1999

69 (1)

M.Tech. (Computer Science) Dissertation Report

SIMULATION OF SECOND GENERATION IMAGE COMPRESSION ALGORITHM

By

Champak Kumar Dutta

Under the Supervision of

Prof. Bhabatosh Chanda

and

Prof. Malay Kumar Kundu

Indian Statistical Institute
203, Barrackpore Trunk Road
Calcutta - 700035

JULY, 1999

Certificate of Approval

This is to certify that the thesis entitled *Simulation of Second Generation Image Compression Algorithm* submitted by *Champak Kumar Dutta*, in partial fulfillment of the requirements for M.Tech. in Computer Science degree of the Indian Statistical Institute, Calcutta, is an acceptable work for the award of the degree.

Date:

(Supervisor)

Prof. Bhabatosh Chanda
Electronics and Communication Sciences Unit
Indian Statistical Institute
203, Barrackpore Trunk Road
Calcutta - 700035

Date:

(Supervisor)

Prof. Malay Kumar Kundu
Machine Intelligence Unit
Indian Statistical Institute
203, Barrackpore Trunk Road
Calcutta - 700035

Date: 28/7/99

(External Examiner)

Acknowledgements

The persons whom I would like to pay my heartfelt thanks and gratitude and whom I feel myself honored to have worked with are my supervisors *Prof. Bhabatosh Chanda* and *Prof. Malay Kumar Kundu* for their guidance, advice, enthusiasm and support throughout the course of this dissertation.

I would like to take this opportunity to thank *Prof. Bhargab B. Bhattacharya, Prof. Aditya Bagchi, Dr. Nabanita Das, Prof. Bhababani Prasad Sinha and Prof. C.A.Murthy* for the excellent and valuable courses they have offered in M.Tech. which help me in this work.

I would also like to thank to all C.S.S.C. staffs for helping me to avail the computing facilities and also giving their valuable advises, whenever I have faced any problem.

Lastly, I should thank my classmates for their helpful suggestions throughout the project and for sharing those moments of joy and frustration making my two years stay at I.S.I. enjoyable.

I.S.I., Calcutta,
July 1999

(Champak Kumar Dutta)

Abstract

Image coding is,in general,a two stage process commonly stated as: modeling the image data and coding of the model parameters.All of the second generation compression schemes mimic the properties of human visual system in their modelling stage.In the segmentation based techniques ,the modeling stage is based on the fact that the eye is good at identifying regions that appear similar and grouping them accordingly.In this work,this phase has been highlighted and also possible further works in this direction has been indicated

Contents

Abstract

Image coding is,in general,a two stage process commonly stated as: modeling the image data and coding of the model parameters.All of the second generation compression schemes mimic the properties of human visual system in their modelling stage.In the segmentation based techniques ,the modeling stage is based on the fact that the eye is good at identifying regions that appear similar and grouping them accordingly.In this work,this phase has been highlighted and also possible further works in this direction has been indicated

Chapter 1

INTRODUCTION

As the amount of information that is needed,desired,and available increases,the need for more efficient ways of storing this information increases as well.The goal of data compression is to provide the most efficient way to represent information.This goal is accomplished by developing techniques to exploit the different kinds of structures that may be present in the data.The information can be in a variety of forms,such as speech,images,text,video and so on.Different forms of information have their own specific type of structures.They also share characteristics that can often be exploited to develop techniques that have relevance to all kinds of informartion.

There exists a variety of methods in literature that use different kinds of structures present in data from different sources to provide efficient representations of data. The majority of the conventional image coding techniques rely on methods based on classical information theory to exploit the redundancy in the images in order to achive compression and in these methods performances at high compression are limited because the processing does not take into account the object geometry or the discontinuities.Even with a lossy technique a higher compression ratio is achieved only at the expense of image quality.Attempts have been made to develop new image compression techniques that outperform the conventional image coding techniques (called first generation methods).These methods identify features within the image in an an elegant way and use these features to achieve compression.These developments are termed as second generation image coding .All second generation techniques incorporate the characteristics of the human visual system into the coding strategy in order to achieve high compression ratio and still maintaining acceptible quality of the image.In other words,most of the techniques are of a lossy nature:however they attempt to identify and seperate visually significant and insignificant areas of the image and apply appropriate coding technoque to each area. For this purpose much effort has been expended in identifying what the human observer considers visually important.It is generally concluded that edge information is vital to the human perception of images.As a result ,a vast majority of the work in the development of image coding techniques has concentrated on methods that preserve edge information and seperate edge and texture information while coding them seperately.

In a nutshell,most of the second generation methods follow three steps :

- The segmentation that splits the data into various homogeneous components corresponding to as much as possible to semantic units(each component may consist of homogeneous regions,predefined visual patterns,directional components,pyramidal components etc.).
- The contour coding that consists of coding the information about the partition of the space.
- The texture coding that deals with the information inside each region.

Chapter 2

COMPRESSION SCHEME

One of the promising second generation technique relies on a hierarchical morphological segmentation algorithm for image coding.The algorithm follows a Top-Down proceedure.It first takes into account the global information and produces a coarse segmentation,that is,with a small number of regions. Then the segmentation quality is improved by introducing regions corresponding to more local information.Each segmentation stage relies on four basic steps: simplification,marker extraction,decision,and quality estimation.The simplification removes the information from the image to make it easier to segment. The marker extraction identifies the presence of homogeneous regions.It is based on constrained flat region labelling .The goal is to precisely locate the contours of regions detected by the marker extraction.This decision is performed by a modified watershed algorithm.Finally the quality estimation concentrates on the coding residue all the information about the regions that have not been properly segmented and therefore coded.This proceedure allows the introduction of the texture and contour coding schemes within the segmentation algorithm.The coding residue is then transmitted to the next segmentation stage to improve the segmentation and coding

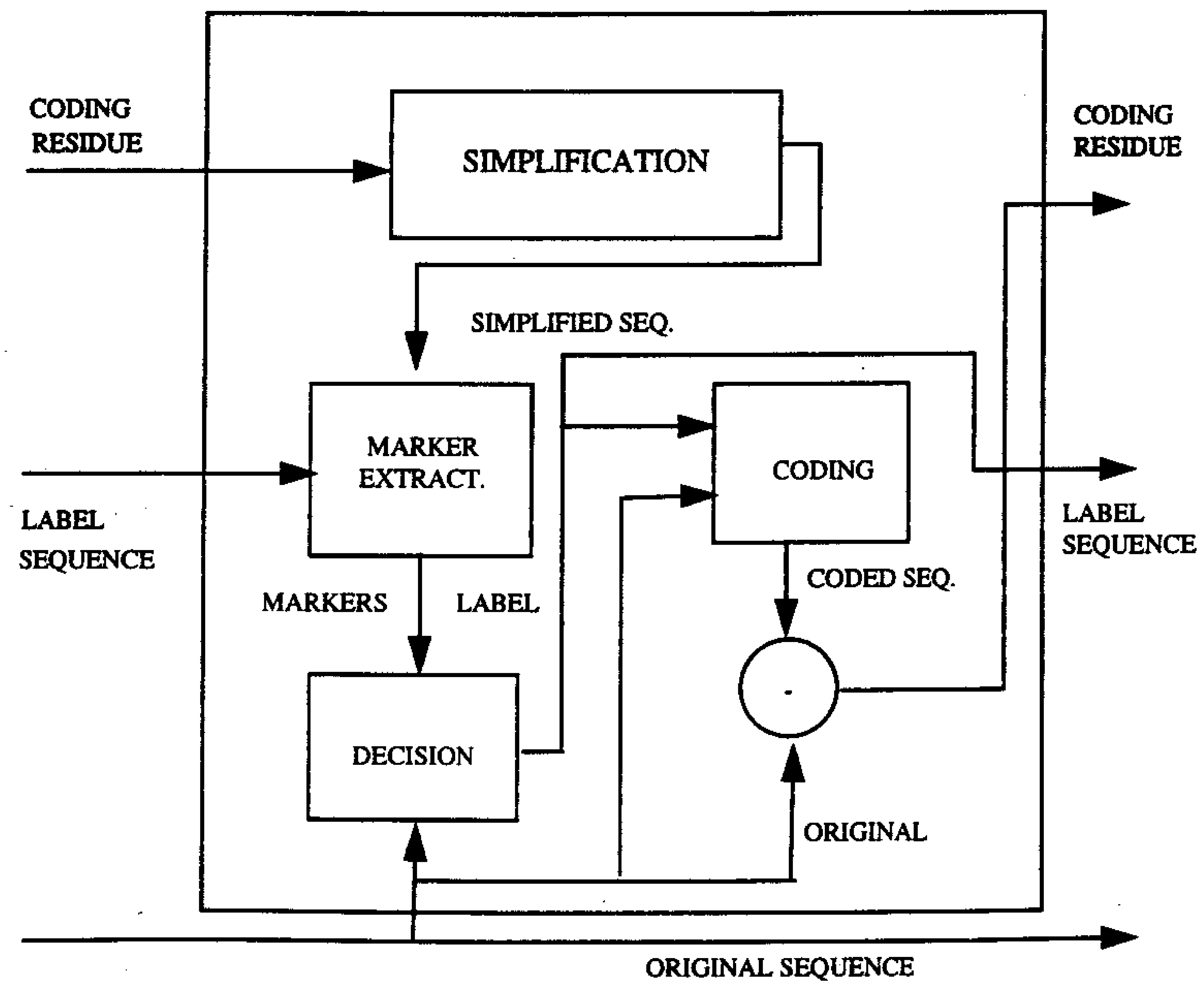


Fig. 1 *Hierarchical structure for segmentation and coding.*

Chapter 3

MORPHOLOGICAL TOOLS

3.1 TOOL1 :FILTERS

The word morphology commonly denotes a branch of biology that deals with the form and structure of living beings.In the context of image processing, mathematical morphology acts as a tool for extracting image components that are useful in the representation and description of region shape,such as boundaries,skeletons and convex hull. For preprocessing and postprocessing of images morphological techniques e.g. filtering,pruning,thining etc. offers a unified and powerful approach.

The language of mathematical morphology is set theory.Sets in mathematical morphology represent the shapes of an object in an image.For example,gray scale digital images can be represented as sets whose components are in three dimensional integer space.In this case ,two components of each element of a set refer to the coordinates of a pixel and the third corresponds to its discrete intensity value.Several operations on these sets are defined to manipulate the images. A detailed theoritical development of these have been done in [5].Some of the related concepts are reviewed in the remaining part of this section.

A few basic notations and their significances are summarised in the following.Here f(x) denotes an input signal and M_n is astructuring element of size n.

Erosion	$\epsilon_n(f)(x) = Min\{f(x+y), y \in M_n\}$
Dilation	$\delta_n(f)(x) = Max\{f(x-y), y \in M_n\}$
Geodesic Dilation of size one	$\delta^{(1)}(f,r) = Min\{\delta_1(f), r\}$
Geodesic Erosion of size one	$\epsilon^{(1)}(f,r) = -\delta^{(1)}(-f,-r)\}$
Reconstruction by dilation	$\gamma^{(rec)}(f.r) = \delta^{(\infty)}(f,r) = \ldots\delta^{(1)}(\ldots\delta^{(1)}(f,r)\ldots,r)$
Reconstruction by erosion	$\phi^{(rec)}(f.r) = \epsilon^{(\infty)}(f,r) = \ldots\epsilon^{(1)}(\ldots\epsilon^{(1)}(f,r)\ldots,r)$
Morphological opening	$\gamma_n(f) = \delta_n(\epsilon_n(f))$
Morphological closing	$\phi_n(f) = \epsilon_n(\delta_n(f))$

Opening by reconstruction of erosion $\gamma^{(rec)}(\epsilon_n(f), f)$

Closing by reconstruction of dilation $\phi^{(rec)}(\delta_n(f), f)$

Opening by partial reconstruction $\gamma^{(rec)}(\epsilon_n(f), \gamma_k(f))$

Closing by partial reconstruction $\phi^{(rec)}(\delta_n(f), \phi_k(f))$

Morphological gradient $g = \delta_1(f) - \epsilon_1(f)$

Dilation and erosion are two operations which act as a basis of the whole set of significant morphological operators.Dilation by disk structuring elements corresponds to isotropic swelling or expansion of an image while the erosion transformation is popularly conceived of as a shrinking of the original image. The choice of the basic dilation of size one defines the notion of connectivity and neighbourhood.

A morphological opening simplifies the original signal by removing the bright componenrs that that do not fit within the structuring element.Similarly closing removes dark regions which are small compared to the size of structuring element.These filters can be used as simplification tool before segmentation,but they do not allow perfect preservation of the contour information.

In order to improve the contour preservation properties filters by reconstruction can be used. In case of opening by reconstruction of erosion ,the simplification is performed by the erosion that eleminates the bright components that are smaller than the size of structuring element. Then,the reconstruction process restores the contours of the components that have not been totally removed by the erosion.Images obtained with this type of filters are good starting point for segmentation step because these aremuch simpler than the original images,but the objects that are present are precisely defined.

But if the moving objects are considered then above mentioned filter may pose a problem because the reconstruction process may artificially connect some point on the background with the object points.This happens due to noise in the background mainly caused by the effect of motion blur. To get rid of this problem partial reconstruction can be used. Here the reference signal in reconstruction process may be smoothed by a small opening with structuring element size k,say.The parameter k allows a smooth tuning from no reconstruction(k=n in the above-mentioned set of definition) to full reconstruction (k=0).

To qualitatively assess the quality of the filters presented above synthetic test sequences have been used. These are composed of moving objects(triangles,polygons etc.) on a constant background and they are corrupted by noise.The test sequemces are first simplified by morphological filter and then segmented (to have coherent results , the segmentation relies on the watershed algorithm, described in the next chapter).The result of optimal segmentation of synthetic sequences is apriori known.Once a test sequence has been segmented ,two parameters are measured.

- Edge localization : This parameter is defined as the number of pixels differing between the current and optimal segmentation divided by the area of the objects to segment. It measures the contour preservation property of the filter.
- Flatness: This parameter is the varience of the simplified signal inside each segmented

region. It measures the efficiency of the filter to produce flat, and therefore,easily segmentable,regions.

Each measure is plotted on a plane(edge localization/flatness).For each filter a set of measures is obtained by modifying the structuring element size.They create a curve in the edge localization/flatness plane. Ideally a good simplification has a low edge localization and flatness parameters.It has been reported in [1] that, the median and open-close filters produce a poor flatness and their contour preservation is [redacted] bad.The morphological opening by reconstruction of erosion achieves a much better flatness .In case of images of moving objects the use of partial reconstruction filters is preferred.

3.2 TOOL2 :WATERSHED ALGORITHM

Segmentation is one of the key problem in image processing and there exist as many techniques as there are specific situations.The techniques related to gray-tone image segmentation may be divided into two groups :the techniques based on contour detection and those involving region growing. An original method of segmentation based on the use of watershed lines has been developed in the framework of mathematical morphology. This technique , which may appear to be close to region growing methods , leads in fact to a general methodology of segmentation and has been applied with success in many different situations.

The watershed algorithm derives from topographic works and it is a segmentation based on signal minima. In image processing this notion has been introduced by considering gray-level values of a picture as an altitude of an imaginary relief.The *catchment basin* associated with a regional minima of a gray level image regarded as a topographic surface refers to the locus of the points ~~textit~~p such that a drop falling at p slides along the surface until it reaches the minima. The *watershed lines* of a gray scale image are the lines which seperate various catchment basins of that image. Hence watershed lines partition the space by associating a catchment basin surrounding each local minimum . A large number of algorithms have been proposed for the efficient computation of watershed.I have followed the immertion simulation stated in [1](See figure later).

Immersion simulation consists in flooding the surface from its local minimum. Starting from the local minimum of lowest altitude ,the water progressively fills up the catchment basins.When the water level reaches the altitude of other minima,these minima start to be active,and the flooding process also originates from the minima.Now,when the water coming from two different minima would merge,an imaginary dam is created to prevent the mixing of water. The proceedure is ended when the water level is higher than the absolute maximum. In this case each minimum is surrounded by water,i.e. its catchment basin,and a dam delimiting its border,i.e. its watershed line.The catchment basins constitute the partition of the space.

Hence the immersion simulation relies on a double ordering.First,a point at a given altitude will be flooded after the points of lower altitude,and second,at a given altitude,points that are close to some flooded points will be flooded before points that are far away. The hierarchical queue structure referred in [1] implements this double ordering in an efficient way. It is a set of queues with different priorities and each queue is a FIFO data structure.The elements processed by the queue are pixel positions.Pixels are put into one of the queues depending on a notion of priority.The first pixel to be pulled out of the queue is the first one that has entered the queue with highest priority.Then successively all pixels in the queue of highest priority are extracted.Finally,if the queues of highest priority are empty,the first pixel to be extracted is the first pixel of the first nonempty queue.

The immersion algorithm can be simulated with these queues.The algorithm works in two distinct steps :

- **Initialization:** This proceedure consists of putting the locations of all local minima in the queue with the gray level value as their priority
- **Flooding:** This proceedure consists of extracting a pixel from the queue.If the pixel does not yet belong to a catchment basin it is guaranted by the filling process that it has at least one neighbour in a catchment basin.Therefore all such neighbours are examined and the pixel is assigned to the catchment basin corresponding to the closest gray-level value. Then if the current pixel has some neighbours that do not belong to any catchment basin and that are not already in any queue, these neighbours are placed in the queue with a priority defined by its gray-level value.

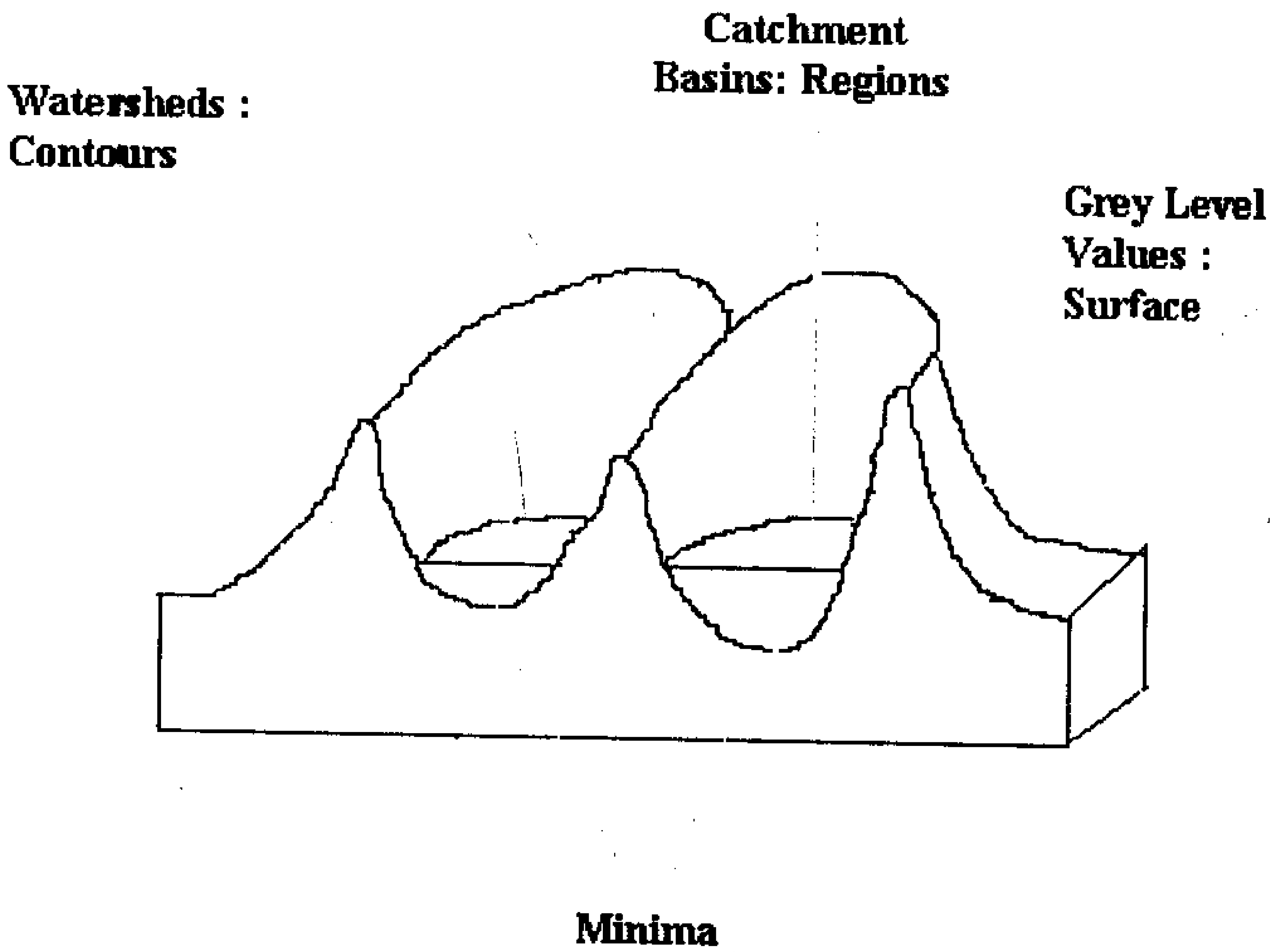


Fig 2: Catchment basin and watershed lines

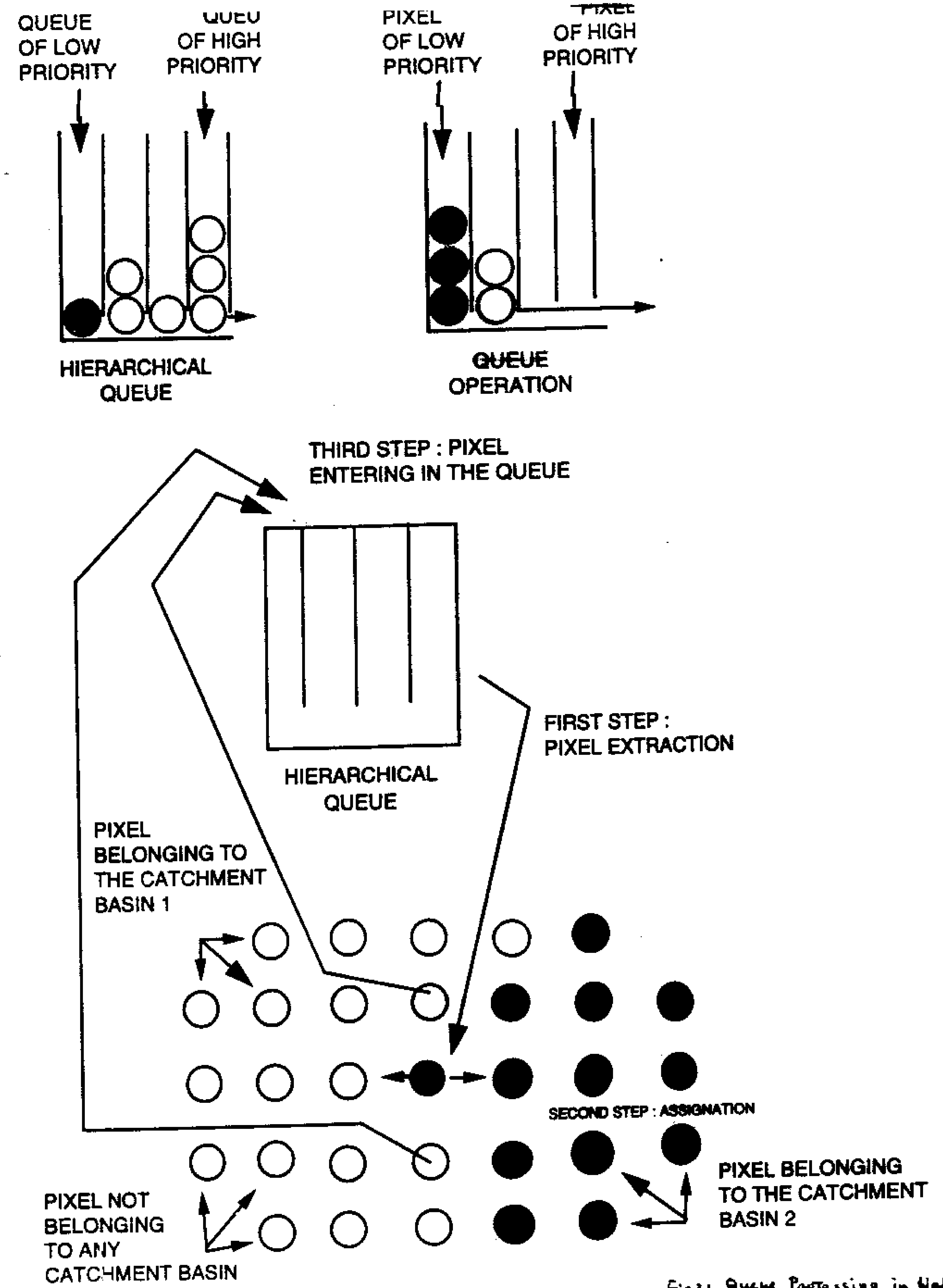


Fig 3: Queue Processing in Watershed algorithm.

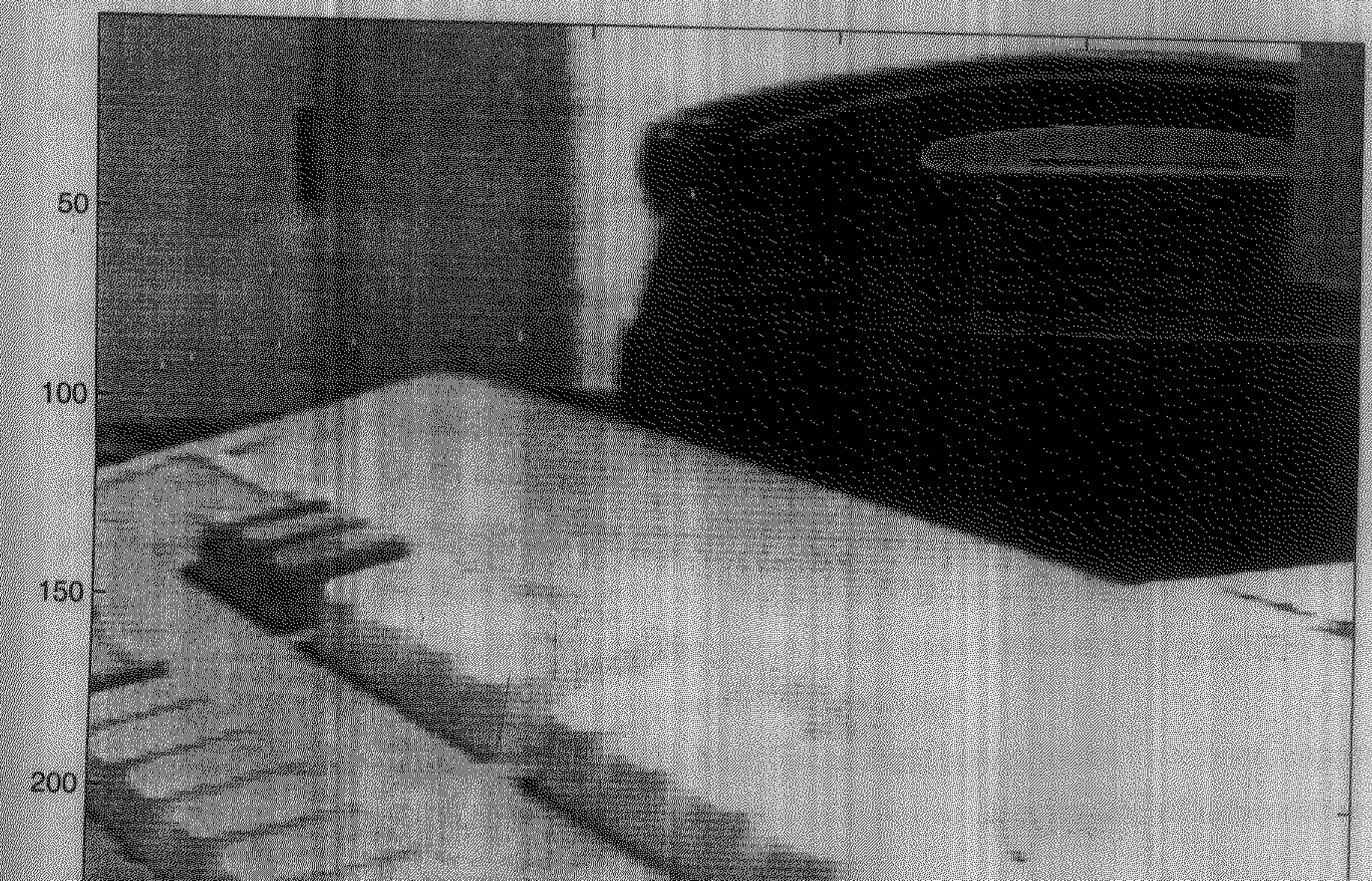
Chapter 4

SEGMENTATION SCHEME

My work consists of segmentation of a gray level image with a review of the algorithm presented in [1]. The image I have selected as my working sample is displayed in the next page.

Original Image

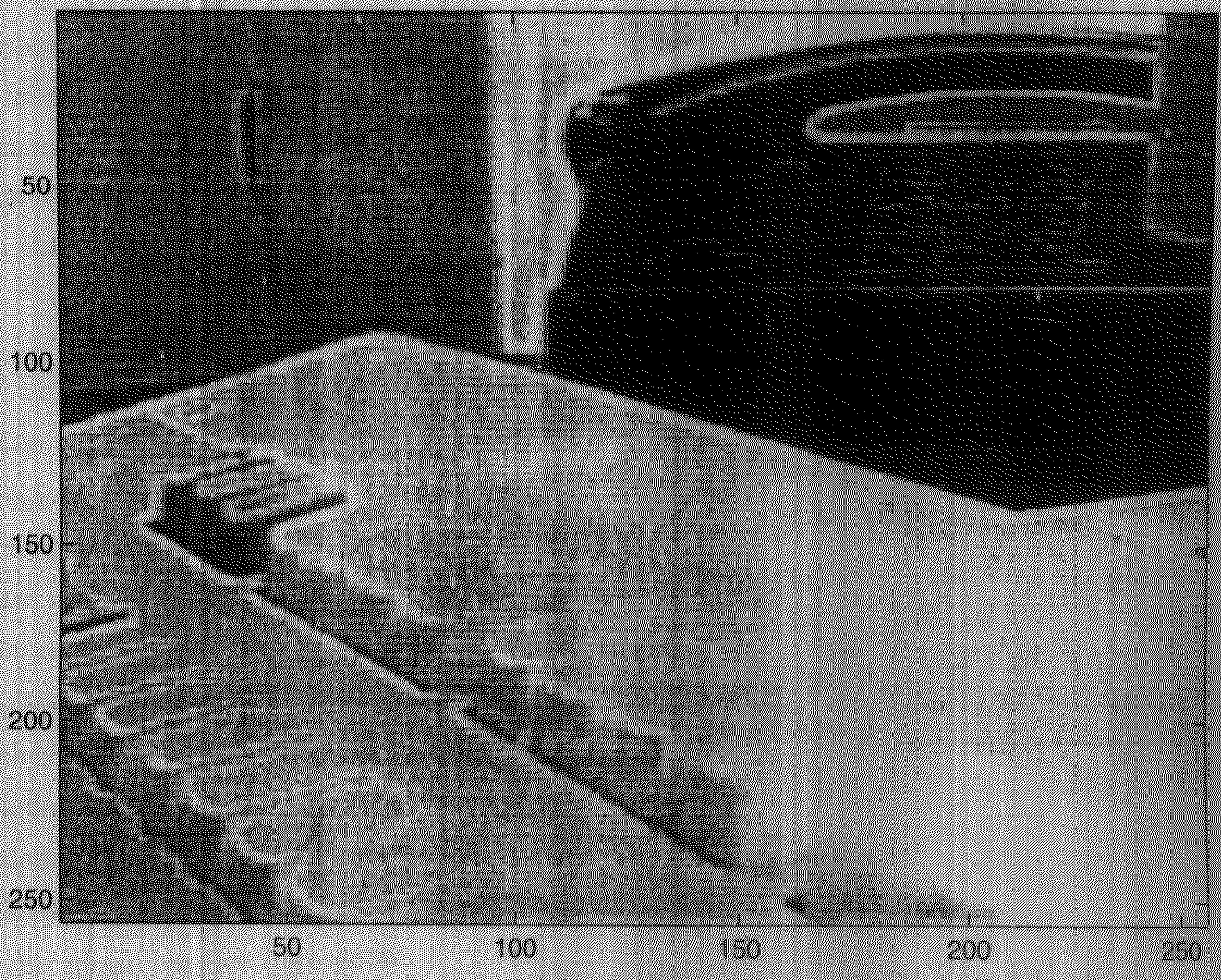
The first step of this work involves the simplification of the original image which controls the type and amount of information that is removed from the image to make the task of segmentation easier. As I am working with the still images I have chosen the opening by reconstruction of dilation as my filter.It is good enough to provide a pretty good result so far as simplification is concerned.The structuring element I have chosen is 10×10 square block.The resultant output is presented here.



Simplified Image

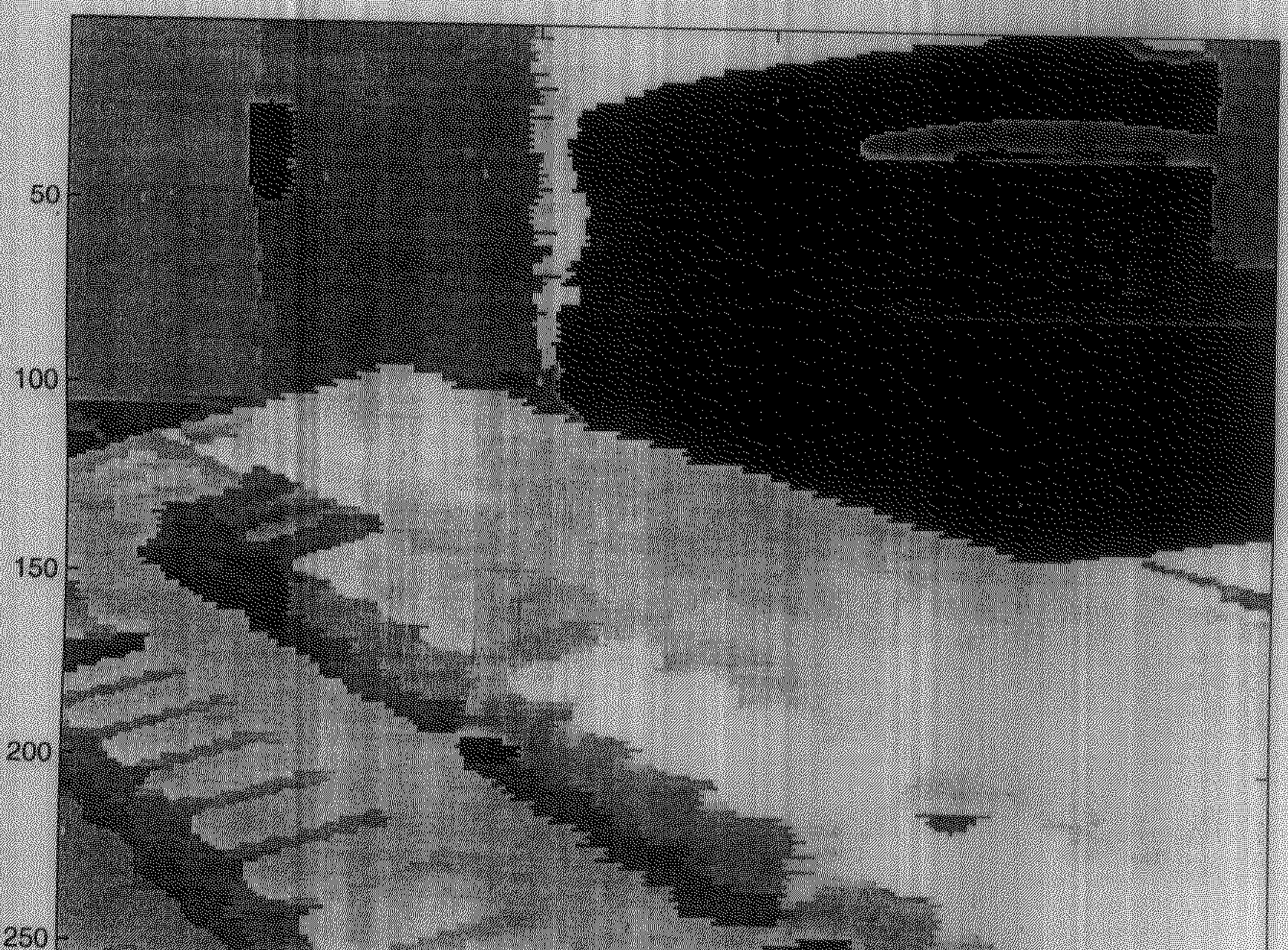
In the next stage the segmentation is carried out with the help of the marker extraction technique which aims at finding homogeneous units in the simplified image. These regions can very simply be identified by labeling flat regions, that is, by labeling the connected components of the space where the function is of constant gray-level value. Labeled markers are gray-level signals identifying tyhe presence of homogeneous regions that will be precisely delimited by the decision step. The interior of each bomogeneous region is "marked" by a label, that is, a constant gray lavel value, which is unique for this region. Moreover, the zones that are not homogeneous or in between two homogeneous regions are not considered as "marked" .

While working with the simplified image as the input to this stage I have found that if the image does not contain sufficient number of minima then the result obtained after segmentation is not satisfactory . For this reason I have constructed a gradient image (as suggested in[2]) from the simplified image which does not suffer from the scarcity of minima but this has not been treated as the input image in the marker extraction step. Use of gradient has been criticized in [1] in the sense that gradient image looses some information which are present in the original image specially if the original image is of a moving object.But in case of still image gradient image does not suffer from the huge amount of loss.Instead,it helps to restore the edge information to some extent which have been lost in the simplification stage.The output of this processing is shown in the next page.
Hence minima has been found from the gradinet image and these minima have been imposed on original image.Original image has been treated as input so that no loss of information is incurred.

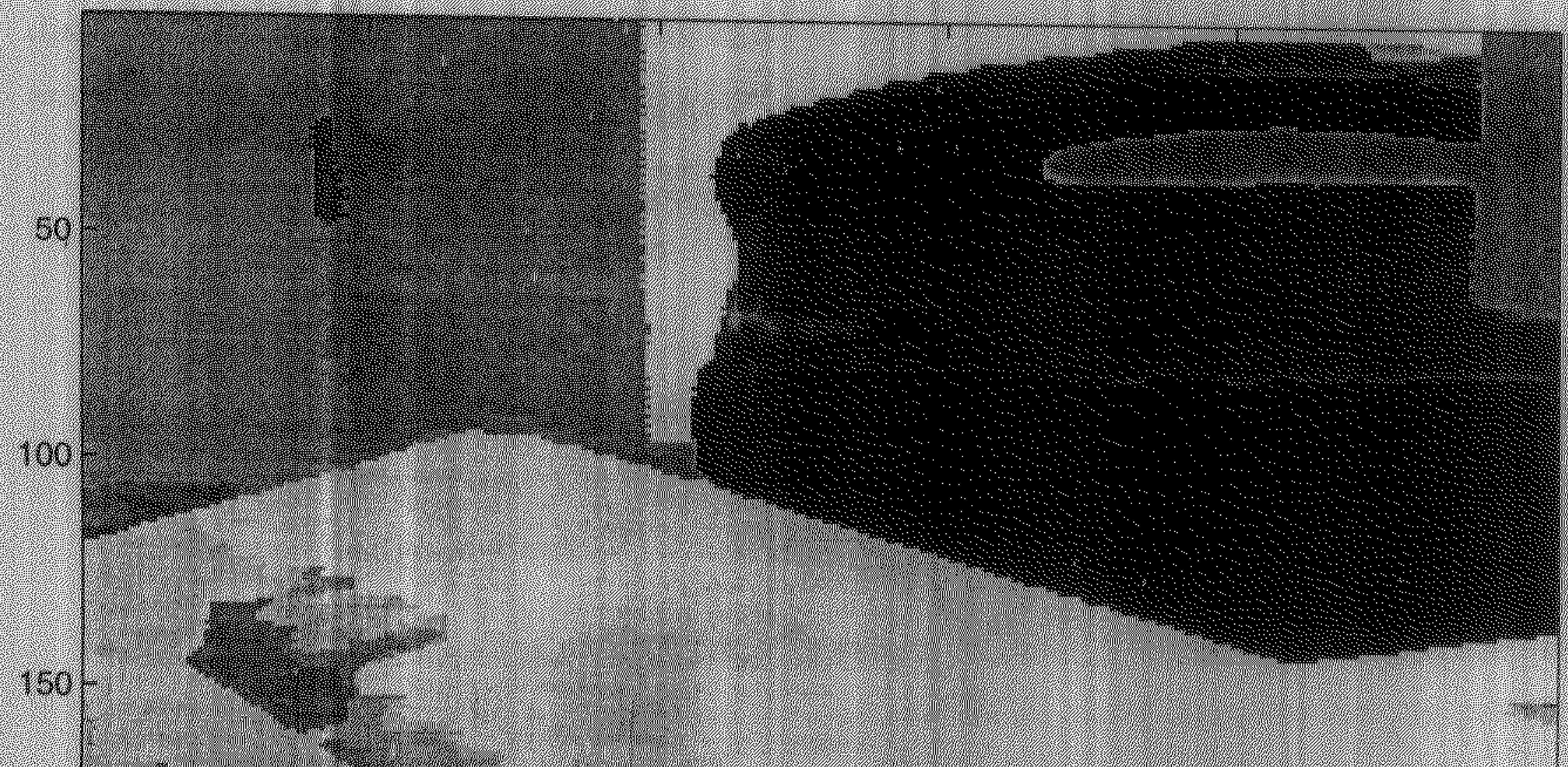

50
100
150
200
250
50
100
150
200
250

The labelling technique I have used, follows the basic principles of watershed algorithm and it is briefly described as follows.
As in the case of watershed, a very efficient implementation of a labeling algorithm uses a queue structure (hierarchical queue with one level of hierarchy): The first pixel of the sequence is put in the queue as will as its neighbors that have the same gray-level value. When a pixel is extracted from the queue, because of the filling procedure, we know that this pixel has at least one neighbor of same gray-level value that has already a lebel. This label is assigned to the current pixel. Then, all its neighbors of the same gray-label value without a label are put in the queue, and the procedure is iterated until the queue is empty. When the queue is empty, it means that the current flat zone has been entriely labeled; therefore, the label number is increased, and the first nonlabled pixel of the signal is searched and put in the queue. Finally, the image has been labeled when all pixels have a label.

At this point it is noted that a flat region can very well be composed of a single pixel. The flat regions of interest are not all flat regions but, because of the simplification filter, flat regions of size larger than a minimum. Indeed, the simplification filter has removed signal components smaller than a given limit. Therefore, flat regions of size smaller than this limit are not of interest and comstitute uncertainty areas, that is transition pixels between two large flat zones. As a result the labeling algorithm has to check not only that the current pifxel and its neighbors are of the same gray lavel value (that is the local flatness) but also that they correspond to the same region with respect to the current segmentation. Moreover, once a flat region has been labeled, if its size (number of pixels) is smaller than the minimum size defined by the simplification, its label is removed and the region is considered as uncertainty area. This labeling algorithm can be viewed as a constrained labeling algorithm. Output of this phase has been presented.



Result after initial labelling (with input image original one)



Result (with Gradient Image as input) after labelling

Once the markers have been defined, the decision can be taken by a modified watershed algorithm.In the algorithm presented earlier the priority of pixel was defined by its gray-level value: A high (low) priority was assigned to a dark (bright) pixel. We note that independently of the time instant a pixel is introduced in the queue, it will always have the same priority. In this modified watershed, the priority is defined as the degree of certainty with which a pixel belongs to a given region. Let us call this distance,the degree of certainty. Of course, various distances and, therefore, priorities can be used. In the following, the distance is defined as the absolute difference between the assigned label value of the pixels and the label values of the pixels belonging to the neighboring region.

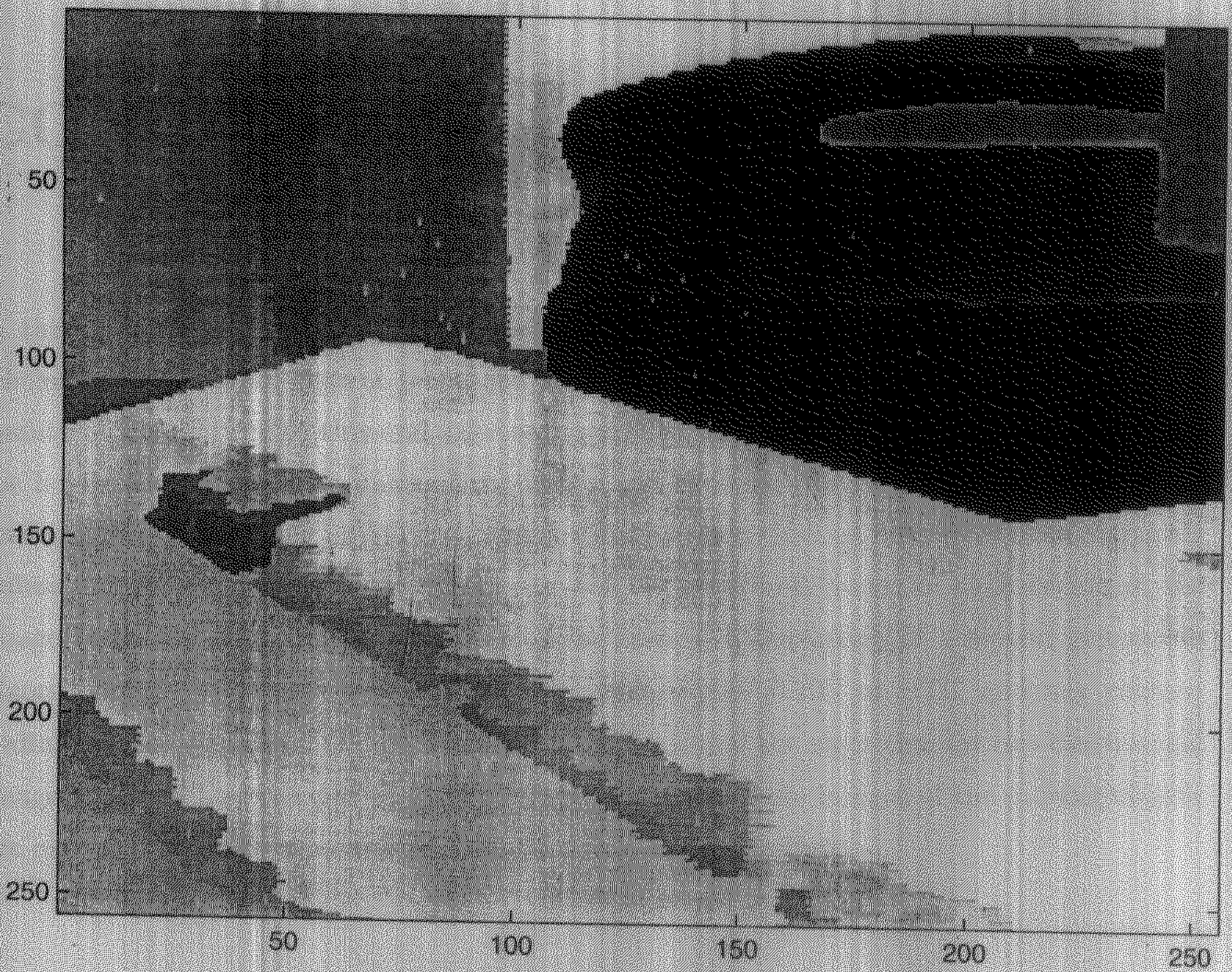
As mentioned earlier the algorithm involves two steps:

- **Initialization:** This step puts in the queue the location of all pixels corresponding to the interior of a region in the labelled marker(pixels not belonging to the uncertainity areas).These pixels have the highest priority(distance 0) because they certainly belong to their respective regions.
- **Flooding:** The flooding assigns pixels to regions following a region growing procedure. The flooding extracts a pixel from the queue. If the pixel does not belong yet to a region, we know that at least one of its neighbors belongs to a region. Therefore, all neighboring regions are examined, the distances between these neighboring regions and the current pixel are assessed, and the pixel is assigned to the region giving the highest certainty. Of course, if there is only one neighboring region, the pixel is directly assigned to it. Once a new pixel has been assigned to a region, the mean label value of the region is updated in order to accurately compute its distance with respect to new pixels. Then, if the pixel that has been assigned to a given region A has some neighbours that do not belong to any region, these neighbors are put in the queue with a priority defined by their distance to the region A. However, these pixels can be

neighbours of pixels previously examined and assigned. As a result they may already be in the queue with an arbitrary priority. This contrasts with the classical watershed algorithm where the pixel's priority is uniquely defined by its gray-level value, and if a pixel is already in the queue, it is not necessary to put it in again. Here, since the priority does not only depend on the pixel gray-level but also on characteristics of one of its neighboring regions, its priority may change. As a result, the pixel should always be introduced in the queue again except if it is already in the queue with a higher (or equal) priority.

In the simulation of this stage there is a difference in the working principle with that has been stated in [1].In [1] the mean gray-level value of the pixels of a region has been taken as the identification of that region while I have used the average value of the labels obtained in the previous step as the identifier .Otherwise it has been found that due to the averaging over a wide range of pixel-intensities, the brightness of the two neighbouring regions come closer resulting in a loss of contrast in the whole output image.The final output is in the next page.

Final Result after decision step (input image is the original one)



Final Result with gradient image as input

Chapter 5

CONCLUSION

In the conclusion, I can recall that the basic aim of the second-generation compression scheme is to devise an elegant technique to code an arbitrary input image maintaining the quality of the picture.To this end , the segmentation of the image is a major step and once the homogeneous semantic units of the image are extracted the contours and textures can scrarately be coded using any first generation method and this can be carried out recursively unless one is satisfied with the quality of the decompressed image of each region of the image.Thus,this work can be extended to the complete simulation of a second generation algorithm using a standard texture coding method like JPEG or BTC and any contour coding method like chain coding.Finally, I have simulated the algorithm for two dimensional images,this can be extended to three dimensional cases where images of the moving objects are expected to be treated more efficiently using this proceedure.

Bibliography

[1] .**P.Selembier and M.pardas,**"Hierarchical Morphological Segmentation for Image Sequence Coding"

[2] .**Luc Vincent and Edward R.Dougherty**, "Morphological Segmentations for Textures and Particles"

[3] .**S.Beucher and F.Meyer**, "The Morphological Approach to Segmentation:The Watershed Transformation

[4] .**M.M.Reid,R.J.Miller and N.D.Black,**"Second Generation Image Coding:An Overview"

[5] .**J.Serra,**"Image Analysis and Mathematical Morphology"

[6] .**Gonzalez and Woods,**"Digital Image Processing"