CHANGE DETECTION AND ANALYSIS IN MULTISPECTRAL IMAGES

Keith Price
Image Processing Institute
University of Southern California
Los Angeles, California 90007

Raj Reddy
Computer Science Department
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213

ABSTRACT*

This paper describes work on the development of symbolic registration and change analysis techniques applied to the problem of the comparison of pairs of images of a scene to generate descriptions of the changes in the scene. Unlike earlier work in change analysis, all the matching and later analysis is performed symbolically. We also discuss techniques for the generation of symbolic descriptions of images, both the segmentation and feature extraction problems. These techniques have been applied to several different scenes and in this paper we present the results for two of these scenes.

INTRODUCTION

This paper describes research toward the development of a general image understanding system. We will describe a system for the analysis of pairs of images of the same scene to generate partial descriptions of the changes in the scene. We directed this work toward general image analysis and matching rather than toward specific, related problems such as motion or stereo analysis.

This work in change analysis differs from earlier computer change detection in the use of symbolic analysis of the image to detect and express the changes which have occurred. Earlier efforts in the change analysis area (Quam, 1971; Lillestrand, 1972; Allen et.al., 1973) used correlation guided matching to establish a set of corresponding point pairs. These point pairs are then used to transform the second image so that it is precisely aligned with the first. The aligned images are subtracted and changes are indicated by a large difference in the intensity value of the point in the two images. That is, two images are processed to produce a third image which indicates possible changes. For further machine analysis of the changes, the change results should be represented symbolically. Rather than generate a symbolic description of the difference image, which is not always reliable, the initial matching should also be done symbolically. This use of symbolic analysis is intended to expand the class of images which can be successfully analyzed for changes compared to the class of images handled by techniques depending on point to point matching and global transformations. Past work on symbolic motion analysis has used completely and correctly segmented images, that is, the segmentation and recognition performed by humans, not by machines (Balder, 1975). We will be using "real" images which will require the generation of the symbolic descriptions in addition to the processing of the symbolic descriptions.

There are several segmentation techniques which might be applied to a wide range of images (Yakimovsky, 1973; Ohlander, 1975), and we have chosen the region splitting system of Ohlander. The processing starts with the pair of input images and first produces a partial or complete segmentation of these images. Feature values are computed for all these segments. The features include size, location, shape, etc. The feature based description of the segmented image constitutes the symbolic representation of the image. The symbolic representations of two images are compared to determine the corresponding regions in the two images, this is called symbolic registration. These results are used to guide further segmentations, to guide further registration, or to analyze changes in the regions which occurred between the two views (changes in color, size, location, shape, etc.).

We have applied this procedure on several diverse scenes: a simple house, a cityscape, satellite images, aerial images, and radar images. For each of these scenes a different task was performed. The task controls the segmentation, feature analysis, symbolic registration, and the final change analysis (Price, 1976). Because of space limitations, we will present results from only two of these scenes.

SUMMARY OF THE TASKS AND IMAGES

Before any analysis is possible a task must be given and any knowledge about the images and the task must be specified. For a system designed to solve a specific problem, the task description and available outside knowledge are used throughout the implementation. But, in a more general system, such as this one, the task and knowledge must be expressed explicitly for each stage of the processing. This outside information will be used

*This work was performed at Carnegie-Mellon University and was supported in part by the Defense Advanced Research Projects Agency under contract number F44620-73-C-0074 and is monitored by the Air Force Office of Scientific Research.

to control the type of regions which are segmented, the portion of the image which is analyzed, which segments are matched, what change information is generated, and the final use of this change information.

The first scene is given in two views of a single house (Figures la and lb). This scene is represented by full color images with about 0, 5 million pixels for each color input. There are originally three colors (red, green, and blue), and these are transformed to generate other color parameters (Kender, 1976). There are few real changes between these two images, so that the primary task is to illustrate the operation of the segmentation, feature extraction, and symbolic registration procedures. For this task, the large, clearly segmented, regions are sufficient; thus there is no need to segment each individual brick.

The other pair is two aerial views of an urban scene (Figures 2a and 2b). These images are monochromatic and contain about 4 million pixels each. This scene contains many, relatively small, distinct objects, and many other smaller objects. There are several global changes between the two images; objects are larger in the first image, there is a translation difference, and there is a sun angle difference so that shadows in the first image do not appear in the second image. The basic task for this scene is to analyze changes in the pier area of the scene, specifically changes in the number of ships. Before this analysis can take place, the corresponding subsections of the two images must be located, therefore the entire image must be partially segmented to locate certain anchor regions which can be used to limit the area of the pier subsections. Additional refinement of this area will be necessary and can be performed by using a description of the pier area: piers jutting out from the land surrounded by water or ships, and the land is not included in this final model. Therefore the total processing of this scene will include the segmentation of anchor regions in the two images (bright regions will be used), the registration of these partial segmentations, the restriction of the area of detailed analysis by using the symbolic registration results, the partial segmentation of this area to aid the refinement of the pier area, the complete segmentation of the refined area, and the analysis of the segmented regions to produce the desired change results.

SEGMENTATION

Our work on segmentation is an extension of the histogram guided region splitting technique developed by Ohlander (1975). This method was originally developed for use on color images. Basically the procedure splits a region into subregions by thresholding one of the spectral inputs. The threshold values are selected by the analysis of the histograms of the values for all pixels in the region (one histogram for each spectral input). The threshold values are selected as the upper and lower bounds of the "best separated" peak which appears in the set of histograms. More details of this basic segmentation method are given in Ohlander (1975) and Price (1976). There are two problems in the use of this technique for the segmentation of our set of images. First, the segmentation method is much too slow to process a large set of images in a reasonably short time. Second, the segmentation technique was developed for multi-spectral images and is not expected to work as well on monochromatic images.

PLANNING: The first problem is solved by the introduction of "planning" (Kelly, 1970; Hanson, 1974). By planning, we mean the generation of an approximation for the final segmentation using a reduced version of the image and the use of this approximation as a plan to more efficiently derive the true segmentation of the image. The use of planning reduces the total time to segment the image by about one order of magnitude compared to segmentation, by this method, without planning. The plan generation is performed by the application of the basic segmentation procedure to the reduced images. An expansion procedure uses the segments generated by the planning procedure to restrict the area in the large image where the threshold is applied, and generates the full size segmentation.

MONOCHROMATIC IMAGES: The segmentation of monochromatic images required additional alterations to the initial segmentation method. The original segmentation method was based on the hope that if one spectral feature cannot provide a reasonable split of the region, then, perhaps, another color parameter will. If the procedure is presented with only one spectral input, there is no other color parameter to turn to when there is only one peak in the histogram. The large monochromatic images also contain many small different objects which cause the histogram to have only one peak. This occurs because the range of intensity for each object overlaps the ranges of intensities for other objects.

We can introduce additional spectral-like features by the use of simple textural operators designed to show specific features such as homogeneous regions, or high contrast areas. We use one feature, the number of micro-edges in the reduction window, to indicate general homogeneous regions (Rosenfeld, 1969). A homogeneous region is one which contains few micro-edges so that these regions can be extracted by using a threshold of zero edges in a reduced image. The regions which are extracted by the edge feature are more sensitive to noise in the image, especially local noise such as scratches. This feature is not useful for the extraction of small detailed regions, but is useful for the extraction of general homogeneous regions.

Many other spatial or textural operators are possible (e.g. Haralick et. al., 1971). For easy incorporation in a general segmentation method, similar to this one, the operators should produce image-like values for all points in the image.

We also used the computation of histogram for portions of the image rather than for the entire image (Chow, 1970). This was intended to approach a solution to the problem of the occurrence of many small similar regions in a single image. The use of partitions means that the number of separate objects which contribute to one histogram is reduced.

Results and Evaluation of Segmentation:

Figures 3a and 3b show the regions segmented in both images: the sky, roof, lawn, four wall regions, several bushes, chimney, door, shadows, and several regions in the window area. There are some differences in the segmentations: differences in the number and size of the bushes, different regions in the window and door areas, and a cloud is segmented in the second image. These images are segmented with the planning procedure outlined above, with the images reduced by a factor of eight in each direction. No textural properties are used in the segmentation because the color parameters are sufficient. The total segmentation required about 24 minutes (465 million instructions on a PDP-KA10) with about 50% of the time for the reduction of the images, 9% for the plan generation, and 40% for expansion of the plan. An approximation of the segmentation time for these images without planning can be generated by scaling the times in the planning step which depend on the image size, by the reduction factor. This gives a segmentation time of about 72 minutes (1300 million instructions), but, in reality, the time required would be much greater due to the extra overhead involved in processing images which are too large to fit in primary memory.

Figures 4a and 4b give the partial segmentations for the entire urban scene images, and Figures 5a and 5b show the complete segmentation for the pier subsections of these images. Because of the task description, no attempt was made to completely segment the entire image, and only the bright regions are extracted. There are some differences between the two segmentations: some regions in one image are not in the other image, and some regions change shape due to the different sun angle. But, there are enough corresponding pairs to determine the scale and location changes and to use as anchor regions. A final refinement of the pier area, the removal of the land portion, is required before the complete segmentation is attempted. In both images the regions generated by the first segmentation step are used, shadows in the first image and water in the second. In the pier subsection segmentations, there are some obvious "errors": single objects such as ships or piers are broken into two or more regions and multiple objects such as adjacent ships appear as single segments. These two segmentations used the textural measures described above, and were generated by the planning procedure with a reduction factor of four.

The segmentation procedure is automated, but includes provisions for operator interaction. The special case segmentation operations to generate general partial segmentations (bright regions only, homogeneous regions only, etc.) may require operator intervention. It is difficult to judge segmentation quality, except to ask if the segmentation is sufficient for the given task. The absolute accuracy of the segmentation could be improved by the use of additional features, but this was not necessary for our tasks.

This section has discussed a segmentation procedure, not a combined segmentation and interpretation procedure as given in Yakimovsky (1973) or Tenenbaum et. al. (1974). This segmentation procedure can be used as a component of a larger system to completely segment a scene before any other analysis is performed, to select specific regions for further analysis, or to further analyze previously segmented regions.

FEATURE EXTRACTION

The matching and change analysis will be performed with a feature based description, symbolic representation, of the images rather than the signal description of the image. This section will describe the type of features which are used.

The features fall into several general classes: size, shape, color (including texture), and location. The exact feature measures are intended to capture some aspect of these feature classes. The measures given here are not intended to be a complete set, just an indication of the type of features which can be used. The features include:

Absolute size	Number of pixels in the region
Colors	Mean of spectral parameters over the region, one for each spectral input and one for each textural measure
Absolute location	Location of the center of mass of the region
Relative position	The regions above, below, to-left, and to-right of the region
Neighbors	The adjacent regions
Orientation	The orientation of the major axis of the region
Length to width ratio	Orientation independent, ratio of minor and major axis
Fractional fill	Fraction of the minimum bounding ellipse filled by the region
Perimeter^2 /Area	Square of the perimeter divided by the area

As we implemented these features, the total computation time for all the feature measures for one image is approximately the same as the time for the segmentation of the image. This would not be the case in a more optimized implementation, since the feature computation times can easily be reduced more than the segmentation times by simple prograrnming changes and the elimination of some generality in the implementation of feature computations.

SYMBOLIC REGISTRATION AND CHANGE ANALYSIS

Earlier systems for change analysis relied on a correlation guided matching procedure to locate corresponding point pairs. The location differences of these point pairs are used either for transforming one image so that it is aligned with the other, or for depth analysis if the system uses stereo pairs (Quam, 1971; Lillestrand, 1972; Allen et. al., 1973; Levine et. al., 1973). The aligned images are subtracted, producing a third, difference, image. This difference image must be analyzed to determine where the changes occurred, and the type of changes that occurred. Special purpose systems have been built to perform these tasks, so that these apparently expensive operations are performed quickly. But, change analysis systems which are intended to operate on uncontrolled image pairs (i.e. not on stereo pairs) encounter several problems. Major changes in the point of view of the observer, especially in oblique views, will cause inaccurate matches when the matches depend on intensity values in a small neighborhood of a point on the edge of the object. These changes are difficult, if possible, to account for in a global warping of the image. Most systems assume a "rubber sheet" warping, so that points adjacent in one image are assumed to be adjacent in the other image. A new object in the scene can cause similar errors in matching, and would usually be indicated as a large difference in the subtracted image.

Symbolic matching is an alternative technique to eliminate these problems encountered by signal based change analysis methods. Since the matching for one region does not necessarily depend on the intensity values adjacent to the region being matched, the change in the relative position of objects should not reduce the chances of a correct match. The knowledge about a scene can specify that the relative position or adjacency relations will change; this indicates that these features are not to be used for symbolic registration. New objects in the scene are indicated by regions in the second image which had no corresponding region in the first image, and missing objects by regions in the first image which fail to match with any region in the second. A symbolic change analysis system describes the changes as changes in the features of regions. Thus there is no need for extensive processing of a difference image to discover the kinds of changes which occurred. Many of these changes are given directly from the symbolic analysis. Symbolic change analysis is composed of symbolic registration: finding the corresponding regions in the two regions, and change analysis: determining what changes in feature values occurred between the two views of the scene. We will now discuss these two problem areas in more detail.

Symbolic Registration: The basic symbolic registration procedure computes a numerical rating for the match between any two regions in different, or even the same, images. This rating procedure combines the differences in each available feature of the two regions and produces an overall rating for the region to region match. The same procedure also produces a rating for each feature to feature match. The feature to feature match rating can be used for later change analysis. If the match is exact (e.g. when matching a region with itself) then the rating will be zero, and as the match worsens, the rating decreases.

The outside knowledge sources can indicate that certain features may change and thus should not be used in the matching procedure. For example, when the task description indicates that there are rotation differences between the two images, the matching procedure should not use the rotation dependent features, such as the absolute position, the orientation, the regions above, the regions below, etc. in the registration process. Rather than eliminate the use of these features altogether, v/e use different strengths for the features which should remain constant and the features which will change. The strengths are selected so that a bad match rating in one feature that should remain constant will have more impact than several bad match ratings in features which may change.

The region to region match procedure is used by the symbolic registration procedure to find the best available match for a given region. To find the region in the second image which corresponds to a given region in the first image, the symbolic registration procedure matches the region in the first image with each region in the second. The best matching region is considered to be the corresponding one. Even if a region does not have a corresponding region in the other image, some region will be selected as the corresponding region. This region will be the most similar region, but these two regions may have differences in features which should remain constant. Also, when this occurs, another region in the first image may correspond to the same region in the second image. New regions are indicated by regions in the second image which had no corresponding region in the first image, but missing regions must be confirmed by matching the regions in the second image with regions in the first.

Change Analysis: Some change analysis can be performed directly on the symbolic registration results and used by the symbolic analysis procedures. For example, in the urban images we are given, through the task and image description, the fact that there is a scale change between images. The amount of the scale change is not given by the outside knowledge, but it can be computed from the size differences found in early matches. This scale change is used to adjust the size measures for regions in the later matches. Since there is a scale difference between the two images, the absolute size and location features will change and cannot be used as constant features in the matching operation. But, with the computed scale difference, the size feature can be used as if it is a constant feature. This use of change results can be extended to the absolute location and orientation features. These adjustments can apply only when the changes are uniform throughout the image, which is not the case when there are perspective changes as in oblique views. But, such adjustments are possible to use on these types of feature changes in most aerial images.

Results and Evaluation:

The house scene is completely segmented, therefore most regions in the first image should have a corresponding region in the second image. Figure 6 shows the results from the symbolic registration procedure. In this figure, the corresponding regions in the two images are given the same colors. The major regions are all matched to the proper corresponding region (such as the sky, roof, lawn, wall sections, bushes, shadows, etc.). The differences in the segmentations and in the size of some regions, because they are adjacent to the edge of the image, did not interfere with the symbolic registration of these images. This pair of images is very straightforward to analyze, and there are few opportunities for mistakes. For example, no other region resembles the sky in terms of color and location, so that this match, and many other matches, could be determined with very few features. Some changes are indicated in the color parameters of most of the regions, and in the size, shape, and position features of some of the regions.

The full urban scene is only partially segmented so that corresponding regions in the two images are found for only a small number of the objects in the scene (Figure 7). The first pairs of corresponding regions were used to adjust the size and location features for the later matches. The size adjustment is 1.484 (for the area of a region). The translation differences are 14 pixels in the "I" direction and -221 in the "J" direction. The regions labeled "B" and "M" were used for this purpose. These regions matched well even with the size and location differences. In regions "C", "D", and "E", where their shapes are different, the use of the adjusted location feature in the matching is very important. Some changes are found between the regions in these two images: The regions in the second image are brighter than the regions in the first image. Some of the regions change size, shape, and location due to differences in segmentations, or actual differences in the bright objects. The location changes are usually induced by the size changes.

The pier subsections were not registered to each other because the desired changes were not changes in the features of objects, but changes in the number of ships. To effectively perform this final task, a recognition procedure would be required to identify the segmented regions. If the recognition procedure were given the segments which were extracted, several errors might be expected because of the segmentations. Several ships are broken into two regions and these may be recognized as two individual ships or as no ships. Also, some of the piers are broken into several pieces and some of these pieces may resemble a ship more than a pier.

CONCLUSIONS

We have presented an initial effort in the use of symbolic techniques for change analysis. The results of this work indicate that symbolic analysis can be applied to a large class of scenes by a general change analysis system. The use of a general analysis system introduces the problem of specifying and incorporating task knowledge which is not encountered in special purpose systems. We are not proposing that symbolic analysis techniques are the best solution to each special problem, but that symbolic analysis techniques offer a chance for the development of more general change analysis systems.

Further work is still required in the area of segmentation, especially the refinement of segments and merging of segments into objects; in the area of feature extraction, especially the use of textural or spatial features; in the use of knowledge in the registration and analysis of the multiple views; and the application of symbolic change analysis to other tasks.

REFERENCES

1. Allen, G.R., L.O. Bonrud, J.J. Cosgrove, and R.M. Stone (1973), "The Design and Use of Special Purpose Processors for the Machine Processing of Remotely Sensed Data," IEEE Symposium on Machine Processing of Remotely Sensed Data, Purdue University, October 1973.

2. Badler, N.I. (1975), "Temporal Scene Analysis: Conceptual Descriptions of Object Movements," Technical Report Number 80, Department of Computer Science, University of Toronto, February 1975.

3. Chow, C.K., and T. Kaneko (1970), "Boundary Detection of Radiographic Images by a Threshold Method," IBM Research Report RC-3203, 1970.

4. Hanson, A.R., and E.M. Riseman (1974), "Preprocessing Cones: A Computational Structure for Scene Analysis," Technical Report 74C—7, Computer and Information Science, University of Massachusetts, Amherst, Massachusetts, September 1974.

5. Haralick, R.M., and D.E. Anderson, "Texture-Tone Study with Applications to Digital Images," ETL-CR-71-14, Center for Research, University of Kansas, Lawrence, Kansas, November 1971.

6. Kelly, M.D. (1970), "Visual Identification of People by Computer," AIM-130, Thesis (Computer Science), Stanford University, Stanford, California, July 1970.

7. Kender, J. (1976), "Saturation, Hue, and Normalized Color: Calculation, Digitization Effects, and Use," Technical Report, Department of Computer Science, Carnegie-Mellon University, Pittsburgh, Pennsylvania, 1976.

8. Levine, M.D., D.A. O'Handley, and G.M. Yagi (1973), "Computer Determination of Depth Maps," Computer Graphics and Image Processing, (1973), 2, pp. 131-150.

9. Lillestrand, R.L. (1972), "Techniques of Change Detection," IEEE Transactions on Computers, Vol. C-21, No. 7, July 1972.

10. Ohlander, R. (1975), "Analysis of Natural Scenes," Thesis (Computer Science), Carnegie-Mellon University, Pittsburgh, Pennsylvania, June 1975.

11. Price, K. (1976), "Change Detection and Analysis in Multi-Spectral Images," Thesis (Computer Science), Carnegie-Mellon University, Pittsburgh, Pennsylvania, December 1976.

12. Quam, L.H. (1971), "Computer Comparison of Pictures, " AIM-144, Thesis (Computer Science), Stanford University, Stanford, California, May 1971.

13. Rosenfeld, A., Picture Processing by Com* puter, Academic Press, New York, 1969.

14. Tenenbaum, J.M., T.D. Garvey, S. Weyl, H.C. Wolf (1974), "An Interactive Facility for Scene Analysis Research," Artificial Intelligence Center Technical Note 87, SRI, Menlo Park, California, January 1974.

15. Yakimovksy, Y. (1973), "Scene Analysis Using a Semantic Base for Region Growing, " AIM-209, Thesis (Computer Science), Stanford University, Stanford, California, July 1973.

Figure la.
First image, house scene.

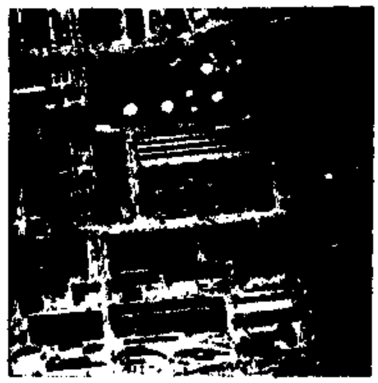

Figure 2a.
First image, urban scene.

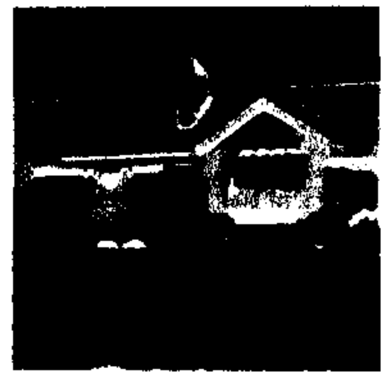

Figure 3a.
Segmentation of first house image. Each separate region is shown at a different intensity. The brightest regions are the areas of the image left unsegmented.

Figure lb.
Second image, house scene.

Figure 2b.
Second image, urban scene.

Figure 3b.
Segmentation of second house image.

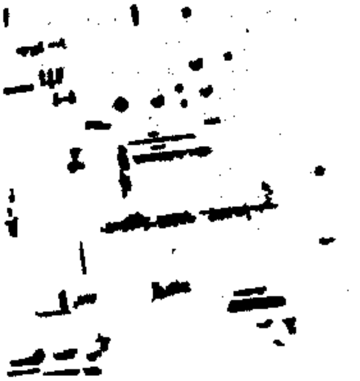

Figure 4a.
Segmentation of first urban image. Only the bright regions are segmented and these are shown as dark areas.

Figure 5a.
Detailed segmentation of pier area of first urban image.

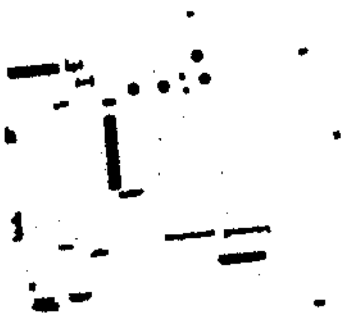

Figure 4b.
Segmentation of second urban image.

Figure 5b.
Detailed segmentation of pier area of second urban image.

Figure 6.
Corresponding regions for the house scene. The first image is on the top and the second image is on the bottom. Corresponding regions are given the same intensity in the two views. These intensities correspond to that used in Figure 3a, the first house image segmentation.

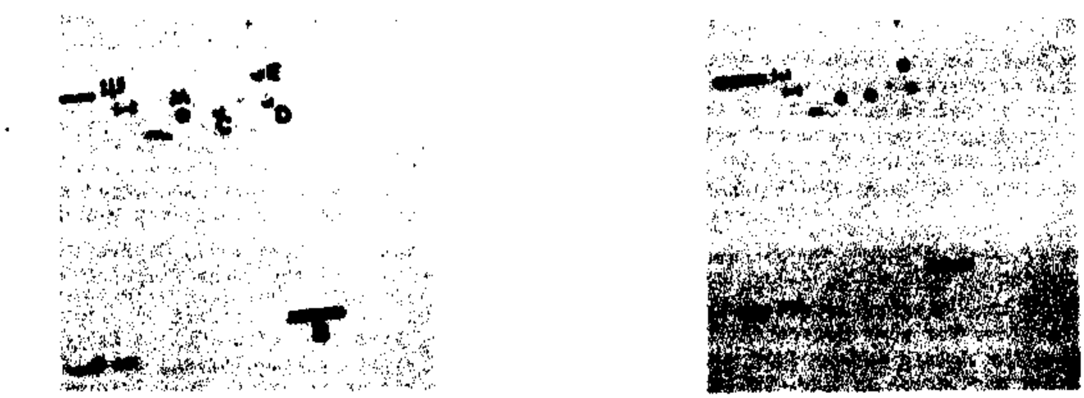

Figure 7.
Corresponding regions for the urban scene. The first image is on the left and the second image is on the right. Each matching pair is shown with a unique intensity.