Consiglio Nazionale delle Ricerche

# ISTITUTO DI ELABORAZIONE DELLA INFORMAZIONE

PISA

# COMPUTATION OF STEREO DISPARITY USING REGULARIZATION

Riccardo MARCH

Nota Interna B4-34

Dicembre 1987

# COMPUTATION OF STEREO DISPARITY USING REGULARIZATION

Riccardo MARCH
C.N.R., Istituto di Elaborazione della Informazione
Via S.Maria 46, I-56100 Pisa, Italy

Abstract

The computation of stereo disparity is a mathematically ill-posed problem. However, using regularization theory it may be transformed into a well-posed problem. Standard regularization can be employed to solve ill-posed problems by using stabilizing functionals that impose global smoothness constraints on acceptable solutions. However, the presence of depth discontinuities causes serious difficulties in standard regularization, since smoothness assumptions do not hold across discontinuities. This paper presents a regularization approach to stereopsis based on controlled-continuity stabilizing functionals. These functionals provide a spatial control over smoothness, allowing the introduction of discontinuities into the solution. An iterative method for the computation of stereo disparity is derived, and the result of a computer simulation with a synthetic stereo pair of images is shown.

Key words: Stereopsis, depth computation, ill-posed problems, regularization.

## 1. Introduction

Many problems in early computational vision are ill-posed in the sense that the existence, uniqueness, and stability of solutions cannot be guaranteed in the absence of additional constraints (4). The computation of binocular disparity from stereo images is a typical example. This paper addresses the problem of stereo matching as an ill-posed inverse problem.

The recognition that early vision problems are ill-posed suggests the use of regularization methods developed in mathematics for this type of problem. In standard regularization, due mainly to Tikhonov (9), and suggested by Poggio (4,5) as a natural mechanism in early vision, the class of admissible solutions of an ill-posed problem is restricted by imposing additional constraints. Using Tikhonov regularization a problem is made well-posed by restricting the acceptable solutions to spaces of smooth functions. The validity of smoothness constraints for early vision reconstruction problems is based on the physical assumption that the coherence of matter tends to give rise to smoothly varying characteristics in a three-dimensional scene (2).

Poggio and others (4) showed how the stereo matching problem can be regularized by imposing a constraint of smooth variation on the disparity field. However, the presence of occlusions between different surfaces in the scene leads to discontinuities in disparity, and smoothness constraints clearly do not hold across visual discontinuities. Consequently, the application of standard Tikhonov regularization theory to stereopsis when discontinuities are involved destroys surface occlusions and this leads to serious difficulties. It is clear that this drawback must be overcome in order to preserve surface depth discontinuities.

Terzopoulos (7) proposed a general method for the regularization of ill-posed problems involving discontinuities, which introduced controlled-continuity constraints. These special constraints provide a spatial control over the smoothness properties of the regularized solution, allowing the selective reconstruction of discontinuities. This method extends standard regularization theory to ill-posed visual problems involving both continuous regions and discontinuities for which global smoothness constraints fail.

This paper examines the possibility of extending the regularization approach to the stereo matching problem following the lines proposed by Terzopoulos. From this approach, an iterative algorithm for the computation of stereo disparity is derived, and the result of a computer simulation on a synthetic stereo pair of images is presented.

## 2. A regularization approach to stereopsis

Most early vision problems are ill-posed in the sense of Hadamard. By definition, a problem is well-posed when its solution (i) exists, (ii) is unique and (iii) depends continuously on the input data. From this definition, we can see that the stereo matching problem is ill-posed since the solution is not unique. In addition, it is an inverse problem, and most inverse problems are ill-posed.

Regularization transforms ill-posed problems into well-posed problems. Tikhonov regularization (9) requires the choice of a penalty functional $\mathscr{P}(u)$ and of a stabilizing functional $\mathscr{S}(u)$ defined on the space of the admissible solutions u(x,y) of the problem. The penalty functional $\mathscr{P}(u)$ measures the discrepancy between the solution and the input data. The stabilizing functional $\mathscr{S}(u)$ measures the degree of regularization (smoothness) of the solution and embodies the additional constraints imposed on the problem.

The common method of regularization looks for the solution to the following minimization problem, which will be referred to as a variational principle:

find a function u(x,y) from the space of the admissible functions which minimizes $\mathscr{P}(u) + \lambda \mathscr{S}(u)$.

Parameter $\lambda$ controls the compromise between the closeness of the solution to the data and the degree of regularization.

In our problem we choose the penalty functional as the error measured in the least square sense given by

$$\mathscr{P}(u) = \iint \left[L(x,y)-R(x+u(x,y),y)\right]^2 dxdy \;, \qquad (1)$$

where R(x,y) and L(x,y) are the left and right image intensities, or a simple function of them (such as a Laplacian), and u(x,y) is the disparity map. For simplicity we assume that vertical disparity is negligible. Within Tikhonov regularization we can take a stabilizing functional of the form

$$\mathscr{S}(u) = \iint w(x,y)(u_x^2 + u_y^2)dxdy \;, \qquad (2)$$

where the subscripts on u denote partial derivatives and w(x,y) is a non-negative continuous weighting function (9). It can be shown that with this choice of functionals the variational principle is well-posed (3). The stabilizing functional imposes a smoothness constraint on the disparity field by penalizing large disparity gradients.

In this form, the regularization method can deal meaningfully with a scene where the depth changes smoothly relative to the viewer. This is usually the case of scenes containing only a single surface. In presence of more than one surface the method attempts to reconstruct a single smooth surface interpolating across the occluding boundaries, which correspond to places where surfaces in the scene occlude one another from the viewer. The result is the destruction of depth discontinuities and this is a severe defect of global smoothness constraints (6). Therefore, this smoothing effect must be prevented across occluding boundaries.

Following the approach proposed by Terzopoulos (7), we modify the Tikhonov stabilizing functional in order to preserve depth discontinuities. The modified functional, called a controlled-continuity stabilizer by Terzopoulos, is obtained by replacing the continuous weighting function w(x,y) in the Tikhonov functional with a discontinuous function, which is allowed to make jump transitions to zero values (7). This makes it possible to introduce specific discontinuities into the solution, a central property for the reconstruction of the depth map when occlusions between different objects are involved.

This spatial control over the smoothness properties of the controlled-continuity stabilizer is exercised in the following way. In regions where depth is continuous (usually corresponding to a single surface), w(x,y)=1, thus the stabilizer reduces to an ordinary Tikhonov stabilizer and generates a continuous surface. Along depth discontinuities (usually corresponding to occluding boundaries between different surfaces), w(x,y)=0, thus deactivating the interpolation effect across different surfaces, and thereby allowing the surfaces to fracture freely. In this way, the modified variational principle yields a piecewise continuous reconstructed surface. A suitable choice of the continuity control function w(x,y) will allow the introduction of specific discontinuities in the reconstructed solution u(x,y).

## 3. Iterative computation of stereo disparity

In regions where depth is continuous, the disparity function u(x,y) must minimize the full functional with w(x,y)=1, and will thus be solution of the following Euler-Lagrange equation

$$\lambda \Delta u + \left[ L(x,y) - R(x+u,y) \right] R_x(x+u,y) = 0 \; , \qquad (3)$$

where $\Delta u$ is the Laplacian of u(x,y) defined as $\Delta u = u_{xx} + u_{yy}$, and $R_x$ denotes the partial derivative of R with respect to x. In equation (3), we have assumed implicitly that image brightness is differentiable.

We now discretize the above partial differential equation on a uniform grid by using finite difference approximations.

Assuming the unit of length equal to the grid spacing interval, we use for the Laplacian of u(x,y) approximation $\Delta u = 4(u^*(i,j) - u(i,j))$, where $u^*(i,j)$ is the local average of u defined as $u^*(i,j) = (u(i+1,j) + u(i,j+1) + u(i-1,j) + u(i,j-1))/4$. The discrete version of the Euler-Lagrange equation is then given by

$$u(i,j) = u^*(i,j) + (1/\lambda) \left[ L(i,j) - R(i+u,j) \right] R_x(i+u,j) \; , \qquad (4)$$

where a factor 4 has been absorbed into $\lambda$ .

Rearranging the equation, a solution can be computed using the Gauss-Seidel iterative method

$$u^{n+1}(i,j)=u^{*n}(i,j)+(1/\lambda)\left[L(i,j)-R(i+u^{n},j)\right]R_x(i+u^{n},j) \quad , (5)$$

where n is the iteration step.

In order to improve the numerical stability of the algorithm (see for instance (1)) we resort to the slightly different relaxation formula

$$u^{n+1}(i,j)=u^{*n}(i,j)+(1/\lambda)\left[L(i,j)-R(i+u^{*n},j)\right]R_x(i+u^{*n},j) \quad , (6)$$

where $u^*(i,j)$ is used in evaluating both R and $R_x$.

Since $u^*(i,j)$ is generally not an integer, we compute the quantity $R(i+u^*(i,j),j)$ using a Lagrange interpolation polynomial in the x coordinate. An estimate of the derivative $R_x(i+u^*(i,j),j)$ is then obtained using the derivative of the interpolation polynomial. In the computer simulation presented in this paper, a third-degree Lagrange polynomial was used, fitted through four adjacent nodes on the same horizontal row of the discrete image.

In regions where $w(x,y)=1$ the iterative scheme (6) yields a continuous reconstructed surface. Without the spatial control over the smoothness properties of the stabilizing functional this type of algorithm will attempt to interpolate a single surface across depth discontinuities because of the presence of the term $u^*(i,j)$. This term, being a local average of $u(i,j)$, is responsible for the undesired smoothing effect across the occluding boundaries between different surfaces.

The controlled-continuity stabilizer deactivates the continuity constraint along the curves with $w(x,y)=0$. Assume for simplicity that all points with $w(x,y)=0$ coincide with nodes in the grid. A convenient numerical implementation of the controlled-continuity stabilizer is then obtained by replacing $u^*(i,j)$ in (6) with the following weighted average:

$$(w(i+1,j)u(i+1,j)+w(i,j+1)u(i,j+1)+w(i-1,j)u(i-1,j)$$
$$+w(i,j-1)u(i,j-1))/w^*(i,j) \ , \qquad (7)$$

where $w^*(i,j)$ is defined by $w^*(i,j)=w(i+1,j)+w(i,j+1)+w(i-1,j)+w(i,j-1)$. The value of disparity is now undefined in the nodes where $w(x,y)=0$.

It is easy to see that with this modification the propagation of the local average of $u(i,j)$, and hence the interpolation effect, is prevented across the curves $w(x,y)=0$. In this way, surfaces on opposite sides of an occluding contour have no influence on one another and are allowed to fracture freely making the reconstruction of depth discontinuities possible.

## 4. A computer simulation

The iterative algorithm using the continuity control function has been implemented and applied to a synthetic stereo pair of images corresponding to a simple pattern. The results shown here are for stereo images of 128X128 pixels.

Figure 1 shows the two images of the stereo pair representing an object shaped as a revolution surface and portrayed against a plane background. The brightness pattern, both of the surface and of the background, is a linear combination of spatially orthogonal sinusoids. The spatial frequency of the sinusoids is chosen to give a reasonably strong brightness gradient such as that usually required for binocular stereo matching. Depth in the image is discontinuous along the occluding boundary between the object surface and the background. The stereo pair is obtained by simulating a simple perspective projection of the object yielding a horizontal disparity inversely proportional to depth, and no vertical disparity.

The algorithm exhibits the typical behavior of local iterative methods: convergence is rapid during the first few iterations, but quickly degenerates to a slow asymptotic rate of progress. The number of iterations performed in the present simulation is 192. This number of iterations was necessary in order to compute the solution over the large depth range of the stereo pair. However, convergence can be accelerated by employing multigrid relaxation algorithms that propagate constraints across the nodes in the grid more quickly (8). The iteration is started with an initial estimate of disparity equal to a constant value. The algorithm shows convergence for a broad range of the initial "guess" of disparity. However, a reasonable initial value is needed to obtain a reliable iterative solution.

Parameter $\lambda$ controls the compromise between the degree of regularization of the solution and its closeness to the input data. This factor plays a significant role in areas where the brightness gradient is small, preventing random adjustments to the estimated disparity map occasioned by the presence of noise. The value of parameter $\lambda$ should be roughly equal to the root-mean-square of the noise in the input data measurements.

In the present case, the input data are not noisy and the only noise present in the reconstructed solution is due to errors in the estimates of the derivatives and in the Lagrange interpolation. The problem of the optimal choice of parameter $\lambda$ will not be examined here. In the computer simulation, the value $\lambda=1$ was chosen on the basis of the results of a number of experiments.

The continuity control function $w(x,y)$ is constructed by setting $w(x,y)=0$ along disparity discontinuities (the occluding contour between the object and the background), and $w(x,y)=1$ in continuous regions.

Figure 2 shows a perspective view of the surface reconstructed in the simulation. The controlled-continuity stabilizer plays an essential role in the reconstruction of the surface along the occluding boundary.

Figure 3 shows the result of a simulation executed without controlled-continuity constraints by setting $w(x,y)=1$ everywhere, as in standard Tikhonov regularization. As can be seen the standard regularization method yields an incorrectly reconstructed depth map in proximity of the occluding contour. In fact, the smoothing effect destroys the occlusion by interpolating indiscriminately across the contour and giving the undesirable impression of a tablecloth thrown between the object and the background. On the contrary, the controlled-continuity constraints, deactivating all continuity along the contour, prevent the interpolation effect and yield the correctly reconstructed depth discontinuity.

## 5. Detection of depth discontinuities

The use of controlled-continuity stabilizers makes it possible to introduce specific discontinuities into the reconstructed solution. However, depth discontinuities are not known in advance, and in the previous discussion we have ignored the important problem of their detection.

Terzopoulos (7) suggested a possible approach to the localization of discontinuities as an integral part of controlled-continuity reconstruction. It is possible to augment the functional in the variational principle in order to detect depth discontinuities as the regularized solution is being computed. The new functional is given by

$$\mathscr{P}(u) + \lambda \mathscr{S}(u,w) + \mathscr{D}(w) \ , \qquad (8)$$

where the additional term $\mathscr{D}(w)$ imposes constraints on the structure of the discontinuities.

The functional $\mathscr{D}(w)$ should penalize each detected discontinuity in order to prevent the formation of an incoherent solution, and should incorporate additional knowledge regarding depth discontinuities. Generally, as a consequence of the coherence of matter, occlusions in depth between different surfaces tend to be spatially continuous giving rise to curves and contours. Hence the additional functional $\mathcal{D}(w)$ should encourage the formation of depth discontinuities along continuous contours.

Computation of depth requires the minimization of the new functional (8) with respect to both disparity u and the continuity control function w. However, the augmented functional is now non-convex, since it contains non-quadratic terms, and may thus have multiple local minima.Iterative methods can be adapted to perform the minimization of non-convex functionals (6). The expression of the functional $\mathcal{D}(w)$ and the choice of a suitable optimization method to find near global minima will be the argument of a forthcoming paper.

## 6. Conclusion

The computation of stereo disparity as an ill-posed problem can be approached using standard Tikhonov regularization theory. However, global smoothness constraints are inadequate near depth discontinuities. This paper has presented a more general regularization approach to stereopsis using controlled-continuity constraints originally proposed by Terzopoulos in the context of regularization theory. By using these generalized constraints, the smoothness properties of the solution may be regulated spatially in order to preserve depth discontinuities along occluding contours.

The constraints were formulated as a

controlled-continuity stabilizer comprising a standard Tikhonov regularization functional combined with a noncontinuous continuity control function. This latter function makes it possible to introduce specific discontinuities into the solution.

This regularization method yields a simple way of dealing with depth discontinuities once they have been marked explicitly. When the continuity control function is not prespecified, the locations of discontinuities may be reconstructed as the regularized solution is being computed.

References

(1) Horn, B.K.P. and K.Ikeuchi (1981). Numerical shape from shading and occluding boundaries. Artificial Intelligence 17, 141-184.

(2) Marr, D. and T.Poggio (1979). A computational theory of human stereo vision. Proc. R. Soc. Lond. B 204, 301-328.

(3) Morozov, V.A. (1984). Methods for Solving Incorrectly Posed Problems. Springer-Verlag, New York.

(4) Poggio, T., V.Torre and C.Koch (1985). Computational vision and regularization theory. Nature 317, 314-319.

(5) Poggio, T. (1985). Early vision: from computational structure to algorithms and parallel hardware. Computer Vision, Graphics and Image Processing 31, 139-155.

(6) Terzopoulos, D. (1985). Computing visible-surface representations. M.I.T. A.I. Lab., Cambridge, MA, AI Memo 800.

(7) Terzopoulos, D. (1986). Regularization of inverse visual problems involving discontinuities. IEEE Trans. Pattern Anal. Mach. Intell. 8, 413-424.

(8) Terzopoulos, D. (1986). Image analysis using multigrid relaxation methods. IEEE Trans. Pattern Anal. Mach. Intell. 8, 129-139.

(9) Tikhonov, A. and V.Arsénine (1976). Méthodes de Résolution de Problémes Mal Posés. Mir, Moscow.

## FIGURE CAPTIONS

Fig. 1: A synthetic stereo pair of images.

Fig. 2: Reconstructed depth map using controlled-continuity constraints.

Fig. 3: Reconstructed depth map using standard regularization.

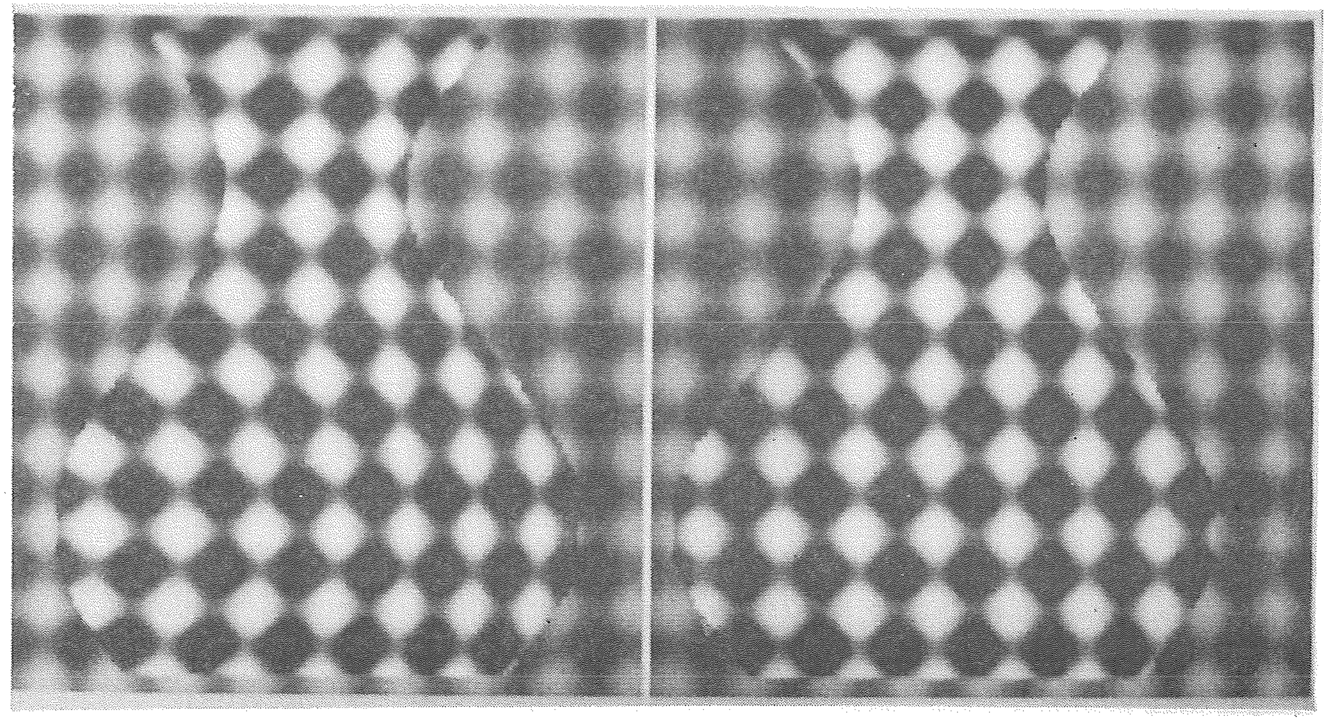

Figure 1

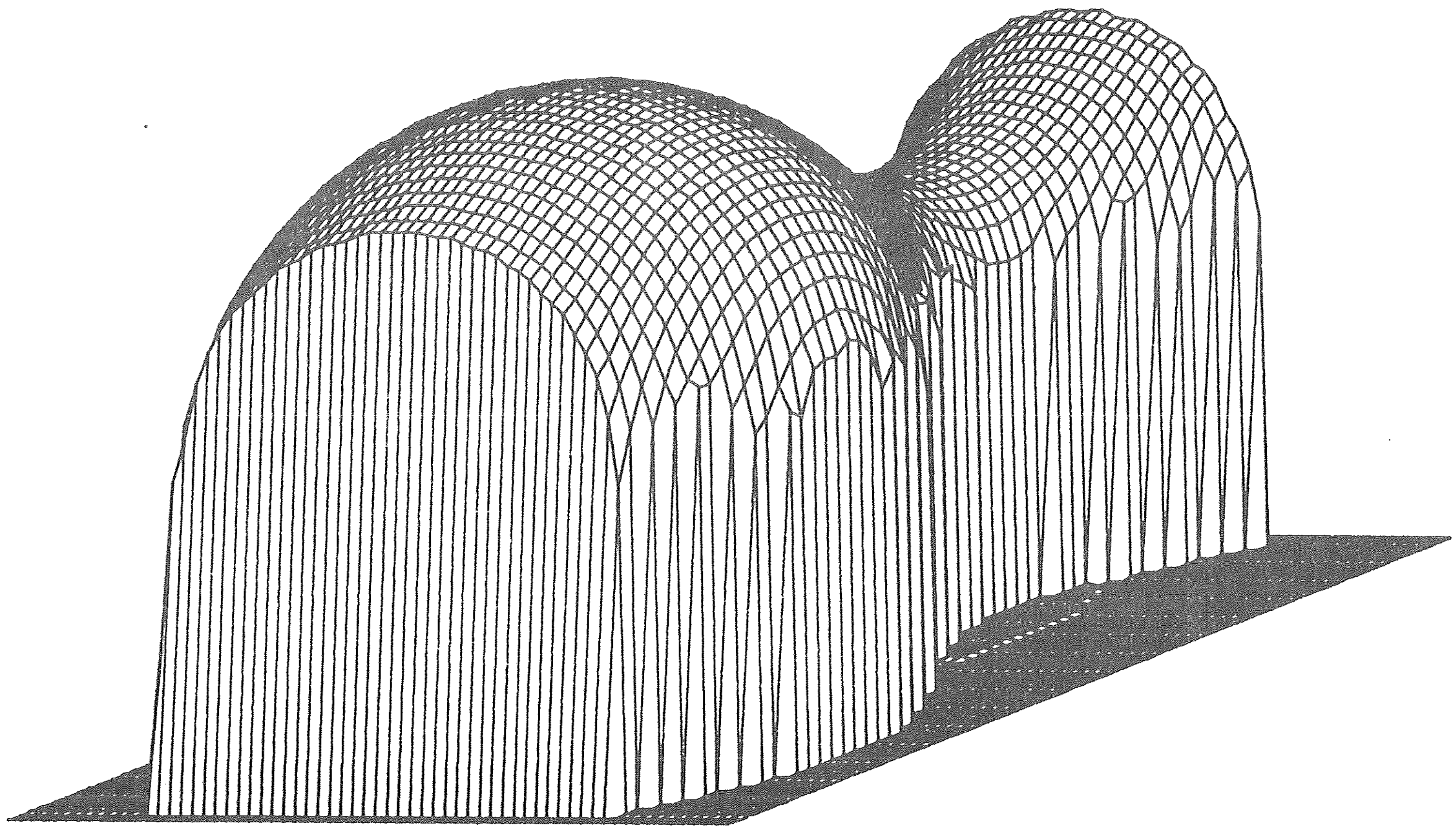

Figure 2

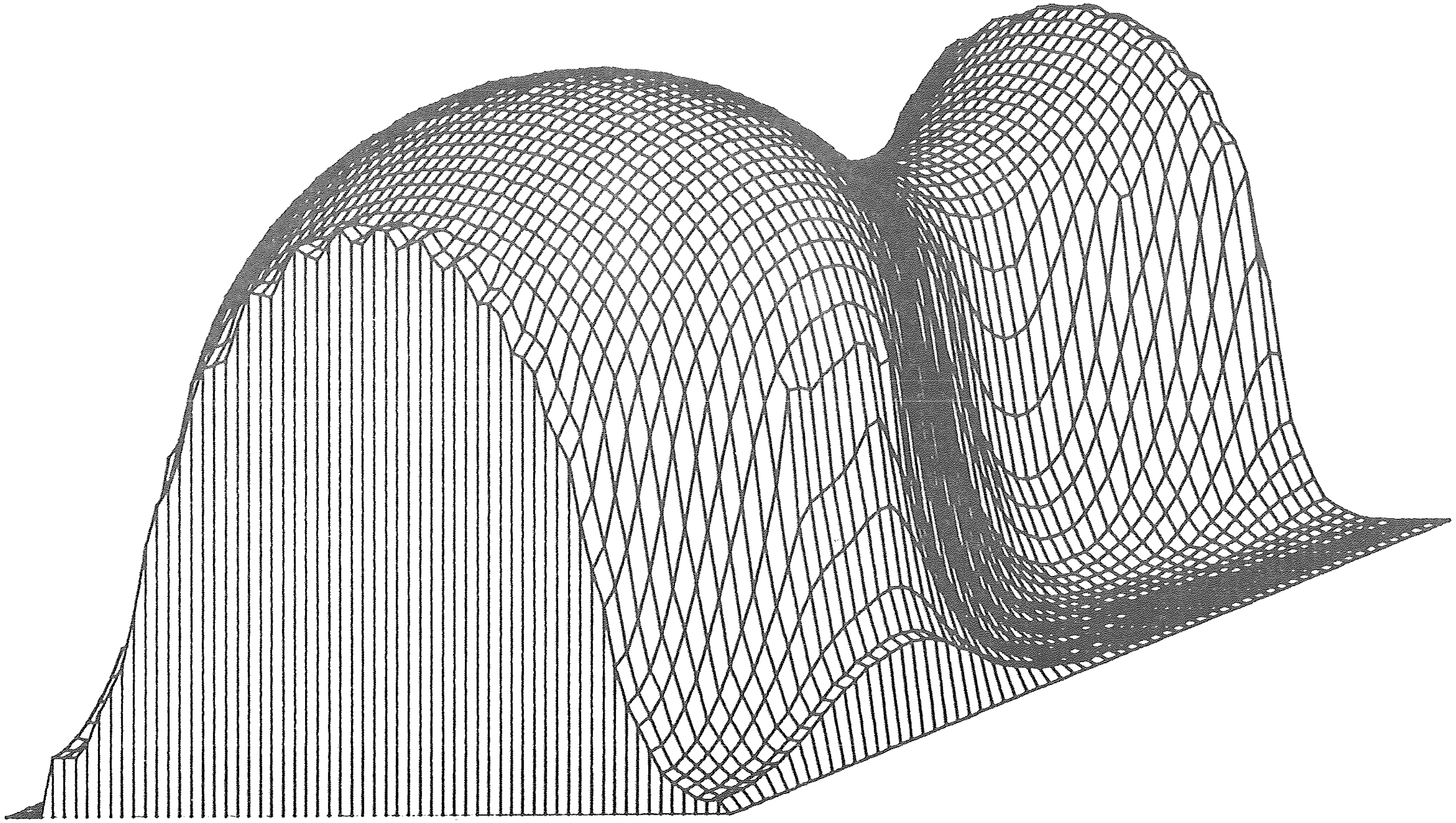

Figure 3