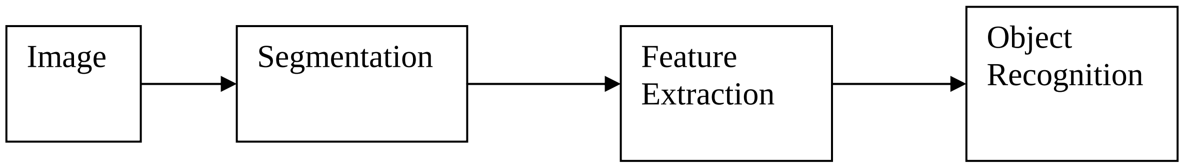
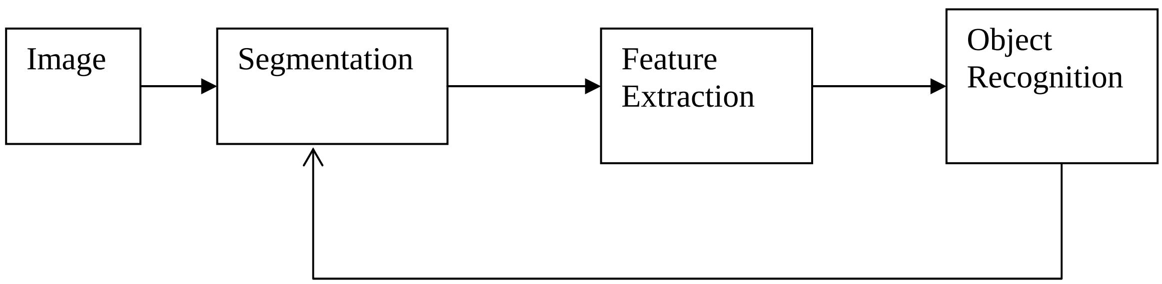
1. Image Segmentation

Real-world vision systems must deal with complicated and cluttered environments. Computer vision is a computational activity that involves constructing representations of an image at successive levels of abstraction. One starts at the pixel level and terminates at a symatically meaning symbolic description of the image. Image segmentation is the process that partitions the image pixels into nonoverlapping regions such that each region is homogeneous and connected [Bhandarkar and Zhang, 1999]. In addition, the union of adjacent regions is not homogeneous. A segmented image is the highest domain-independent or data-driven abstraction of an image. Semantic or world information is not used. One uses the basic properties of the image data such as uniformity. This implies that the methods may apply to a wide variety of image classes. They will in many cases, not in themselves, give completely satisfactory segmentation of the image. Image processing models which apply to any image class are called general purpose processing models. A segmented image may be the input to a higher vision system that utilizes domain-specific information to further analyze and interpret the image.

Most vision systems are not robust in that they do not allow for the variations that may occur in the input images. Vision system often progress from segmentation, to feature extraction from the regions, to object recognition. This approach may be called an open-loop approach. Errors made in early steps progress through the recognition process. Object recognition is a process of making a sequence of decisions. It may not be possible to determine the affect of a decision until one makes the final decision as to the confidence of object recognition. Also the segmentation and feature extraction methods may be complex and elaborate. To achieve good results the different steps of the object recognition process must interact. A knowledge of the object can aid in the segmentation process. The might be called a closed-loop process in control theory terminology [Peng and Bhanu, 1998].



Open-Loop Object Recognition Process



Closed-Loop Object Recognition Process

Figure 1. Object Recognition Process

Image segmentation refers to a process of dividing the image data into regions whose pixels have some common property such as uniform gray-level. A region is a group of pixels with similar properties. Regions often correspond to or have a link to objects [Jain, Kasturi, and Schunck, 1995, pp. 73]. Regions are in the image-domain and objects are in the world-domain. A complex object such as a person may be subdivided into subparts that correspond to different regions with different properties.

The common property is often a property of region uniformity. The uniformity property may be defined over the pixel attributes such as intensity, color, range, or texture. An simple example of a homogenity property would be that the pixels have a uniform gray-level. Each region in the segmentation needs to be uniform and connected. A uniformity predicate is often used to express the uniformity concept. Let $R \subseteq \Omega$ where R is a region and Ω is the image grid. Predicate P(R) assigns a value of True or False according to the values of g(x,y) in region R. It should have the property that if $R_1 \subseteq R$ then $\mathrm{P(R)=True} \Rightarrow \mathrm{P(R_1)=True}$. A mathematical statement of a set of regions which segment an image follows [Levine, 1985, pp.385, 386; Pavlidis, 1977, pp. 65; Weeks, 1996, pp. 445, Jain, Kasturi, and Schunck, 1995, pp. 76]]. The regions are connected.

$\{R_i\}$ is a segmentation of Ω if

$\Omega = \bigcup_i R_i$, the regions cover Ω;

$R_k \bigcap R_j = \varphi$ the empty set, there is no overlap of the regions;

$P(R_i) = True$, P is a uniformity predicate; and

$P(R_k \bigcup R_j) = False$ when R_k and R_j share a common boundary, the regions are as big as possible and still have the predicate be true.

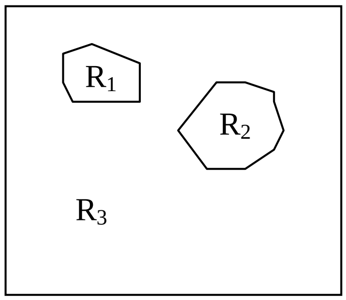


Figure 2. Regions that Partition an Image

Examples of possible of possible predicates are P(R)=True if $|g(x_1, y_1) - g(x_2, y_2)| \le \varepsilon$ for all $(x_1, y_1), (x_2, y_2)$ in R and $P(R) = True$ if $T_1 \le g(x, y) < T_2$ for all (x,y) in R where T_1 and T_2 are thresholds that define the region..

Computing an image partition is a problem of high combinatorial complexity. For this reason an exhaustive enumeration of all possible image partitions is not possible. Most image segmentation methods have been developed for specific image classes of a well defined application. There is no unified approach to image segmentation [Bhandarkar and Zhang, 1999]. One can put the image segmentation problem in the context of a constrained optimization problem. In this situation one defines an objective function in addition to the predicate P above. One must then find the partition $\{R_i\}$ that satisfies the segmentation requirements above for predicate P and also optimizes the objective function. The solution to the search problem is in general difficult.

Over segmentation sometimes occurs as one is attempting to find the best segmentation of an image. Over segmentation refers to the fact that the objects of the scene have been partitioned in the segmentation process. Many insignificant boundaries have been created. In this situation one may use the segmentation as an initial step and then refine it to find a segmentation that better matches the objects in the image.

1.1 Region Representation

Regions are often given different representations [Jain, Kasturi, and Schunck, 1995, pp. 86]. Regions can be represented by closed contours. This indicates a correspondence between boundary (edge) and region determination. One can represent a region with arrays, hierarchical representations, and lists. They can be represented in arrays by attaching a coordinate system to the array and giving a coding to each region. The code in the array gives the pixels in the region. For example, a 1 might represent region one in the array. Hierarchical representation are discussed in the next sections.

There are many different ways to describe a region with lists. The simplest would be to have a list of all the points in the region, or a list of the enclosing boundary points of the region.

The following figures show image data and two regions represented in array form coded with 0 and 1 for the two regions. The image function is denoted g(x,y) and the region membership function is denoted mem(x,y). The region could also be represented with a list as Region One = { (2,4), (3,4), (4,4), (2,3), (3,3), (4,3), (3,2), (4,2) }. This is the region corresponding to the 1 code in the array representation. This list is not a very efficient representation. There are others that are more efficient.

7	0	0	0	0	0	0	0	0
6	1	0	0	0	0	0	0	0
5	0	0	0	0	0	0	1	0
4	0	0	7	8	8	0	0	0
y 3	0	0	7	8	7	0	0	0
2	0	0	0	7	8	0	0	0
1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	2	0
	0	1	2	3 x	4	5	6	7

g(x,y)

Figure 3. Image Data

7	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0
4	0	0	1	1	1	0	0	0
y 3	0	0	1	1	1	0	0	0
2	0	0	0	1	1	0	0	0
1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
	0	1	2	3 x	4	5	6	7

mem(x,y)

Figure 4. Array Region Representation

1.2 Segmentation Strategies.

A number of different strategies have been developed for image segmentation. A number of the methods use region properties [Sonka, Hlavac, and Boyle, 1993, pp. 166]. One approach is region merging. The approach is outlined as follows:

1. Define a starting method to segment the image into small atomic regions satisfying the basic segmentation criterion using a suitable uniformity predicate.
2. Define a method for merging adjacent regions.
3. Merge adjacent regions which satisfy the merging conditions and the basic segmentation criterion. If no regions can be merged then stop

Another approach uses region splitting. A region splitting method begins with the entire image in one region. The regions are then split to form regions which satisfy the segmentation uniformity predicate. Both methods use a measure of region homogeneity to define the uniformity predicate used in the segmentation process. There are also segmentation procedures which combine these two approaches. A split and merge procedure will use a picture tree to guide the application of the split and merge processes.

There are also segmentation strategies that look for boundaries between regions using edge detection methods. Sometimes these approaches are combined with the region based approaches.

Ballard, D. B., and Brown, C. M. (1982). *Computer Vision*, Prentice Hall, Englewood Cliffs, N. J.
Bhandarkar, S. M., and Zhang, H. (1999). “Image Segmentation Using Evolutionary Computation.” *IEEE Transactions on Evolutionary Computation*, VOL 3, No 1, April 1999, pp. 1-21.
Gonzalez, R. C., and Woods, R. E. (1992). *Digital Image Processing*, Addison Wesley.
Horn, B. K. P. (1986). *Robot Vision*, MIT Press, Cambridge, Massachusetts,.
Jain, R., Kasturi, R., and Schunck, B. G. (1995). *Machine Vision*, McGraw-Hill Inc.
Klette, R., and Zamperoni, P. (1996). *Handbook of Image Processing Operators*, John Wiley & Sons.
Levine, M. D. (1985). *Vision in Man and Machine*, McGraw-Hill Book Company.
Mallat, S. (1998). *A Wavelet Tour of Signal Processing*, Academic Press.
Pavlidis, T. (1977). *Structural Pattern Recognition*, Springer-Verlag.
Pavlidis, T. (1982). *Algorithms for Graphics and Image Processing*, Computer Science Press, Inc.
Peng, J., and Bhanu, B. (1998). “Closed-Loop Object Recognition Using Reinforcement Learning.” *IEEE Transactions on Pattern analysis and Machine Intelligence*, VOL 20, No 2, February 1998, pp. 139-154.
Rosenfeld, A., and Kak, A. C. (1982). *Digital Picture Processing*, Academic Press, Inc.
Rosenfeld, A., and Kak, A. C. (1982). *Digital Picture Processing*, Academic Press, Inc.
Schalkoff, R. J. (1989). *Digital Image Processing and Computer Vision*, John Wiley & Sons.
Sonka, M., Hlavac, V., and Boyle, R. (1993). *Image Processing, Analysis, and Machine Vision*, Chapman and Hall Computing.
Stucki, P. (1978). *Advances in Digital Image Processing*, Plenum Press.
Wang, F. (1991). “Relational-Linear Quadtree Approach for Two-Dimensional Spatial Representation and Manipulation.” *IEEE Transactions on Knowledge and Data Engineering*, Vol. 3, No. 1, March 1991, pp. 118-122.
Weeks, A. R., Jr. (1996). *Fundamentals of Electronic Image Processing*, SPIE Optical Engineering Press.