

Segmentation using intensity and range data

Jay K. Hackett
Mubarak Shah, MEMBER SPIE
University of Central Florida
Computer Science Department
Orlando, Florida 32816

Abstract. Perception is the integration of various sensor outputs into a unified vision of the world. A major part of research in machine vision has been limited to the use of individual sensors for building machine vision systems. The use of multisensor data is advocated for object recognition systems, and an algorithm for fusion of intensity and range data for segmentation is proposed. The algorithm consists of two steps. First, the initial seed segmentation is achieved by using the most dominating sensor at a given time. For this purpose the distributions of the intensity and range data are considered, and the image is segmented recursively by using the most significant peak in both histograms. Second, the initial segmentation is refined by using region merging; the regions are merged if the combined strengths of range and intensity boundaries are low. The experimental results for synthetic and real scenes are presented.

Subject terms: multisensor integration; multisensor fusion; segmentation; range data; region merging.

Optical Engineering 28(6), 667-674 (June 1989).

CONTENTS

1. Motivation
2. Segmentation
 - 2.1. Our approach
 - 2.2. Related work
3. Segmentation using range and intensity
 - 3.1. Seed segmentation
 - 3.2. Region merging
4. Experimental results
5. Summary and conclusions
6. Acknowledgments
7. References

1. MOTIVATION

Perception is the integration of various sensor outputs into a unified vision of the world. We humans have the ability to sense the environment concurrently with several senses. To recover the three-dimensional structure, for example, we use stereo, motion, texture gradients, and shape from contours. Sometimes the information from these sources is comple-

mentary—depth information can be obtained from both the binocular stereo and motion—but the combined information is helpful in resolving any ambiguity in the weaker input. Most research in machine vision has been limited to the use of individual sensors for building machine vision systems. Therefore, few existing systems exploit multiple sensor capabilities. Our aim is to design a machine vision system for object recognition and tracking that uses multiple sensors. This system will be useful for many automated industry and robotics applications.

The advantage in using several sensors is clear from the observation that different sensors are sensitive to different signals, each of which can reveal a particular set of properties of the sensed environment. The video camera gives intensity information about the light reflected from the patches of the surfaces, the laser range finder measures the distance of the objects from the observer, the infrared sensor records the temperature of the objects in the scene, and the tactile sensor gives information about the force with which the object comes into contact with the sensor.

We believe that the fusion of various sensor outputs should be done at three levels: At the *lower* level the raw sensor outputs are fused. If the vantage points for all sensors are the same or if the position of one sensor with respect to

Paper 2577 received May 31, 1988; revised manuscript received Jan. 20, 1989; accepted for publication Jan. 20, 1989.
©1989 Society of Photo-Optical Instrumentation Engineers.

the other sensors is known, the sensor outputs can be considered as a vector measurement, where each scalar component in a vector corresponds to a particular sensor output. With the vector representation, some of the manipulation becomes easier because the developed theory of vector analysis can be used. The second level of fusion is the *intermediate* level. For example, the intensity and range information are processed independently to achieve the depth map. The depth map obtained from the two binocular intensity images will not be accurate, owing to poor stereo algorithms. Similarly, the range finder may not give accurate information owing to sensor noise and quantization effects. We fuse two depth maps in a suitable manner to achieve a better assessment of the depth. The final level of fusion is the *cognitive* level, where fusion is done between the symbols. Each sensor is treated as an expert in its modality. Given a scene, each expert goes through the processing of its information independently or with the cooperation of one or two other sensors at lower and intermediate levels. Each sensor expert makes a hypothesis about the presence of a particular object and its geometrical characteristics, orientation, motion, and structure. The final decision about a given hypothesis is achieved by the consent of all of the experts, and the hypothesis is either rejected or accepted.

This paper deals with *lower level fusion*. We present an algorithm for segmenting scenes that uses multiple sensors as inputs. Although we show our results with range and reflectance data, our approach is more general and will be useful for a variety of sensors.

2. SEGMENTATION

A basic goal of low level computer vision research is object segmentation. Ideally, segmentation is used to extract from an image those regions or edges that correspond to separate objects in a scene. Some of the popular segmentation methods are edge detection, region growing, region clustering, and split-and-merge.

Edge-based segmentation is based on the idea that a set of edges enclosing a region defines that region. This means that we can label that region using a *flood fill* algorithm or some other labeling algorithm. However, the edges surrounding regions are difficult to find accurately and completely since edge linking must be performed to create closed contours.

Region growing is a method of extracting initial regions from an image and merging those regions on the basis of some merging predicate. This predicate varies depending on the type of data being considered. Region clustering¹ is a method by which regions are labeled based on grouping of features in measurement space. One of the widely used clustering schemes is the histogram-threshold method because it requires only one pass of the image data.

Split-and-merge segmentation² is a combination of a splitting scheme and a region-growing scheme. The method depends on performing successive image splitting and merging based on the predicate that the final segmentation can no longer be split or merged to make it a more optimal segmentation.

Much research has been done with the above segmentation techniques. The most popular input images are intensity or range. The intensity image segmentation suffers mainly because of its dependence on scene lighting and upon an object's color variations. For instance, irregular lighting or

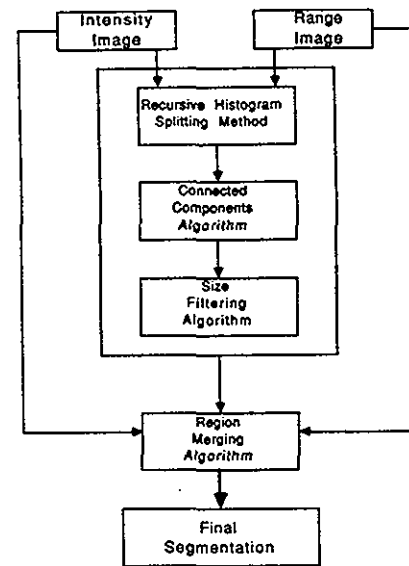


Fig. 1. Summary of our approach.

shadows cast upon an object may cause that object to be segmented into many parts, and an object having different colors may be segmented incorrectly. The range image segmentation segments incorrectly in the following instances: If an object has a sloping surface, the sloping surface has varying range values and may be segmented into many regions. Two objects that meet at a common boundary such as a box sitting on the floor or a lamp on a table may be segmented as the same object.

2.1. Our approach

The segmentation method examined in this paper performs an intensity and range segmentation without the inadequacies of intensity segmentation or range segmentation alone. We feel that intensity or range data alone are not sufficient to produce a good segmentation. An intuitive feeling is that a better segmentation would result if both forms of data could be combined. The method presented here recursively uses the *dominant*, or best, data from the intensity or range data to produce a seed segmentation, then uses both forms of data again to refine the segmentation by using region-merging techniques. Figure 1 is a diagram summarizing our approach. The figure shows that both an intensity and a range image are the inputs, and a final segmentation is the output. First, both data sources are fused by a histogram-splitting method, which we call a seed segmentation method. Second, a connected components algorithm and a size-filtering algorithm are applied to prepare the seed segmentation for further processing. Then a region-merging algorithm is applied that uses the original intensity and range data as well as the seed segmentation image. In this way, the region-merging algorithm does not start over but continues to improve the segmentation.

2.2. Related work

Duda, Nitzan, and Barrett³ deal with eliminating the projective distortion in the intensity data of a planar surface to obtain the normal view of the surface. They first convert the

range data into Cartesian coordinates by multiplication of translation, rotation, and projection transformation matrices. Next, they find a set of pixels in range data whose depth value falls within some tolerance, uniformly sample the x,y Cartesian plane to find the image coordinate j,i values corresponding to $x-y$ values, and determine the intensity values at x,y locations by interpolating among the four pixels surrounding j,i . They claim that their method may be applicable to recognition of skewed patterns, such as characters and labels on boxes that are positioned and oriented randomly.

Hebert⁴ proposes the use of data fusion to combine color data and range data. However, no research was completed on the actual fusion aspect; he only describes a method for registration of range and color data.

Gil, Mitiche, and Aggarwal⁵ combine range and intensity edge maps by using the *global and* operator to locate edge boundaries common to both sources of data. They find that a pixel-to-pixel *and* of the edge map is not possible since different algorithms are used to compute each edge map and thus the two maps are not registered. Therefore, they prefer the *global* operation to trace and match edge segments instead of edge points. Their method of segmentation fails when a gap is present between edge pixels in the edge map. The gap will cause two apparently different regions to be the same region since the regions are connected.

Zuk et al.⁶ use a range and a reflectance sensor. In their method, images of road scenes are segmented so that navigation of the road can occur. The fact that roads tend to have a smoother texture than the surrounding environment is used to separate road from nonroad regions. Range pixels are neglected when the corresponding reflectance value is very small. Small reflectance values usually indicate that the range data cannot be trusted. The texture is then measured on each scan line of the revised range image. Thresholds are chosen that separate weak from strong textures, and then another algorithm is used to surround weak texture areas by a boundary. This in effect separates or segments the image. Navigation occurs by crossing only the smooth texture regions that correspond to the road.

Nandhakumar and Aggarwal⁷ use thermal and intensity data to estimate surface heat fluxes that can be used in segmentation of outdoor scenes. The infrared value obtained at the sensor is dependent upon an object's surface temperature as well as the surface emissivity, which is shown to be fairly constant for outdoor scene objects. Objects that are present in normal outdoor scenes are examined for conductive heat fluxes that emanate from their surfaces. These heat fluxes are estimates of average heat flux for a type of object. The intensity image is then used to compute the solar absorptivity and the difference in angle between the projection of the surface normal to the heat source. A calibration is performed on the camera to yield this surface normal angle for each region (pixel) in the intensity image. The solar absorptivity and surface normal angles are used to compute the amount of heat absorbed by the surface of an object. The ratio of heat conducted from the surface into the object to the heat absorbed by the surface is computed for each pixel. It is shown that this value varies widely for objects in outdoor scenes. This ratio is used to directly *segment* the scene. Errors occur where shadows occur in the scene because very little heat emanates from objects in a shadow.

Therefore, the method's accuracy for indoor scenes, as well as for night vision, must be questioned.

Hu and Stockman⁸ fuse sparse 3-D information obtained from a light striping projector and intensity data. Contour types are defined, by considering properties of both the 3-D data and the intensity data, that identify all of the possible image contours. The fusion is actually built into the contours since they are defined only in terms of the sensor data that provide the contour type. Rules are triggered based upon the type of surface found at a contour in the image. *The triggering of rules is simply a segmentation or labeling process.* When a rule is invoked, the contour being examined is assigned the rule name. A relaxation procedure is used to select the correct name for the contour in case a conflict arises in assigning rule names. Since sparse 3-D data are used, it is expected that many errors will occur in selecting the correct rule.

3. SEGMENTATION USING RANGE AND INTENSITY

In this section, we discuss our algorithm for fusion of range and intensity images in order to segment a scene. We assume that we are given a discrete vector field $f(x,y) = [R(x,y), I(x,y)]$, where R and I represent the range and intensity images, respectively, and we also assume that these two images are registered with each other. In our experiments so far we have used range and reflectance data obtained from the same sensor. However, one can also use the intensity data obtained from a video camera for the same scene. There are easy calibration methods for registering the video data with the range image, e.g., see Ref. 4. Our algorithm consists of two steps. During the first step, the initial seed segmentation is created using global information from the dominating sensor. During the second step, this crude segmentation is refined by the use of local information from both sensors.

3.1. Seed segmentation

Because the raw sensor data are noisy, imprecise, and unreliable, one cannot rely fully on any one particular sensor's data. Also, the combination of all sensor output at this lowest level may become more complex and require complete information about the sensors and underlying noise models. We believe that, at a given instant, one sensor's data may be more useful than the other's.

Our seed segmentation algorithm looks at the pool of information from both sensors and decides which is dominating or more reliable. First, two histograms are created, one for intensity and one for range. The histograms are then smoothed. Next, the most significant peak from each histogram is located through the use of a peakiness test. A peakiness test is some rule for locating the most significant peak in a histogram. A good peakiness test will remove the effects of small peaks caused by noise in the original images. The more significant of the two peaks and their corresponding valley points are used to cut the histograms and segment its corresponding images. This process is repeated until there are no more significant peaks. When a point in an image is segmented, it is no longer considered in the future histograms.

The peakiness test decides which sensor is the dominating sensor by choosing to use the data that provide the largest peakiness value at a particular time. Thus, as the segmenta-

tion continues, the peakiness test may in effect switch from using one sensor's data to using data from the other sensor as many times as necessary. The seed segmentation image is created by using the best pool of information from both sensors.

This approach to segmentation employs recursive region splitting and is similar to the recursive histogram-splitting method by Ohlander et al.⁹ However, they used only color camera data. The seed segmentation achieved by using their method is not of good quality when used with intensity and range data. Thus, we refine the segmentation by using region merging.

We smooth both histograms with a Gaussian filter to reduce the effects of simple noise. We have used a Gaussian with a large one-dimensional mask size so that a fair amount of smoothing is performed. The size of the mask can be changed to control the amount of noise that is present in the original images. Our initial peakiness test was that the tallest peak in a histogram is the most significant. This test performed fairly well on synthetic intensity and range images but failed when noise was added to both images. The major drawback of this initial test was that it did not consider the number of pixels in the peak; i.e., a short, wide peak would not be more significant than a very narrow, tall peak, even though the short peak might contain more pixels. We needed to find a peakiness test that would consider both the width and the height of the peaks. To capture the width of a peak, we use the ratio of area under the peak and the area of the enclosing rectangle of a peak. The product of the height and width of a peak is used as the final term in the peakiness test.

Since our algorithm is recursive, it continues to segment until there are no more significant peaks. Thus, we needed a peakiness test threshold that would ignore insignificant peaks. These insignificant peaks, if used, would result in a segmentation having many small regions, which for the most part are not significant for our purposes. A threshold of 10% of maximum peakiness was chosen since the resulting segmented image most closely reflected the actual number of objects present in the scene.

By examination of the histograms constructed, we decided that processing of peaks that occurred in the zero frequency area in both the intensity and range images could be ignored because the peaks at zero correspond to the background in both types of images. If intensity has a peak at zero, then the peak corresponds either to background or to an object that is too poorly lit to be seen. If the range image has a peak at zero, then the peak corresponds either to background or to a very distant object. In either case, this information does not need to be considered.

Once the initial seed segmentation is created using the above algorithm, it is improved upon. Preliminary processing includes assigning unique labels to every region by using a four-connected component algorithm.¹⁰ A size-filtering algorithm is then applied to remove any regions of area less than 20 pixels. This helps to remove small regions that were caused by noise in the original images as well as to remove holes in regions.

3.2. Region merging

In Sec. 3.1, we presented a recursive method for initial segmentation of images using intensity and range data. This

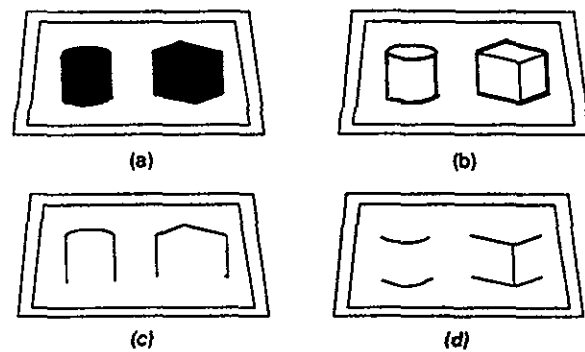


Fig. 2. Boundaries in range and intensity images. (a) Original scene; (b) Intensity (gradient) boundaries; (c) jump boundaries; (d) roof boundaries.

method is considered a *global* method. The drawback is that it considers only one source, either intensity or range, for a particular region at one time. This is because only the most significant peak in the histograms is being chosen. A method is needed to incorporate some *local* pixel information from the original intensity and range images. Our approach is to merge regions based on weak boundaries between regions. The idea is to iteratively merge the regions in the image until no common regional boundaries are weak. Weakness is defined to be large when a small difference is present between pixels that lie across a regional boundary.

Previously, this method has been shown to work for intensity images.² However, we are dealing with both intensity and range images, and we need to somehow incorporate both intensity and range information into our merging technique. We are using a gradient magnitude $S_1 = (I_x^2 + I_y^2)^{1/2}$ along region boundaries to determine how weak a boundary is in the intensity image.

In the range image one can identify at least two types of boundaries: *jump* and *roof* boundaries. Jump boundaries are range discontinuities and roof boundaries are the edges of a cross-line between two plane regions. Figure 2 shows a scene of a bottle and a box on a tabletop and its corresponding gradient, jump boundary, and roof boundary images. Although this is an ideal situation, it does point out that the combination of jump and roof boundaries produces the same edges as does the intensity gradient image.

We also propose that a form of jump boundaries will work satisfactorily for range images. Jump boundaries³ are present where both first and second derivatives are sufficiently large in range images. By using this method we can create a binary jump boundary image. However, for our purposes we need a quantitative measure of edge strength. We have chosen to sum both the first and second derivatives to produce a quantitative edge strength image since if both derivatives are large, their sum will be large. Therefore, edge strength in a range image for a jump boundary is given by $S_R = |R_x| + |R_y| + |R_{xx}| + |R_{yy}|$.

The two strengths of a boundary in range and intensity images are combined as

$$S_T = \omega_1 S_R + \omega_2 S_1, \quad (1)$$

where ω_1 and ω_2 are predetermined constants such that $\omega_1 + \omega_2 = 1$. It is obvious from the equation that we are us-

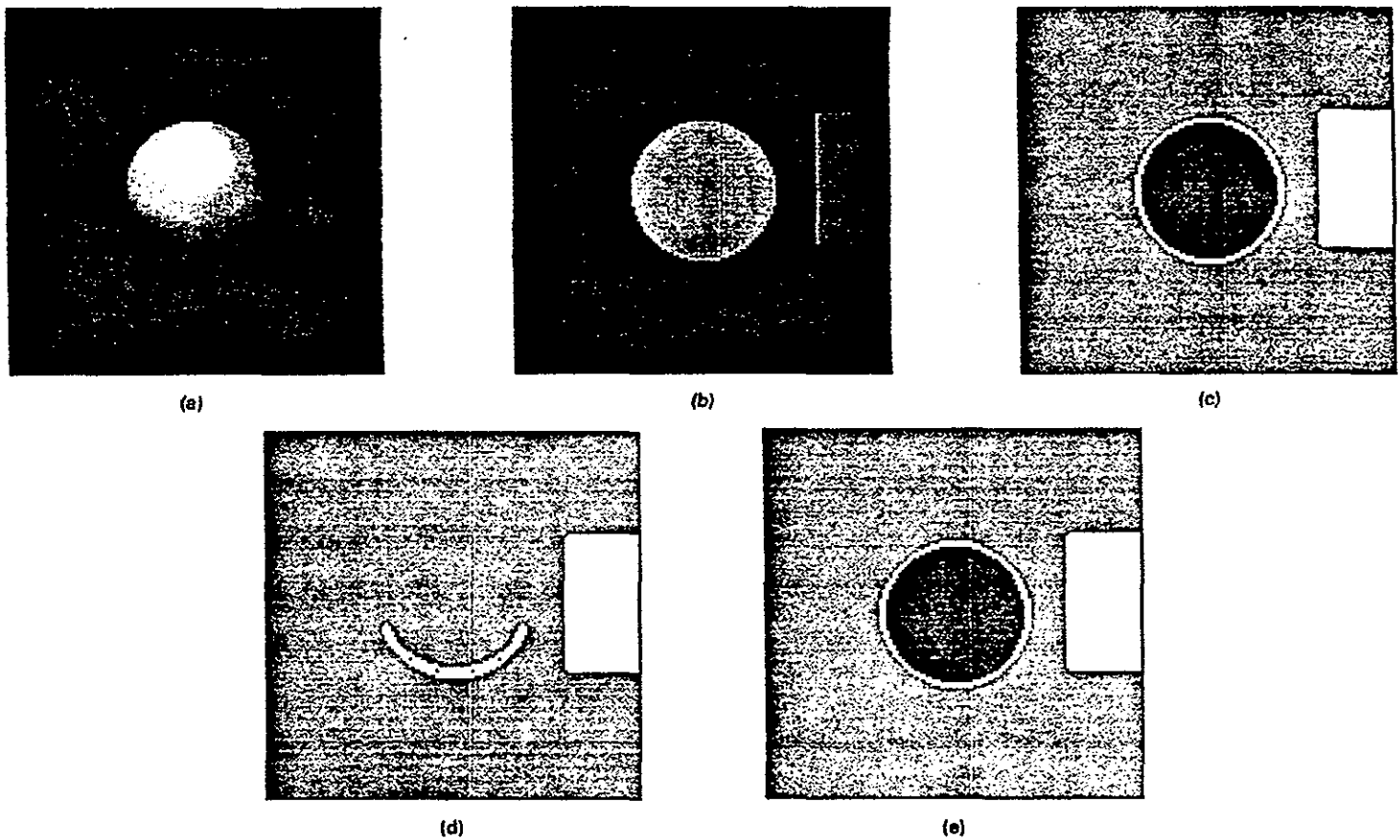


Fig. 3. Results of seed segmentation for "sphere and rectangle" synthetic scene. (a) Intensity image; (b) range image; (c) segmented image using range and intensity data; (d) segmented image using intensity only; (e) segmented image using range only.

ing *dual data* in computing the total strength of region boundaries. However, dual data have also been used in the seed segmentation process, since histograms of both sensors are considered so that the best data at a particular time are used.

It may appear that we are neglecting the seed segmentation, but in truth we are using it to provide a preliminary segmentation so that region boundaries can be located (see Fig. 1). Thus, we compute edge strengths only for these boundaries.

We determine when regions should be merged by combining average boundary strength information along with the sizes and boundary lengths of adjacent regions. Regions are merged if the boundary, on the average, is weak or if both regions have small areas and large perimeters. The method is defined mathematically as follows:

MERGE if

$$\frac{S_T}{N} \leq T_1$$

or

$$\frac{\sqrt{A(R_1)} + \sqrt{A(R_2)}}{P(R_1)P(R_2)} \leq T_2,$$

where R_1 and R_2 are regions under consideration for merging, S_T is total strength along the boundary between two

regions, N is the number of pixels along a common boundary, and $P(R_i)$ and $A(R_i)$ are respectively the perimeter and area of region R_i .

The first predicate is data driven and the second is a heuristic. The second predicate is applied only if the first predicate fails. The reason for using both a data-driven and a heuristic criterion for region merging is that the data-driven criterion is not sufficient in all cases to merge long, narrow regions. We have chosen not to use roof boundaries in calculating the range edge strength since these boundaries are much more difficult to compute. Thus, we have chosen to add the heuristic criterion that long narrow regions be merged since the absence of roof boundaries will tend to increase the merging performed near areas where roof boundaries occur. At the same time we have adjusted our merging threshold T_1 so that less merging is done by the first predicate. The average strength along a boundary is a good measure of how strong an edge is. It is also nearly oblivious to noise in both the intensity and range images since the effects of noise will be spread out over the entire length of the boundary.

4. EXPERIMENTAL RESULTS

In this section, we demonstrate the segmentation results obtained with our method and attempt to show that the segmentation obtained is better than that achieved by intensity or range data alone. Figures 3(a) and 3(b) are synthetic intensity and range images of a sphere and a rectangle

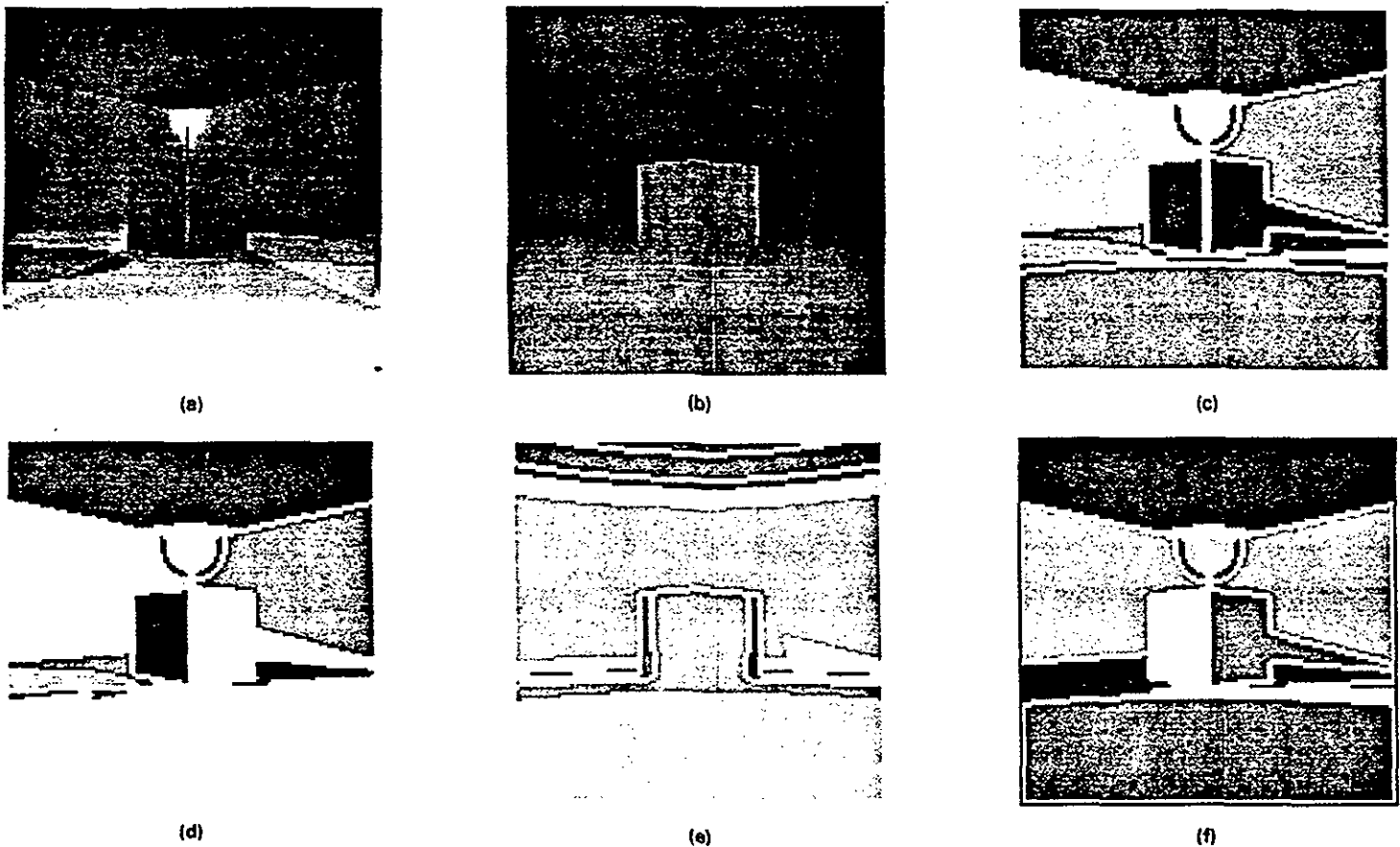


Fig. 4. Results of segmentation for "box in a room" synthetic scene. (a) Intensity image; (b) range image; (c) segmented image using range and intensity data; (d) segmented image using intensity only; (e) segmented image using range only; (f) segmented image after region merging.

suspended in space. The rectangle slopes from right to left, and both objects are the same color. The only light source is located at the camera viewpoint. Figure 3(c) shows the seed segmentation results obtained with our method. The segmentation is correct even though both objects are the same color and each is a different distance from the sensors. Figure 3(d) shows the results of segmentation by using intensity data only; the results show that intensity information is not sufficient to segment this scene using ordinary histogram segmentation. Figure 3(e) shows the results obtained by using range data only. These results are the same as for our method. The reason is that there are no features in this scene that can create problems for range segmentation, as described in Sec. 2.

Figures 4(a) and 4(b) are the synthetic intensity and range images of a box in a room. The scene is lit by two light sources, one opposite the camera viewpoint, i.e., just above the box, and the other at the camera viewpoint. The camera is at one corner of the room and is pointed toward the opposite corner. Note that the box is not a true box but is created by placing two planes next to each other with a 1 pixel gap between them. Thus, the box is actually two regions, as will be shown in the segmentation results. Figure 4(c) is the seed segmentation produced from our method. An accurate segmentation is produced; that is, the floor, ceiling, walls, and "box" are all segmented correctly. Figure 4(d) shows results from intensity data only. The results are

fairly accurate, but problems are evident where the wall meets the floor. Figure 4(e) shows results for range data only. As expected, these results contain many errors: the ceiling is segmented into many regions, the box is segmented with the floor, and the sloping surfaces of the box as seen by the sensor are broken up into many different regions. Figure 4(f) improves upon the seed segmentation of Fig. 4(c). Several of the regions on the floor are merged together. These regions were the result of shadows in the intensity image and are also located where the back wall meets the floor. Also, the box was left untouched as two regions.

Figures 5(a) and 5(b) are real images of a person standing in an office. The office is lit by many overhead fluorescent lights. A desk and a blackboard are present. Figure 5(a) is not an intensity image but a reflectance image obtained from the range sensor as the amplitude of the returning laser light. We feel, however, that it should complement the range data just as intensity data would. Figure 5(c) shows the results obtained with our method. It is surprising that the person is segmented as one region. We were also pleased that the segmentation almost always set aside one and only one region for each object in the scene. Any errors that did occur we hope will be fixed in our region-merging method. Figure 5(d) shows that intensity segmentation is very bad. Many of the objects were segmented with the background, perhaps because of their similar color or because of lighting conditions. Figure 5(e) demonstrates segmentation by use of

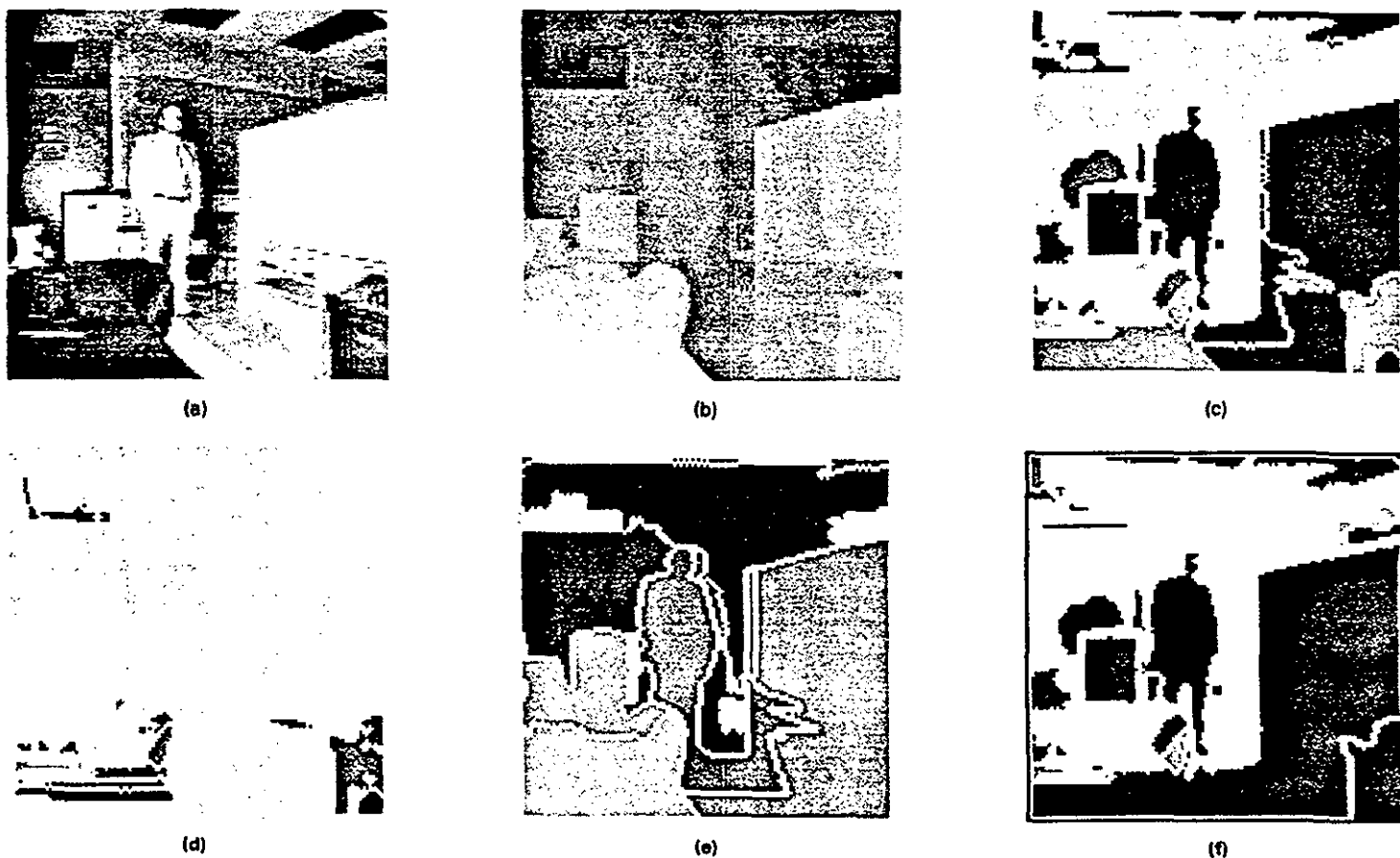


Fig. 5. Results of segmentation for "person in office" real scene. (a) Intensity image; (b) range image; (c) segmented image using range and intensity data; (d) segmented image using intensity only; (e) segmented image using range only; (f) segmented image after region merging.

range data only. At first it appears that the segmentation is good because it is easy to recognize the person in the scene. However, the person is segmented with the floor, as is usually the case when segmentation by range is used. Many of the other objects are segmented together as well. For example, the objects on the desk are missing and the fluorescent lights are missing, as expected. Figure 5(f) shows the region merged image of Fig. 5(c), which is the segmentation of the person standing in an office. Many of the fluorescent lights were merged into the background. The blackboard on the right side of the image originally had two regions, but it is now merged into one region. Also note that the person was not merged into any of the other regions.

Figures 6(a) and 6(b) are also real images, depicting a hallway with a man at the end. Note that Fig. 6(a) is not an intensity image but a reflectance image. Figures 6(c), 6(d), and 6(e) are similar to the corresponding results of Fig. 5. Figure 6(f) demonstrates the results obtained from the long hallway scene of Fig. 6(c). Many of the narrow wall and ceiling regions were merged together, and the person was again left untouched. We believe the results obtained with these images are typical of those to be obtained from most synthetic and real images.

5. SUMMARY AND CONCLUSIONS

We have discussed our approach toward multisensor fusion at the lower level. We assumed that the sensor outputs are registered with each other and are given in the form of a

discrete vector field. We considered the problem of segmentation. Although we demonstrated our results by using range and intensity data, our approach is more general and will be useful for a variety of sensors. Our algorithm has two stages. First, initial segmentation is achieved by using the most dominating sensor at a time in a recursive fashion. Second, the initial segmentation is refined by using the region-merging method, where each sensor is given an equal opportunity to contribute toward a merging decision. Future study will relate to the use of other sensors besides range and intensity and to fusion at intermediate and cognitive levels.

6. ACKNOWLEDGMENTS

This research was supported by the Center for Research in Electro Optics and Lasers (CREOL), of the University of Central Florida under grant 20-52-043 and by the National Science Foundation under grant IRI 87-13120. The authors thank Paul Besl of General Motors for providing the image pairs.

7. REFERENCES

1. R. M. Haralick and L. G. Shapiro, "Survey: image segmentation techniques," *Comput. Graph. Image Proc.* 29, 100-132 (1985).
2. T. Pavlidis, *Structural Pattern Recognition*, Springer-Verlag, New York (1977).
3. R. O. Duda, D. Nitzan, and P. Barrett, "Use of range and reflectance data to find planar surface regions," *IEEE Pat. Anal. Machine Intell. PAMI-1*, 259-271 (1979).

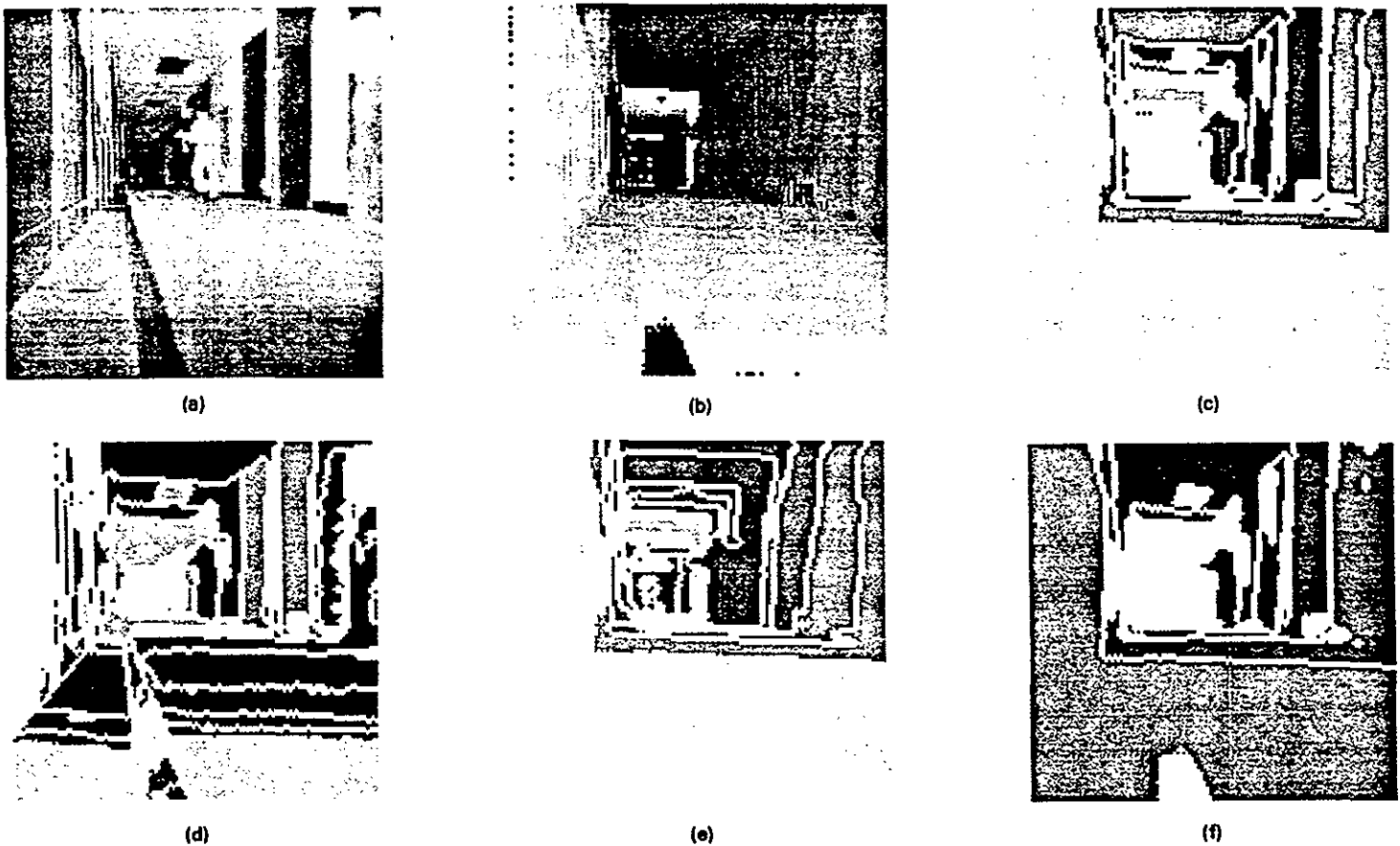


Fig. 6. Results of seed segmentation for "hallway" real scene. (a) intensity image; (b) range image; (c) segmented image using range and intensity data; (d) segmented image using intensity only; (e) segmented image using range only; (f) segmented image after region merging.

4. M. Hebert, "Outdoor scene analysis using range data," in *IEEE Conf. on Robotics and Automation* 3, 1426-1432 (1986).
5. B. Gil, A. Mitiche, and J. Aggarwal, "Experiments in combining intensity and range edge maps," *Comput. Graph. Image Proc.* 21, 395-411 (1983).
6. D. Zuk, F. Pont, R. Franklin, and M. Dell'Eva, "A system for autonomous land navigation," *Tech. Rep.*, Environmental Research Inst. Michigan, Ann Arbor (Nov. 5, 1985).
7. N. Nandhakumar and J. K. Aggarwal, "Multisensor integration — experiments in integrating thermal and visual sensors," *IEEE Int. Conf. on Computer Vision*, 83-92 (1987).
8. G. Hu and G. Stockman, "3-D scene analysis via fusion of light striped image and intensity image," in *Proc. 1987 Workshop on Spatial Reasoning and Multi-Sensor Fusion*, Morgan Kaufmann Publishers, Inc., Los Altos, Calif. 138-147 (1987).
9. R. Ohlander, K. Price, and D. Raj Reddy, "Picture segmentation using a recursive region splitting method," *Comput. Graph. Image Proc.* 8, 313-333 (1978).
10. D. H. Ballard and C. M. Brown, *Computer Vision*, Prentice-Hall, Englewood Cliffs, N.J. (1982).



Jay Hackett received a BS degree in computer science from the University of Central Florida in 1987. He is currently a research assistant and is pursuing his MS degree in computer science at the University of Central Florida. His interests include image processing, computer vision, computer graphics, and mathematics. He is a student member of IEEE Computer Society.



Mubarak Shah was born in 1954 in Sind, Pakistan. He completed his BE in electronics from National College of Engineering, Karachi, Pakistan, in 1979. In 1980 he spent one year with Philips International Institute at Eindhoven, The Netherlands, where he obtained his electronic design engineering diploma. Dr. Shah completed his MS and Ph.D. degrees in computer engineering at Wayne State University, Detroit, in 1982 and 1986, respectively.

Since 1988 he has been with the Computer Science Department of the University of Central Florida. His research interests in computer vision are edge detection, scale space, dynamic scene analysis, multisensor fusion, and hardware implementation of vision algorithms. He has published in major conferences and journals.

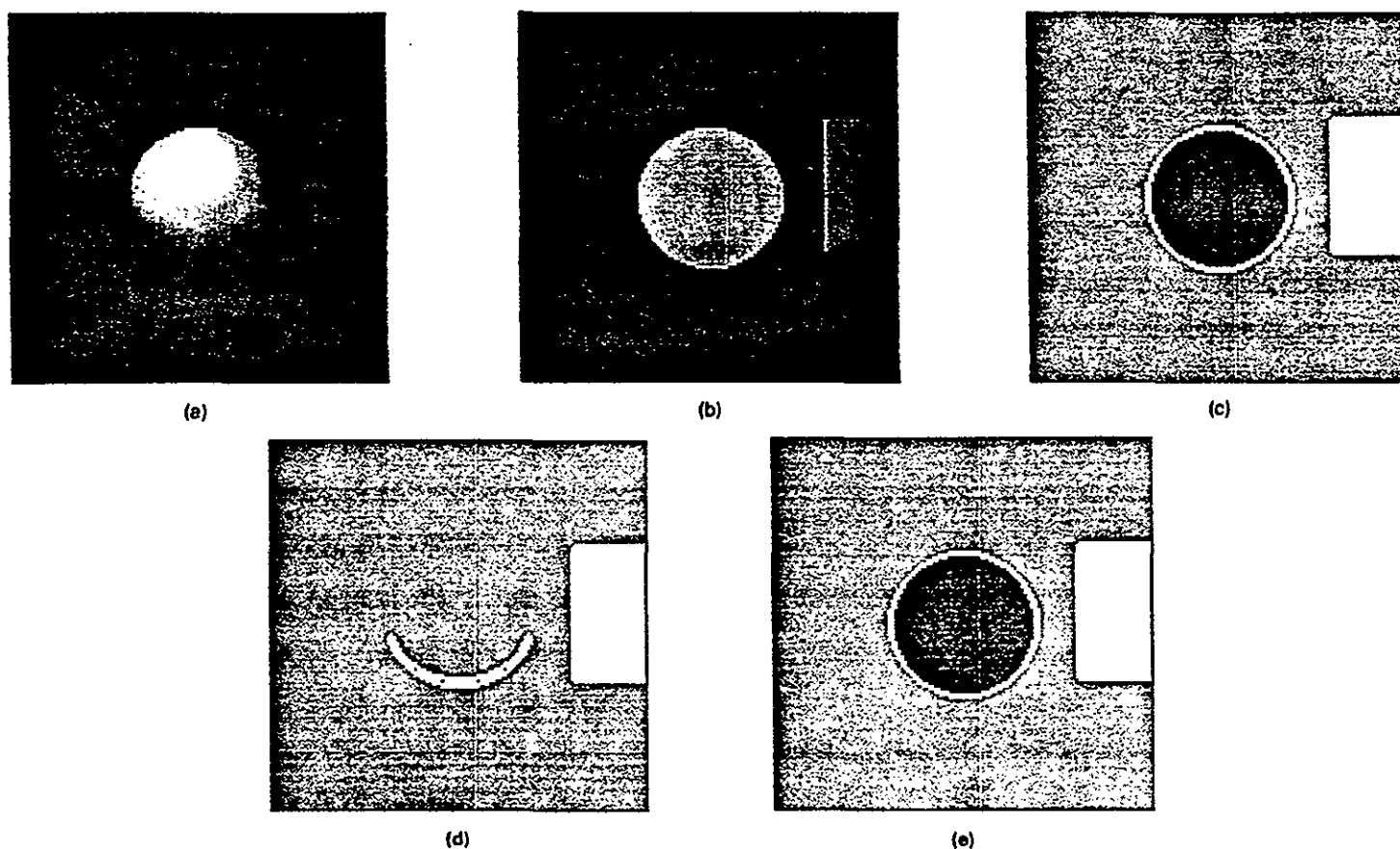


Fig. 3. Results of seed segmentation for "sphere and rectangle" synthetic scene. (a) Intensity image; (b) range image; (c) segmented image using range and intensity data; (d) segmented image using intensity only; (e) segmented image using range only.

ing *dual data* in computing the total strength of region boundaries. However, dual data have also been used in the seed segmentation process, since histograms of both sensors are considered so that the best data at a particular time are used.

It may appear that we are neglecting the seed segmentation, but in truth we are using it to provide a preliminary segmentation so that region boundaries can be located (see Fig. 1). Thus, we compute edge strengths only for these boundaries.

We determine when regions should be merged by combining average boundary strength information along with the sizes and boundary lengths of adjacent regions. Regions are merged if the boundary, on the average, is weak or if both regions have small areas and large perimeters. The method is defined mathematically as follows:

MERGE if

$$\frac{S_T}{N} \leq T_1$$

or

$$\frac{\sqrt{A(R_1)} + \sqrt{A(R_2)}}{P(R_1)P(R_2)} \leq T_2,$$

where R_1 and R_2 are regions under consideration for merging, S_T is total strength along the boundary between two

regions, N is the number of pixels along a common boundary, and $P(R_i)$ and $A(R_i)$ are respectively the perimeter and area of region R_i .

The first predicate is data driven and the second is a heuristic. The second predicate is applied only if the first predicate fails. The reason for using both a data-driven and a heuristic criterion for region merging is that the data-driven criterion is not sufficient in all cases to merge long, narrow regions. We have chosen not to use roof boundaries in calculating the range edge strength since these boundaries are much more difficult to compute. Thus, we have chosen to add the heuristic criterion that long narrow regions be merged since the absence of roof boundaries will tend to increase the merging performed near areas where roof boundaries occur. At the same time we have adjusted our merging threshold T_1 so that less merging is done by the first predicate. The average strength along a boundary is a good measure of how strong an edge is. It is also nearly oblivious to noise in both the intensity and range images since the effects of noise will be spread out over the entire length of the boundary.

4. EXPERIMENTAL RESULTS

In this section, we demonstrate the segmentation results obtained with our method and attempt to show that the segmentation obtained is better than that achieved by intensity or range data alone. Figures 3(a) and 3(b) are synthetic intensity and range images of a sphere and a rectangle

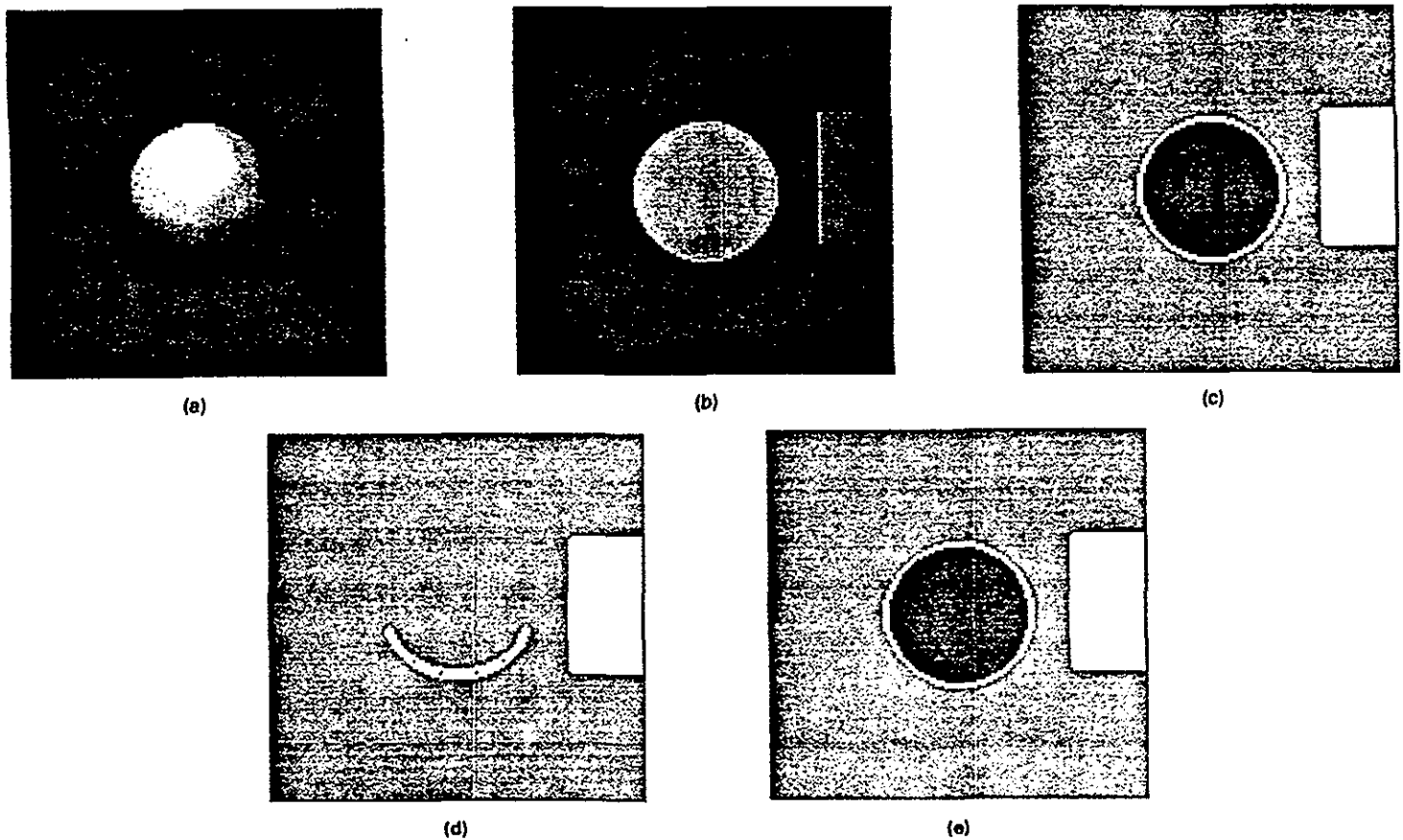


Fig. 3. Results of seed segmentation for "sphere and rectangle" synthetic scene. (a) intensity image; (b) range image; (c) segmented image using range and intensity data; (d) segmented image using intensity only; (e) segmented image using range only.

ing *dual data* in computing the total strength of region boundaries. However, dual data have also been used in the seed segmentation process, since histograms of both sensors are considered so that the best data at a particular time are used.

It may appear that we are neglecting the seed segmentation, but in truth we are using it to provide a preliminary segmentation so that region boundaries can be located (see Fig. 1). Thus, we compute edge strengths only for these boundaries.

We determine when regions should be merged by combining average boundary strength information along with the sizes and boundary lengths of adjacent regions. Regions are merged if the boundary, on the average, is weak or if both regions have small areas and large perimeters. The method is defined mathematically as follows:

MERGE if

$$\frac{S_T}{N} \leq T_1$$

or

$$\frac{\sqrt{A(R_1)} + \sqrt{A(R_2)}}{P(R_1)P(R_2)} \leq T_2,$$

where R_1 and R_2 are regions under consideration for merging, S_T is total strength along the boundary between two

regions, N is the number of pixels along a common boundary, and $P(R_i)$ and $A(R_i)$ are respectively the perimeter and area of region R_i .

The first predicate is data driven and the second is a heuristic. The second predicate is applied only if the first predicate fails. The reason for using both a data-driven and a heuristic criterion for region merging is that the data-driven criterion is not sufficient in all cases to merge long, narrow regions. We have chosen not to use roof boundaries in calculating the range edge strength since these boundaries are much more difficult to compute. Thus, we have chosen to add the heuristic criterion that long narrow regions be merged since the absence of roof boundaries will tend to increase the merging performed near areas where roof boundaries occur. At the same time we have adjusted our merging threshold T_1 so that less merging is done by the first predicate. The average strength along a boundary is a good measure of how strong an edge is. It is also nearly oblivious to noise in both the intensity and range images since the effects of noise will be spread out over the entire length of the boundary.

4. EXPERIMENTAL RESULTS

In this section, we demonstrate the segmentation results obtained with our method and attempt to show that the segmentation obtained is better than that achieved by intensity or range data alone. Figures 3(a) and 3(b) are synthetic intensity and range images of a sphere and a rectangle

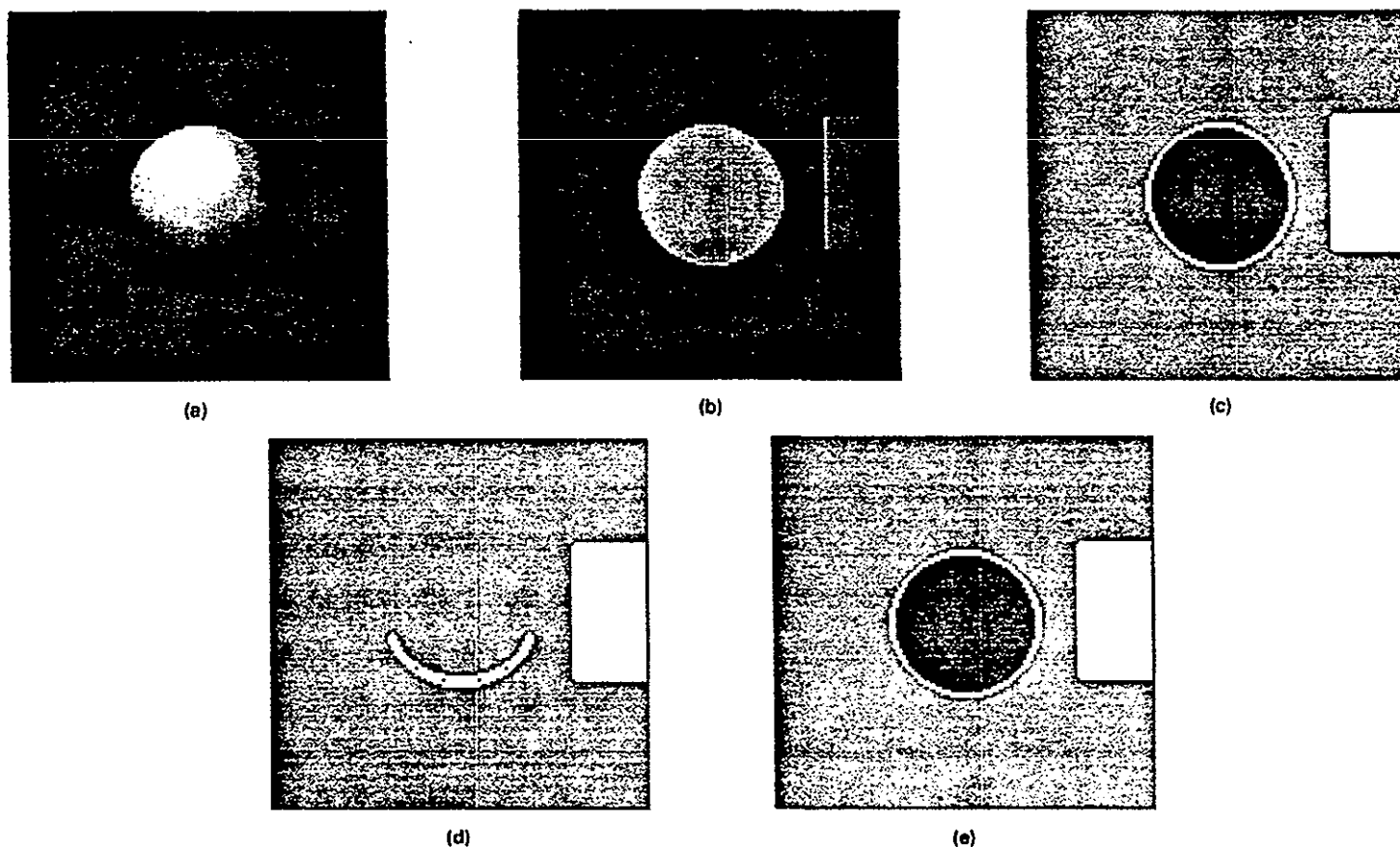


Fig. 3. Results of seed segmentation for "sphere and rectangle" synthetic scene. (a) Intensity image; (b) range image; (c) segmented image using range and intensity data; (d) segmented image using intensity only; (e) segmented image using range only.

ing *dual data* in computing the total strength of region boundaries. However, dual data have also been used in the seed segmentation process, since histograms of both sensors are considered so that the best data at a particular time are used.

It may appear that we are neglecting the seed segmentation, but in truth we are using it to provide a preliminary segmentation so that region boundaries can be located (see Fig. 1). Thus, we compute edge strengths only for these boundaries.

We determine when regions should be merged by combining average boundary strength information along with the sizes and boundary lengths of adjacent regions. Regions are merged if the boundary, on the average, is weak or if both regions have small areas and large perimeters. The method is defined mathematically as follows:

MERGE if

$$\frac{S_T}{N} \leq T_1$$

or

$$\frac{\sqrt{A(R_1)} + \sqrt{A(R_2)}}{P(R_1)P(R_2)} \leq T_2,$$

where R_1 and R_2 are regions under consideration for merging, S_T is total strength along the boundary between two

regions, N is the number of pixels along a common boundary, and $P(R_i)$ and $A(R_i)$ are respectively the perimeter and area of region R_i .

The first predicate is data driven and the second is a heuristic. The second predicate is applied only if the first predicate fails. The reason for using both a data-driven and a heuristic criterion for region merging is that the data-driven criterion is not sufficient in all cases to merge long, narrow regions. We have chosen not to use roof boundaries in calculating the range edge strength since these boundaries are much more difficult to compute. Thus, we have chosen to add the heuristic criterion that long narrow regions be merged since the absence of roof boundaries will tend to increase the merging performed near areas where roof boundaries occur. At the same time we have adjusted our merging threshold T_1 so that less merging is done by the first predicate. The average strength along a boundary is a good measure of how strong an edge is. It is also nearly oblivious to noise in both the intensity and range images since the effects of noise will be spread out over the entire length of the boundary.

4. EXPERIMENTAL RESULTS

In this section, we demonstrate the segmentation results obtained with our method and attempt to show that the segmentation obtained is better than that achieved by intensity or range data alone. Figures 3(a) and 3(b) are synthetic intensity and range images of a sphere and a rectangle