Pyramid Segmentation in 2D and 3D Images Using Local Optimization

Alan Kalvin †*   Shmuel Peleg ‡   Robert Hummel †

† Dept. of Computer Science, New York University, 251 Mercer St, New York, NY 10012, U. S. A.

* Inst. of Reconstructive Plastic Surgery, New York University Medical Center, 550 First Ave, New York, NY 10016, U. S. A.

‡ Dept. of Computer Science, The Hebrew University of Jerusalem, 91904 Jerusalem, ISRAEL

ABSTRACT

We present a segmentation algorithm which uses pyramid data structures as a means for achieving computational efficiency and for improving the quality of the segmentation. Results of experiments on synthetic and actual computed tomographic (CT) data are presented, and future enhancements to the algorithm are discussed.

INTRODUCTION

Given an image, the segmentation problem consists of partitioning the image pixels into a number of different regions that correspond as accurately as possible to the underlying physical structures represented in the image.

Of particular use in the design of segmentation algorithms are multiresolution pyramids [1]. We present a segmentation algorithm that uses two pyramids, a gray level pyramid, that is built bottom-up, and a label pyramid that is built top-down. The algorithm incorporates a fixed resolution segmentation method which it applies at each level in the pyramids. The motivation for using pyramids is to reduce computational costs through a coarse-to-fine segmentation strategy, and to improve the quality of the segmentation.

LOCAL COST FUNCTION OPTIMIZATION

Nathan and Peleg [2], and Nathan [3] describe an iterative algorithm for optimal discrete labeling of a set of objects and apply it to 2-D image segmentation. Given an image $I(x,y)$, the segmentation of the images is defined by a labeling $L(x,y)$ of the image pixels in such a way as to minimize a global heuristic cost function

$$H = \sum_{x,y} R(x,y) + \alpha \sum_{x,y} D(x,y),$$

where R is a roughness measure, D is a discrepancy measure (both of which depend on the labeling L), and α is a weighting factor for D. $R(x,y)$ is defined to be equal to the number of 4-neighbors of pixel (x,y) whose labels differ from $L(x,y)$, and $D(x,y) = (I(x,y) - G(\lambda))^2$ where $\lambda = L(x,y)$, and $G(\lambda)$ is the average gray level value of all pixels labeled λ. In order to compute this global function H separately at each pixel, it is broken into the sum of local functions $h(x,y) = R(x,y) + \alpha D(x,y)$ for every pixel in the image. Of course, the local value $h(x,y)$ depends upon the labels in a neighborhood of (x,y).

The basic step in the segmentation process involves examining a pixel (x,y) and relabeling it with the label that minimizes $h(x,y)$. To attempt to get a global improvement in H, this local updating is performed at every pixel in the image. This process of updating every image pixel is iterated until it converges to some final labeling at which stage each pixel has its best label. It is possible that a label that is optimal in one configuration may not be optimal in the updated one, however in our experiments we have found that this process does reduce the global value H.

Choosing an initial labeling

An interesting feature of the local optimization procedure is that it works well even when the initial labeling is quite poor — it will still converge quite quickly to a good segmentation. It is useful though to speed up convergence by staring with an initial labeling that is a reasonable approximation to the desired segmentation, and inexpensive to compute.

One good method is *optimal quantization* [4] which quantizes the different gray levels in the image into a smaller number of discrete levels (equal to the number of labels) in a way that minimizes the total squared quantization error.

MULTIRESOLUTION SEGMENTATION

Our multiresolution segmentation algorithm incorporates the fixed resolution method into a pyramid data structure consisting of a gray level pyramid and a label pyramid. This algorithm is described for 2-D images; the extension to 3-D is straightforward.

First the gray level pyramid is constructed up to some appropriately chosen level k. The fixed resolution segmentation method is applied to the gray level image I_k at this level, to produce L_k, the labeling at level k in the label pyramid. The initial labeling is done using optimal quantization.

L_k is passed down to level $k-1$ of the label pyramid as follows:

$$L_{k-1}(x,y) = L_k(\lfloor x/2 \rfloor, \lfloor y/2 \rfloor)$$

The segmentation of the gray level image I_{k-1} now proceeds using the labeling passed down as its initial labeling. Since this initial labeling is a reasonable approximation to the correct segmentation of image I_{k-1}, a high proportion of the pixels will already be correctly labeled, with most of the incorrect labels occurring along segment boundaries. So just a few iterations of the label updating procedure are now required to converge to the final segmentation.

This process is continued, working down the pyramids, until the original gray level image I_0 at the base of the gray level pyramid has been segmented.

The rationale for this strategy is that at each level m of the gray level pyramid below the starting level k, the number of iterations required to segment the image I_m will be far smaller than the number that would be needed if the segmentation were started from scratch, using optimal thresholding, or some other method of assigning the initial labels. Consequently, when the procedure terminates at the base of the pyramid, the total amount of computation will be less than the amount that would be needed to segment the original image I_0 using a direct application of the fixed resolution segmentation procedure.

GAUSSIAN vs. MEDIAN-FILTERED PYRAMIDS

Two different versions of the gray level pyramid were used in our experiments, the gaussian pyramid, and the median-filtered pyramid. In the gaussian pyramid, at a given level in the pyramid, the gray level of each pixel is equal to the weighted average of the gray levels of a 5 x 5 block of pixels in the level below. For the median-filtered pyramid, the median of the gray levels in the 5 x 5 block of pixels is used instead of their weighted average.

The median-filtered pyramid proved to give better results, primarily since it preserved the original gray level values. The problem with the gaussian pyramid is that if the gray level gradient between two adjacent segments is very low, these neighboring segments tend to be blended together. Conversely, if the gradient is high, a spurious segment is produced along the boundary between the 2 true segments.

EXPERIMENTS

The multiresolution segmentation was applied to synthetic 3-D images which are simplified models of medical CT scan data of the human head. These images consist of 64 x 64 x 20 blocks of data, made up of of five separate segments, the background, an outer layer of skull bone, the brain, the eyes, and the remainer of the skull cavity which is considered "fluid". levels for these objects were chosen to match those in real CT scan images.

Figure 1 shows the results of applying the fixed segmentation method and the multiresolution pyramid segmentation method to a noisy version of this synthetic head. (Only one transversal slice of the head is shown, but the actual segmentation is done on the full 3-D image, not on separate slices). Noise has been added to the image as follows. First the head is constructed with sampling noise having a gaussian distribution with a standard deviation of 0.1. Additive gaussian noise with a standard deviation of 1.0 is then introduced, followed by multiplicative gaussian noise with a standard deviation of 0.05.

For both segmentation methods, α, the discrepancy weighting factor, equals 0.05. Figure 1(a) shows the result of applying the optimal quantization procedure to the image data, dividing them into 5 classes. Figure 1(b) shows the segmentation achieved with the Peleg/Nathan algorithm after the 10th and final iteration.

The results of applying the pyramid segmentation algorithm are shown in figures 1(c) - 1(e). Here 2 levels of the pyramid were used. Figure 1(c) shows the initial and final segmentation at level 1 in the pyramid. The initial segmentation is by optimal quantization, and the final result is reached in 3 iterations. The final segmentation of the image, reached in 4 iterations is shown in figure 1(d).

In terms of the number of local label updating steps required, the total savings achieved with the pyramid segmentation is 52.5%. It is also clear that this method gives a better quality segmentation, the reason being that the reduction in the resolution in the construction the gray level pyramid filters out high frequency noise.

Figures 2 and 3 illustrate the effect of changing the discrepancy weighting factor α. Here a noisier version the synthetic head is used; the standard deviations for the sampling, additive and multiplicative noise are 0.5, 2.0 and 0.05 respectively.

In the pyramid segmentation algorithm, α is increased by a factor of 5 every time we move up a level in the pyramid. By increasing α in this manner, the permissible degree of roughness in the image segmentation increases with the height of the level at which it is performed. This results in a better segmentation, since smaller details in the original image are not smoothed away (i.e. given incorrect labels) at the higher levels of the pyramid.

In figure 2 $\alpha = 0.001$. Figure 2(a) shows the final (9th) iteration of the Peleg/Nathan method. The pyramid segmentation algorithm, using 3 levels, reduces total computation time by 32.6% ; it converges after 4 iterations at level 2, 3 iterations at level 1, and 5 at level 0. Figure 2(b) shows the final iteration at the bottom level of the pyramid.

For $\alpha = 0.01$ as shown in figure 3, the Peleg/Nathan method requires 13 iterations, and the pyramid segmentation converges after 4, 8 and 7 iterations at levels 2, 1 and 0 respectively to give 30.3% saving in computation costs.

The final figure, figure 4, shows the pyramid segmentation method applied to real CT scan data. The original image is a 256 x 256 slice produced by a Philips 310 scanner. This slice is divided into 5 segments, using 3 pyramid levels. with $\alpha =$ 0.001 at level 0, and increasing by a factor of 8 with every level up the pyramid. The number of iterations required for convergence at levels 2, 1, and 0 of the pyramid are 4, 10 and 7 respectively. The reduction in computation time, compared with that required by the Peleg/Nathan algorithm is 19.3%.

FUTURE DEVELOPMENTS

One problem with the segmentation method is that even though it can produce a good segmentation starting with a poor initial labeling, if the initial labeling is bad enough, it is possible that the process will converge to a local minimum that represents an unsatisfactory solution. This is particularly likely when there are large differences in the sizes of the objects in the image. A small object can get "lost" in the higher levels of the pyramid, and if its average gray level is close to that of a much larger neighboring object, it may not be recovered lower down in the pyramid. We are investigating methods for recovering information at finer scales that has been lost at coarser ones.

The current measure of roughness assigns the same penalty cost for each neighbor of a pixel that does not have the same label as the pixel. We plan to introduce domain dependent information so that we can evaluate roughness with the aid of a *compatibility matrix* where the penalty cost reflects some feasibility of two objects being neighbors.

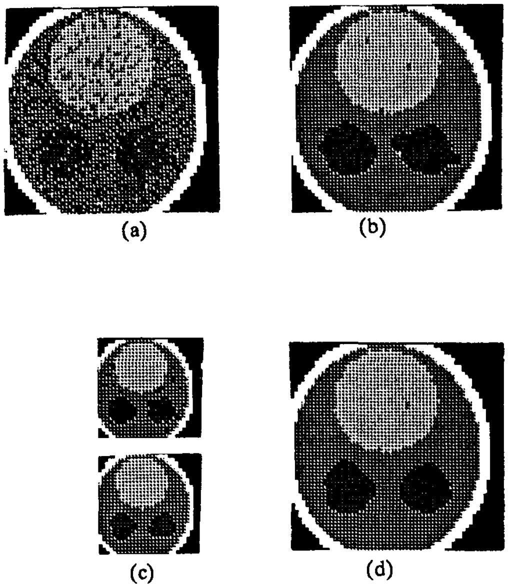
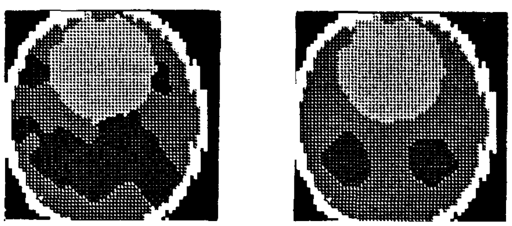
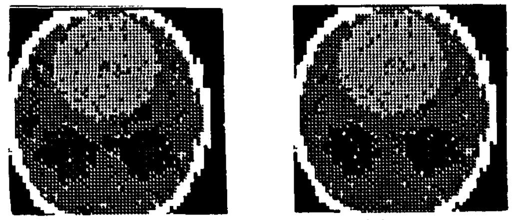


Figure 1



(a) (b)

Figure 2



(a) (b)

Figure 3

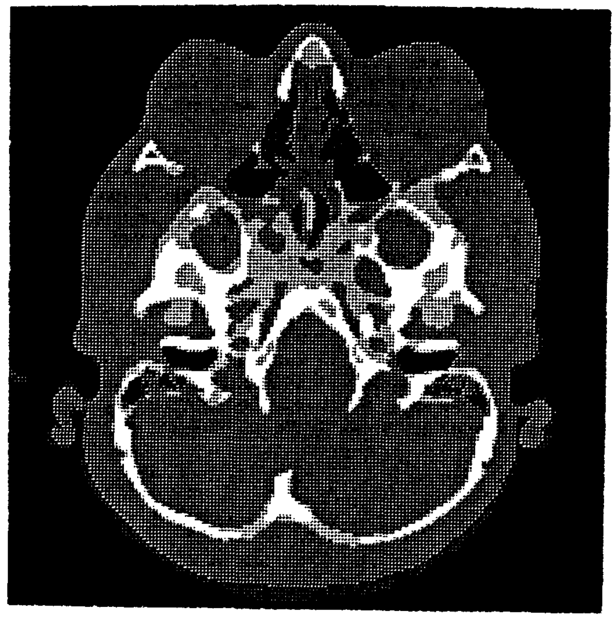


Figure 4

At present there are three parameters affecting the quality of the segmentation that have to be determined *a priori*, namely the number of segments, the pyramid height and α, the discrepancy weighting factor. We are considering two approaches to allow the parameters to be determined dynamically. The first way is to build this into the procedure itself; the algorithm will start off with some given parameters, evaluate the quality of the segmentation, and on the basis of this result make adjustments to the parameters. A second approach is to allow user interaction to guide the segmentation. This will be particularly useful in the situation in which the segmentation algorithm is part of a larger system that provides interactive display and manipulation of 3-D images.

References

1. A. Rosenfeld (ed.), *Multiresolution Image Processing and Analysis*, Springer-Verlag (1984).
2. S. Peleg and A. Nathan, "Classification and tracking using local optimization," *IEEE Trans. Sys., Man, Cybern.* **13** pp. 1158-1162 (December 1983).
3. A. Nathan, "Image segmentation and sequence analysis," M.S. thesis, Computer Science Department, Hebrew University of Jerusalem, Israel (1975).
4. A. Rosenfeld and A.C. Kak, *Digital Picture Processing (2nd. edition)*, Academic Press (1982).