MENTATION PAGE

Form Approved
OMB No. 0704-0188

REPORT DATE	3. REPORT TYPE AND DATES COVERED
1/30/92	Final 6/1/91 - 11/30/91

4. TITLE AND SUBTITLE
"A Real Time System for Multi-sensor Image Analysis through Pyramidal Segmentation" (U)

5. FUNDING NUMBERS
6550 2F
3005/A1

6. AUTHOR(S)
L. Rudin, S. Osher, G. Koepfler, J.M. Morel

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Cognitech, Inc.
2800 28th Street Suite 101
Santa Monica, CA 90405

8. PERFORMING ORGANIZATION REPORT NUMBER
AFOSR-TR- 92 0021

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR/NM
Bldg. 410
Bolling AFB DC 20332-6445

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
F49620-91-C-0036

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approve for Public Release;
Distribution unlimited

12b. DISTRIBUTION CODE
UL

13. ABSTRACT (Maximum 200 words)

A state of the art, fully functional, multi-scale and multi-channel segmentation tool has been developed. It is based on the recently developed computational theory of the 2-normal segmentations. A fast multi-scale pyramidal algorithm has been designed and applied to the theoretical variational segmentation model of Mumford-Shah. This algorithm has a multi-channel capability, as well as a much more general class of solutions. Namely, a piecewise polynomial segmentation is natural to the pyramidal multi-channel framework. The piecewise affine segmentation has been implemented and tested. Application specific channels include: gray scale information, two-dimensional wavelet channels for texture discrimination, and multi-scale singular feature channels. The accuracy of the pyramidal segmentation algorithm has been experimentally compared to the accuracy of two other modern segmentation algorithms. The performance of the pyramidal algorithm has shown an average four-fold reduction in error measurements.

Computational experiments with reconnaissance photography, multi-sensor satellite imagery, medical CT and MRI multi-band data have shown a great practical potential of this novel technique. Preliminary experimentation in clutter removal via multi-channel segmentation points to a totally new class of feature-preserving decluttering algorithms.

14. SUBJECT TERMS

15. NUMBER OF PAGES
30

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAR

A real-time system for multi-sensor image analysis through pyramidal segmentation.

January 1992

Research sponsored by the Air Force Office of Scientific Research (AFSC), under Contract #F49620-91-C-0038.

Cognitech Inc.
Carlsberg Corporate Center
2800, 28th Str.
Santa Monica, CA 90405
Effective date of contract: 6/1/91
Contract exparation date: 11/30/91

Principal Investigator:
Dr. Leonid I. Rudin

Phone: (310) 314-3587

DISCLAIMER

"The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either express or implied, of the Defense Advanced Research Project Agency or the U.S. Government."

A General Multi-Sensor Image Analysis System Through Pyramidal Segmentation.

Principal Investigator	: Dr. Leonid Rudin
Co-investigator	: Dr. Stanley Osher
Project scientist	: Dr. Georges Koepfler
Consultant	: Dr. Jean-Michel Morel

Cognitech, Inc., Santa Monica, Ca. 90405

January 1992

1 Introduction.

Image processes today are often presented with multiple data of the same scene obtained using various sensors. Typical examples include satellite pictures using infrared, optical and other multispectral data or medical data such as MR images taken under different protocols (proton, T2, ...) which give complementary physiological and anatomical information in brain images. It thus seems quite natural to include this information in image processing algorithms. This not only should improve analysis of the data but should also help to get more stable algorithms as there is more information available.

Segmentation algorithms are basic tools in the extraction of global features out of digitized data. This report presents the first fully functional multi-channel segmentation tool based on a fast multiscale and pyramidal algorithm. The input for the different channels is obtained by using data from different sensors and by preprocessing some of these channels with the two dimensional wavelet transform combined with a texture dicrimination algorithm and a multiscale edge detector.

This report presents the mathematical background of Cognitech's segmentation algorithm. The algorithm itself is discussed in section 3. Computational examples are presented and discussed in sections 4 and 5.

2 Computational theory of segmentation devices

General principles upon which the method used at Cognitech is based are stated and a variational formulation of the segmentation problem will be given. A sketch of the recently developed computational theory of the so-called "2-normal" segmentations will be given. More (theoretical) details are discussed in [MorS, KoeMS, Koe].

1. Cognitech has first of all designed a *universal boundary detection device*, definable and analyzable independently from the number and kind of channels (grey level, colour, texture, features, etc...) to be used as input for the segmentation ([BecSI, MaliP2, Ju]). (By "boundary in an image", the boundary in the topological sense is meant: it is the boundary between regions of a partition of the picture. Thus boundaries are different from edges obtained by some local filtering.)

2. A second principle, adopted in the following, is that an algorithm for segmentation should be multiscale. Indeed boundaries may appear on some scale and not be relevant on another. It follows that a sequence of scale parameters λ_i defines a corresponding sequence of segmentations and each segmentation captures objects of a specific geometrical scale.

3. Finally a *comparison principle* is adopted. It states that given two different segmentations of a datum, it is always possible to decide which one of them is considered as better than (or equivalent to) the other. Thus the existence of some total ordering over all possible segmentations is assumed, and this can be simply achieved if this ordering is reflected by some real functional E. Namely if $E(K_1) < E(K_2)$, then the segmentation K_1 has to be considered "better" than the segmentation K_2. For instance, this principle is satisfied by segmentation devices based on Gibbs energy functional as in [GemG]. It is not satisfied by the region growing methods based on thresholds [Pav1, Pav2, PavL, Zu], nor by the edge local detection devices [Marr, MaliP1].

One defines an image g as a scalar or vector function, defined on the image domain Ω (generally a rectangle). One seeks a segmentation, that is, a partition of this rectangle into a finite set of regions, each of which corresponds to a part of the image where g is as regular as possible. Moreover, one wishes to compute explicitly the region boundaries and, of course, to control their regularity and location. Since by the comparison principle, all of these criteria must be taken into account in the functional E, this functional necessarily contains terms which control :

- the similarity of each region with respect to the chosen channels.

- the size, location and regularity of the boundaries.

The model is a special case of the Mumford and Shah (MS) variational model (see[MumS1, MumS2]). According to this model the segmentation (u, K) should be obtained by minimizing the functional

$$E(u, K) = \int_{\Omega \setminus K} (u - g)^2 dx dy + \lambda \ell(K), \tag{1}$$

where K is a union of boundaries in Ω with Hausdorff length $\ell(K)$, and u is piecewise constant on $\Omega \setminus K$. The constant λ represents the scale parameter of the functional and measures the length of the boundary : if λ is small, many boundaries are allowed and one gets a fine segmentation. As λ increases, the segmentation gets coarser and coarser.

Mumford & Shah [MumS1] and Morel & Solimini [MorS] have proved that either the boundary K is regular (at least C^1), or it has singular points of two types, namely, triple points where three branches meet at 120° angles or boundary points where K meets the boundary $\partial\Omega$ of Ω at a 90° angle. Similar functionals have also been studied by [DeG, AmbDeG, CarDeGL] in order to model physical phenomena of phase transitions [FoT] and liquid crystals. We follow a computational approach of [MorS] for the reduced formulation (1). Presently the full (MS) model has no constructive solution in multi-dimensions.

Functional (1) represents a compromise between accuracy of the regions and parsimony of the boundaries. In case of grey level segmentation a pure region growing would simply put together the pixels with similar grey levels [Zu, HarS]. The above procedure generates very nonsmooth boundaries, too "small" or too "thin" regions. Thus, if no control is made on the boundaries, one needs additional criteria and thresholding to achieve a decent segmentation. A functional such as (1) is designed to avoid this kind of ad-hoc treatment. One hopes to have all the criteria assembled in the same functional. Many early functionals in image analysis had only two dimensional energy terms, coupled with some thresholding criteria (see [Pav1, Zu]) and no global approximation and convergence results were available until now.

It is well known that functionals such as (1) may have many local minimizers. A search for global extrema leads to an NP-complete problem.
One has thus to choose between two strategies :

- The global minimization is performed by simulated annealing methods, which leads to huge computational complexity, but ensures that the global minimum is attained [GemG].
- Alternatively, one can define a concept of a "good" local minimum which should be more accessible to fast computations.

Recently, even users of simulated annealing have developed a tendency to define a fast approximation, which need not find a global minimum [Az]. The homotopy method of Blake and Zisserman [BlakZ] also seeks a "good" local minima.

The following notions are basic to the analysis of the algorithm. A *subsegmentation* of a segmentation K is a segmentation obtained by merging an arbitrary number of adjacent regions of K.

Definition 1 *A segmentation K will be called 2-normal if for every pair of regions O_i and O_j, the new segmentation K' obtained by merging these regions satisfies $E(K') > E(K)$.*

Theorem 1 *The set of 2-normal affine segmentations has the following compactness property: for every sequence K_n of such segmentations, there exists a subsequence converging to a segmentation K such that*

$$E(K) \leq \liminf_n E(K_n).$$

K is not necessarily 2-normal, but it has a 2-normal subsegmentation with lower energy.

Compactness is a weak form of uniqueness : it indicates that the set of 2-normal segmentations is much smaller than the set of all possible segmentations (which is not compact).

It also provides a very important computational property : If one defines an algorithm computing better and better solutions, that is solutions with decreasing energy, these solutions must accumulate around "good" local minima of the functional (1).

3 A pyramidal algorithm constructing 2-normal affine segmentations.

Consider now the problem of defining and computing a sequence of recursive mergings resulting in a good 2-normal segmentation. Notice that not all 2-normal segmentations are interesting; for instance, the empty segmentation, where Ω is the single region is clearly a 2-normal segmentation.

Assume that the datum g is defined on a rectangle. This rectangle is divided into small squares of constant size (the pixels) and g is assumed to be constant on each pixel. The properties required for the segmentations computed by a region growing algorithm, defined as an application associating to g and λ a segmentation (u, K) are:

a) "Correctness" (Fixed point property) : Assume that g is piecewise constant on some areas of the rectangle. Then there exists a value λ_0 of the parameter λ such that for every $\lambda < \lambda_0$, the segmentation (u, K) obtained by the algorithm satisfies $u = g$ and K is the union of the boundaries of the areas where g is constant.

b) "Causality" (Pyramidal segmentation property) : If $\lambda > \lambda'$, then the boundaries provided by the algorithm for λ are contained in those obtained for λ' and the areas of the segmentation associated with λ are the unions of some of the areas obtained for λ'.

The last property ensures that a fast pyramidal algorithm can be implemented, by computing a hierarchy of segmentations from fine to coarse scales. Moreover the coarser segmentation will be deduced from the finer by "merging" operations, with a pyramidal structure for the computation. Note, that as a consequence of the fixed point property, if λ is very small, the computed segmentation is attained with (u_0, K_0) where $u_0 = u$ and K_0 consists of all the boundaries of all the pixels, and therefore coincides with the global minimum for λ equals zero. We shall call this segmentation, where each pixel is a region, the "trivial segmentation". The recursive merging algorithm used at Cognitech satisfies all the above properties.

3.1 Principles of the segmentation algorithm

The algorithm is based on the classical technique of region growing, or merging. Starting with an initial partition (for example pixel by pixel, which will be called the trivial segmentation) one merges those regions which satisfy an homogeneity criterion. One hopes to get in this manner homogeneous regions with respect to the criterion.

The algorithm starts with a partition of the images into square regions which can be of $1\times1, 2\times2, 4\times4$ pixels. At this stage, all boundaries between regions are therefore horizontal or vertical. Presently the final segmentation consists of patches of pixels. The boundary K falls in between image pixels

The initial segmentation has a certain number of channels which are computed in each square region. A channel contains some information which is associated with each region, such as its area, its mean grey level, its mean gradient, or (for texture analysis) its mean oscillation at a given scale. Boundaries and vertices of the segmentation can also be associated with several channels. For a boundary, the channel can be the length, a measure of the contrast (i.e. jump across the boundary), ... For a vertex, the channel can be a contrast measure, which e.g. gives a measure of whether the point can be a crossing point or not. All of the channel information will contribute to give a measure of the homogeneity of regions.

An important thing to keep in mind is that once channels have been associated with a region, the single pixels of the region are "forgotten" and only the value of the channels in the region will influence the merging process. This explains the strength and speed of the algorithm, since there are many fewer channels than pixels. Moreover, the only operations to be made on the channels when merging are additions. Also, this is true for all kinds of channels involved in piecewise polynomial or piecewise textural segmentation. Thus the algorithm is universal and independent of the meaning and number of the channels.

3.2 Algorithmic model and data structure design.

The representation of a partition of a picture is based on the following structures:

- *Region* : described by its channels (a list of real numbers), and its borders with adjacent regions.
- *Border* : described by its own channels and the two adjacent regions, and by a list of its connected components.
- *Connected component* of border : given by its two vertices and by a chain of pixels joining them.
- *Vertex* : described by the list of its adjacent borders, its coordinates and its channels.

Every object is defined by some proper information (channels, coordinates) and its relations to other objects. For example Ω, in figure 1(a), is a picture made of three regions Ω_1, Ω_2 and Ω_3.
Let us recall that $\partial\Omega_i$ denotes the boundary of a region Ω_i and $\partial(\Omega_1,\Omega_2)$ is the common boundary of Ω_1 and Ω_2. We assume that the external boundary of Ω (i.e. $\partial\Omega$) is never part of the boundaries of the regions which partition the image.
Denote by $c_i, i = 1,\ldots,4$ connected boundary elements, then $\partial\Omega_1 = (c_1 \cup c_4) \cup c_2$, $\partial\Omega_2 = (c_1 \cup c_4) \cup c_3$ and $\partial\Omega_3 = c_1 \cup c_4$.
With the above conventions the *region* Ω_1 has two *borders* $\partial(\Omega_1,\Omega_2)$ and $\partial(\Omega_1,\Omega_3)$, with $\partial(\Omega_1,\Omega_2)$ having two *connected components* (e.g. c_1, c_4) and $\partial(\Omega_1,\Omega_3)$ just one (e.g. c_2). Associated to these are *vertices*, to c_2 vertices S_2 and S_3. Vertex S_1 has one adjacent border

$\partial(\Omega_1, \Omega_2)$, whereas vertex S_2 has two adjacent borders $\partial(\Omega_1, \Omega_2)$ and $\partial(\Omega_1, \Omega_3)$.

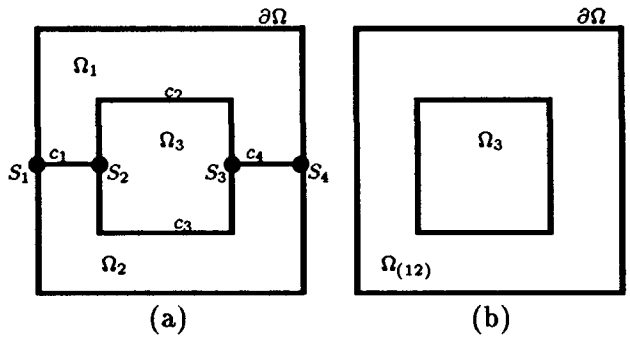


Figure 1: Example of regions before (a) and after (b) merging.

The result of merging regions Ω_1 and Ω_2 is represented in figure 1(b). To achieve this one has to erase $\partial(\Omega_1, \Omega_2)$, which means the two connected components c_1 and c_4 and the vertices S_1, S_2, S_3, S_4. The frontier of the new region $\Omega_{(12)}$ with Ω_3 will have a single connected component obtained by merging c_2 and c_3.

The implementation in the programming language C uses structured data types with dynamical memory allocation. A region will be a variable of strucured type `REGION`, made of an array of real numbers and pointers to a list of variables of type called `BORD`, which represents the boundary between two regions.

3.3 Description of the algorithm.

Initialization. After calculating the channels and, as stated above, taking an initial segmentation made of a single pixel, the entire data structure is set up. The regions are put into a linked list. The present choice has been to create a list of all initial regions as follows: left-right and up-down. This structure, convenient from a programming point of view, does introduce some bias. For a multiscale segmentation the series $\lambda_1, \lambda_2, \ldots, \lambda_L$ handles the growth of the scale parameter.

The criterion. The decision to proceed to a merging of two regions O_i and O_j depends on the sign of $E(\tilde{u}, K \setminus \partial(O_i, O_j)) - E(u, K)$. The algorithm looks for a decrease of global energy by merging these regions. This is the criterion of 2-normal segmentations introduced in section 2. The simplified Mumford and Shah model is implemented by choosing the following energy functional :

$$E(u, K) = \int_\Omega \|u - g\|^2 + \lambda \ell(K).$$

Here g is a vector valued function, whose components are different channels, defined on the rectangle Ω, u is the approximating vector function, and K is the set of boundaries with total length $\ell(K)$. As in the piecewise constant case u is the mean value of g on each region O.

The above functional is just denoted by $E(K)$, i.e. to obtain (u, K) one needs to know only K. Then the fusion criterion is

$$E(K \setminus \partial(O_i, O_j)) - E(K) = \frac{|O_i| \cdot |O_j|}{|O_i| + |O_j|} \cdot \|u_i - u_j\|^2 - \lambda \cdot \ell(\partial(O_i, O_j)),$$

where $|.|$ is the area measure and u_i the approximation of g on O_i. When g is scalar the norm $\|.\|$ is just the absolute value. For multichannel data a weighted norm $\|.\|$ is used. It is specific to each application and to the meaning of the different channels. This will be emphasized in the next section.

To get the necessary data for evaluating the criterion the following information has to be used: Suppose $g = {}^t(g^1, \ldots, g^n)$, then every region O has a channel with its area $|O|$ and n channels $c_O^l = \int_O g^l$, $(l = 1, \ldots, n)$. These yield the values for u restricted to O : $u_O = {}^t(u_O^1, \ldots, u_O^n)$ by simply computing $u_O^l = \frac{c_O^l}{|O|}$ to get the mean value.

The channels of region O_{new} obtained by merging O_i and O_j are given by $|O_{\text{new}}| = |O_i| + |O_j|$ and $c_{O_{\text{new}}}^l = c_{O_i}^l + c_{O_j}^l$, $(l = 1, \ldots, n)$.

Thus a merging of two regions just implies adding the corresponding channels and updating the data structure (i.e. pointer/address affections).

It is straightforward to use the same model for higher order polynomials just by changing the criterion. To get a fast algorithm one just ensures the same, additive, property of the channels (see next section).

The algorithm.

(i) Take the segmentation (u_0, K_0) resulting from the initialization and $\lambda_i = \lambda_1$ as a scale parameter.

(ii) Scan the list of regions and for every candidate region look for the adjacent region which yields the best merging score (i.e. the maximal energy decrease). If such a region exists proceed to merge and update the data. The next region in the list becomes a candidate for merging.
For $\lambda = \lambda_i$ fixed, repeat the scanning of the picture until there is no merging possible. After this step, a 2-normal subsegmentation (u_i, K_i) of the initial segmentation for scale parameter λ_i is achieved.

(iii) For every λ_i, $i = 1, \ldots, L$, calculate a 2-normal segmentation by iterating step (ii). The algorithm stops if there is just one region left or after computing the last scale parameter λ_L a 2-normal segmentation (u_L, K_L).

Result. The intermediate segmentations obtained at each scale λ_i is saved before sweeping to λ_{i+1}. Each ouput yields the following objects : connected boundaries K, the approximation u, perimeter and area of each region O. Each of the above computed objects represents an important feature oriented information about the analyzed picture. See the numerical experiments below to illustrate the applications of these objects.

Remark. This way of sweeping a list of regions is not optimal because the best merging is done by searching over the whole list. This slight bias is however compensated by the multiscale structure of the merging process. The algorithm begins with a small value of λ and increases it when all possible mergings have been done. This multiscale loop orders the mergings by order of priority with respect to the functional. A time expensive track keeping of a chart with the best merging out of all adjacent regions would be prohibitively expensive. The compactness property assures a "good" local minima solution.

3.4 The use of multiple-channels.

The piecewise constant channel case has been explained above. Similar local average channels can be used to get a piecewise affine segmentation of the original data. Recall that input data consists of region area channels $\{|O_i|\}$ and channels $\left\{c^l_{O_i} = \int_{O_i} g^l dx\right\}$.

Piecewise affine segmentation.

One keeps the length term constant and considers how the integral term changes in this case. Let $u(x,y) = ax + by + c$ be the function which approximates g on region O in a sense of the least square mean error

$$\int_O [g(x,y) - ax - by - c]^2 dx dy \,. \tag{2}$$

Accordingly the coefficients a, b and c (depending on O) are solution of the following normal linear system:

$$\begin{pmatrix} \int x^2 & \int xy & \int x \\ \int xy & \int y^2 & \int y \\ \int x & \int y & \int 1 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \int gx \\ \int gy \\ \int g \end{pmatrix} \,.$$

The information needed to compute a, b and c is contained in the following 10 integral terms: $\int_O 1\ (= |O|)$, $\int_O x$, $\int_O y$, $\int_O x^2$, $\int_O xy$, $\int_O y^2$, $\int_O g$, $\int_O gx$, $\int_O gy$, $\int_O g^2$.

Note that the pyramidal nature of the merging is still preserved here. Simply declare 10 channels for each region and use these channels to represent the necessary data for piecewise affine segmentation. It is not necessary to compute (2) on a new region obtained by a merging. The moments listed above will give the necessary information without "going back" to the original picture, hence the pyramidal structure. The coefficients a, b and c can be computed if one needs the reconstruction u, but, for obtaining getting the partition of the picture, coefficients a, b, c are not even needed. In figure 2 the result of piecewise constant segmentation is compared to a piecewise affine segmentation on a synthetic original picture made of affine elements. The original piecewise affine image (i) yields a piecewise constant segmentation, represented by its boundaries (ii) and the approximation (iii). A piecewise affine segmentation gives more accurate partition (iv) and approximation (v).

The first line represents the original picture, and the second line shows the boundaries K and the reconstruction u using a piecewise constant model. The third line shows both K and u for a piecewise affine model.

Improvements of this program and for higher degree polynomials may be developed using ENO-subcell methods which are successfully used for image compression at Cognitech.

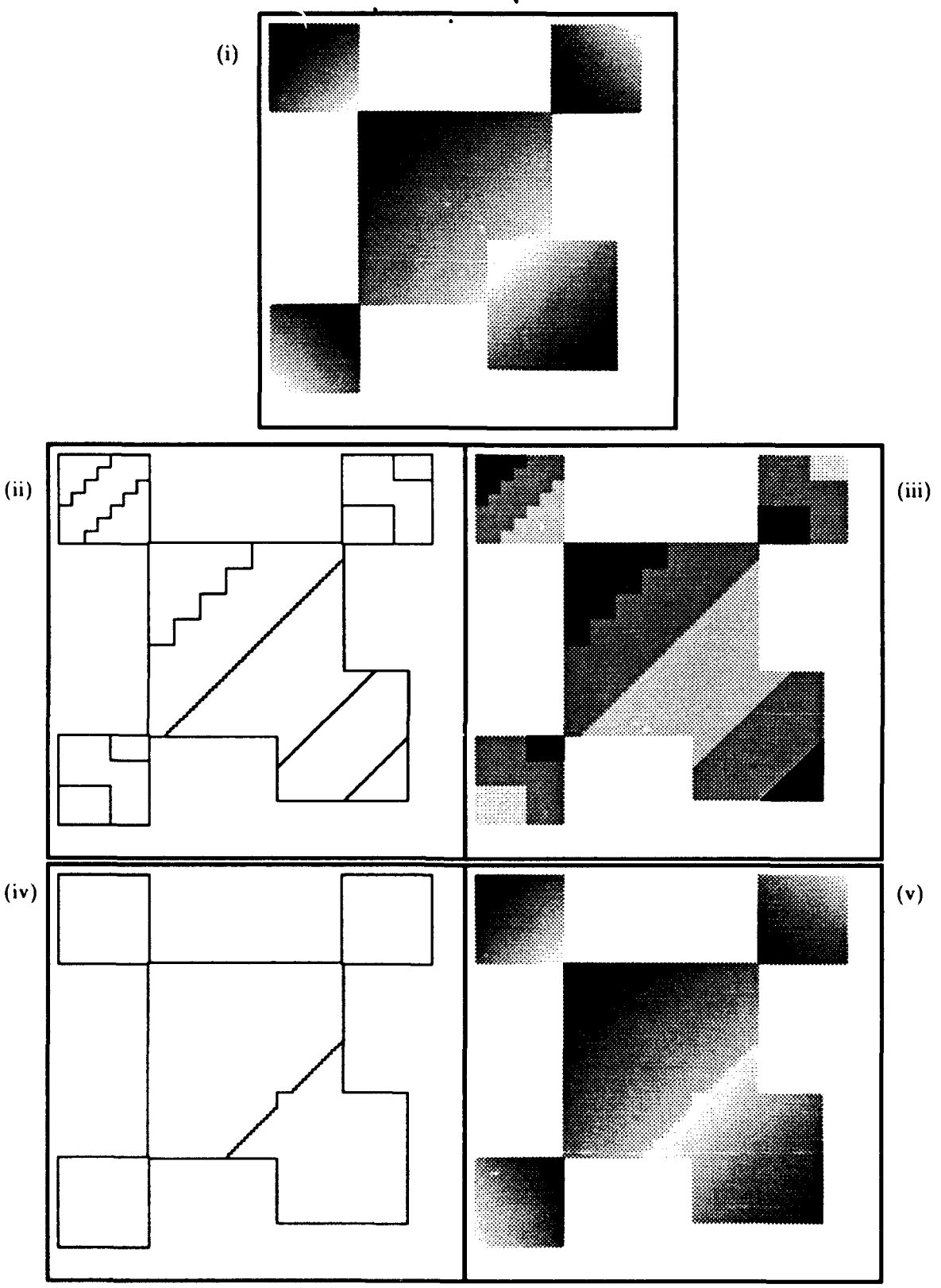


Figure 2: Piecewise Constant versus Piecewise Affine.

Applying a wavelet transform.

Differential or semilocal integrodifferential operators provide useful analysis of pictoral information. The 2-D wavelet transform is used here for such analysis. It is based on the pyramidal algorithm of Stéphane Mallat [Mall] using the multiresolution approximation (see [Mey]) formulation. The bidimensional decomposition uses two monodimensional decompositions (following the X, Y axis).
Suppose V_j is a multiresolution of $L^2(\mathbb{R})$, with associated orthonormal basis $(\phi_k^j)_k$, and W_j the orthogonal complement of V_j in V_{j+1}, with orthonormal basis $(\psi_k^j)_k$. Then $\tilde{V}_j = V_j \overline{\otimes} V_j$ is a multiresolution analysis of $L^2(\mathbb{R}^2)$. Let us denote by $\tilde{W}_j$ the orthogonal complement of $\tilde{V}$ in $\tilde{V}_{j+1}$,

- $(\phi_n^j(x)\phi_m^j(y))_{n,m}$ is an orthonormal basis of $\tilde{V}_j$.
- $(\phi_n^j(x)\psi_m^j(y), \psi_n^j(x)\phi_m^j(y), \psi_n^j(x)\psi_m^j(y))_{n,m}$ is an orthonormal basis of $\tilde{W}_j$.

The approximation of picture g at resolution 2^j is
$S_j = (S_j(n,m))_{n,m}$, with $S_j(n,m) := 2^j < g, \phi_n^j \phi_m^j >$.

The difference of information from S_{j+1} to S_j is given by three "details"
$D_j^1 = (D_j^1(n,m))_{n,m}$, with $D_j^1(n,m) := 2^j < g, \phi_n^j \psi_m^j >$,
$D_j^2 = (D_j^2(n,m))_{n,m}$, with $D_j^2(n,m) := 2^j < g, \psi_n^j \phi_m^j >$,
$D_j^3 = (D_j^3(n,m))_{n,m}$, with $D_j^3(n,m) := 2^j < g, \psi_n^j \psi_m^j >$.

The S_j, D_j^1, D_j^2 and D_j^3 can be deduced from S_{j+1} as follows

$$
\begin{aligned}
S_j(n,m) &= \sum_{k,l \in \mathbb{Z}} S_{j+1}(k,l) h(k-2n) h(l-2m) \\
D_j^1(n,m) &= \sum_{k,l \in \mathbb{Z}} S_{j+1}(k,l) h(k-2n) f(l-2m) \\
D_j^2(n,m) &= \sum_{k,l \in \mathbb{Z}} S_{j+1}(k,l) f(k-2n) h(l-2m) \\
D_j^3(n,m) &= \sum_{k,l \in \mathbb{Z}} S_{j+1}(k,l) f(k-2n) f(l-2m)
\end{aligned}
$$

The functions f and h are quadratic mirror filtors, related to Fourier transforms of ϕ and ψ and satisfy $f(n) = (-1)^{(1-n)} h(1-n)$.

The X, Y separation of the decomposition in the 2-D case implies that D_j^1 represents the loss of information in the Y direction by passing from S_{j+1} to S_j, D_j^2 corresponds to the horizontal information loss and D_j^3 is the crossed information loss at scale 2^j. Figure 3 shows the classical representation of a wavelet decomposition, using Daubechies' compact supported wavelets, computed at level $j = -1, -2$.

The filters used are typically the Haar basis or Daubechies' basis which have a finite number of non-zero cefficients.

Texture channel computation.

Denote by F_i the output obtained by wavelet filter, nonlinear rectification by taking negative and positive part of F_i yields two channels $R_{2i} = F_i^+$ and $R_{2i+1} = F_i^-$. This step is required

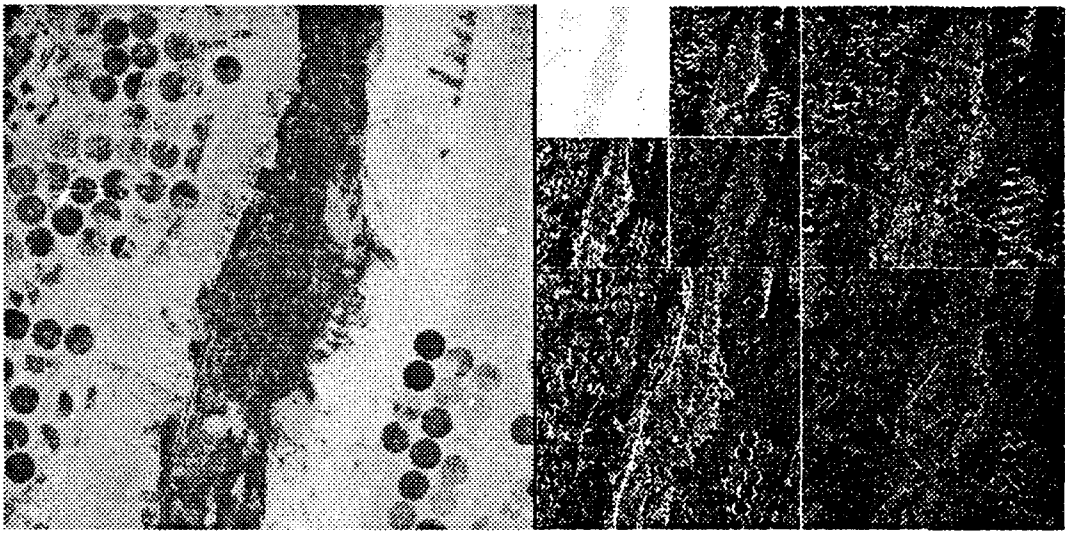


Figure 3: 2-D wavelet Transform.

for texture segmentation. The particularity of the presented algorithm is that all the information computed from the initial picture is kept and not just the most important (oscillant) channel for each region.

Indeed the preceding steps, whose aim it is to obtain texture discriminating information, are under different guises quite widely spread in texture segmentation algorithms. However most of these algorithms are unable to obtain a segmentation from multi-channel input, so they have to restrict themselves to one information channel. Moreover they choose the maximum responding channel to characterize a region. It seems quite obvious that two adjacent regions can have the same "maximum" channel without having comparable responses in other channels so that they are discriminable but "maximum-response-algorithms" will not notice the different behaviour in non extremal channels.

Figure 4 shows an example of texture discrimination using only information provided by wavelet-channels. The input is a pair of textures ("Brodatz-textures") and the output shows the boundary between the textures discriminated by the wavelet and the scales which have been used (3 scales, Haar basis).

Multiscale edge detection.

To analyse the effect of local features, via the application of multiscale singularity detectors (see [Rud]), additional channels are created. This class of channels provides information which is application specific. For example the user may search for regions whose edges satisfy special geometric properties, such as linearity or circularity on a particular scale. The feature channel can take a bit map form as presented in figure 5 (see also figure 12(ii) for the original data).

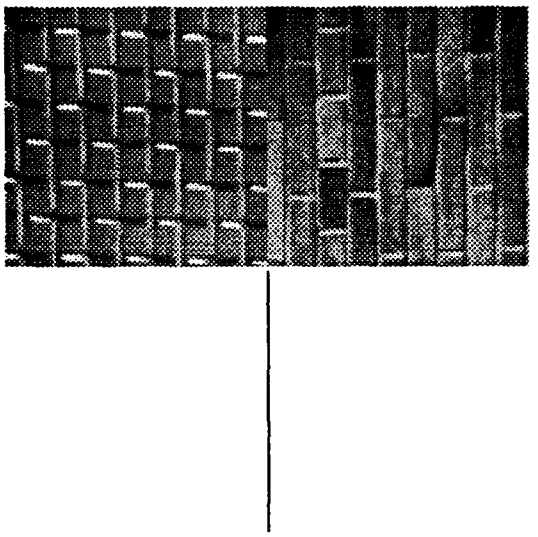


Figure 4: Brodatz-texture discrimination using Haar basis and 3 scales.

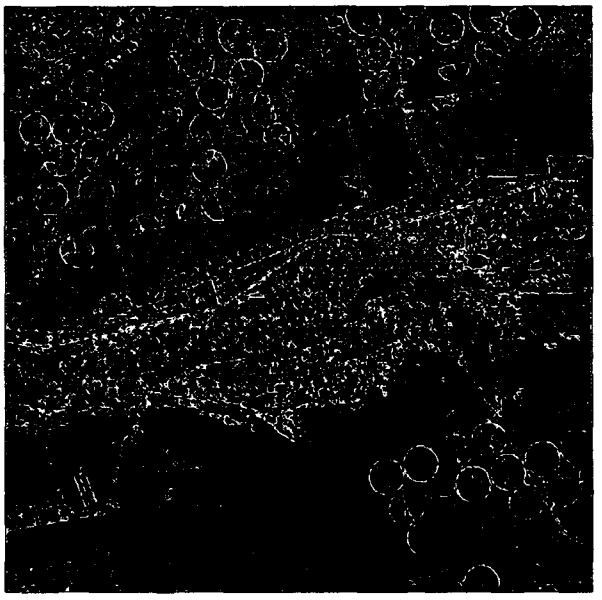


Figure 5: Example of edge map.

3.5 Details on implementation.

The program is written in C, using dynamic representation for the data. The file `StructSeg.h` contains the definitions of all data structures, `InitSeg.h` and `CanalSeg.h` contain all procedures necessary for the initialization of the topological model on the initial grid, as well as the computation of the channels. `MergeSeg.h` contains all functions updating the topological structure of regions, borders and vertices when a merging is activated.
The output of the obtained segmentation is made by functions specific to an application.
The program uses MEGAWAVE[1] and IMAGE[2] libraries for image file handling.

Size of the code and of the used RAM during the run : around 60 kB (2000 lines). The choice being to have a fast procedure, but only if a lot of RAM is avalaible, several error messages are included in the files to warn against a lack of available RAM. Furthermore, a procedure called `Estimation` gives the estimated memory size necessary at the beginning of execution.

The data structure. (Confer file `StructSeg.h`)
In order to represent computationally the topological relations of regions, borders and vertices, structured variables available in the C language are used. The data structure has been partially represented in figure 6. One sees that this data structure is absolutely faithful to the topological relations.

Since the number of borders corresponding to a given region or a given vertex changes during the region growing, all border handling is made by the use of a linked list. Thus the datatype `Li_boundaries` (list of borders) is introduced. The type `Li_pixels` (List of pixels) allows to define a path from a vertex to another.

A remark on the orientation of borders.
Each border is given by its adjacent regions, the right region `rd` and the left region `rg`. It is thereafter decomposed into its connected components. Each one of them is oriented in such a way that the left region `rg` be always left and the right region `rd` always right of the path when the path is scanned from vertex `sa` towards vertex `sb`.

Let us however recall that borders and vertices are in fact virtually defined between the pixels of the image. Thus no pixel is associated with a boundary or a vertex. In order to print or display a border, one has to make a choice for the pixels which will represent it. One allows therefore a 1/2 pixel error, since boundary and vertex are put into the right below adjacent region. Another possibility is to create an edge map of size $(2dx-1, 2dy-1)$, where the "border" pixels are plotted in between the "image" pixels (both are implemented).

Model initialization.
It has two different parts. The first deals with the construction of all initial regions, borders, connected components of borders, and vertices, and with their connections. This initialization is fully independent of the merging criteria. The second part of the initialization is to compute all channels for initial regions, borders and vertices. This part is easy to change and guarantees are large variety of applications.

[1] © of CEREMADE, University of Paris-Dauphine
[2] © of Cognitech, Inc. Santa Monica

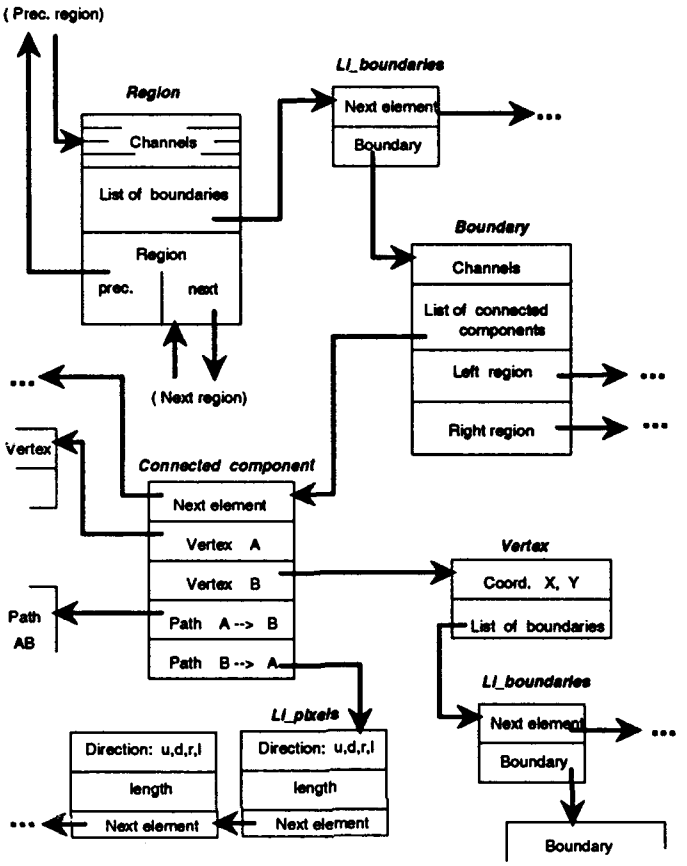


Figure 6: Representation of the data structure.

Requirements about the RAM and the computing time necessary for a fast execution. Assuming an initial grid of dimensions (gx, gy), there are $gx \cdot gy$ initial regions, $gx(gy-1)+gy(gx-1)$ borders and $(gx+1)(gy+1)-4$ vertices. In general $s = gx = gy = 128$, 256 or 512, therefore a square image with approximately s^2 regions, $2s^2$ borders and $(s+1)^2$ vertices. Denote by NCR, NCB and NCS the number of channels by region, border and vertex respectively. This is the memory involved for the different structures :

Regions	$(\text{NCR}+3)s^2$
Borders	$8(\text{NCB}+3)s^2$
Connected components	$40s^2$
Vertices	$4(s+1)^2 \cdot (\text{NCS}+2)$
List of borders	$64s^2$
List of pixels	$32s^2$

Thus around $200s^2 + 4(NCR + 2NCB + NCS)s^2$ bytes for the initial segmentation. If one

wants to process a 512×512 pixels image with an initial grid 2×2 and 48 texture channels, one obtains $400s^2$, that is, with $s = 256$, 25 megabyte. Estimates on the computation time with 48 channels and $s = 256$ give a time of 2 minutes on Cognitech's 76 MIPS HP730 workstation. This is good for an experimental device.

4 Measuring the accuracy of pyramidal segmentation.

Modern applications of image processing require high degree of accuracy during the feature extraction stage. This should be also true for segmentation analysis. Instances of the importance of accurate region segmentation are found in such distinct fields, as the nondestructive material testing via imaging, automated imaging quality control and the construction of accurate physical models of patient anatomy for surgical planning. Such accuracy analysis is notably missing from the current computer vision literature works on segmentation techniques. However, some study was done by the medical imaging researchers [MRE]. Here we shall follow their method of benchmarking segmentation accuracy through experiments with the Robertson phantom, as explained below.

Robertson's phantom on the figure 7 is a 125mm diameter cylinder containing 4 different materials as depicted on this figure. Eight machined cylindrical steps are cut through the innermost material. Prior to assembly, the true inner radii were measured. This experiment uses images obtained by the CT cross sectioning of the phantom.In all, eight cross sections images of the annular phantom were segmented with a three different segmentation techniques. First and second segmentation algorithms are variants of the commonly used contour tracing technique [Pav3]. The accuracy in localization and radius estimation of these methods were compared to the accuracy of Cognitech's pyramidal segmentation.

The experiment is illustrated by showing two of the cross-sections on the figure 8(i),(ii) with different inner radii. The 8(iii), (iv) show corresponding segmentations. Two inner radii are measured and compared to a known true radii. The plot of the radius versus the angular position for all three techniques is presented on the figure 9. Note here that while the pyramidal segmentation oscillates around the true mean value radius, the competing algorithms are consistently above or below the true value. The performance of the pyramidal algorithm demonstrated an average four-fold reduction in error measurement of the inner circle! The bar diagram on page 19 visualizes this comparison of the error in inner and outer radius.

Modifications of the present pyramidal algorithm will be proposed in order to obtain accuracy in the variance of the geometrical boundaries of extracted regions.

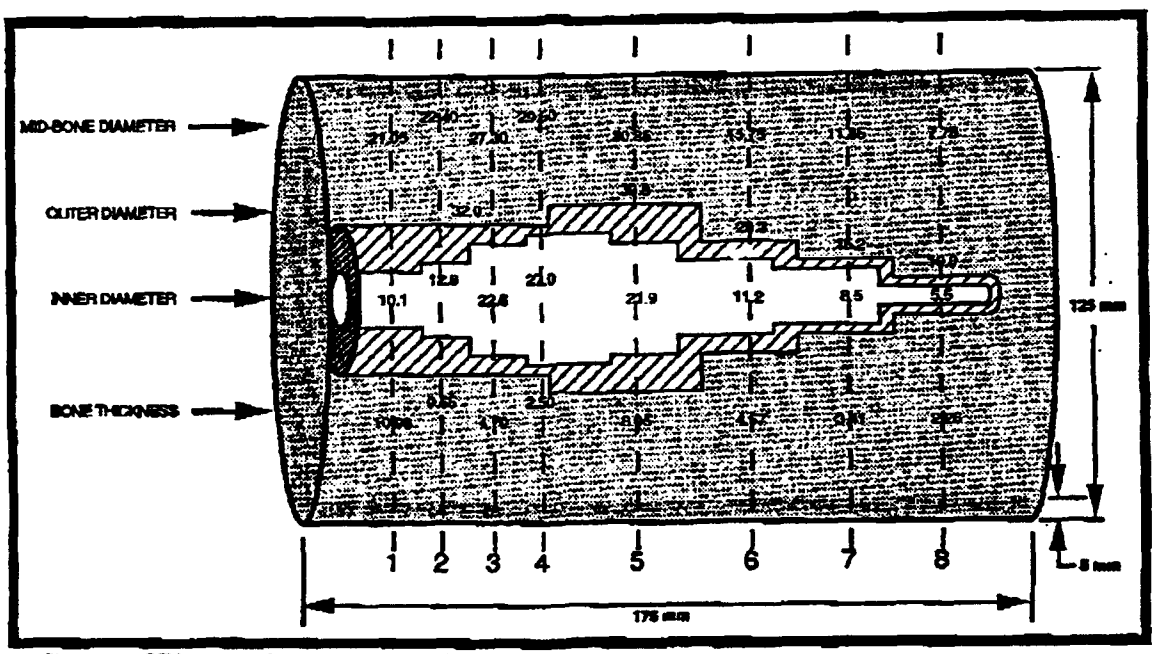


Figure 7: Scheme of the phantom used for accuracy tests.

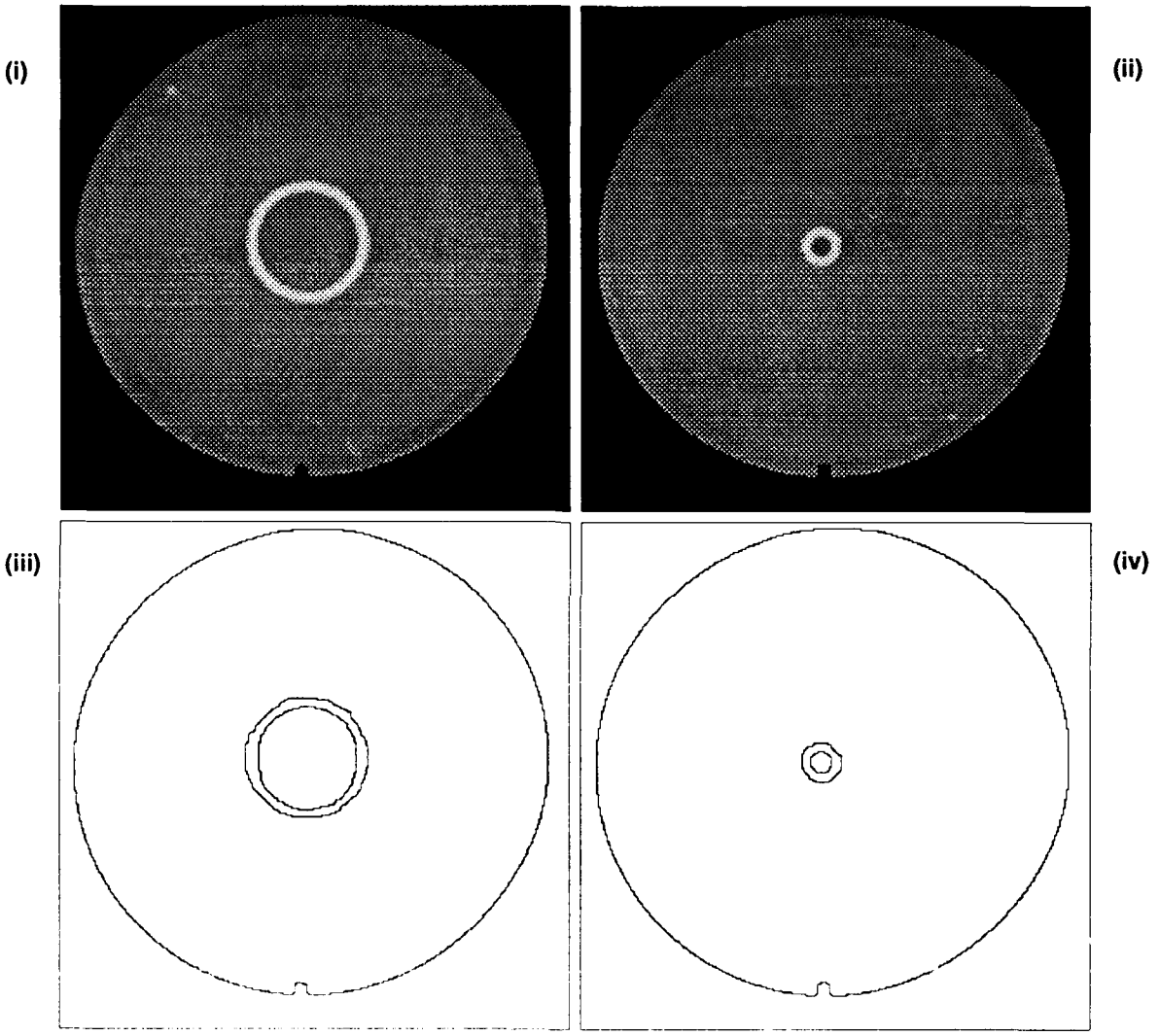
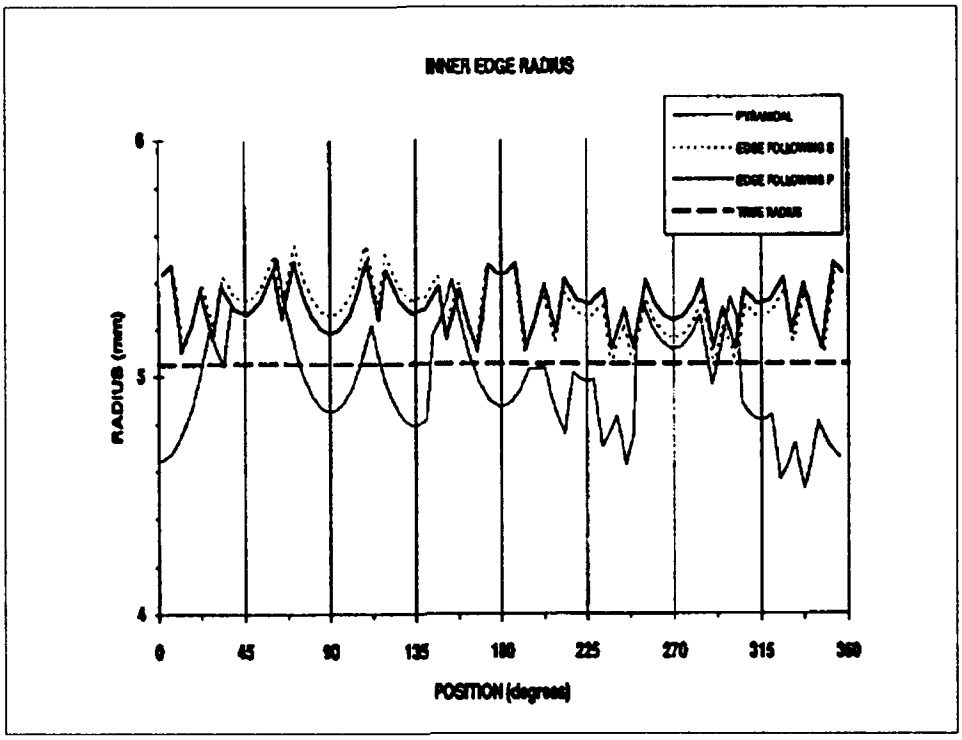


Figure 8: Two different sections of the phantom with respective segmentation.

(i)



(ii)

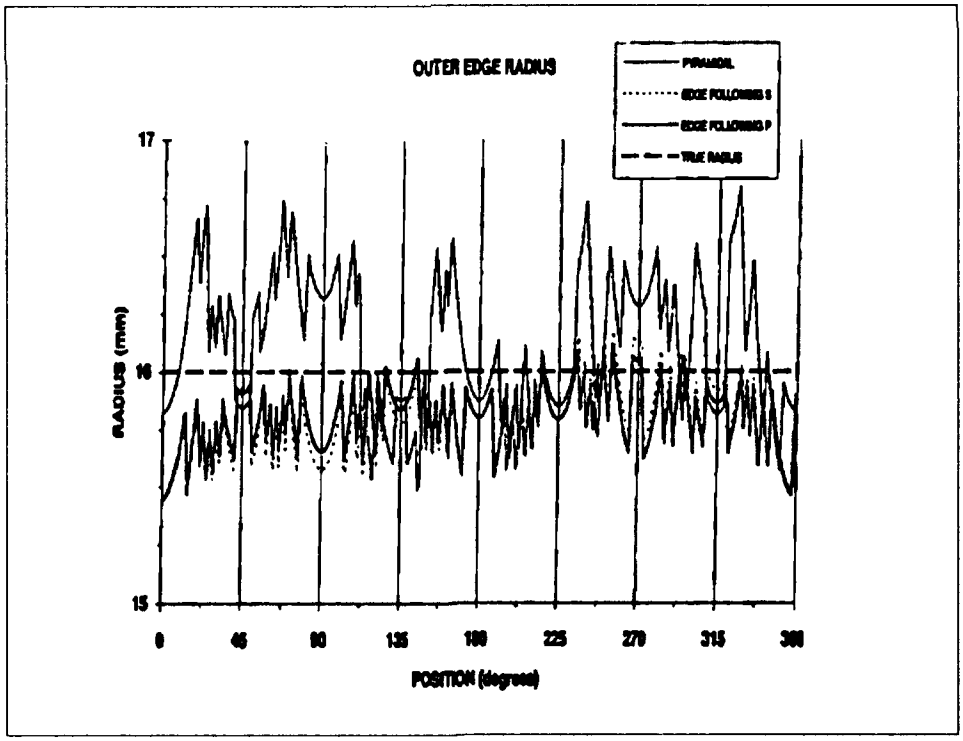
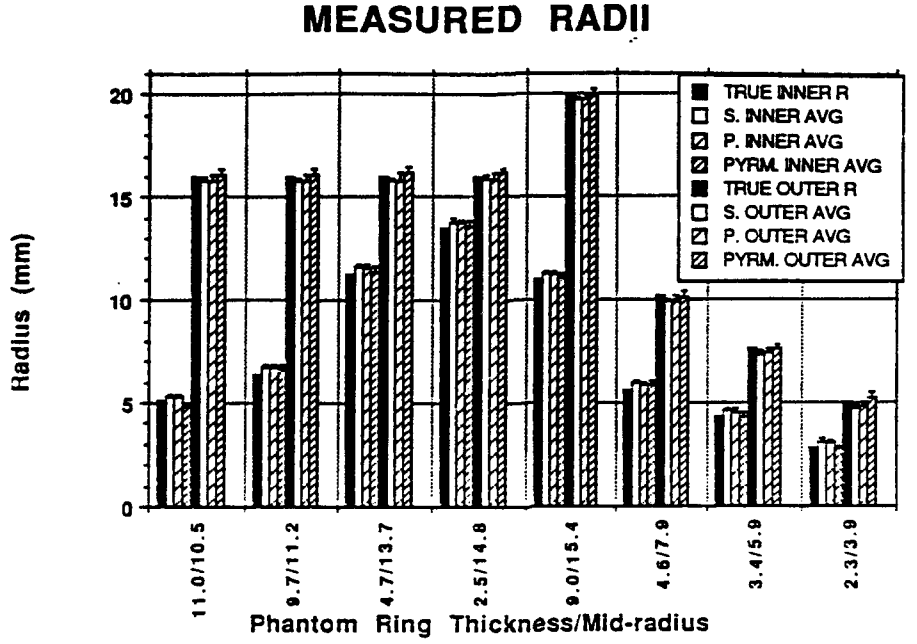
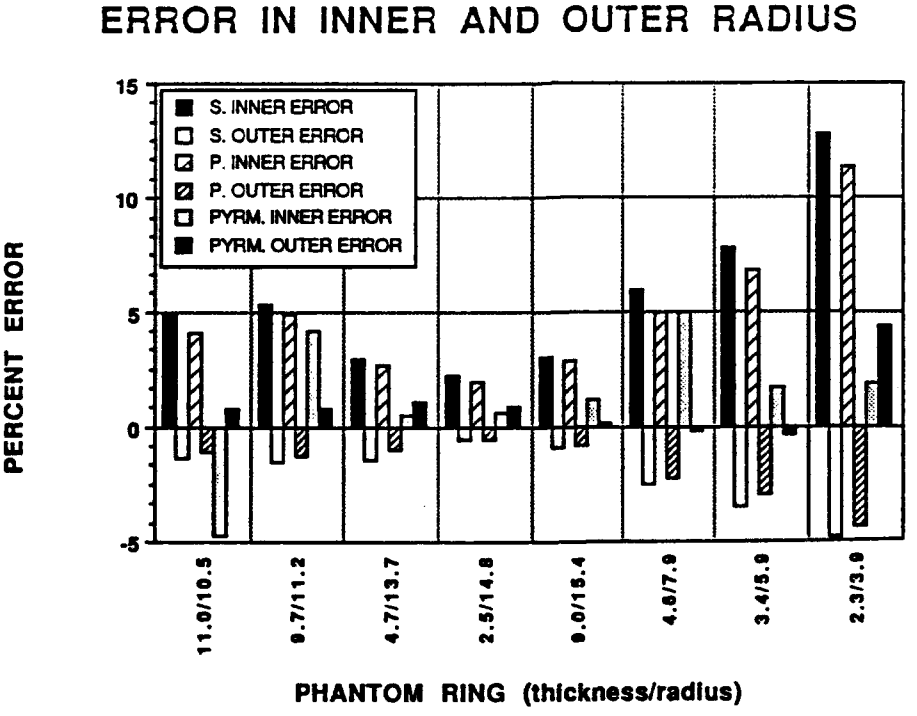


Figure 9: Accuracy comparison between different estimation methods of the inner (i) and the outer (ii) radius.


MEASURED RADII
Radius (mm)
20
15
10
5
0
TRUE INNER R
S. INNER AVG
P. INNER AVG
PYRM. INNER AVG
TRUE OUTER R
S. OUTER AVG
P. OUTER AVG
PYRM. OUTER AVG
11.0/10.5
9.7/11.2
4.7/13.7
2.5/14.8
9.0/15.4
4.6/7.9
3.4/5.9
2.3/3.9
Phantom Ring Thickness/Mid-radius


ERROR IN INNER AND OUTER RADIUS
PERCENT ERROR
15
10
5
0
-5
S. INNER ERROR
S. OUTER ERROR
P. INNER ERROR
P. OUTER ERROR
PYRM. INNER ERROR
PYRM. OUTER ERROR
11.0/10.5
9.7/11.2
4.7/13.7
2.5/14.8
9.0/15.4
4.6/7.9
3.4/5.9
2.3/3.9
PHANTOM RING (thickness/radius)

5 Experimental results.

Reconnaissance photography.

Figure 10 depicts multiscale segmentation of the single frame 256*256 B&W 8-bit "Chemical plant" (i). The sequence of three scales progressively remove small scale features, thus uncovering such objects of interest as roads and buildings in the figures 10(ii) and 10(iii). At the end only the main storage facility remains as a region in the figure 10(iv).

To illustrate effects of segmentation on the p-w constant reconstruction, consider figure 11(i) of the "Desert airplane". A choice of the scale gives the segmentation of figure 11(ii). Figure 11(iii) is the corresponding p-w constant approximation. Such reconstruction is useful if a shading has to be removed. Notice that this segmentation draws a box around the aircraft.

To introduce a true multichannel segmentation, a multisensor satellite set "Cairo" was obtained via Los Alamos Nat. Lab. Figure 12(i) and 12(ii) were collected with two distinct spectra devices. A fusion of a multisensor data is required since complementary information was captured. Figure 12(iii) is the two channel fusion. If in addition edge channels are fused, a quite different segmentation figure 12(iv) results (the edge map of 12(ii) is shown in figure 5). The implication of this example is a new way to fuse such distinct channels as infrared, optical, laser range.

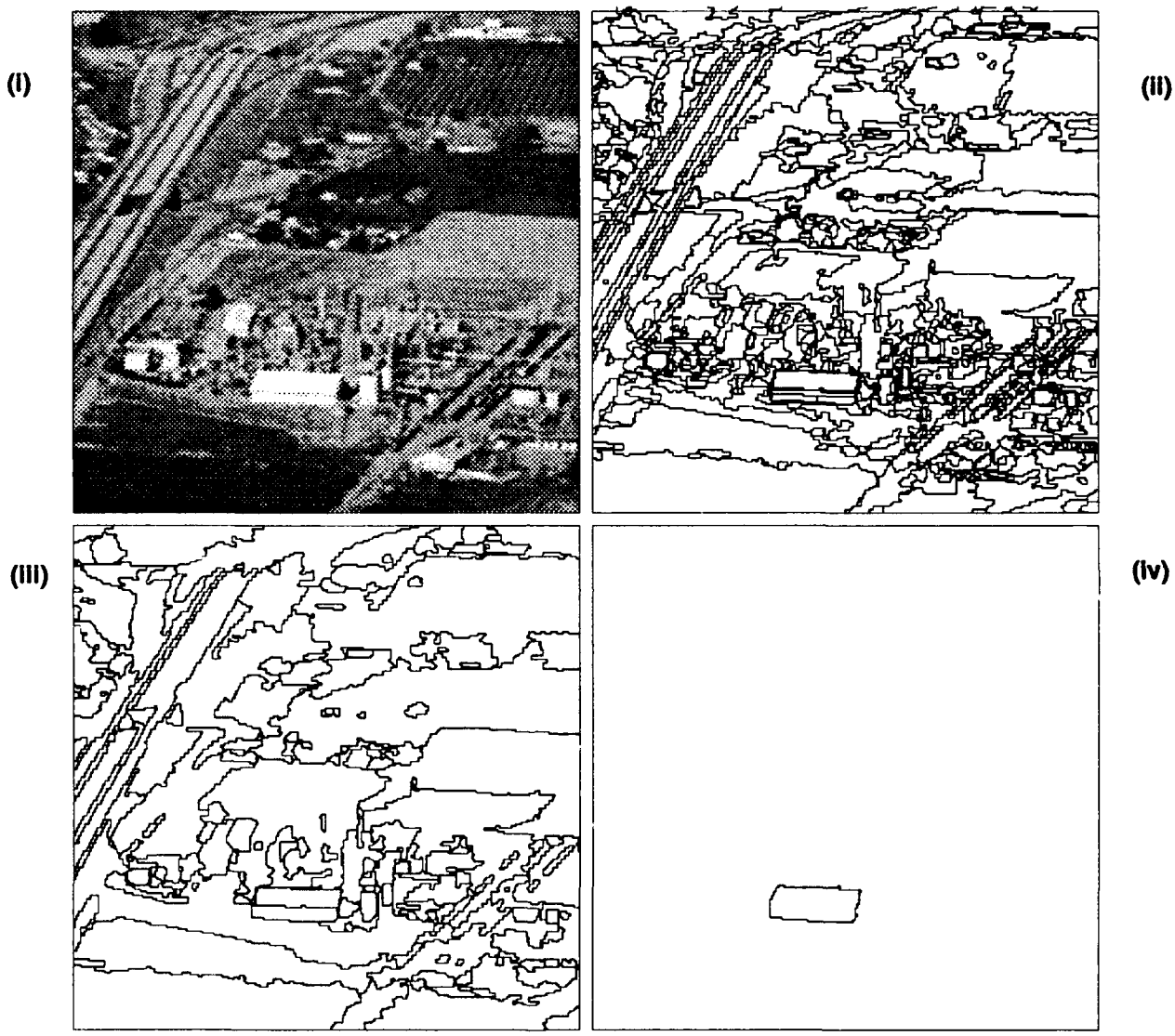


Figure 10: Different scales for "Chemical Plant" Segmentation.

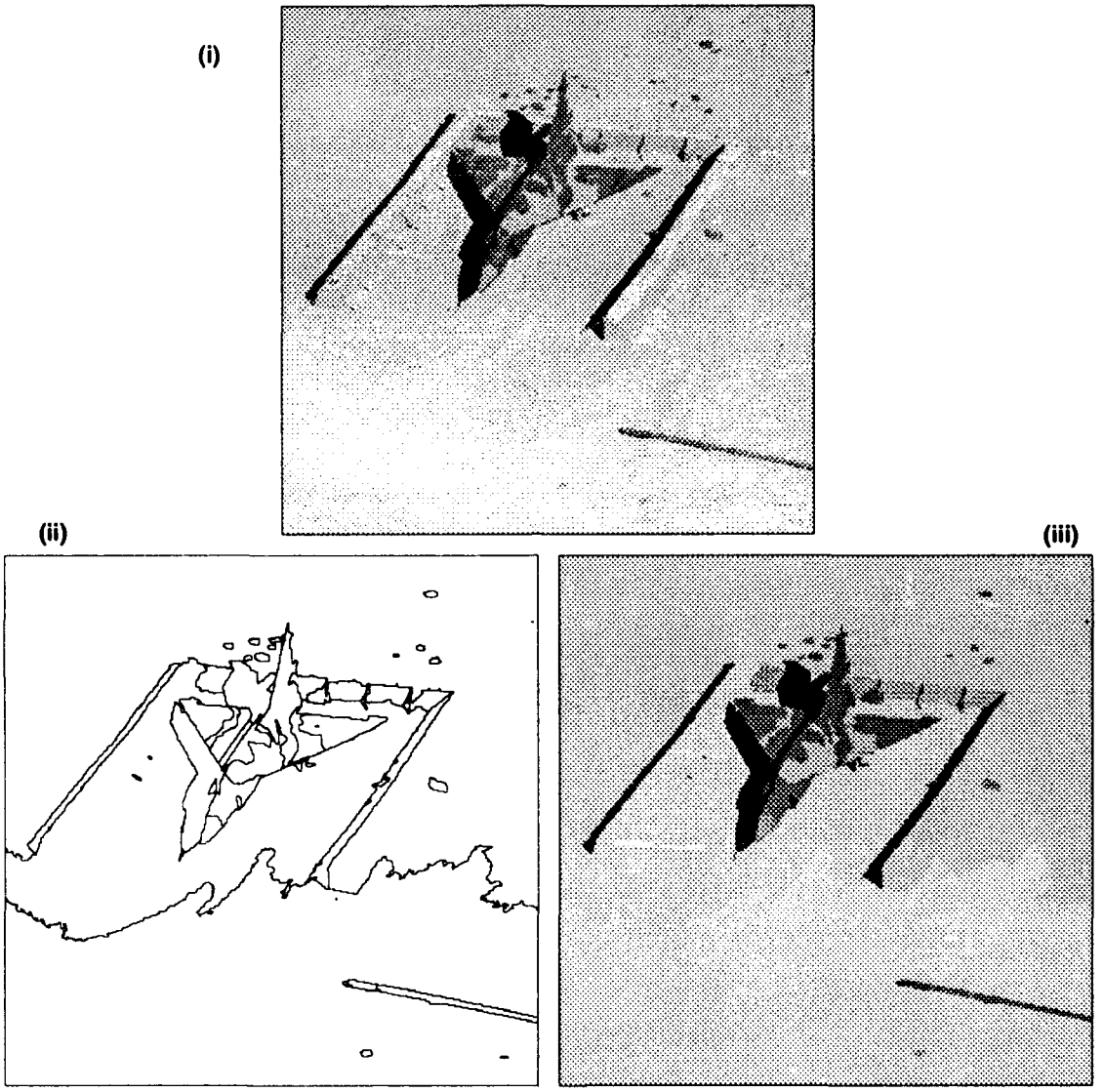


Figure 11: Results for the "Desert airplane".

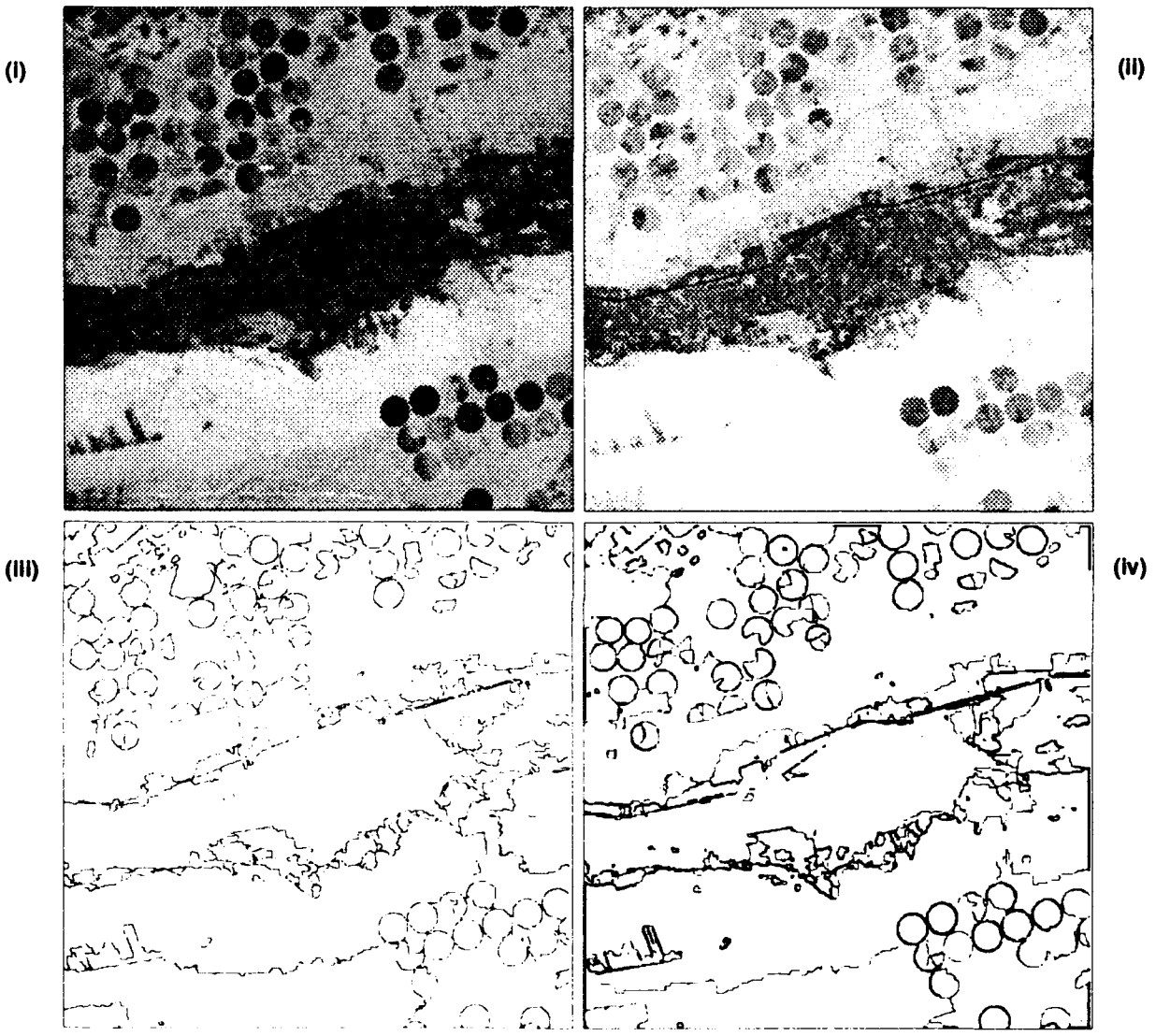


Figure 12: Example of multiple channel segmentation of the multisensor satellite image pair "Cairo".

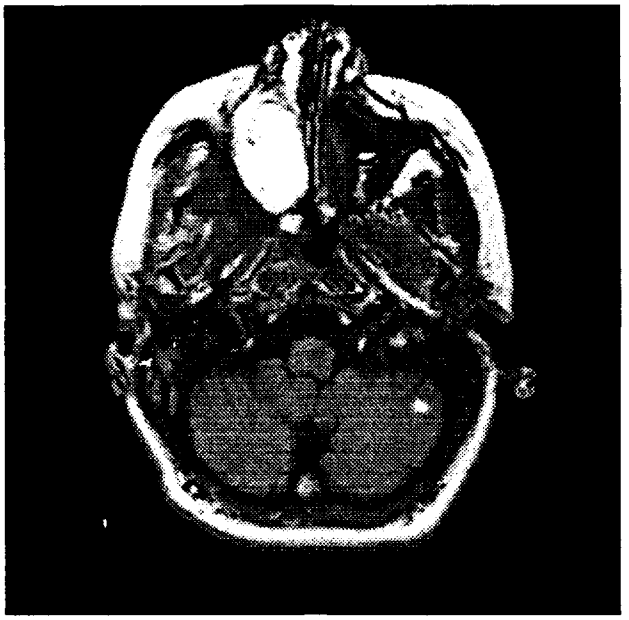
Medical imaging applications.

Rapid, accurate image segmentation can be applied to radiological imaging. Cognitech's algorithm offers a high degree of noise and threshold immunity and allows multi-band data , thus making it ideal for the tasks of multi-echo MRI sequences.

Figure 13(i) is 256*256 grey scale MRI reconstruction of the slice of brain of the patient with tumor. The pyramidal segmentation is applied to determine quantitative brain anatomy at a specified scale. The measurements of brain structures (area and surface) pointed with arrows on the figure 13(ii) are presented . This analysis yielded an accurate location and estimation of the tumor, shown here as a shaded area. In the future, Cognitech, Inc. will develop automated neuroanatomical tool based on the presented algorithm.

The set of figure 14(i) and 14(ii) is composed of MR sectional images. They represent image pairs taken under different MR protocols. The goal of segmentation is to combine this two sensor input in order to separate three anatomical strata: centrally located "butterfly" of the CSF (ceribro-spinal fluid) from the brains white matter and the gray matter.
Figures 14(iii) and 14(iv) show two independent segmentations, each having complementary as well as conflicting information. On the figure 14(v) a two-channel segmentation gives anatomically meaningful partition as well as accurate length/area measurements. These results are presented in Cognitech's [MRKMO] SPIE paper.

(i)



(ii)

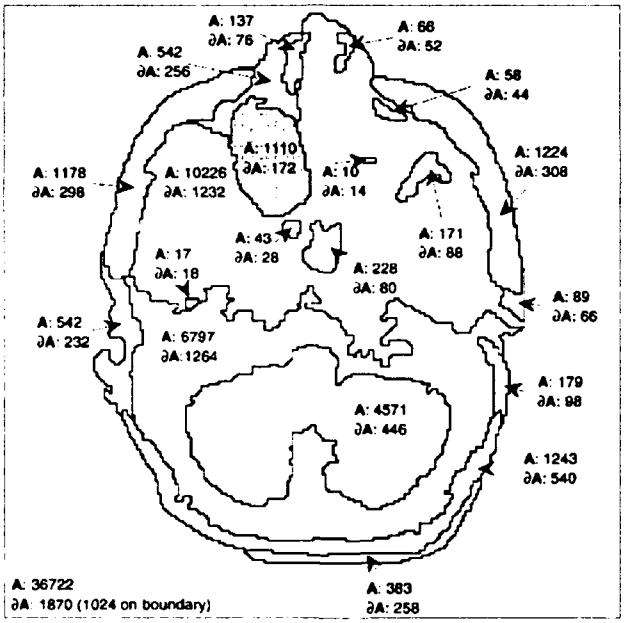


Figure 13: Segmentation of brain with tumor.

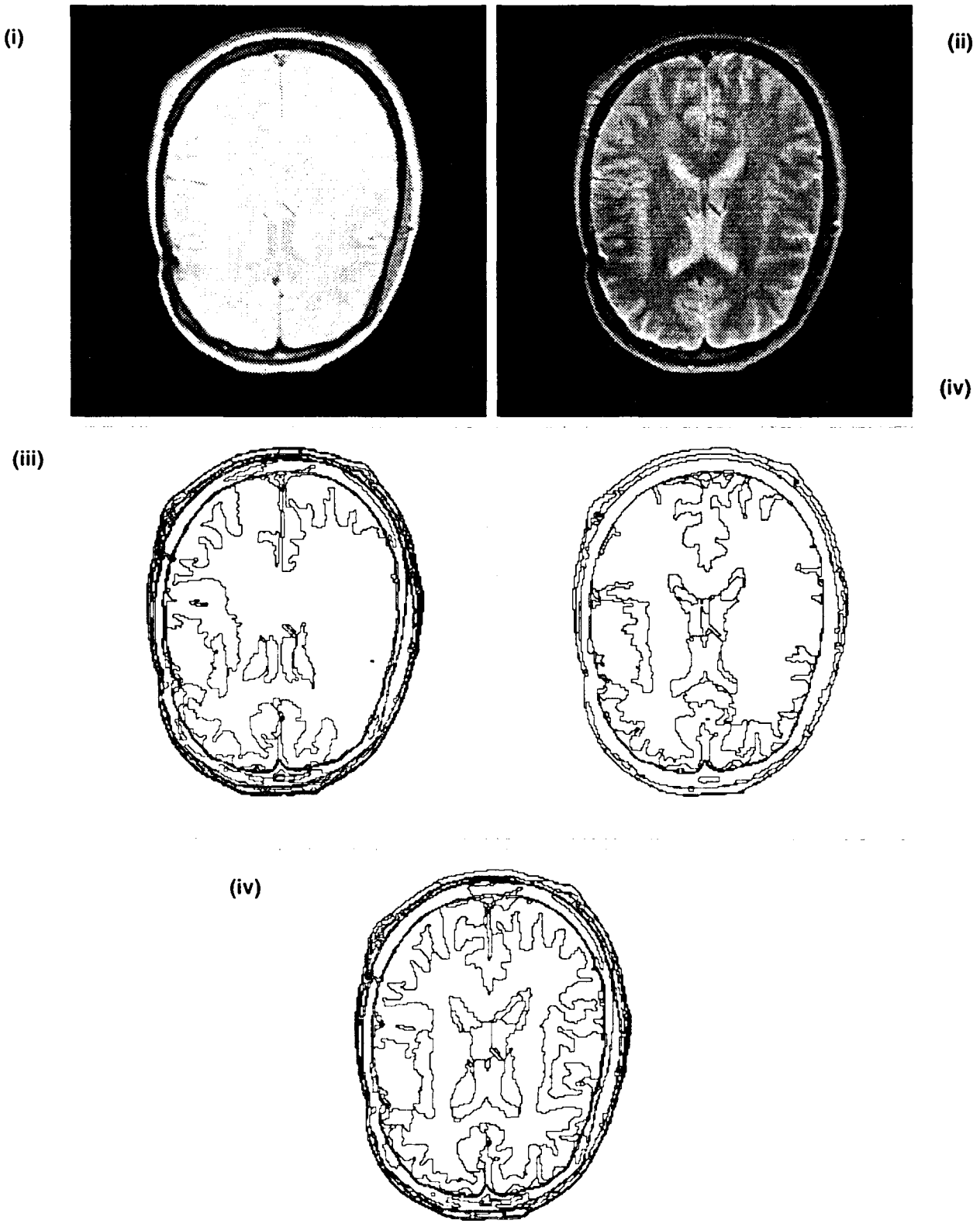


Figure 14: Multichannel segmentation of medical data.

Towards clutter removal algorithms.

Finally, to give some feeling for the future applications of the pyramidal, multichannel image analysis, the single gray scale channel segmentation of the figure 15 (512×512 "Tank") is compared to the segmentation based on 18 wavelet channels. A grey level segmentation and reconstruction was computed on figure 16(i) and 16(ii) correspondingly. The scale parameter was chosen large enough to eliminate the grassy background. The reconstruction 16(ii) shows that unfortunately most of the tank's features are also gone. Only gross scale outline features were preserved. This is the price of clutter removal with a single channel analysis.
In contrast the Haar basis wavelet decomposition as in section 3.4 above followed by the separation of a positive and negative parts yields 18 texture/feature channels. The fusion of these 18 channels is illustrated on the figure 16(iii). Remarkable tank reconstruction and simultaneous background clutter removal is presented on the 16(iv). Clearly, the multichannel algorithm was capable of filtering out the grassy terrain, while keeping the essential tank features, e.g treads. To our knowledge, this represents the state of the art.

Figure 15: Original grey level image "Tank".

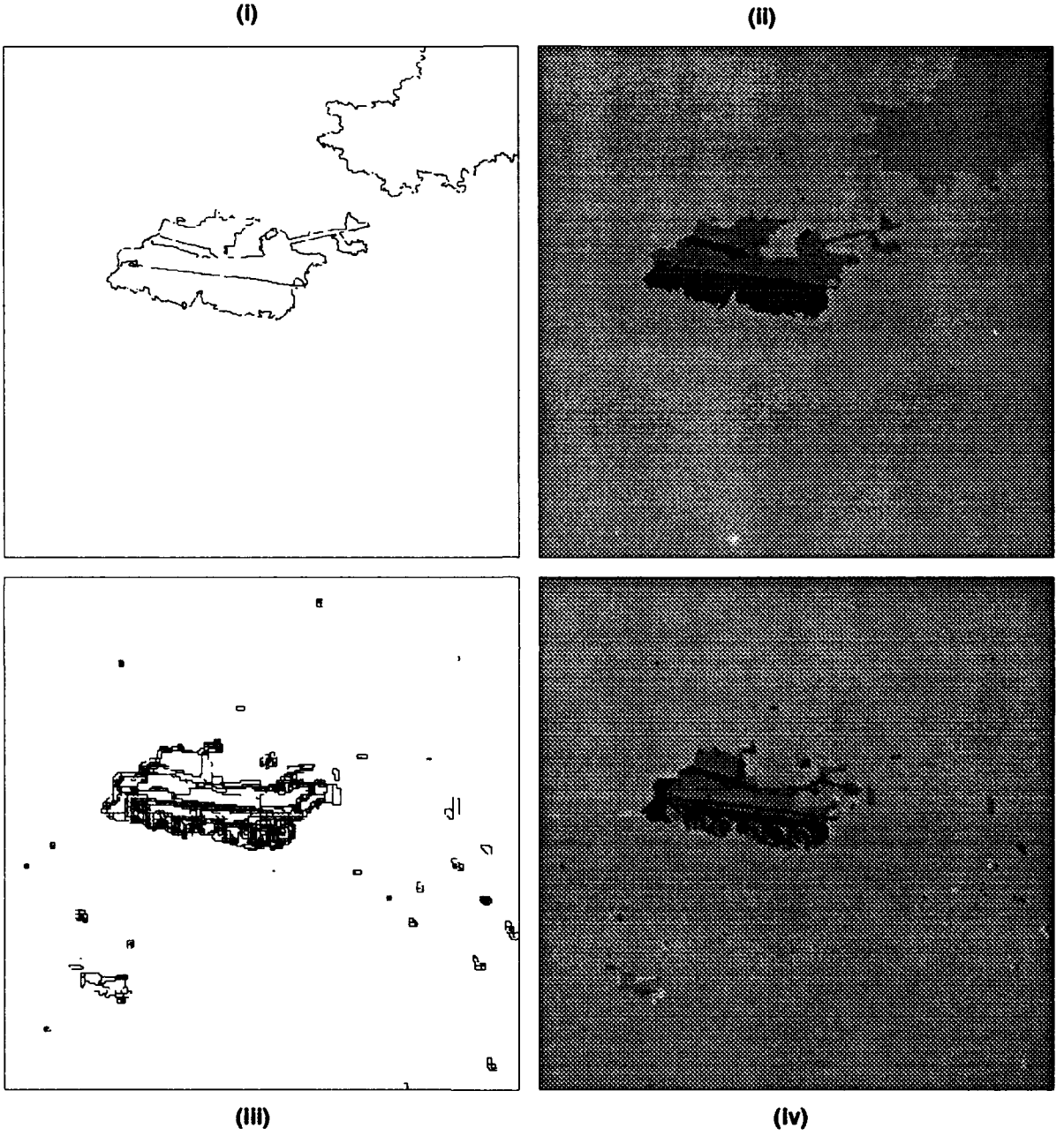


Figure 16: Non linear clutter removal via segmentation with a single channel and with multichannel based on wavelet analysis.

References

[AmbDeG] E. De Giorgi, and L. Ambrosio. *Un nuovo tipo di funzionale del calcolo delle variazioni.* To appear in Atti Acad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur.

[AmbT] L. Ambrosio, and V. M. Tortorelli. *Approximation of functionals depending on jumps by elliptic functionals via G-convergence.* Preprint. Scuola Normale Superiore di Pisa.

[Az] R. Azencott. *Synchronous Boltzmann machines and Gibbs fields; Learning algorithms.* To appear, Springer Verlag 1989, (NATO Series, Les Arcs Congress Proceedings).

[BecSI] J. Beck, A. Sutter and R. Ivry. *Spatial frequency channels and perceptual grouping in texture segregation.* Computer Vision, Graphics and Image Processing, 37, 299-325, 1987.

[BlakZ] A. Blake and A. Zisserman. *Visual Reconstruction.* MIT Press, 1987.

[CarDeGL] M. Carriero, E. DeGiorgi and A. Leaci. *Existence theorem for a minimum problem with free discontinuity set.* Arch. Rat. Mech. Anal., vol. 108, 1989.

[DeG] E. De Giorgi. *Free discontinuity problems in calculus of variations.* To appear, Proc. of the meeting in J.L. Lions's honour, Paris, 1988.

[FoT] I. Fonseca and L. Tartar. *The gradient theory of phase transitions for systems with two potential wells.* Carnegie Mellon Univ. To appear.

[GemG] S. Geman and D. Geman. *Stochastic relaxation, Gibbs distributions and the Baysian restoration of images.* IEEE PAMI 6, 1984.

[HarS] R.M. Haralick and L.G. Shapiro. *Image segmentation techniques.* Computer Vision Graphics and Image Processing, 29, 100-132, 1985.

[Ju] B. Julesz. *Texton gradients : the texton theory revisited.* Biological Cybernetics, 54 : 245-251, 1986.

[Koe] G. Koepfler. *Formalization and numerical analyses of image segmentation.* Ph.D. Thesis University of Paris-Dauphine, 75775 Paris, France.

[KoeMS] G. Koepfler, J.M. Morel and S. Solimini. *Segmentation by minimizing a functional and the "merging" methods.* Cahiers de Mathematiques de la Décision,n°9022,1990; presented in september 1991 at the '*GRETSI* Colloque' in Juan-les-Pins (France).

[MaliP1] J. Malik and P. Perona. *A scale space and edge detection using anisotropic diffusion.* Proc. IEEE Computer Soc. Workshop on Computer Vision, 1987.

[Mall] S. Mallat. *A theory for multiresolution signal decomposition : the wavelet representation.* IEEE PAMI, 674-693, July 1989.

[MaliP2] J. Malik and P. Perona. *Preattentive texture discrimination with early vision mechanisms.* Journ. of the Opt. Society. of America. A, 923-932, vol. 7, n°5, 1991.

[MRE] N.J. Mankovich, D.D. Robertson and J. Essinger. *Measuring the accuracy of CT bone geometry for orthopedic implants.* Preprint: Computer Assisted radiology, Berlin, July 1991.

[MRKMO] N.J. Mankovich, L.I. Rudin, G. Koepfler, J.M. Morel and S. Osher. *Application of a new pyramidal segmentation algorithm to medical images.*SPIE Medical Imaging VI: Image Processing.

[Marr] D. Marr. *Vision.* Freeman and Co. 1982.

[Mey] Y. Meyer. *Ondelettes et opérateurs.* Hermann, 1990.

[MorS] J.M. Morel and S. Solimini. *Segmentation of images by variational methods: a constructive approach.* Rev. Matematica de la Universidad Complutense de Madrid, Vol.1, n°1,2,3, 169-182, 1988.

[MumS1] D. Mumford and J. Shah. *Optimal Approximations by Piecewise Smooth Functions and Associated Variational Problems.* Communications on Pure and Applied Mathematics. vol. XLII No.4, 1989.

[MumS2] D. Mumford and J. Shah. *Boundary detection by minimizing functionals.* Image Understanding, 1988, Ed. S. Ullman and W. Richards.

[Pav1] T. Pavlidis. *Segmentation of pictures and maps through functional approximation.* Comp.Gr. and Im.Proc. 1, 360-372, 1972.

[Pav2] T. Pavlidis. *Structural Pattern Recognition.* Springer, New York 1977.

[Pav3] T. Pavlidis. *Algorithms for graphics and image processing.* Computer Science Press, 1982.

[PavL] T. Pavlidis and Y.T. Liow. *Integrating Region Growing and Edge Detection.* Proc. of the IEEE Conf. Comp. Vision and Patt. Recognition 1988.

[Rud] L. Rudin. *Images, Numerical analysis of singularities and shock filters.* CALTECH Tech. Report 5250, 1987.

[Zu] S.W. Zucker. *Region growing: Childhood and Adolescence (Survey).* Comp. Graphics and Image Proc. 5, 382-399, 1976.