INTERNATIONAL WORKSHOP ON INDUSTRIAL APPLICATIONS OF
MACHINE INTELLIGENCE AND VISION (MIV-89), Tokyo, April 10–12, 1989

Low Level Image Segmentation with High Level "Emergent Properties": Color Based Segmentation.

Roberto Battiti
Caltech Concurrent Computation Program
158-79 California Institute of Technology, Pasadena CA 91125, USA

Abstract.

The presented method incorporates a discontinuity detection process into a multigrid relaxation algorithm, with the goal of recovering "significant" discontinuities at different scales. Line processes are activated in a deterministic way, depending on local properties of both neighboring line processes (at different scales) and neighboring continuous variables. Computational complexity is $\mathcal{O}(n)$ for an image with n pixels and convergence time is a small multiple of that required by one relaxation step at the finest grid.

The suggested scheme is applied to the problem of image segmentation based on color differences. These dissimilarities are detected by considering changes in the relative intensity of the red, green and blue components of the pixels adjacent to a given discontinuity.

A final relaxation step restricted within the detected boundaries is then suggested as a way of "coloring" the delineated regions in a uniform way.

The algorithm has been implemented with high efficiency on a MIMD parallel computer with distributed memory. A *coarse grain* decomposition is found to be useful for this and other multiscale problems.

1 Introduction.

If one were allowed to use a Procrustean bed, one could separate two possible approaches to the image segmentation problem. In the first approach the problem is attacked *from above*. High level concepts (relational graphs, heuristic search, interest measures, expert systems, focus of attention) are used to guide processing. In the second approach instead the problem is attacked *from below* using cooperative computation done by a usually large number of simple and densely interconnected computing elements (neural networks, regularization, relaxation, simulated annealing). Examples of the two approaches are in the cited references ([2]- [4], [6]-[7], [9]-[12], [14]- [19], [21]-[23]).

Part of the difference is naturally related to definition of the segmentation problem, or to the desired internal representation produced as final result of the process. If, for example, a segmentation system is designed for a dedicated job where domain dependent knowledge is a necessary component of segmentation and the desired output is a symbolic description of parts, a certain amount of high level knowledge will be necessary to accomplish the task. Nonetheless, since the representational gap between raw visual data and desired internal representation is large, even in these cases a large fraction of the processing can usually be done in a *parallel* and general fashion, postponing the high level stage until it is really needed.

It is in fact evident from many cases in different fields ([13]) that "high level", structured, low entropy results do not necessarily imply high level forces driving the evolution of the system [1].

The advantages of the second approach are its generality and its preattentive and bottom-up nature. From a technological point of view this allows design of digital (or analog) VLSI vision circuits employing a large amount of *parallelism* with close to real-time computational capabilities.

2 Multigrid Algorithm with Line Processes.

Multiscale algorithms have recently been proposed to speed up the solution of partial differential equations encountered in various image analysis problems ([21], [22]).

In a previous work dedicated to the 3D surface reconstruction problem based on the multigrid technique ([3]), it is showed that one can combine the smoothing and discontinuity detection steps *in time and scale*. Both phases are executed concurrently on grids at different scales; information flows between different grids and between computing elements detecting discontinuities (henceforth called *line processes*, abbreviated as LP's) or representing continuous values.

Let's consider first the discontinuity updating process. Activation of a given LP is based on the presence of other active LP's at the same scale and *at finer and coarser scales*. This effect is embodied into a *cost* function, where for example the cost for a given LP is low if this leads to a better local LP structure [2]. This is a function of binary variables and it is useful to define it first for a small set of typical configurations of the LP neighborhood and then extend it using rotational symmetry ([3]).

Furthermore activation is based on the discrepancy between points at different sides of the given LP, measured by a *dissimilarity* function $\mathcal{D}(P,Q)$ pertinent to the segmentation.

The activation rule is the following: a LP becomes active if and only if the *dissimilarity* of the separated points is greater than its *cost*.

$$LP \leftarrow 1 \quad \text{iff} \quad \mathcal{D}(P,Q) > cost \qquad (1)$$

A table look-up approach is used to find the influence of the neighborhood at the same scale on the *cost*, as shown in Figure 1.

[1] Taking an example from Physics: the low level interaction between atoms in the active material and electromagnetic field, combined with an appropriate geometry, is responsible for the laser effect. The ultimate example is naturally the collective result of putting together a large amount of neurons, i.e. our mind.

[2] For example one may favor continuous, non intersecting lines.

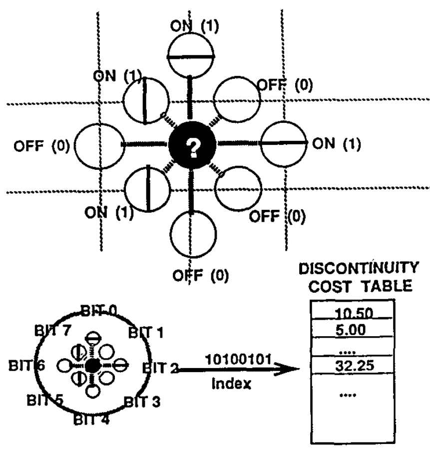

Figure 1: Table look-up for discontinuity updating (influence at the same scale).

Activation values (0 or 1) of LP's in a neighborhood of the undecided LP are used as bits to form an index into a table of costs. This provides maximum flexibility in the definition of the *cost* function and speed during the computation.

The influence of LP's at different scales is illustrated in Figure 2.

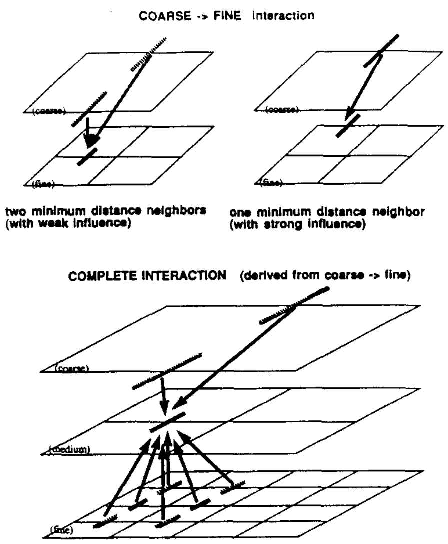

Figure 2: "Spreading of activation" between discontinuities in different scales.

The effect is excitatory (existence of an active LP hints at the presence of similar LP's in contiguous scales) and it is taken into account by "discounting" the *cost* by a selected factor for each active connected LP. Given the rectangular tesselation of the image plane and the doubling of the grid step going to coarser grids, the number of connected LP's varies depending on the precise location on the grid. This effect is corrected by scaling the discount factors so that the global effect does not depend on the precise number of connected LP's but only on their average activation.

Considering now the cooperative smoothing process, the updating rule for the continuous field $\phi(x,y)$ is derived applying variational calculus to the following functional:

$$E(\phi(x,y)) = \int_{Image} \beta(\phi(x,y) - d(x,y))^2 + (\phi_x^2 + \phi_y^2)dxdy \quad (2)$$

A physical analogy is that of fitting the data (x,y) with a membranous sheet pulled by strings connected to the data. A given ϕ value is updated as follows:

$$\phi(x,y) \leftarrow \frac{\phi_{sum} + \beta \times h^2 \times d(x,y)}{n_{sum} + \beta \times h^2} \quad (3)$$

where $h \equiv$ grid step.

$\phi_{sum} \equiv$ sum of neighboring ϕ's not separated by active LP's;

$n_{sum} \equiv$ number of terms in this sum;

$$\phi_{sum} = \sum_{dx=\pm h; dy=\pm h} \mathrm{LP}(x+dx, y+dy) \times \phi(x+dx, y+dy);$$

$$n_{sum} = \sum_{dx=\pm h; dy=\pm h} \mathrm{LP}(x+dx, y+dy);$$

The two phases are interlaced and computation evolves on the different levels of the pyramid as illustrated in Figure 3.

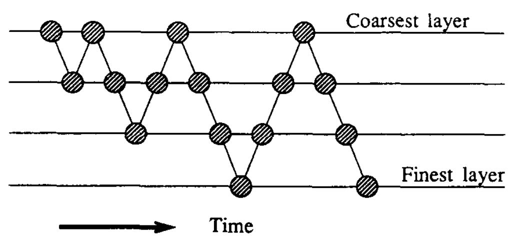

Figure 3: Evolution of multigrid computation on the different levels. Each blob corresponds to a relaxation and discontinuity detection step.

3 Color Based Segmentation.

In the present work the *dissimilarity* function previously introduced is defined in order to apply the algorithm to the problem of segmenting an image with borders suggested by color differences. Color information is only one of the clues provided by a complex visual system and limiting consideration to it is done only to study this effect in isolation.

Dissimilarity in color between points P and Q is defined as the square of the difference in the *orientation* of the vectors $\mathbf{\Phi}$ and $\mathbf{\Psi}$ containing the red, green and blue components of the two pixels.

$$\mathcal{D}(P,Q) \equiv \left\| \frac{\mathbf{\Phi}}{\|\mathbf{\Phi}\|} - \frac{\mathbf{\Psi}}{\|\mathbf{\Psi}\|} \right\|^2 \quad (4)$$

The motivation for this definition is that the light intensity at a point in the image is the product of the reflectance and the illumination ([14]), therefore a uniform change in illumination that multiplies the rgb components by the same factor will *not* be detected as color dissimilarity by the algorithm.

Now, from a computational point of view, one would like to avoid the floating point computation and calculation of square roots implied by eqn. 4. With this goal in mind the following expression is easier to deal with:

$$\mathcal{D}(P,Q) \equiv \frac{\|\Phi \times \Psi\|^2}{\|\Phi\|^2\|\Psi\|^2} \tag{5}$$

If θ is the angle between the two color vectors ($\theta = \arccos\left(\frac{\Psi \cdot \Phi}{\|\Psi\|\|\Phi\|}\right)$), one can show that:

$$\left\| \frac{\Phi}{\|\Phi\|^2} - \frac{\Psi}{\|\Psi\|} \right\|^2 = 4\sin^2(\theta/2) \tag{6}$$

$$\frac{\|\Phi \times \Psi\|^2}{\|\Phi\|^2\|\Psi\|^2} = \sin^2(\theta) \tag{7}$$

Since the qualitative behavior of the functions in eqn. 6 and eqn. 7 is the same, one can obtain the same segmentation in both cases by appropriately scaling the *cost* values.

The expression of eqn. 5 in terms of the rgb components is:

$$\mathcal{D}(P,Q) = \frac{(r_P g_Q - g_P r_Q)^2 + (r_P b_Q - b_P r_Q)^2 + (g_P b_Q - b_P g_Q)^2}{(r_P^2 + g_P^2 + b_P^2)(r_Q^2 + g_Q^2 + b_Q^2)} \tag{8}$$

For the more general case of two "fibers" of intensity values in N spectral bands, eqn. 8 is generalized as follows:

$$\mathcal{D}(P,Q) = \frac{1}{2}\frac{\sum_{i,j=1}^{N}(\phi_i\psi_j - \phi_j\psi_i)^2}{(\sum_{i=1}^{N}\phi_i^2)(\sum_{j=1}^{N}\psi_j^2)} \tag{9}$$

4 Tests.

Some tests of the presented method have been made using images taken from television, converted to rgb format and loaded into the finest grid of the pyramid. This consists of four grids, with 129×129 pixels in the finest one. Data for coarser grids are obtained by averaging those in the finer grid using a 3×3 window, parameters defining the *cost function* ([3]) have been tuned to the chosen images.

Tests on different images show an acceptable detection of the relevant *color* discontinuities after the simple multigrid algorithm with "V cycles" described in section 2 (see Figure 3).

For some tests, after the initial LP detection and relaxation cycle, a final relaxation cycle has been executed. In this the parameter β defined in eqn. 3 is set equal to zero and LP's maintain their activation values and serve as delimiters for the smoothing action.

During this final step, coloring in the delineated regions tends to converge to the local average color, as shown by the "thinning" of the histogram peaks in Figure 4. This would for example facilitate extraction of domains (not necessarily connected) based on color similarity. Currently we are studying the quantitative details of this effect.

Figure 5 shows the obtained segmentation on the three coarsest scales (intensity is proportional to the red component, LP's are white). To the left is the result after the relaxation and discontinuity detection cycle, while to the right is the result after applying a final relaxation cycle.

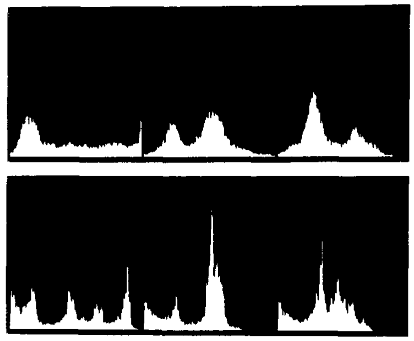

Figure 4: Histogram thinning due to relaxation confined within the delineated regions: red, green and blue components.

5 Parallel Processor Implementation.

If one defines as one *work unit* the amount of computation required by a complete relaxation and discontinuity detection on the finest grid, execution of the algorithm (V cycles, 4 grids) requires 3.43 work units [3].

Given the regularity of the algorithm and the locality of communication between different computational elements, it can be parallelized in a straightforward manner. Assuming that a two-dimensional grid can be embedded in the parallel architecture and that the processors are capable of containing enough data, a two-dimensional *domain decomposition* assigns to every processor a rectangular patch of the image with its "slice" of pyramidal structure (containing elements at all scales corresponding to the assigned patch).

Every processor operates at *all levels* of the pyramid, alternating computation and communication steps to exchange the data on the borders of the assigned domain, as illustrated in Figure 6.

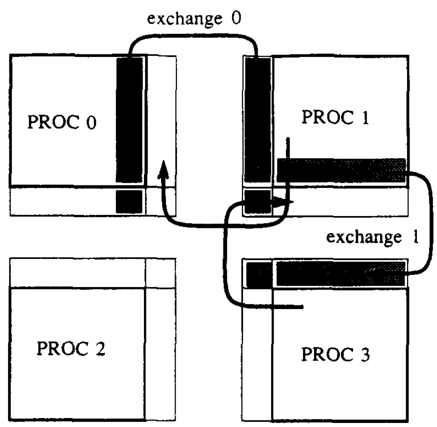

Figure 6: Communication strategy for two-dimensional domain decomposition. Data of the assigned domain are bordered by data received from nearby processors. Two exchanges are sufficient.

[3]Using a *SUN 386i* workstation and C language, this corresponds to approximately one minute, if memory is large enough to contain the entire pyramidal structure.

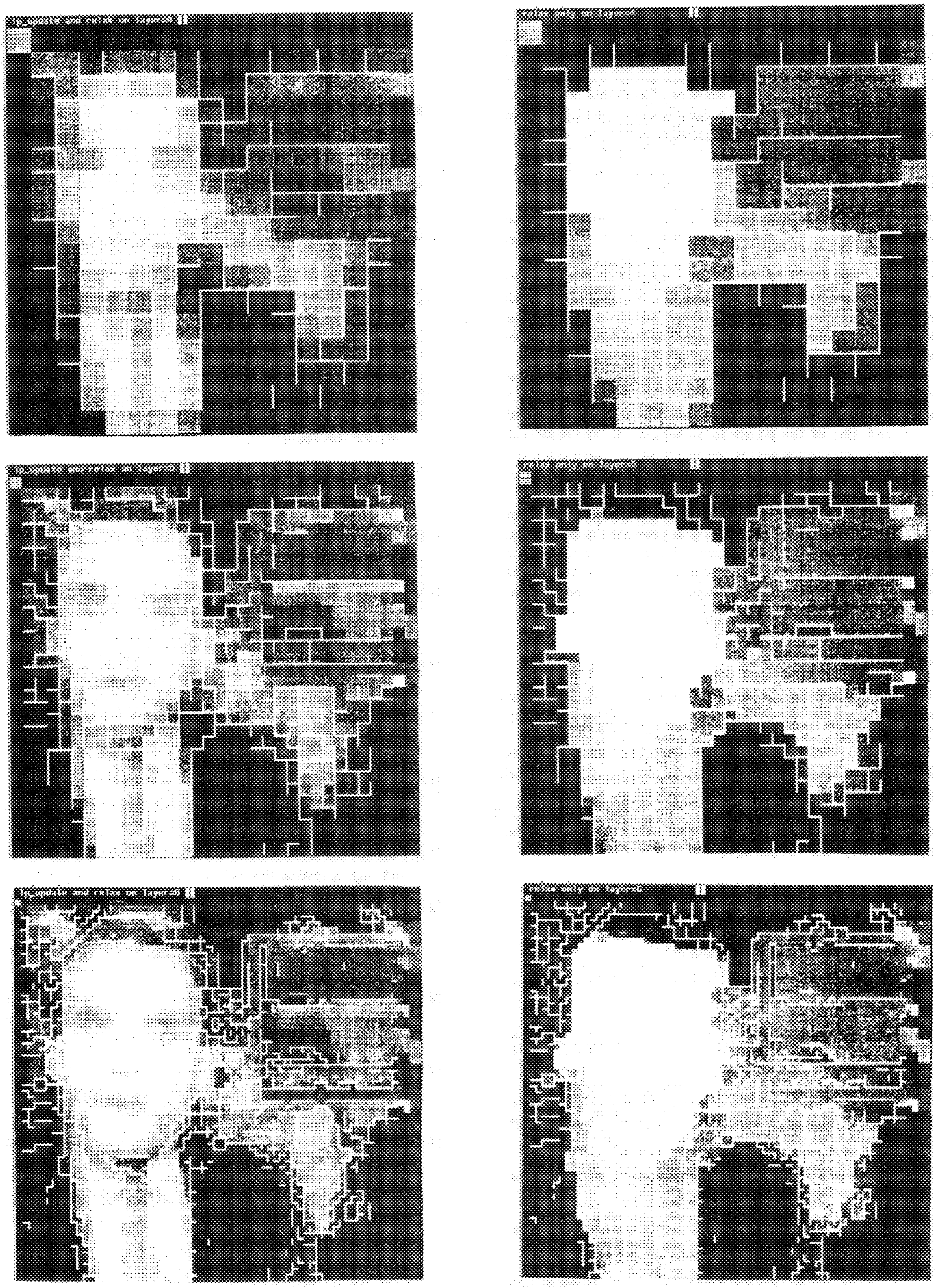

Figure 5: Results of test on the three coarsest levels, before and after the final relaxation cycle.

Since all processors are active most of the time and since the communication overhead is a "surface effect" proportional to $1/\sqrt{n}$, where n is the number of pixels assigned to a given processors, the parallel implementation brings a speed-up that is approximately linear in the number of available processors.

Preliminary tests have been done with up to 16 processors and confirm the theoretical predictions.

6 Conclusion.

A method has been presented that provides a uniform framework for different types of segmentation and relaxation, where the problem dependent information comes in the form of a choice for the *dissimilarity* and *cost* functions. The last one is implemented with a look-up table. The suggested multiscale computation not only speeds up convergence but also provides segmentation with varying details on the different scales, to be used in subsequent processing.

In particular the method has been applied to color based segmentation of rgb images after defining a *dissimilarity* function based on differences between the relative amount of red, green and blue of the points to be separated. For each scale one layer of line processors is coupled with the three rgb layers and defines the limits of the interlaced smoothing procedure, that operates on the three layers separately.

Although preliminary tests have been done only for standard rgb images, we believe that the presented approach can be useful for general applications using multispectral images, like for example satellite imagery or for artificial environments where color information can be appropriately defined and controlled.

Acknowledgment.

This work was done in the Caltech Concurrent Computation and Neural Networks Program and benefited in many ways from the advice of Geoffrey Fox.

I am also pleased to acknowledge useful suggestions and discussions from Paul Messina, Wojtek Furmanski, Christof Koch and Demetri Terzopulos.

Work supported in part by DOE grant DE-FG-03-85ER25009, the National Science Foundation with grant IST-8700064 and by IBM.

References

[1] Ballard D.H., Brown C.M. *Computer Vision* (Prentice Hall,Englewood Cliffs,1982)

[2] Battiti R. "Collective Stereopsis on the Hypercube" *The III Conf. on Hypercube Conc. Comp. and Appl.* **Vol II** 1000-1006 (ACM Press,New York,1988)

[3] Battiti R. "A Multiscale Approach to Surface Reconstruction with Discontinuity Detection." *Caltech C3P-Report***676 B**(1988)

[4] Blake A., Zisserman A. *Visual Reconstruction* (MIT Press, Cambridge,Massachusetts,1987)

[5] Brandt A. "Multi-level adaptive solutions to boundary-value problems" *Math. Comput.* **31**,333-390 (1977)

[6] Burt P.J., Hong T.H., Rosenfeld A. "Segmentation and Estimation of Image Region Properties Through Cooperative Hierarchical Computation." *IEEE Trans. Syst. Man Cybern.* **SMC-11**,802-809 (1981)

[7] Derin H. , Won C. "A Parallel Image Segmentation Algorithm Using Relaxation with Varying Neighborhoods and its Mapping to Array Processors." *Comp.Vis.Graph. and Im. Proc.* **40**,54-78(1987)

[8] Fox G., Johnson M., Lyzenga G.,Otto S., Salmon J., Walker D. *Solving Problems on Concurrent Processors* (Prentice Hall,New Jersey,1988).

[9] Furmanski W. and Fox G. C. "Integrated vision project on the computer network" *Caltech C3P report* **623** (1988)

[10] Geman S., Geman D. "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images." *IEEE Trans. Pattern Analysis Machine Intelligence* **6**,721(1984)

[11] Grossberg S. and Mingolla E. "Neural Dynamics of Surface Perception: Boundary Webs, Illuminants, and Shape-from-Shading." *Comp.Vis.Graph. and Im. Proc.* **37**,116-165(1987)

[12] Gross A.D., Rosenfeld A. "Multiresolution Object Detection and Delineation" *Comp.Vis.Graph. and Im. Proc.* **39**,102-115(1987)

[13] Hopfield J.J. "Neural networks and physical system with emergent collective computational abilities." *Proc.Natl.Acad.Sci.USA* **79**,2554-2558(1982)

[14] Horn B.K.P. "Determining Lightness from an Image." *Comp. Graph. and Image Proc.* **3**,277-299(1974)

[15] Koch C., Marroquin J., Yuille A. "Analog neuronal networks in early vision" *Proc. Natl. Acad. Sci. USA* **83**, 4263-4267 (1986)

[16] Marroquin J.L. "Surface Reconstruction Preserving Discontinuities" *MIT A.I. Memo* **792** (1984).

[17] Nazif A. M. and Levine M.D. "Low Level Image Segmentation: an Expert System." *IEEE Trans. P.A.M.I.* **6**, 555-577(1984).

[18] Poggio T. and Koch C. "Ill-posed problems in early vision: from computational theory to analogue networks" *Proc. R. Soc. Lond. B* **218** ,303-323 (1985)

[19] Poggio ,T., Torre ,V. and Koch,C. "Computational vision and regularization theory" *Nature* **317**,314-319 (1985).

[20] Stüben K. and Trottenberg Multigrid Methods: Fundamental Algorithms, Model Problem Analysis and Applications" *Multigrid Methods Proc.*,1-176 (Springer-Verlag,Berlin,1982)

[21] Terzopulos, D. "Multilevel Computational Processes for Visual Surface Reconstruction" *Comp. Vis., Graph., and Image Proc.* **24**,52 (1983)

[22] Terzopulos, D. "Image analysis using multigrid relaxation methods" *IEEE Trans. Pattern Analysis Machine Intelligence* **8**,129 (1986)

[23] Terzopulos, D. "Regularization of inverse visual problems involving discontinuities" *IEEE Trans. Pattern Analysis Machine Intelligence* **8**,413 (1986)