

A Framework for Robust 3-D Change Detection

Aaron J. Heller, Yvan G. Leclerc, Quang-Tuan Luong

Artificial Intelligence Center, SRI International
333 Ravenswood Ave., Menlo Park, CA 94025 USA

ABSTRACT

We present an application of our framework for 3-D object-centered change detection¹⁻⁴ to combined satellite and aerial imagery. In this framework, geometry is compared to geometry, allowing us to compare image sets with different acquisition conditions and even different sensors. By working in this framework, we do not encounter the restrictions and short-comings of conventional image-based change detection, which requires that the images being compared have similar acquisition geometry, photometry, scene illumination, and so forth.

The contributions of our framework are: (1) using a geometric basis for change detection, allowing image sets acquired under different conditions to be compared; (2) explicit modeling of image geometry to be able to numerically characterize significant and insignificant change. The contributions of this paper are: (1) the algorithms are embedded in an integrated cartographic modeling and image processing system, which can ingest and make use of a variety of government and commercial imagery and geospatial data products; (2) experimentation with a variety of imagery and scene content.

Modifications to the algorithms specific to their use with satellite imagery are discussed and the results from several experiments with both aerial and satellite images urban domains are described and analyzed.

Keywords: Change detection, performance metrics, satellite imagery

1. INTRODUCTION

Most techniques for image-based change detection involve warping one image into alignment with another and then comparing pixel values. This requires that the images being compared have similar acquisition geometry, photometry, scene illumination, and so forth.

We present a framework for 3-D object-centered change detection in high-resolution earth-observing imagery, in which geometry is compared to geometry. By working in this framework, we do not encounter the restrictions outlined above and can compare image sets with different acquisition conditions and even different sensors (e.g., satellite vs. aerial).

The scene geometry can be obtained from a variety of automated 3-D reconstruction methods, such as stereo-correlation-derived dense elevation models or 3-D meshes derived from a collection of two or more views. Each stereo match or 3-D mesh vertex is evaluated using a score based on Minimum Descriptor Length (MDL) theory. This score has proven to correlate well with the likelihood that a stereo match or 3-D mesh vertex is correct.³ Self-consistency criteria are then applied to the scored results to quantitatively distinguish differences that represent actual changes in scene geometry from those arising due to errors in the reconstruction process or uncertainty in the imaging geometry.

Although the methods presented are quite general, the experiments we have conducted are based on two simplifying assumptions.

First, we use a specific class of objects: terrain (both rural and urban) viewed from above using aerial imagery. We take advantage of the special nature of terrain to simplify the problem. The 3-D shape is modeled as a single-valued function $z = f(x, y)$, where z represents elevation above a vertical datum as a function of the horizontal position (x, y) . Second, we assume that all the camera parameters for all the images are known and represented in a common Cartesian coordinate system (A.1 discusses how we handle non-central perspective sensors, such as those found in satellites). Together, these two assumptions reduce the problem of detecting changes in 3-D shape to that of finding point-by-point significant differences in elevation.

Further author information:

A.J.H.: E-MAIL: heller@ai.sri.com ; Y.G.L.: E-MAIL: leclerc@ai.sri.com ; Q.-T. L.: E-MAIL: luong@ai.sri.com

2. PREVIOUS WORK

Change detection is an important task in computer vision and was one of the first topics pursued in the field at the image intensity level^{5,6} and remains a topic of interest.⁷ However, comparing intensity values is not very effective because such changes do not necessarily reflect actual changes in scene content, but are often caused by changes in pose, illumination, and sensor properties.

A typical approach is to estimate the homography of an image pair and then warp one image into pixel alignment with the other, normalize the intensities, and examine pixel or area differences. In image analysis workstations, a flicker-wipe-fade function is often provided to allow a human operator to compare the warped images visually. An exact solution exists only for planar scenes, and as the amount of relief in a scene increases and the viewpoints become more disparate, the quality of matches achieved by warping becomes poorer. For use with high-resolution satellite imagery, such as that from SpaceImaging's IKONOS, this places such tight constraints on tasking that it is not practical. It also precludes carrying out comparisons between existing image archives, such as the USGS National Aerial Photography Program (NAPP)^{*}, and modern satellite images due to differences in sensor geometry models and photometric properties.

For man-made objects, such as buildings, higher-level comparisons have been proposed, based on feature organization⁸ and 3-D models.⁹⁻¹¹ These specialized approaches are the most successful, but are not suited to complex urban scenes or more general objects like natural terrain.

3. OBJECT-CENTERED CHANGE DETECTION

In our approach, we compare 3-D geometry to 3-D geometry, which addresses the problems outlined above. The geometry can be derived from the images with a variety of automated methods, such as the production of "dense digital elevation models (DEM)" using stereo correlation or optimization of a 3-D deformable mesh.¹²⁻¹⁶

We observe, however, that simple thresholding of the normalized point-by-point difference of the derived elevation is unreliable because it does not account for errors in the reconstructions. While each of these techniques produces a quantity that can be interpreted as a figure-of-merit (e.g., the correlation score or the value of the objective function in an optimization), in practice it is impossible to set thresholds for these quantities that work reliably across even a single scene. This is illustrated in Figure 1, which compares the results of the method described in the following sections with a simple thresholding of differences in normalized elevations.

Typical evaluations of the performance of 3-D reconstruction algorithms involve comparisons against independent "ground truth," however in the case at hand, if ground truth measurements were available, we could compare them and not have to use image-derived measurements. We require a method to characterize and distinguish significant from insignificant changes without reference to independent measurements. It is therefore necessary to have a quantitative measure of the uncertainty of the reconstruction algorithm's output.

3.1. Self-consistency

The human visual system has a truly remarkable property: given a static natural scene, the perceptual inferences it makes from one viewpoint are almost always consistent with the inferences it makes from a different viewpoint.[†] We call this property *self-consistency*.

The self-consistency methodology has been described elsewhere,^{4,17} so we will just give a brief summary. This methodology makes it possible to measure the expected variation in the output of a computer vision algorithm as a function of viewing geometry and contextual measures, for a given algorithm and a given class of scenes. The expected variation can be expressed as a probability distribution that is called the self-consistency distribution and is computed from sets of images by independent application of the algorithm to subsets of images.

In our current process, a stereo algorithm attempts to find matches consisting of a pair of points, one in each image, that correspond to the same surface element in the scene. If the image geometry models[‡] are known for the two images, we can estimate the 3-D coordinate of the surface element by triangulation.

^{*}For further information about NAPP, see <http://edcwww.cr.usgs.gov/webglis/glisbin/guide.pl/glis/hyper/guide/napp>

[†]In fact, we often use the terms *optical illusion* and *trompe l'oeil* to describe scenes from which it does not make consistent inferences.

[‡]We use the term *image geometry model* to denote the relationship between 3-D world points and 2-D image points for a particular image. This is commonly referred to as the image's *camera model*.

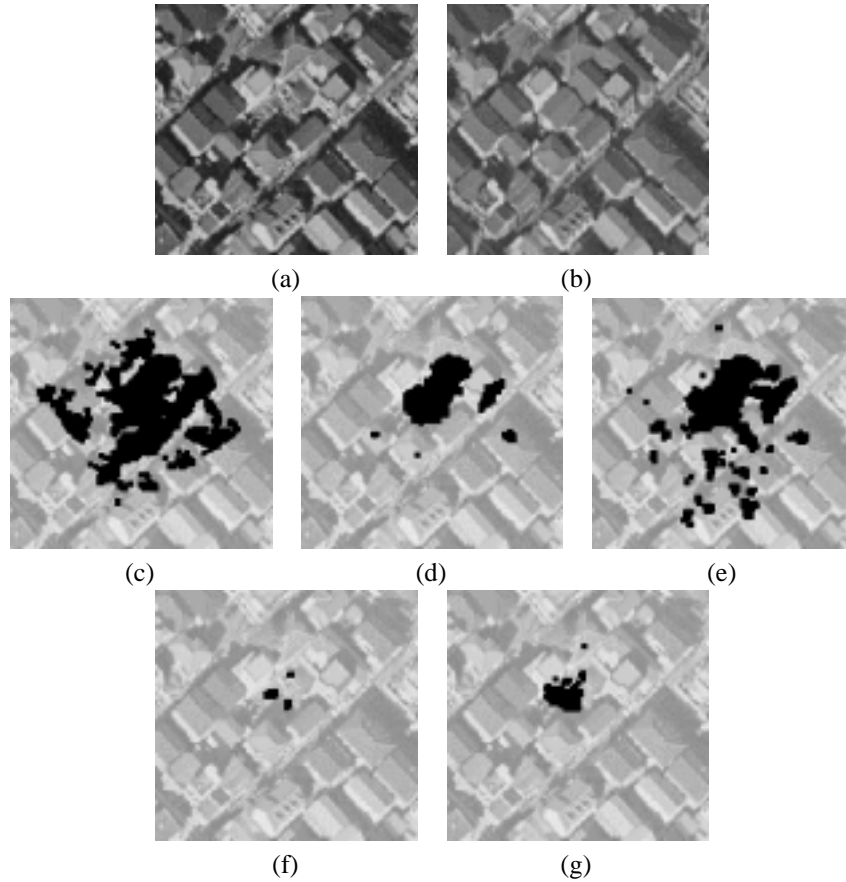


Figure 1. An illustration of the difficulty of using a simple threshold to detect changes in elevation in an urban scene. (a) One of four images taken at time 1. (b) One of three images taken at time 2 (approximately 3 months later). Note that there is a new building near the center of the image. (c) A simple thresholding of the normalized difference in z coordinates, showing differences greater than 3 units. (d) Same as c, but with differences greater than 6 units. (e) Union of all differences from all pairs of images, thresholded at 6 units. (f) The significant differences at the 99% confidence level found in matches between two pairs of images, using the self-consistency algorithm. (g) The union of all significant differences in matches from all pairs of images.

If several images of a static scene are available, we apply the same reconstruction algorithm to all pairs of image. We expect to get results that are similar in some places (e.g., where there is adequate texture and the surfaces are locally planar) and quite different in others (e.g., a incorrect match due to repeated structure in the image). The self-consistency distribution in this case is the distribution of differences in the 3-D coordinates for matches that belong to the same surface element.

Due to the special form of the surface we are estimating, $z = f(x, y)$, it is easy to find the matches corresponding to the same surface element, since they have the same (x, y) coordinate. The histogram of the differences, between the z coordinate for the matches, normalized by image geometry model, is called the *common-xy-coordinate self-consistency distribution*. (See^{4,17} for greater detail.) After normalization, one unit of vertical difference corresponds to a one pixel disparity at that location in a particular pair of images.

3.2. Summarizing self-consistency with the MDL score

Ideally, we would like to be able to use scene reconstruction created by as few as two images each. However, since two images produce only a single match for a given point in the scene, there is not enough data to compute the self-consistency distribution.

We address this by: (a) computing self-consistency distribution of a particular algorithm when applied to a particular scene class, and (b) computing self-consistency as a function of an appropriate score.

For a sufficiently constrained class of scenes (e.g., rural vs. suburban vs. urban), the self consistency distribution remains reasonably constant over many scenes taken at the same time. We can then use the average of these distributions to represent the self-consistency distribution of new images of a new scene within the same class. Furthermore, we can use this average distribution to predict the expected variation in reconstruction with we have only a single image pair of the new scene within the same class, by using the score as a predictor of self-consistency.

For this work, we use a score based on Minimum Descriptor Length (MDL) theory. It has a stronger correlation with self-consistency than other scores we have examined,³ in particular the SSD residual.

3.3. The change detection algorithm

We now have all the components for a robust object-centered change detection algorithm.

In the first stage, we establish baseline distribution for a given class of scene, by running the 3-D recovery algorithm, in this case a stereo algorithm, on a large number of subsets of images of the same class as those in which we want to perform change detection. A bucketing method is used to find all the common-xy-coordinate matches. Each pair is accumulated in a 2-D scatter matrix, in which the x-coordinate is the larger of the MDL scores for the two matches, and the y-coordinate is the image-geometry normalized difference between their triangulated z coordinates.

We then extract the significance level curves for the values of confidence, $s\%$, which we plan to use (say 99%). For a given value of the score, this is the normalized difference below which 99% of the common-xy-coordinate match pairs with that score lie.

In the second stage, we use the significance level curves to judge whether a given pair of matches derived from images taken at different time is significantly different. We find, using the same technique as before, the common-xy-coordinate matches where each match originates from a different point in time, compute the larger of their scores and the normalized difference between the recovered z coordinates.

If, for that score, this distance is greater than the distance at significance level 99%, then the pair of matches is deemed to be a difference with confidence 99%.

4. EXPERIMENTAL RESULTS

To carry out experiments with this technique, we used two 1-meter GSD panchromatic stereo pairs of a scene over downtown San Diego, CA, taken on 30 January and 7 February 2000 by the IKONOS high-resolution imaging satellite. The stereo pairs were controlled by SpaceImaging and image geometry models were supplied as the rational function coefficients in the Rapid-Positioning Capability (RPC) format.

For the baseline images, we used three Color Infrared (CIR) frames covering the same area of San Diego that were taken on 10 October 1996 as part of the USGS NAPP collection. These were scanned at 10 microns, yielding a GSD of approximately 0.4 meters and then minified by a factor of two to obtain images roughly the same scale as those from IKONOS. Two of the three images are shown in Figure 3. The NAPP images were bundle adjusted to ground control points triangulated from the IKONOS images. After adjustment, the resulting projection errors were less than 1 pixel. Unfortunately, the NAPP frames are noticeably less sharp than the IKONOS images, probably due to the fact they were scanned from 2nd generation copies of the original film. This turned out to be the major limiting factor in this study and required lowering the confidence threshold for change from the default 99% to 90% for these experiments.

We selected twenty areas, fifteen with visible changes and five with no changes for testing. Due to space limitations we illustrate our results with five representative areas in Figures 3 to 8.

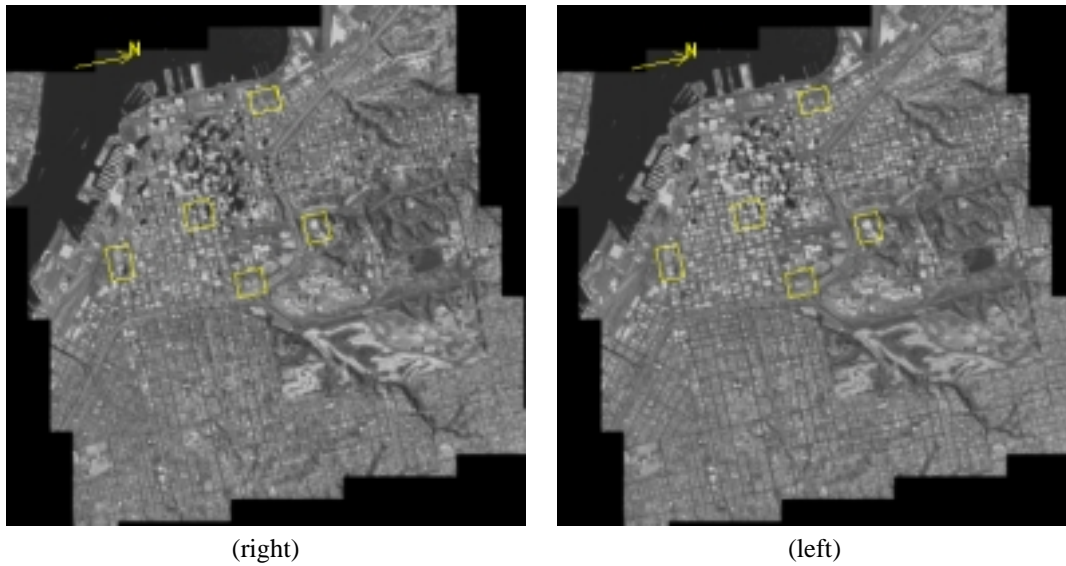


Figure 2. One of the two IKONOS stereo pairs used for the study. Footprint of areas used for experiments are shown.

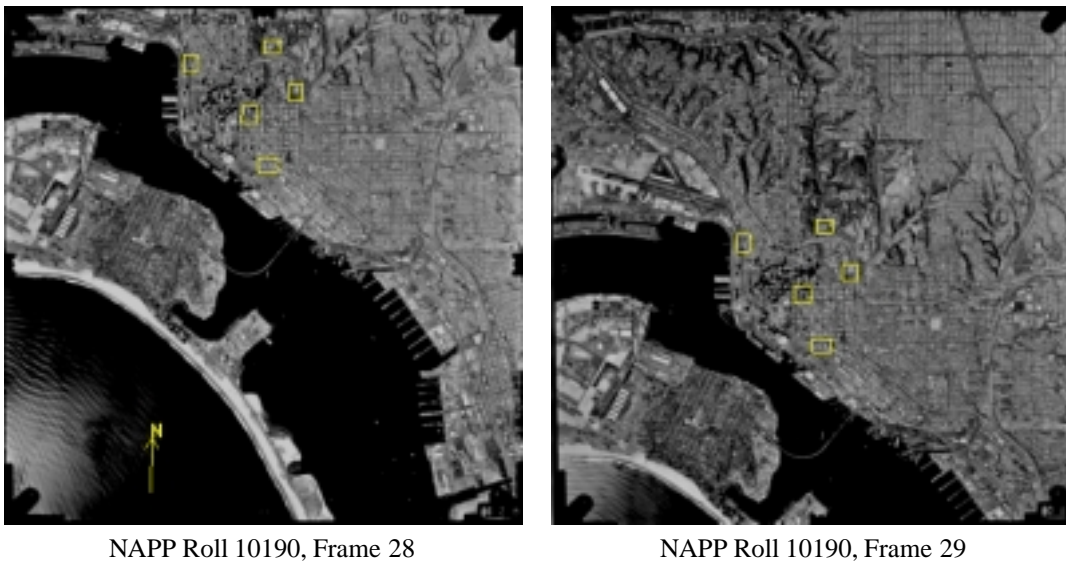
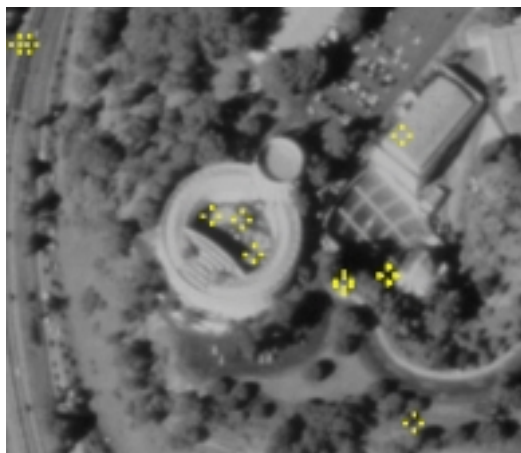


Figure 3. Two of the USGS NAPP images used as the baseline. Footprints of the areas used for experiments are shown.

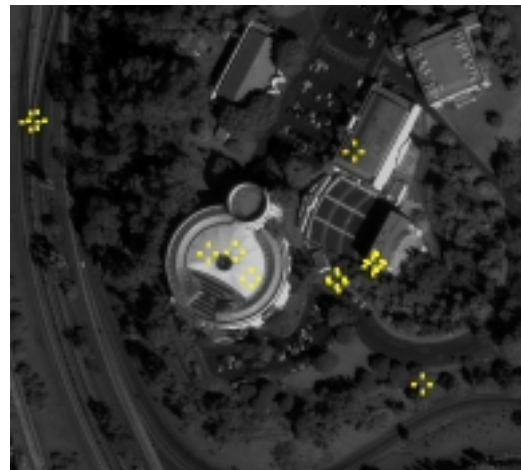
5. SUMMARY AND CONCLUSIONS

We have applied our 3-D change detection algorithm to a mixture of aerial and satellite imagery, which had not been done before. In spite of the many limitations imposed by the poor quality of the NAPP imagery and having only a single pair of IKONOS images taken a few days apart, we demonstrated that our algorithm could correctly detect changes in areas where significant visible change had occurred, while detecting very few false changes where no visible changes could be seen.

The ability to reliably detect changes over time using a mixture of sensors and vantage points, with as few as two images per time, makes it possible to detect changes in a wide variety of situations. For example, it is now possible to detect change across long periods of time, where it is often impossible to have many images of a scene all taken with the same sensor or from the same vantage point. This greatly enhance the utility of change detection.

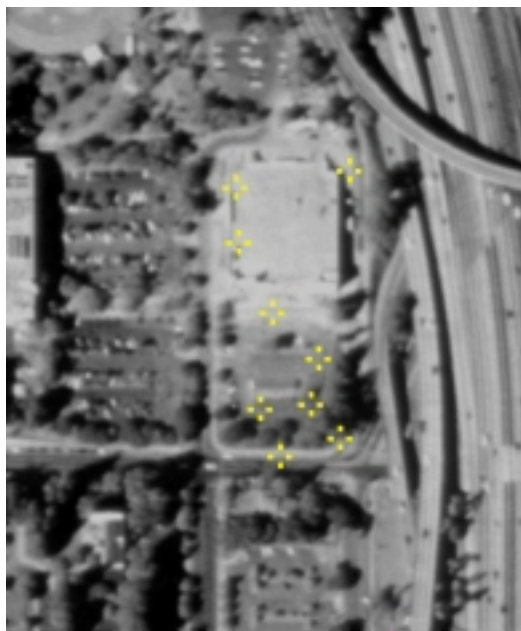


(a)



(b)

Figure 4. Stadium scene. (a) One of the NAPP images of this scene, taken in 1996. (b) One of the IKONOS images of this scene, taken in 2000. (This layout is used in the following 3 figures.) Note that the stadium's roof has closed from 1996 to 2000. The yellow cross-hairs indicate points where a change in elevation has been detected by our algorithm. Even though there are very few features on the closed roof, our algorithm managed to detect a change in elevation for a few points on the roof. Most of the detected changes outside of the roof area appear to be due to changes in vegetation.

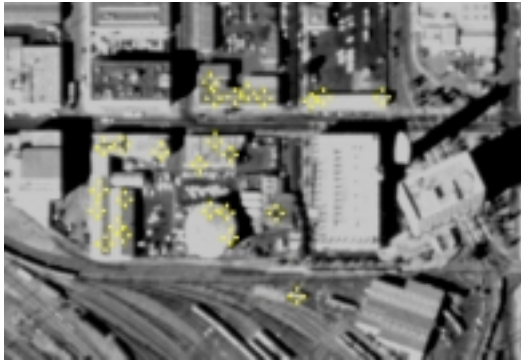


(a)



(b)

Figure 5. Warehouse scene. Note that buildings have been added in the center of this scene, and that there are now cars in the parking lot. Even though the change detection algorithm was applied to the entire scene, all of the detected changes lie within this central area.



(a)



(b)

Figure 6. Demolition scene. Note that many buildings have been demolished in the center of this scene, and that virtually all of the detected changes lie in this central area. The remaining change appears to be due to a new railroad car.



(a)



(b)

Figure 7. New buildings scene. Note that new buildings have been added near the center of this scene, and that virtually all detected changes lie in this central area.

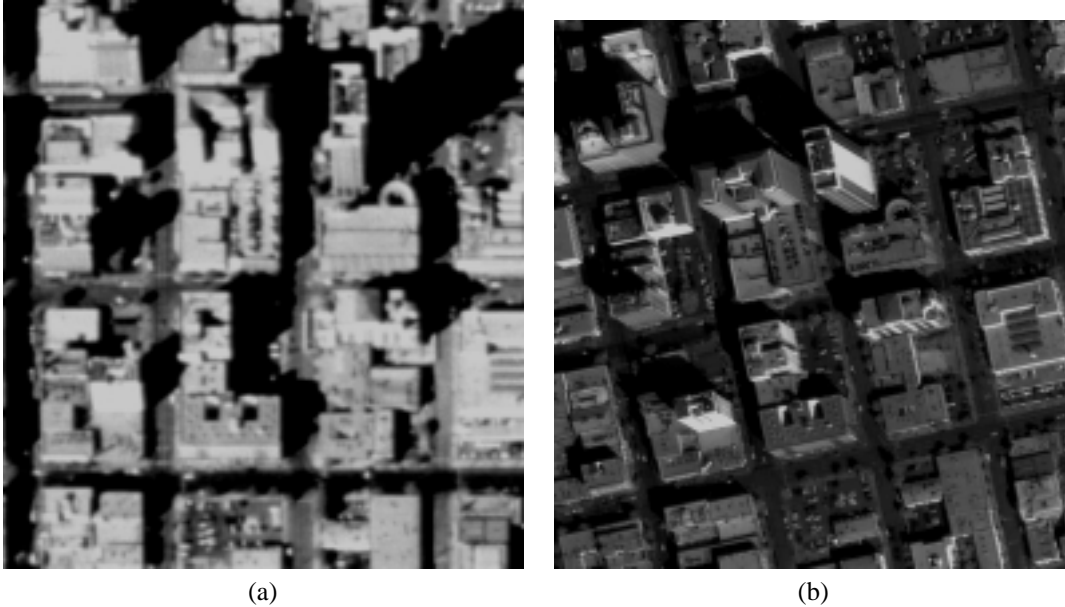


Figure 8. No change scene. Careful visual inspection of these images indicates that no visible changes have taken place over the entire scene. Note that the algorithm has detected no changes, in spite of the large differences in the vantage point and illumination, such as the large shadow cast by the tall building.

APPENDIX A. SENSOR GEOMETRY MODELS

In the context of our work, the term *sensor geometry model* refers to a set of functions that define the relationship between points in 3-D “object space” and their projection in 2-D “image space.” They have a number of parameters whose values are particular to a given image. In US DOD systems, the parameters are often called *support data* and a particular format for representing them in an image file’s header is called a *support data extension*. When a sensor geometry model is instantiated with a set of parameters corresponding to a particular image, we call it an *image geometry model*. This is informally referred to as the image’s sensor model or camera model. The term “math model” is frequently used interchangeably with the term geometry model.

Most commonly encountered cameras (e.g., 35mm, video) are classified by photogrammetrists as *stationary, central-perspective, framing cameras*. This means that all of the light rays that form the image pass through a single point, called the *center of projection* (which is the rear-nodal point in a compound lens), and the entire photographic frame is exposed at a single instant in time, during which the camera is held in a stationary position. These are informally called “pinhole cameras” and can be characterized with eleven parameters, six *exterior orientation parameters* (three position: x, y, z ; and three rotation: ω, ϕ, κ) and five *interior orientation parameters* (focal length: f ; principal point: pp_u, pp_v ; and aspect-ratio and skew).

These are the parameters of the *rigorous math model* of the stationary, central-perspective, framing camera, that relates points in 3-D object space to points in 2-D image space coordinates. We call it a *rigorous* math model because the parameters have a direct physical interpretation. It is also possible to model the projective relationship of a pinhole camera with a 3×4 matrix in homogeneous coordinates (often called a *direct linear transform*), and while it is a mathematically equivalent representation, this is not considered to be a rigorous math model because the entries in the matrix do not directly correspond to physical parameters of the imaging system.

Because of the difficulty in fabricating 2-D Charge-Coupled Device (CCD) arrays larger than a few megapixels, very-high resolution, electro-optical sensors are typically built with a 1-D CCD array. The image is formed by moving either the CCD array (e.g., a document scanner), parts of the optical system (e.g., a swinging-lens panoramic camera), or the entire sensor itself (e.g., LandSat). Since the entire image is not formed at a single instant in time, these are non-framing sensors, however, in the case of a panoramic camera it is possible to arrange for the rear-nodal point of the lens to lie on the axis of the pivot, and hence still form a central-perspective image.

With high-resolution imaging satellites, such as IKONOS, the position and orientation of the satellite is also changing during the exposure, so, additionally, they are not central-perspective sensors. In general, sensors of this type are called *dynamic sensors*. It is possible to develop rigorous math models for these sensors. They are very complex, often taking into account atmospheric refraction and other phenomena and typically have 20 or more parameters, involving, among other things engineering parameters of the sensor and optical system, the satellite’s orbital elements, and the ephemeris during the exposure (regular measurements of its position and orientation in space and their derivatives). Space Imaging has developed rigorous math model for IKONOS, however they consider this to be proprietary information. This model is used by Space Imaging for their processing and production, neither it or the parameters for a particular image are available to their customers.

With the proliferation of digital photogrammetric workstations and mapping tools, it soon became apparent that it was impractical (or in some cases, impossible) to include rigorous math models for each sensor. An additional consideration is that the rigorous math model for most satellite sensors is too slow for an interactive workstation. Furthermore, these math models typically compute geodetic longitude (Λ) and latitude (Φ) as a function of image coordinates (u, v) and geodetic height (H), which is not convenient for 3-D image understanding and modeling systems, that typically need to operate in a vertical Cartesian coordinate system.

SRI pioneered the use of fast, invertible, generic approximations to rigorous math models in its Cartographic Modeling Environment and 3DIUS systems, which implement an object-oriented application programmer interface (API) for projective sensor models.[§] The approximations are generated by using the rigorous math model to generate a grid of corresponding 3-D and 2-D points and then fitting the approximation to those points.

The most widely used sensor math model approximation is the *rational function* model, as implemented in the Rapid Positioning Capability used in some US DOD image products. The specific format of this data is called the Commercial Support Data Extension (RPC00A) and is described in the US DOD document *The Compendium of Controlled Extensions for the National Imagery Transmission Format* [STDI-0002].

The IKONOS stereo pair image product is supplied with sensor geometry model parameters in the RPC format. In NITF, the parameters are supplied in the RPC00A support data extension. In TIFF, the parameters are supplied in a separate file with the suffix *_rpc.txt*. As part of this project, we have written code for 3DIUS that can read both formats and instantiate rational function image geometry models.

In addition to central perspective and rational-function sensor models, 3DIUS supports the use of orthographic, block-interpolated, polynomial projections, as well as compound models, comprised of sequences of transformations and projections.[¶]

Due to the object-oriented nature of the system, there are no restrictions on the types of sensor geometry models that can be intermixed in a particular site and the generic sensor model API allows code to be written without concern for the specific types of the sensor models it will be used with. This allows, for example, images with central-perspective and RPC geometry models to be used simultaneously.

A.1. Rapid Positioning Capability (RPC)

The RPC defines a non-rigorous sensor geometry model that is comprised of two rational functions, that give image line and sample as a function of geodetic longitude (Λ), latitude (Φ), and height (H), referenced to WGS-84. The numerator and denominator of the rational functions are third-degree polynomials in Λ , Φ , and H , having 20 coefficients each. Ten offset and scale parameters, normalize the domain and range of the rational functions to the closed interval $[-1..+1]$. The coefficients of each polynomial are scaled so that they can be written as single-precision floats without loss of precision. In total, an RPC sensor geometry model has 90 parameters.

$$r_n = \frac{\sum_{i=1}^{20} LNC_{it_i}(\Lambda_n, \Phi_n, H_n)}{\sum_{i=1}^{20} LDC_{it_i}(\Lambda_n, \Phi_n, H_n)} \quad (1)$$

[§]See <http://www.ai.sri.com/~apgd/v1/docs/gensen.pdf> and <http://www.ai.sri.com/~apgd/v1/docs/generic-sensor-model.pdf> for further details.

[¶]See Heller and Quam, “The Radius Common Development Environment,” for further explanation. It is available from <http://www.ai.sri.com/ajh/papers/rcde-corrected.pdf>

$$c_n = \frac{\sum_{i=1}^{20} SNC_{it_i}(\Lambda_n, \Phi_n, H_n)}{\sum_{i=1}^{20} SDC_{it_i}(\Lambda_n, \Phi_n, H_n)} \quad (2)$$

Table 1. The ordering of the terms in each of the four polynomials.

i	t_i	i	t_i	i	t_i	i	t_i
1	1	6	$\Lambda_n H_n$	11	$\Lambda_n \Phi_n H_n$	16	Φ_n^3
2	Λ_n	7	$\Phi_n H_n$	12	Λ_n^3	17	$\Phi_n H_n^2$
3	Φ_n	8	Λ_n^2	13	$\Lambda_n \Phi_n^2$	18	$\Lambda_n H_n^2$
4	H_n	9	Φ_n^2	14	$\Lambda_n H_n^2$	19	$\Phi_n^2 H_n$
5	$\Lambda_n \Phi_n$	10	H_n^2	15	$\Lambda_n^2 \Phi_n$	20	H_n^3

$$\Lambda_n = (\Lambda - \Lambda_{off}) / \Lambda_{scale} \quad (3)$$

$$\Phi_n = (\Phi - \Phi_{off}) / \Phi_{scale} \quad (4)$$

$$H_n = (H - H_{off}) / H_{scale} \quad (5)$$

$$c_n = (c - c_{off}) / c_{scale} \quad (6)$$

$$r_n = (r - r_{off}) / r_{scale} \quad (7)$$

A drawback of the RPC geometry model is that when it is used to approximate low-order image geometry models, singularities may arise due to common factors in the numerator and denominator if simple fitting techniques are used. We have developed proprietary fitting algorithms that find the lowest order rational polynomial model consistent with the accuracy requirements, avoiding common factors.

When an image with an RPC geometry model is ingested in 3DIUS, a new rational function model that uses object space coordinates in the site's local vertical Cartesian system (rather than the RPC's original geodetic coordinates), is fit to the original model using SRI's proprietary techniques. By avoiding singularities in the original model, it is possible that the new geometry model has lower RMS error than the original.

This same technique can be used to fit a central perspective sensor geometry model to small areas of an image with a rational function geometry model.

Another widely held belief about rational function image geometry models is that they cannot be adjusted because they are not rigorous (i.e., the parameters do not have a physical interpretation) and have too many under-constrained parameters. To address this, SRI has developed a proprietary 13-parameter *quasi-rigorous* sensor geometry model that is suitable for use with IKONOS and similar satellite imaging systems. We call it "quasi-rigorous" because while the parameters have a physical interpretation, they do not necessarily correspond to those of any particular satellite. However, we can still make *a priori* estimates of the parameter covariances and carry out block adjustments in a least-mean square framework. After they are adjusted, they can be used directly or new rational function model can be fit to them and used.

APPENDIX B. ADAPTING THE CHANGE DETECTION SYSTEM FOR USE WITH IKONOS IMAGERY

A major technical challenge of this project was to adapt a system that was originally written to work with central perspective geometry models, to work with large (e.g., 10000 × 10000 pixels), 11-bit/pixel images with rational-function geometry models typical of IKONOS products. Since a full rewrite of the system is out of the scope of this project, we elected to create a preprocessor within 3DIUS, for the original system that modifies the images and their geometry models into suitable form for the existing system. Our approach exploits locality constraints in the following ways:

1. Over small regions of the images (e.g., 512 × 512), the movement of the satellite is small enough that we can fit an accurate central perspective geometry model to a given region.

2. Over small regions of the images, the dynamic range is typically much smaller than that of the entire image. This allows us to compute photometric transform that maps the pixel-value range of the region in the original image into an 8-bit range.

Once the imagery has been preprocessed, it is very similar to the aerial mapping photography that was used for the initial tests of the change detection system and we would expect similar performance. As mentioned in the introduction, at this point in time we cannot test the full change detection system since we have received only a single IKONOS stereo pair from Space Imaging. We can however run the initial shape estimation code and qualitatively evaluate the results.

ACKNOWLEDGMENTS

This work was sponsored in part by the Defense Advanced Research Projects Agency under contract F33615-97-C-1023 monitored by Wright Laboratory. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency, the United States Government, or SRI International.

REFERENCES

1. Y. G. Leclerc, Q.-T. Luong, and P. Fua, "A framework for detecting changes in terrain," in *Proceedings of the DARPA Image Understanding Workshop*, (Monterey, California), Nov 1998.
2. Y. Leclerc, Q.-T. Luong, and P. Fua, "Characterizing the performance of multiple-image point-correspondence algorithms using self-consistency," in *Proceedings of the Vision Algorithms: Theory and Practice Workshop (ICCV99)*, (Corfu, Greece), Sep 1999.
3. Y. Leclerc, Q.-T. Luong, and P. Fua, "Detecting changes in 3-d shape using self-consistency," in *Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR2000)*, (Hilton Head, South Carolina), Jun 2000.
4. Y. Leclerc, Q.-T. Luong, and P. Fua, "Measuring the self-consistency of stereo algorithms," in *Proceedings of the European Conference on Computer Vision (ECCV2000)*, (Dublin, Ireland), Jun 2000.
5. L. Quam, *Computer Comparison of Pictures*. PhD thesis, Stanford University, 1971.
6. R. Lillestrand, "Techniques for change detection," *TC* **21**, pp. 654–659, July 1972.
7. P. Rosin, "Thresholding for change detection," in *ICCV98*, pp. 274–279, 1998.
8. S. Sarkar and K. Boyer, "Quantitative measures of change based on feature organization: Eigenvalues and eigenvectors," *CVIU* **71**, pp. 110–136, July 1998.
9. M. Bejanin, A. Huertas, G. Medioni, and R. Nevatia, "Model validation for change detection," in *WACV94*, pp. 160–167, 1994.
10. A. Huertas and R. Nevatia, "Detecting changes in aerial views of man-made structures," *IVC* **18**, pp. 583–596, May 2000.
11. T. Boulton, R. Micheals, A. Erkan, P. Lewis, C. Powers, C. Qian, and W. Yin, "Frame-rate multi-body tracking for surveillance," in *DARPA98*, pp. 305–313, 1998.
12. P. Fua and Y. G. Leclerc, "Combining stereo, shading and geometric constraints for surface reconstruction from multiple views," in *SPIE Workshop on Geometric Methods in Computer Vision*, (San Diego, CA), Jul 1993.
13. P. Fua and Y. G. Leclerc, "A unified framework to recover 3-d surfaces by combining image-based and externally-supplied constraints," in *Proceedings of the DARPA Image Understanding Workshop*, (Monterey, CA), Nov 1994.
14. P. Fua and Y. G. Leclerc, "Using 3-dimensional meshes to combine image-based and geometry-based constraints," in *European Conference on Computer Vision*, pp. 281–291, (Stockholm, Sweden), May 1994.
15. P. Fua and Y. G. Leclerc, "Object-centered surface reconstruction: Combining multi-image stereo and shading," *International Journal of Computer Vision* **16**, pp. 35–56, Sep 1995.
16. P. Fua and Y. G. Leclerc, "Taking advantage of image-based and geometry-based constraints to recover 3-d surfaces," *Computer Vision and Image Understanding* **64**, pp. 111–127, Jul 1996.
17. Y. G. Leclerc, Q. T. Luong, and P. Fua, "Self-consistency, a novel approach to characterizing the accuracy and reliability of point-correspondence algorithms," in *One-day Workshop on Performance Characterisation and Benchmarking of Vision Systems*, (Las Palmas de Gran Canaria, Canary Islands Spain), 1999.