

# Detecting change in images with parallax

Mark J. Carlotto<sup>1</sup>

General Dynamics, Advanced Information Systems

## Abstract

The ability to detect significant change in images is reduced when there is parallax. Fixed objects above or below the ground plane displace from their true position by an amount that depends on their height and the look angle of the sensor, reducing background cancellation and limiting the effectiveness of change detection. A method for reducing the effect of parallax shifts using dynamic time warping (DTW) to align images along the epipolar direction is described. A performance model is developed for predicting the processing gain of change detection (CD) over single image object detection as a function of the parallax. The model is used to predict CD performance with, and without DTW. A 10x improvement in processing gain is demonstrated on optical imagery, which results in a significant reduction in the number of false alarms.

**Keywords:** Change detection, dynamic time warping, performance modeling

## 1. INTRODUCTION

Change detection generally works best for images acquired under similar conditions – at around the same time of day so shading and shadowing are comparable, and from roughly the same location in space to minimize parallax differences. Based on a recent survey of change detection techniques [1] there has been relatively little work on the problem of detecting changes taken under different illumination/sensor geometries. One exception is the use of site modeling [2], which involves creating a detailed 3-D geometric representation of the imaged scene, and model-supported change detection, which uses the site model to orthorectify images acquired from different sensor angles. Although this approach works well when the changes are large compared to the size of the false alarms caused by parallax, a site model may not be detailed or accurate enough to allow smaller changes to be detected. The reason is that residual parallax errors caused by inaccuracies in the site model can produce false changes that obscure the true changes in the scene.

This paper proposes the use of dynamic time warping (DTW) for detecting small changes in images. Originally developed for stereo image matching [3], DTW is used to precisely align images along the epipolar direction for change detection. Section 2 begins by deriving a model for change detection in the presence of parallax. The model predicts the processing gain of image subtraction as a function of the residual (uncorrected parallax). Section 3 reviews DTW and discusses its application to change detection. An image example is presented in Section 4 that shows a 10 db increase in processing gain (background cancellation) using DTW, which is in agreement with the performance model.

## 2. A MODEL FOR CHANGE DETECTION WITH PARALLAX

Parallax causes local image displacements that affects background cancellation and limits the effectiveness of change detection algorithms. The degree of background cancellation achievable depends on the correlation structure of the image, which depends on the surface and on the illumination. Assume two images taken under similar lighting and

---

<sup>1</sup> mark.carlotto@gd-ais.com

background conditions, but from different positions in space. Video images satisfy this requirement as well as stereo images acquired for topographic mapping. Let  $r(i, j)$  be the reference image, and  $n(i, j)$  be a new image taken at a later time. Assume that an object of constant brightness has appeared in the new image. Let  $\mu^2$  be the signal power and  $\sigma_n^2$  the noise power in the new image. The ability to detect the object in the new image depends on the input signal to noise ratio,  $\gamma_n = \mu^2 / \sigma_n^2$ . Let  $e(i, j) = r(i, j) - n(i, j)$  be the difference image with noise power  $\sigma_e^2$ . The output SNR (i.e., after change detection) is  $\gamma_e = \mu^2 / \sigma_e^2$ , and the processing gain is  $\gamma_e / \gamma_r$ . We now develop a model to predict the processing gain as a function of the imaging geometry and the correlation structure of the surface.

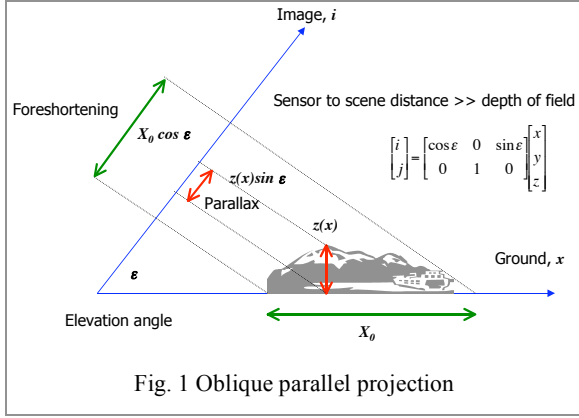


Fig. 1 Oblique parallel projection

Images must be scaled to correct for foreshortening, and rotated before they can be compared. This formula corrects for foreshortening

$$\begin{aligned} r\left(\frac{[x \cos \epsilon_r + z \sin \epsilon_r]}{\cos \epsilon_r}\right) &= r(x + z \tan \epsilon_r) \\ n\left(\frac{[x \cos \epsilon_n + z \sin \epsilon_n]}{\cos \epsilon_n}\right) &= n(x + z \tan \epsilon_n) \end{aligned} \quad (2)$$

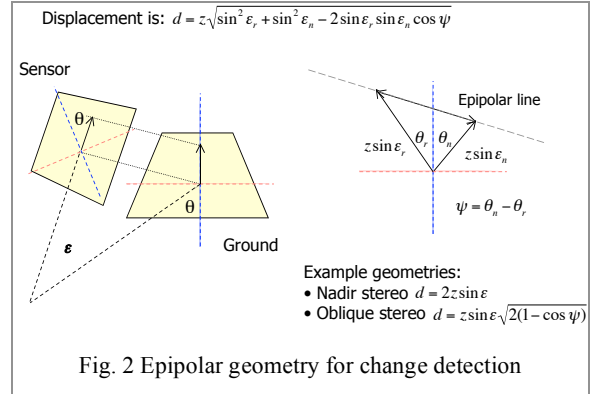


Fig. 2 Epipolar geometry for change detection

Depending on the image azimuth angles  $\theta_r$  and  $\theta_n$ , the images must also be rotated (Fig. 2). The resultant displacement term due to parallax is

$$d(z) = z \sqrt{\tan^2 \epsilon_r + \tan^2 \epsilon_n - 2 \tan \epsilon_r \tan \epsilon_n \cos \psi} \quad (3)$$

where  $\psi$  is the azimuth angle difference.

**Surface Model** - Assume the surface can be modeled as a 1st-order discrete Gauss-Markov process

$$r(i+1) = ar(i) + \eta \sqrt{1-a^2} \quad (4)$$

where  $a$  is the horizontal pixel correlation (Markov coefficient), and  $\eta$  are identical, independent Gaussian random variables with variance  $\sigma^2$ . The correlation between the new and reference images varies as a function of pixel displacement; i.e.,  $A = a^{|d(z)|}$ . In this paper it is assumed that the new and reference image backgrounds are acquired

under the same illumination so that the only difference is due to parallax. The new image as a function of the reference image is

$$n = ra^{|d|} + \eta\sqrt{1 - a^{2|d|}} \quad (5)$$

where the indices have been dropped for brevity. The difference image is

$$r - n = r(1 - a^{|d|}) - \eta\sqrt{1 - a^{2|d|}} \quad (6)$$

Assuming texture and displacement are independent random variables, the difference image variance is

$$E[(r - n)^2 | d] = 2\sigma^2(1 - a^{|d|}) \quad (7)$$

Since displacement is also a random variable

$$E[(r - n)^2] = 2\sigma^2(1 - E[a^{|d(z)|}]) \quad (8)$$

The processing gain is thus

$$\gamma = \frac{E[r^2]}{E[(r - n)^2]} = \frac{1}{2(1 - E[a^{|d(z)|}])} \quad (9)$$

As the Markov coefficient approaches unity, the processing gain approaches infinity (perfect background cancellation). At the other extreme when the backgrounds are uncorrelated, the gain is reduced by two.

**Processing Gain** – The processing gain can be estimated numerically from the histogram or statistics of the elevation surface. The key term in (9) is

$$E[a^{|d(z)|}] = \sum_z a^{|d(z)|} p(z) \Delta z \quad (10)$$

where  $p(z)$  is the elevation histogram. Fig. 3 shows predicted processing gains as a function of various parameters.

### 3. IMAGE ALIGNMENT USING DYNAMIC TIME WARPING

Accurately aligning images prior to change detection reduces residual parallax and increases processing gain. DTW was originally developed to align 1-d signals for speech recognition. Ohta and Kanade [3] describe a dynamic programming solution to the problem of searching for the best (minimum cost) correspondence between stereo images along scan-lines (epipolar lines). Let  $r(s)$  and  $n(t)$  be 1-D brightness functions along an epipolar line in the reference and new images, respectively. Define  $c(s, t)$  to be the cost in mapping a point in the new image to a point in the reference image. For  $c(s, t) = 1 + |r(s) - n(t)|$  points that have the same brightness have minimum cost (which must be non-zero). Each mapping corresponds to a different path through the cost array  $c(s, t)$ . The “best” mapping corresponds to the path with minimum cost.

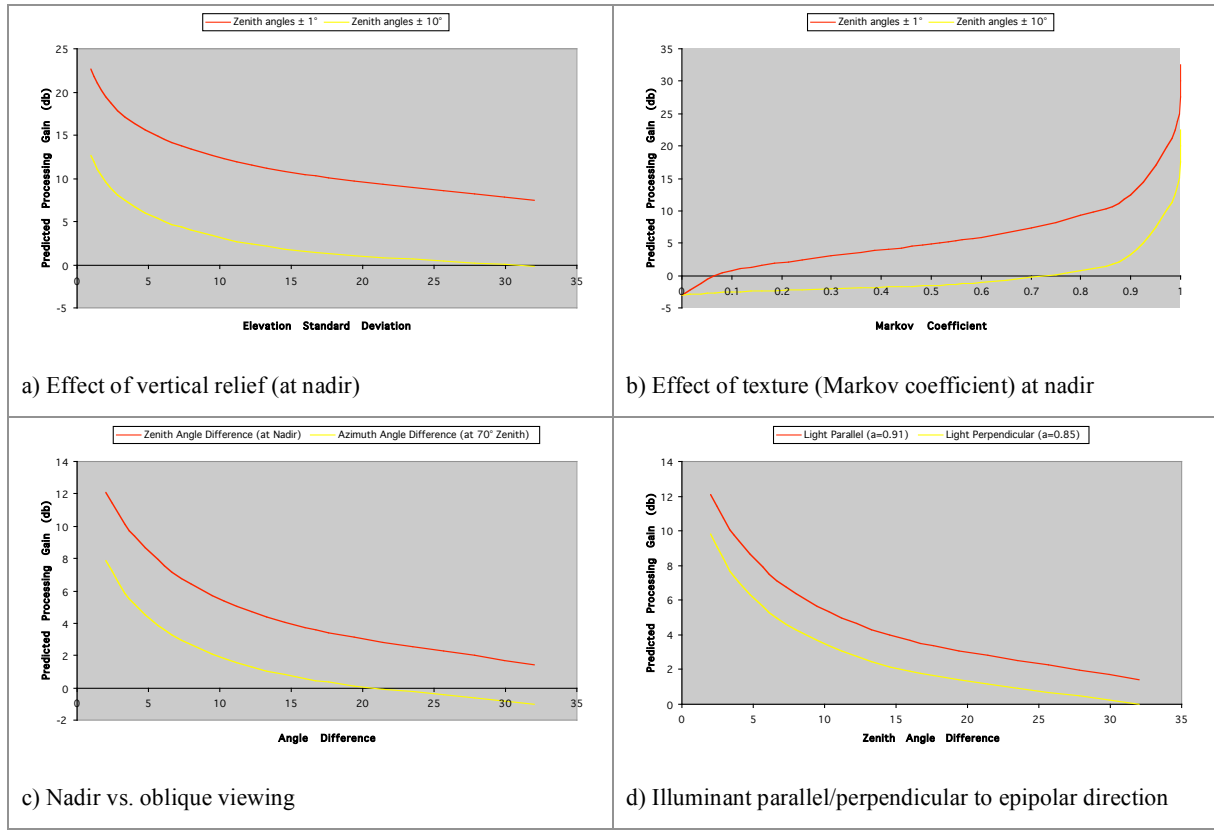


Fig. 3 Predicted change detection processing gains

A graphical depiction of the DTW algorithm is shown in Fig. 4. Assume the reference image corresponds to the right image of a stereo pair, and the new image is the left image, where a) is the part of an object (e.g., side of a building) visible in the new image and obscured in the reference, b) is the part of an object visible in the reference image but obscured in the new, and c) is a change. Only those paths whose indices increase monotonically are valid (i.e., reversals in the direction of a path are not allowed). Search is limited to a band around the diagonal whose width is twice the maximum displacement.

The displacement of a point in the new image relative to the same point in the reference image is  $d(s) = s - t$ . Instead of computing a height (elevation) map, we use the displacement information to warp the images in the epipolar direction to reduce parallax. Fig. 5 shows the method is capable of significantly reducing local misregistration due to parallax. Taken a short time apart the only significant changes are several vehicles that move between frames. Small changes caused by a car moving are evident the time warped difference

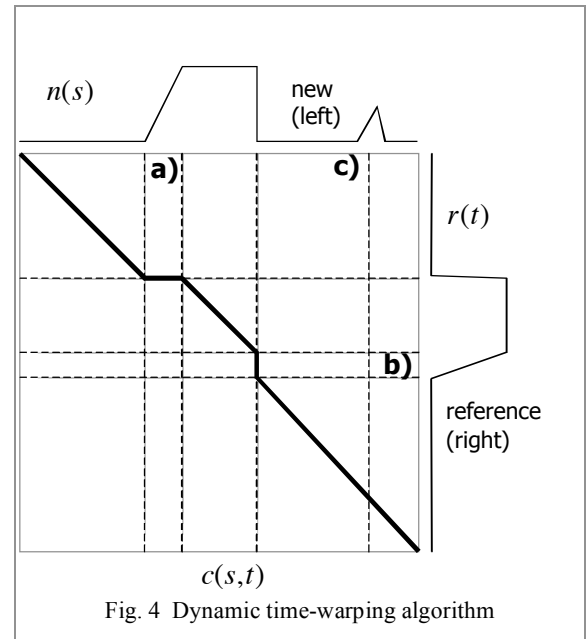


Fig. 4 Dynamic time-warping algorithm

image, while those same changes are lost among larger changes caused by parallax differences (and glint) between the two images.

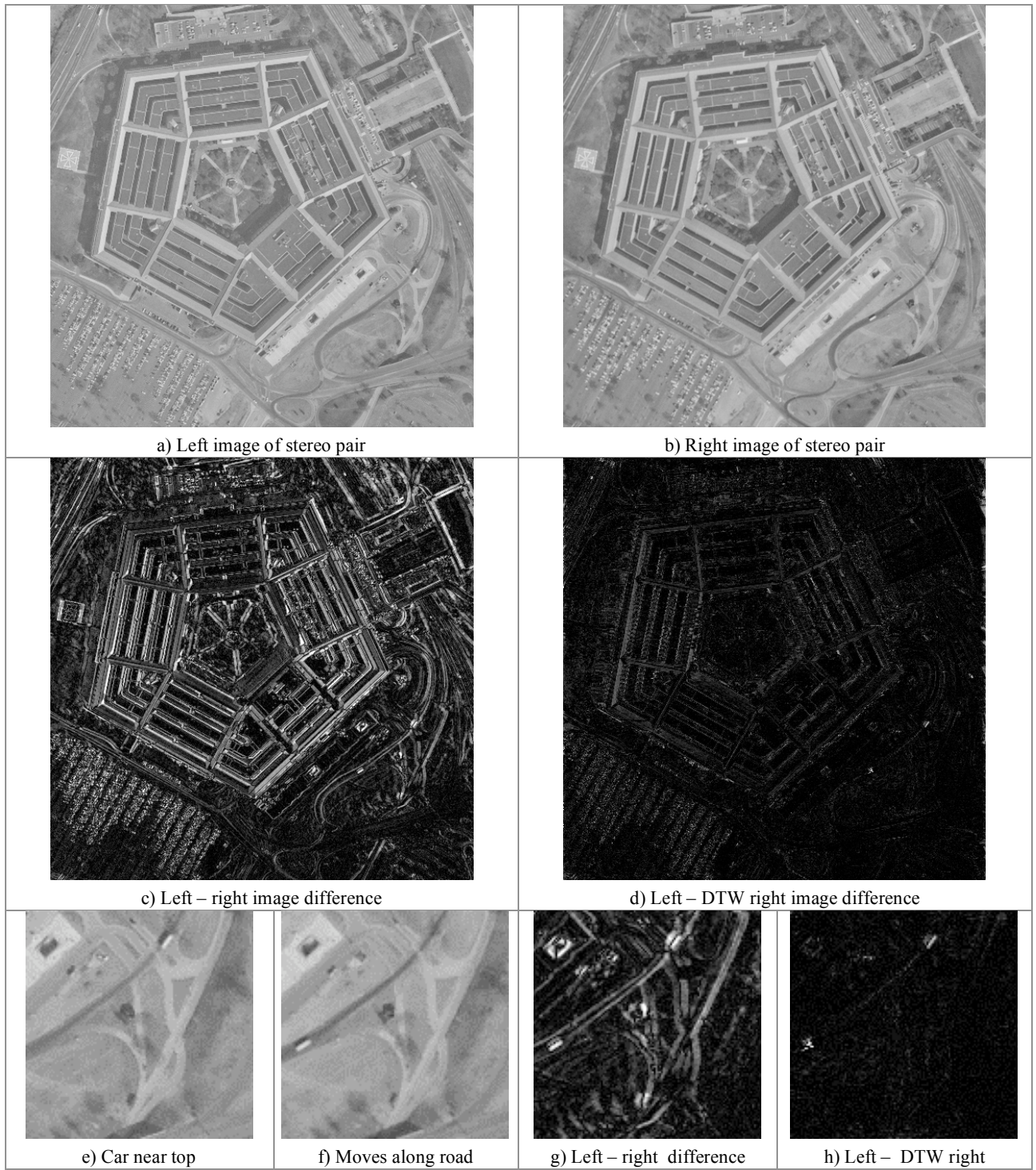


Fig. 5 Stereo image over the Pentagon (CMU/VASC Image Database)

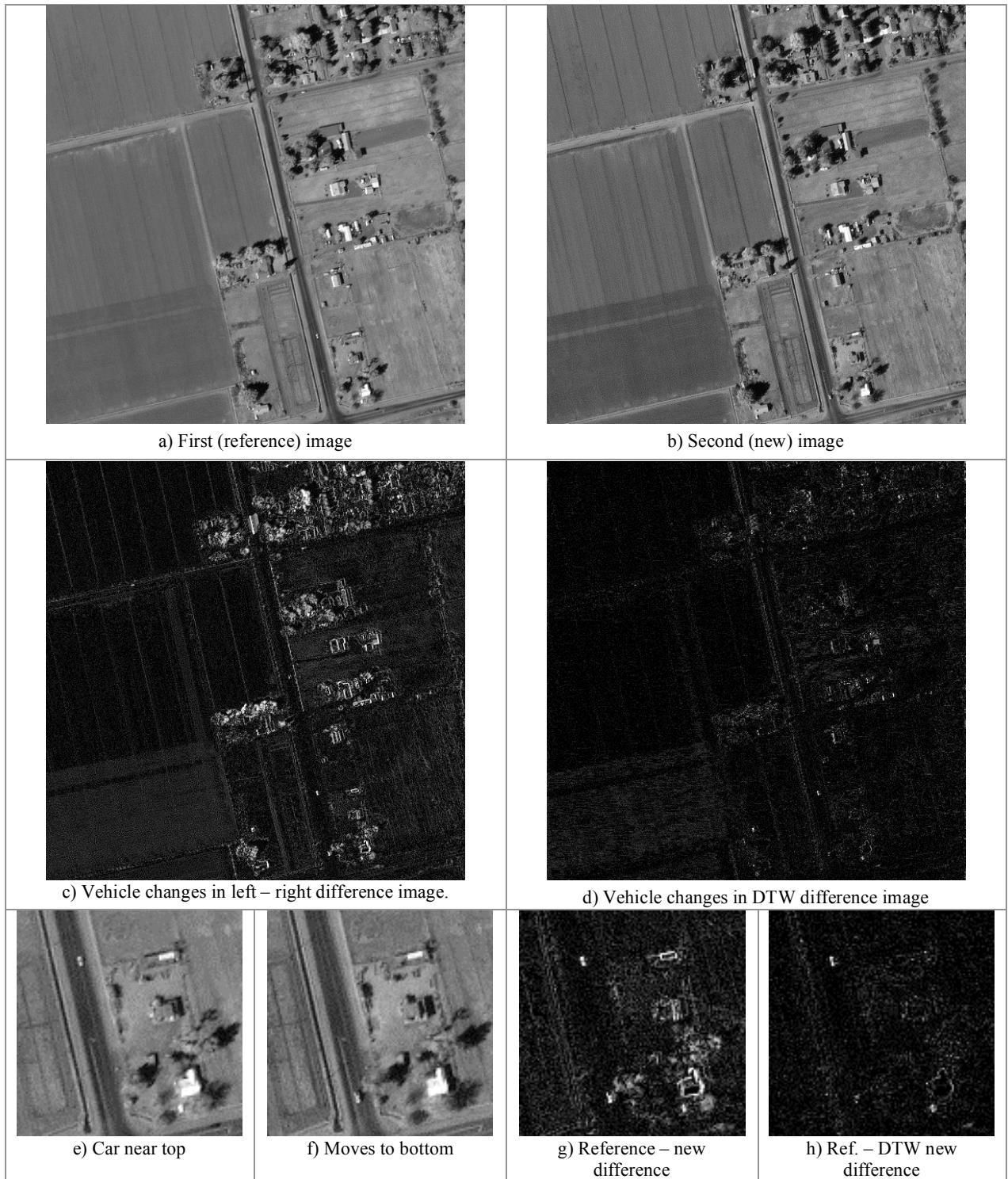


Fig. 6 Ikonos image pair (Imagery courtesy GeoEye).

## 4. IKONOS IMAGE EXAMPLE

We now use the performance model to estimate the processing gain of DTW on a pair of Ikonos images. The image pair (Fig. 6) was taken on the same pass about 1 minute apart and contains several vehicles in motion. One can thus assume relatively constant illumination between the two images (a,b). After coarse registration using an affine transform, residual parallax differences from buildings and trees remain (c), which are greatly reduced after warping (d).

Using DTW to estimate the displacement map, we obtain a mean value of zero (approximately), and a standard deviation of 2.57 pixels. The autocorrelation function of the reference image was computed and an exponential fit of measurements along the epipolar direction performed. A value of 0.879 was obtained for the correlation coefficient, which yields a predicted processing gain of 3.7 db. The input (new) image standard deviation = 26.17, and difference image stdev. = 15.4 gives a measured gain of  $10 \log 684 - 10 \log 237 = 4.6$  db. The model thus under predicts the processing gain by 0.9 db.

Knowing the displacements we warp the new image to the reference image. The warped image difference stdev. = 8.12, gives a measured gain  $28.8 - 18.2 = 10.6$  db. Running the DTW algorithm on the warped image pair produces a displacement map with a residual standard deviation of 0.39 pixels. This in turn leads to a predicted processing gain of 17.3 db. Our constant brightness model over predicts performance by 6.7 db. Actual performance is lower than predicted due to non-modeled changes in the background caused by non-Lambertian scattering effects such as glint that change with viewpoint. Thus the accuracy of our model is limited by the constant brightness assumption.

## 5. DISCUSSION

Using DTW to align images prior to change detection reduces residual parallax and increases processing gain. Measured performance agrees with predictions as long as parallax effects dominate. As we saw in the second part of the previous example, the accuracy of the prediction, which is based only on parallax, is reduced when other (non-modeled) effects dominate.

A problem with the method is that DTW tends to blur objects in the new (warped) image that are not present in the reference image. This can be seen to some extent in Fig. 5h. The effect is greater in Fig. 6h where the car in the new image moves close to a bright feature. In situations where objects either appear or disappear, this effect can be reduced by warping the image without the objects. Future work will address this problem more fundamentally for images that contain moving objects, as well as the extension of our model to account for other kinds of background change.

## References

- [1] R.J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam., "Image change detection algorithms: A systematic survey," *IEEE Transactions on Image Processing*, Volume 14, Issue 3, March 2005, 294 – 307.
- [2] R. Chellappa, Q. Zheng, L.S. Davis, C.L. Lin, X. Zhang, C. Rodriguez, A. Rosenfeld, and T. Moore, "Site-model based monitoring of aerial images," *1994 DARPA Image Understanding Workshop*, Monterey CA, 13-16 November 1994.
- [3] Y. Ohta and T. Kanade, "Stereo by Intra- and Inter-Scanline Search Using Dynamic Programming", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 7, pp 139-154, 1985.