomalous* change image in which every pixel constitutes an anomalous change – in this case, the *anomalous* change image is obtained from the *base* image simply by scrambling its pixels. The anomalous change detection algorithm is “trained” (which is to say that the covariance matrices are computed) using the *base-normal* pair. Applying the algorithm, at a given threshold, to the *base-normal* pair provides an estimate of the false alarm rate. Applying the same algorithm, with the same threshold, to the *normal-anomalous* pair, one can estimate the detection rate. By varying the threshold, a receiver operator characteristic (ROC) can be generated: this provides detection rate as a function of false alarm rate. This is equivalent to running the conceptual scheme once for every pixel in the image, but a lot cheaper.

3) *Spatio-Spectral scheme*: When spatial pre-processing is built into the ACD algorithm, as it is in the LCRA algorithms, the simulation framework requires additional complexity. It is possible to avoid this complexity for the specific case where the algorithm is asymmetrical, *and* the anomalies are known to be on a particular image; that is the case we investigated in our earlier work [12]. In order to address the more general case, we employed a more appropriate simulation framework.

This framework incorporates the ideas used in earlier work on spatial processing for anomalous change detection [26]. We begin by introducing a *target* mask, which is a binary image with spatially isolated 1's surrounded by 0's. Anomalous changes are introduced only at locations in the image where the *target* mask is 1. That is, the *anomalous* change image differs from the *normal* image only at those locations where the *target* mask is 1. Those anomalous pixels are chosen at random from the rest of the image.

As before, we apply the algorithm at a given threshold to the *base-normal* pair to provide an estimate of the false alarm rate. We apply the same algorithm at the same threshold to the *base-anomalous* pair, but only consider the pixels where the *target* mask is 1. This provides an estimate of the detection rate.

Because the anomalous changes are spatially isolated, we can perform spatial processing without having the individual anomalies interfere with each other. It is of course required that the distance between the anomalies be larger than the diameter of the spatial processing window.

REFERENCES

- [1] M. T. Eismann, J. Meola, A. D. Stocker, S. G. Beaven, and A. P. Schaum, “Airborne hyperspectral detection of small changes,” *Applied Optics*, vol. 47, pp. F27–F45, 2008.
- [2] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, “Multi-modality image registration by maximization of mutual information,” in *Proc. Workshop on Mathematical Methods in Biomedical Image Analysis*, June 1996, pp. 14–22.
- [3] W. M. Wells III, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, “Multi-modal volume registration by maximization of mutual information,” *Medical Image Analysis*, vol. 1, no. 1, pp. 35–51, 1996.
- [4] H.-M. Chen, P. K. Varshney, and M. K. Arora, “Performance of mutual information similarity measure for registration of multitemporal remote sensing images,” *IEEE Trans. on Geoscience and Remote Sensing*, vol. 41, no. 11, pp. 2445–2454, November 2003.
- [5] J. Inglada and A. Giros, “On the possibility of automatic multisensor image registration,” *IEEE Trans. Geoscience and Remote Sensing*, vol. 42, pp. 2104–2120, 2004.
- [6] A. Wong and D. A. Clausi, “ARRSI: Automatic registration of remote-sensing images,” *IEEE Trans. Geoscience and Remote Sensing*, vol. 45, no. 5, pp. 1483–1493, 2007.
- [7] J. P. Kern and M. S. Pattichis, “Robust multispectral image registration using mutual information models,” *IEEE Trans. on Geoscience and Remote Sensing*, vol. 45, no. 5, pp. 1494–1505, 2007.
- [8] H. Goncalves, L. Coret-Real, and J. A. Goncalves, “Automatic image registration through image segmentation and SIFT,” *IEEE Trans. on Geoscience and Remote Sensing*, vol. 49, no. 7, pp. 2589–2600, 2011.
- [9] G. Mercier and J. Inglada, “Change detection with misregistration errors,” in *Proc. IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 3, 2008, pp. III–154–III–157.
- [10] J. Meola and M. T. Eismann, “Image misregistration effects on hyperspectral change detection,” *Proc. SPIE Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XIV*, vol. 6966, p. 69660Y, 2008.
- [11] J. Theiler, “Sensitivity of anomalous change detection to small misregistration errors,” *Proc. SPIE Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XIV*, vol. 6966, p. 69660X, 2008.
- [12] B. Wohlberg and J. Theiler, “Improved change detection with local co-registration adjustments,” in *Proc. IEEE Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing*, Grenoble, France, Aug. 2009.
- [13] —, “Symmetrized local co-registration optimization for anomalous change detection,” in *Proc. SPIE Computational Imaging VIII*, vol. 7533, San Jose, California, USA, Jan. 2010, p. 753307.
- [14] A. Schaum and A. Stocker, “Long-interval chronochrome target detection,” in *Proc. International Symposium on Spectral Sensing Research*, 1998.
- [15] C. Clifton, “Change detection in overhead imagery using neural networks,” *Applied Intelligence*, vol. 18, pp. 215–234, 2003.
- [16] A. Schaum and A. Stocker, “Linear chromodynamics models for hyperspectral target detection,” *Proc. IEEE Aerospace Conference*, pp. 1879–1885, 2003.
- [17] A. A. Nielsen, K. Conradsen, and J. J. Simpson, “Multivariate alteration detection (MAD) and MAF post-processing in multispectral bi-temporal image data: new approaches to change detection studies,” *Remote Sensing of the Environment*, vol. 64, pp. 1–19, 1998.
- [18] J. Theiler and S. Perkins, “Proposed framework for anomalous change detection,” *ICML Workshop on Machine Learning Algorithms for Surveillance and Event Detection*, pp. 7–14, 2006.
- [19] J. Theiler, “Subpixel anomalous change detection in remote sensing imagery,” *Proc. IEEE Southwest Symposium on Image Analysis and Interpretation*, pp. 165–168, 2008.
- [20] J. Theiler, C. Scovel, B. Wohlberg, and B. R. Foy, “Elliptically contoured distributions for anomalous change detection in hyperspectral imagery,” *IEEE Geoscience and Remote Sensing Letters*, vol. 7, no. 2, pp. 271–275, Apr. 2010.
- [21] L. Gueguen, P. Soille, and M. Pesaresi, “Change detection based on information measure,” *IEEE Trans. on Geoscience and Remote Sensing*, vol. 49, no. 11, pp. 4503–4515, 2011.
- [22] J. Theiler, “Quantitative comparison of quadratic covariance-based anomalous change detectors,” *Applied Optics*, vol. 47, pp. F12–F26, 2008.
- [23] “Target detection blind test,” Center for Imaging Science, RIT, <http://dirsapps.cis.rit.edu/blindtest/>.
- [24] D. Snyder, J. Kerekes, I. Fairweather, R. Crabtree, J. Shive, and S. Hager, “Development of a web-based application to evaluate target finding algorithms,” in *Proc. IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 2, 7-11 2008, pp. II–915–II–918.
- [25] J. Inglada, V. Muron, D. Pichard, and T. Feuvrier, “Analysis of artifacts in subpixel remote sensing image registration,” *IEEE Trans. Geoscience and Remote Sensing*, vol. 45, pp. 254–264, 2007.
- [26] J. Theiler, N. R. Harvey, R. Porter, and B. Wohlberg, “Simulation framework for spatio-spectral anomalous change detection,” *Proc. SPIE Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XV*, vol. 7334, 2009.