Image Analysis Using Multigrid Relaxation Methods

DEMETRI TERZOPOULOS, MEMBER, IEEE

Abstract—**Image analysis problems, posed mathematically as variational principles or as partial differential equations, are amenable to numerical solution by relaxation algorithms that are local, iterative, and often parallel. Although they are well suited structurally for implementation on massively parallel, locally interconnected computational architectures, such distributed algorithms are seriously handicapped by an inherent inefficiency at propagating constraints between widely separated processing elements. Hence, they converge extremely slowly when confronted by the large representations of early vision. Application of multigrid methods can overcome this drawback, as we showed in previous work on 3-D surface reconstruction. In this paper, we develop multiresolution iterative algorithms for computing lightness, shape-from-shading, and optical flow, and we examine the efficiency of these algorithms using synthetic image inputs. The multigrid methodology that we describe is broadly applicable in early vision. Notably, it is an appealing strategy to use in conjunction with regularization analysis for the efficient solution of a wide range of ill-posed image analysis problems.**

Index Terms—**Image analysis, inverse problems, lightness. multigrid relaxation, multiresolution algorithms, optical flow, parallel algorithms, partial differential equations, shape from shading, variational principles.**

I. Introduction

VARIATIONAL principles and partial diffential equations have played a significant role in the mathematical formulation of early visual information processing problems (representative examples include [1], [5], [7], [12], [17], [19], [23], [25]–[29], [33], [34], [41]–[44], [46]). An attractive feature of variational and differential formulations, once discretized, is the possibility of computing their solutions by numerical relaxation methods. These iterative methods require only local computations, which can usually be performed in parallel by many locally intercommunicating processors distributed in computational networks or grids.

Local, parallel algorithms are appealing in the context of early vision. They offer a means of attaining high performance in this extremely compute-bound task, while at a certain level of abstraction they do not appear incompatible with the apparent structure of sophisticated biological vision systems [3], [38], [46]. Moreover, they are suitable for implementation on massively parallel computers. Such computers, which will certainly proliferate as VLSI technology advances, promise great processing power through the concurrent use of numerous simple and locally interconnected processing elements, rather than through the sequential use of a few fast, general purpose processors [4], [24].

The desired solutions to many image analysis problems possess certain global characteristics (e.g., consistency, coherence, etc.) that are formally manifested in the variational principle or associated partial differential equation formulations.[1] Given only local interconnections among processors, however, global characteristics must evolve *indirectly*, typically through the iterative propagation of visual constraints across the grid network. Indirect propagation can lead to substantial computational inefficiency in early vision applications, where the computational grids tend to be extremely large. Convergence of the iterative process is often slow, and this severely erodes the computational power of massive parallelism.[2] Indeed, for fine discretizations on large grids, excruciatingly slow convergence rates have handicapped iterative algorithms for computing lightness ([5]; see also [25]), shape-from-shading [29], [39], optical flow [27], [33], 3-D surfaces [19], [41], [42], and other visual reconstruction problems.

Since spatial locality of computation is dependent on spatial resolution, local (e.g., nearest neighbor) computations on a coarse grid over a given region are analogous to more global computations on a fine grid over the same region. This suggests the possibility of counteracting the sluggishness of global interactions by deploying local iterative processes over a multiresolution hierarchy of grids. This is the basis of the *multigrid relaxation methods* which are gaining popularity in applied numerical analysis [20]. The computational structure of multigrid methods bears an interesting analogy to the multiresolution nature of spatial frequency channels in the human early visual system [6], and it is similar to pyramid structures for image processing [37].

Manuscript received December 20, 1984; revised July 10, 1985. Recommended for acceptance by S. Tanimoto. This work was supported in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research Contract N00014-80-C-0505 and by the System Development Foundation. The author was supported by the Natural Sciences and Engineering Research Council of Canada and the Fonds F.C.A.C., P.Q., Canada.

The author is with the M.I.T Artificial Intelligence Laboratory, 545 Technology Square, Cambridge, MA 02139.

IEEE Log Number 8405787.

[1]Variational and differential formulations can be related through the Euler–Lagrange equations of the calculus of variations, given appropriate continuity and symmetry (or self adjointness) conditions [14].

[2]It is possible to accelerate basic Jacobi (parallel) and Gauss–Seidel (sequential) relaxation methods so that fewer iterations are generally required to obtain solutions. However, practical accelerated methods such as successive overrelaxation, Chebyshev semi-iteration, alternating direction implicit methods, and even the conjugate gradient method, use global computations to determine the acceleration parameters [21]. Aside from the greater per-iteration complexity of these globally accelerated methods, in a local, parallel implementation the communications costs of performing the global operations neutralize any potential gains from acceleration.

In earlier work, we developed an efficient surface reconstruction algorithm based on multigrid relaxation methods [41], [42]. We suggested, as has Glazer [18], that multigrid methods are broadly applicable in low-level computer vision. After a brief overview of multigrid methodology, we apply it to three other vision problems: the well-known problems of computing lightness, shape-from-shading, and optical flow from images. Novel multiresolution algorithms are developed for each problem, and the algorithms are shown empirically to offer order-of-magnitude gains in efficiency over their conventional, single-level counterparts. Finally, we discuss the approach of the paper within a broader theoretical perspective which unifies the applications.

II. Multigrid Methodology

Pioneering investigations into multigrid methodology include the work of Fedorenko [15], Bakhvalov [2], Brandt [9], [10], and Nicolaides [35]. It has been applied to many boundary value problems (see [8] for an extensive bibliography) and there has also been some development in the context of variational problems [11], [35].

A. Multigrid Relaxation Methods

Multigrid relaxation methods take advantage of multiple discretizations of a continuous problem over a range of resolution levels. The coarser levels trade spatial resolution for direct communication paths over larger distances. Hence, they effectively accelerate the global propagation of information to amplify the overall efficiency of the iterative relaxation process.

The inherent computational sluggishness of local iterative algorithms can be studied from a spatial frequency perspective. A local Fourier analysis of the error function (or, more conveniently, the dynamic residual function) from one iteration to the next shows that high-frequency components of the error—those components with wavelengths on the order of the grid spacing—are short-lived, whereas low-frequency components persist through many iterations [10]. Hence, common (L_2 or L_∞) error norms decrease sharply during the first few iterations, so long as there are high-frequency components to be annihilated, but soon degenerate to a slow, asymptotic diminution when only low-frequency components remain (see Fig. 1). This suggests that while relaxation is inefficient at completely annihilating the error function, it can be very efficient at smoothing it. From this point of view, the grid hierarchy enables the efficient smoothing properties of relaxation to be exploited over a wide range of spatial frequencies.

Empirical studies of model problems (Poisson's equation in a rectangle) indicate that multigrid methods can converge in essentially order $O(N)$ number of operations, where N is the number of nodes in the grid [10]. This can be compared to typical complexities of $O(N^3)$ operations for the solution of model problems by standard (single level) relaxation. Therefore, in image analysis applications, where N tends to be very large (order 10^4–10^6, or more), multigrid methods offer potentially dramatic increases in efficiency over standard relaxation methods.

The multiresolution visual algorithms developed in this paper combine several features: 1) multiple visual representations covering a range of spatial resolutions, 2) local iterative relaxation processes that propagate constraints within each representational level, 3) local coarse-to-fine *extension* processes that allow coarser representations to constrain finer ones, 4) fine-to-coarse *restriction* processes that allow finer representations to constrain and improve the accuracy of coarser ones, and 5) (recursive) coordination schemes that enable the hierarchy of representations and component processes to cooperate towards increasing efficiency.

In multigrid methods, the intralevel processes usually are basic relaxation methods such as Gauss–Seidel or Jacobi relaxation, the extension processes involve local Lagrange (polynomial) interpolations, and the restriction processes involve local averaging operations. The exact form of these operations is problem-dependent.

B. Discretization

The intralevel relaxation processes can be derived by local discretization of the continuous image analysis problems. The *finite element method* [40] can be applied directly to variational principle formulations, while the *finite difference method* [16] may be applied directly to the partial differential equation formulation.

The basic idea behind the finite element method is that a global approximation can result from interactions among many very simple local approximations. This is accomplished by tessellating the continuous domain into a number of polynomial elements, whose dimensions depend on a fundamental size h. The degrees of freedom of the local polynomials define a set of element nodal variables. The continuous functional is then expressed as a discrete summation over all the element contributions. In the finite difference method, typically a grid of nodes with spacings proportional to a parameter h is set up over the domain. The differential operator is then replaced by finite difference equations involving nodal variables at neighboring nodes. The collection of finite difference equations defines a discrete system which approximates the given differential equation.

Note that for the case of quadratic variational principles and, consequently, linear partial differential (Euler–Lagrange) equations, the discrete problems resulting from both the finite element and finite difference methods take the form of large and sparse systems of linear equations $\boldsymbol{A}^h\boldsymbol{u}^h = \boldsymbol{f}^h$, where $\boldsymbol{u}^h$ is the vector of nodal variables. This case covers a wide range of vision applications [41]–[43]. A great deal of effort in numerical analysis has been directed to the solution of large, sparse systems—they turn out to be especially suited to solution by relaxation methods [21].

C. Multigrid Structure and Coordination

In this paper, a spatially uniform discretization of the continuous visual problems is employed to obtain uniform

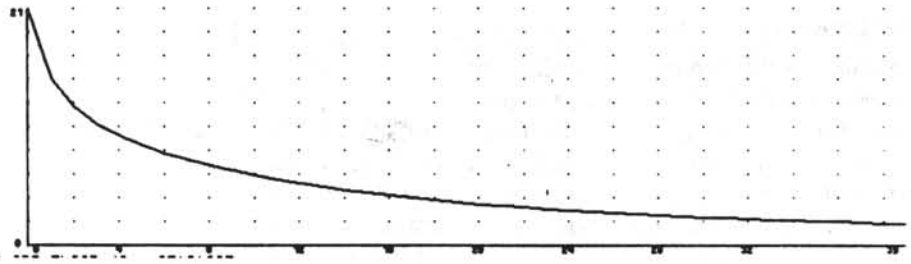

Fig. 1. Asymptotic reduction by error relaxation. The mean square (dynamic residual) error is plotted as a function of the iteration number for a sequence of (Gauss–Seidel) relaxation iterations of a surface reconstruction algorithm. The curve exhibits a typical behavior of local iterative methods: convergence is rapid during the first few iterations, but quickly degenerates to slow asymptotic error reduction.

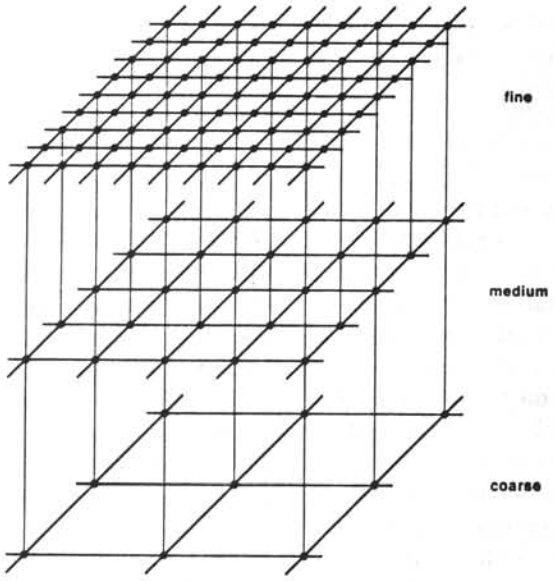

Fig. 2. Possible grid organization of a multiresolution algorithm. A small portion of three levels of the 2 : 1 multigrid hierarchy is shown. Only nearest-neighbor interprocessor connections are included.

grids on each level of the multigrid hierarchy. Multigrid implementation can be further simplified by having a 2 : 1 decrease in grid resolution between adjacent levels[3] and by having the grid nodes of coarser grids coincide with grid nodes on adjacent finer grids. The resulting regular hierarchy, a small portion of which is illustrated in Fig. 2, maps readily, in principle, onto regularly interconnected VLSI architectures. In a fully parallel implementation each node represents a separate processing element.

The multiresolution visual algorithms to be described utilize simple injection $I_{l \Rightarrow l-1}$ for the fine-to-coarse restrictions, bilinear interpolation $I_{l-1 \Rightarrow l}$ for the coarse-to-fine extension, and an *adaptive* multigrid coordination scheme which was employed successfully in the surface reconstruction algorithm (see [41]–[43] for details). The scheme first performs a sufficient number of relaxation iterations to solve the coarsest level discrete system $\boldsymbol{A}^{h_1}\boldsymbol{u}^{h_1} = \boldsymbol{f}^{h_1}$ to desired accuracy (procedure SOLVE), and then proceeds to the finest level $l = L$ according to

[3]The 2 : 1 resolution ratio also appears to be near optimal with regard to multigrid convergence rates [10].

```
procedure FMG
    u^{h_1} ← SOLVE (1, u^{h_1}, f^{h_1});
    for l ← 2 to L do
        begin
            v^{h_l} ← I_{l-1⇒l} u^{h_{l-1}};
            MG (l, v^{h_l}, f^{h_l})
        end;
```

applying the multigrid algorithm

```
procedure MG (l, u, g)
    if l = 1 then u ← SOLVE (1, u, g)
        else
            begin
                for i ← 1 to n_1 [while ... ] do u ← RELAX (l, u, g);
                v ← I_{l⇒l-1} u;
                d ← A^{h_{l-1}} v + I_{l⇒l-1} (g − A^{h_l} u);
                for i ← 1 to n_2 [while ... ] do MG (l − 1, v, d);
                u ← u + I_{l-1⇒l} (v − I_{l⇒l-1} u);
                for i ← 1 to n_3 do [while ... ] u ← RELAX (l, u, g)
            end;
```

After n_1 relaxation iterations (procedure RELAX) have been performed at level l, MG performs a restriction to the next coarser level $l - 1$. It then calls itself recursively on the coarser level n_2 times. Finally, it performs an extension from the coarser level back to level l, following up with n_3 more iterations on level l. The equations on the coarsest level $l = 1$ may be solved to desired accuracy with sufficiently many iterations (procedure SOLVE). One can readily show that when MG is invoked on level l it calls RELAX a total of $n_2^{l-\lambda}(n_1 + n_3)$ times on level $\lambda \neq 1$ and it calls SOLVE n_2^{l-1} times on level 1. In general, most of the relaxation iterations are performed on the coarser levels [22].

The optional [while ...] clauses denote conditions that may be checked during the computation and used to terminate some iterations. Dynamic conditions, typically convergence rates measured by error norms, are incorporated into adaptive coordination schemes, whereas *fixed* schemes are controlled only by the constants n_1, n_2, and n_3 [10]. Although adaptive schemes tend to be more efficient in practice, fixed schemes lend themselves better to theoretical analysis and, moreover, they are easier to implement on distributed local-interconnect architectures due, in part, to the absence of error norms which require global computations.

III. The Lightness Problem

The lightness of a surface is the perceptual correlate of its reflectance. Irradiance at a point in the image is proportional to the product of the illuminance and reflectance at the corresponding point on the surface. The lightness problem is to compute lightness from image irradiance, without any precise knowledge of either reflectance or illuminance.

A. Analysis

The retinex theory of lightness and color proposed by Land and McCann [31] is based on the observation that illuminance and reflectance patterns differ in their spatial properties. Illuminance changes are usually gradual and give rise to smooth illumination gradients, while reflectance changes tend to be sharp, since they originate from abrupt pigmentation changes and surface occlusions. Horn [25] proposed a two-dimensional generalization of the Land–McCann algorithm for computing lightness in *Mondrian* scenes, consisting of planar areas divided into subregions of uniform matte reflectance.

Let $R(x, y)$ be the reflectance of the surface at a point projecting to the image point (x, y) and let $S(x, y)$ be the illuminance at that point. The irradiance at the image point is given by $E(x, y) = S(x, y) \times R(x, y)$. Denoting the logarithms of the above functions as lowercase quantities, we have $e(x, y) = s(x, y) + r(x, y)$. Applying the Laplacian operator Δ gives $d(x, y) \equiv \Delta e(x, y) = \Delta s(x, y) + \Delta r(x, y)$. In a Mondrian, illuminance is assumed to vary smoothly so that $\Delta s(x, y)$ is finite everywhere, while $\Delta r(x, y)$ exhibits pulse doublets at intensity edges separating neighboring regions. A thresholding operator T can be applied to discard the illuminance component: $T[d(x, y)] = \Delta r(x, y) \equiv f(x, y)$. Hence, the reflectance R is given by the inverse logarithm of the solution to Poisson's equation

$$\Delta r(x, y) = f(x, y), \qquad \text{in } \Omega,$$

where Ω is the planar region covered by the image.

Horn solved the above partial differential equation by convolution with the associated Green's function. We instead pursue a local, iterative solution based on the finite difference method.

Suppose that Ω is covered by a uniform grid with spacing h. We can approximate $\Delta r = r_{xx} + r_{yy}$ using the order h^2 approximations $r_{xx} = (r^h_{i+1,j} - 2r^h_{i,j} + r^h_{i-1,j})/h^2$ and $r_{yy} = (r^h_{i,j+1} - 2r^h_{i,j} + r^h_{i,j-1})/h^2$ to obtain a standard discrete version of Poisson's equation $(r^h_{i+1,j} + r^h_{i-1,j} + r^h_{i,j+1} + r^h_{i,j-1} - 4r^h_{i,j})/h^2 = f^h_{i,j}$. This denotes a system of linear equations with sparse coefficient matrix.

Rearranging, the Jacobi relaxation step is given by

$$r^{h\,(n+1)}_{i,j} = \tfrac{1}{4}(r^{h\,(n)}_{i+1,j} + r^{h\,(n)}_{i-1,j} + r^{h\,(n)}_{i,j+1} + r^{h\,(n)}_{i,j-1} - h^2 f^h_{i,j}),$$

where the bracketed superscripts denote the iteration index. Jacobi relaxation is suited to parallel synchronous hardware, whereas the Gauss–Seidel relaxation step given by

$$r^{h\,(n+1)}_{i,j} = \tfrac{1}{4}(r^{h\,(n)}_{i+1,j} + r^{h\,(n+1)}_{i-1,j} + r^{h\,(n)}_{i,j+1} + r^{h\,(n+1)}_{i,j-1} - h^2 f^h_{i,j})$$

is more suitable on a serial computer and, moreover, requires less storage.

Note the Poisson's equation $\Delta r = f$ is the Euler–Lagrange equation for the variational principle associated with a membrane problem. The solution can be characterized as the deflection $v(x, y) = r(x, y)$ of a membrane subject to a load $f(x, y)$, and it minimizes the potential energy functional $\mathcal{E}(v) = \int\int_\Omega \frac{1}{2}(v_x^2 + v_y^2) - fv\, dx\, dy$ [14]. Blake [5] offers an alternative variational principle for lightness. Posing the lightness problem as a variational principle permits the direct application of the finite element discretization method, which for instance does not require a uniform discretization of Ω.

B. Results

A four-level multiresolution lightness algorithm (with grid sizes 129×129, 65×65, 33×33, and 17×17) was tested on a synthesized Mondrian scene consisting of patches of uniform reflectance, subjected to an illumination which increases quadratically from left to right. The original image, which is 129×129 pixels in size, and three coarser-sampled versions are shown in Fig. 3. All images are quantized to 256 irradiance levels. The grid function $f^h_{i,j}$, shown in Fig. 4, was computed by maintaining only the peaks in the Laplacian of $r^h_{i,j}$. Zero boundary conditions were provided around the edges of the images, and the computation was started from the zero initial approximation $r^h_{i,j} = 0$.

Fig. 5 shows the reconstructed Mondrian which now lacks most of the illumination gradient. Reconstruction of the image from the functions shown in Fig. 4 required 33.97 work units.[4] The total number of iterations performed on each level from coarsest to finest, respectively, is 142, 100, 62, and 10. In comparison, a single-level lightness algorithm required about 500 work units to compute a solution of the same accuracy at the finest level in isolation. The single-level algorithm requires at least as many iterations for convergence as there are nodes across the surface, since information at a node propagates only to its nearest neighbors in one iteration. The multilevel lightness algorithm is much more efficient because it propagates information more effectively at the coarser scales.

IV. The Shape-From-Shading Problem

In general, image irradiance depends on surface geometry, scene illuminance, surface reflectance, and imaging geometry. The shape-from-shading problem is to recover the shape of surfaces from image irradiance. By assuming that illuminance, reflectance, and imaging geometry are constant and known, image irradiance can be related directly to surface orientation.

[4]For comparative complexity analyses, the total computational expense of multigrid methods may be measured in convenient machine independent units. The basic *work unit* is defined as the amount of computation required to perform one iteration on the finest grid in the hierarchy.

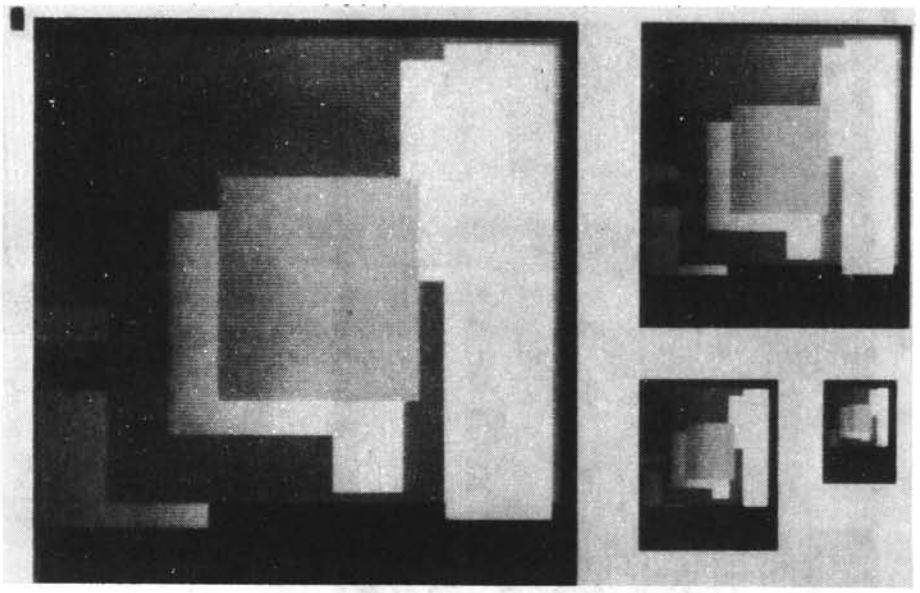

Fig. 3. Synthesized Mondrian images. These images, input to the algorithm, contain patches of uniform reflectances and a left-to-right illumination gradient. The three smaller images are increasingly coarser sampled versions of the largest image which is 129 × 129 pixels, quantized to 256 irradiance levels.

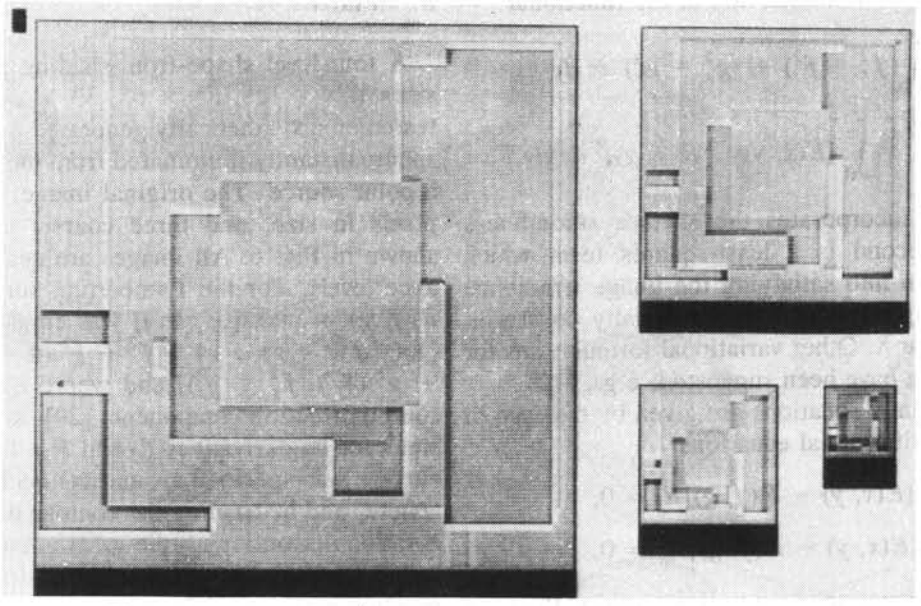

Fig. 4. The grid function $f^h_{i,j}$ on each level. These functions were obtained by maintaining only the peaks in the Laplacian of $r^h_{i,j}$ at each level.

A. Analysis

Let $u(x, y)$ be a surface patch with constant albedo defined over a bounded planar region Ω. The relationship between the surface orientation at a point (x, y) and the image irradiance there $E(x, y)$ is denoted by $R(p, q)$, where $p = u_x$ and $q = u_y$ are the first partial derivatives of the surface function at (x, y). The shape-from-shading problem can be posed as a nonlinear, first-order partial differential equation in two unknowns, called the image-irradiance equation: $E(x, y) - R(p, q) = 0$ [26]. Surface orientation cannot be computed strictly locally because image irradiance provides a single measurement, while surface orientation has two independent components. The image irradiance equation provides only one explicit constraint on surface orientation.

Ikeuchi and Horn [29] proposed an additional surface smoothness constraint and the use of surface occluding contours as boundary conditions. Since the p–q parameterization of surface orientation becomes unbounded at occluding contours, however, surface orientation was reparameterized in terms of the (bounded) stereographic mapping: $f = 2ap$, $g = 2aq$, where $a = 1/(1 + \sqrt{1 + p^2 + q^2})$.

These considerations are formalized by a variational

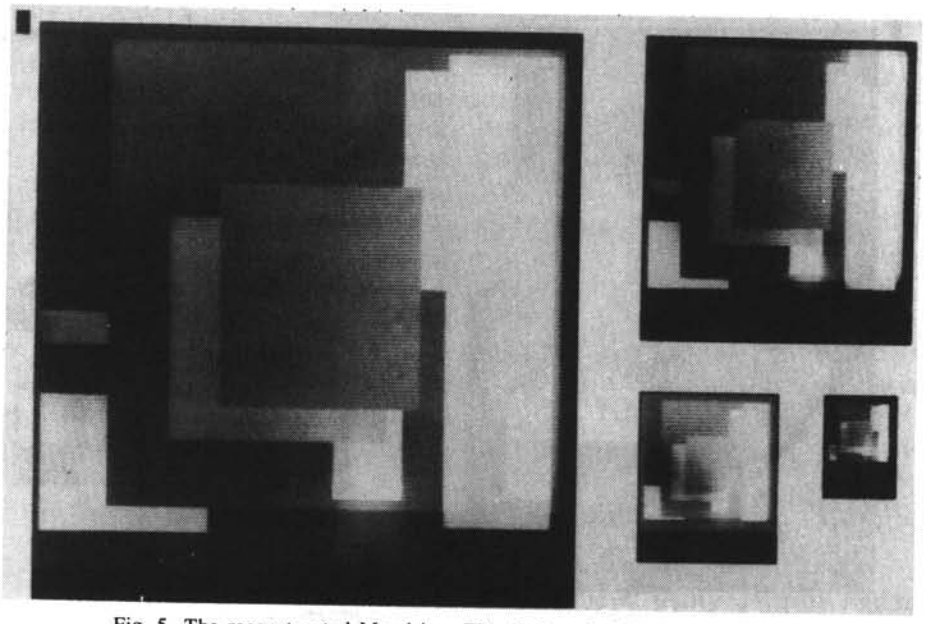

Fig. 5. The reconstructed Mondrian. This is the solution computed after 33.97 work units by the four-level lightness algorithm. Most of the illumination gradient in Fig. 3 has been eliminated.

principle involving the minimization of the functional

$$\mathcal{E}(f, g) = \int\int_{\Omega} (f_x^2 + f_y^2) + (g_x^2 + g_y^2)\, dx\, dy + \frac{\lambda}{2} \int\int_{\Omega} [E(x, y) - R(f, g)]^2\, dx\, dy.$$

The first integral incorporates the surface smoothness constraint. The second is a least-squares term which coerces the solution into satisfying the image irradiance equation by treating the equation as a penalty constraint weighted by a factor λ. Other variational formulations for shape-from-shading have been suggested, e.g., [12].

The Euler–Lagrange equations are given by the system of coupled partial differential equations

$$\Delta f - \lambda[E(x, y) - R(f, g)]R_f = 0,$$

$$\Delta g - \lambda[E(x, y) - R(f, g)]R_g = 0.$$

Discretizing these equations on a uniform grid with spacing h using standard finite difference approximations yields the Jacobi relaxation scheme.

$$f_{ij}^{h\,(n+1)} = \bar{f}_{i,j}^{h\,(n)} + \lambda[E_{i,j} - R(f_{i,j}^{h\,(n)}, g_{i,j}^{h\,(n)})][R_f]_{i,j}^{(n)},$$

$$g_{i,j}^{h\,(n+1)} = \bar{g}_{i,j}^{h\,(n)} + \lambda[E_{i,j} - R(f_{i,j}^{h\,(n)}, g_{i,j}^{h\,(n)})][R_g]_{i,j}^{(n)},$$

where $\bar{f}_{i,j}^h = [f_{i-1,j}^h + f_{i+1,j}^h + f_{i,j-1}^h + f_{i,j+1}^h]/4$ and $\bar{g}_{i,j}^h = [g_{i-1,j}^h + g_{i+1,j}^h + g_{i,j-1}^h + g_{i,j+1}^h]/4$ are local averages of f^h and g^h at node (i, j) (a factor of 1/4 has been absorbed into λ), $R_f = \partial R/\partial f$, and $R_g = \partial R/\partial g$. On a sequential computer, it is preferable to use the analogous Gauss–Seidel relaxation in the multilevel algorithm, due to its greater stability, faster convergence, and reduced memory requirements. Appropriate boundary conditions can be specified at occluding contours in the image.

B. Results

A four-level shape-from-shading algorithm (with grid sizes 129 × 129, 65 × 65, 33 × 33, and 17 × 17) was tested on a synthetically generated image of a Lambertian sphere distantly illuminated from the viewing direction by a point source. The original image, which is 129 × 129 pixels in size, and three coarser-sampled versions are shown in Fig. 6. All images are quantized to 256 irradiance levels. For the Lambertian surface, the expression $R(f, g) = \max [0, \cos i]$ was employed, where $\cos i = [16(f_s f + g_s g) + (4 - f^2 - g^2)(4 - f_s^2 - g_s^2)]/[(4 + f^2 + g^2)(4 + f_s^2 + g_s^2)]$ and where f_s and g_s are the light source direction components [29], and analogous expressions for its derivatives R_f and R_g. The orientation of the surface was specified around the occluding contour of the sphere, and by treating the contour itself as a possible orientation discontinuity, the grid functions f and g were allowed to make discontinuous transitions across it. Computation was started from the zero initial approximation $f = g = 0$.

The solution computed at the four levels after 6.125 work units are shown in Fig. 7. The total number of iterations performed on each level from coarsest to finest respectively is 32, 10, 4, and 4. In comparison, a single-level algorithm required close to 200 work units to obtain a solution of the same accuracy at the finest level in isolation. As in the case of the lightness problem, the single-level algorithm requires at least as many iterations for convergence as there are nodes across the surface, since information at a node propagates only to its nearest neighbors after each iteration. Convergence is somewhat faster, however, because shading information is available at every node inside the occluding contour to constrain surface shape according to the image irradiance equation. In any

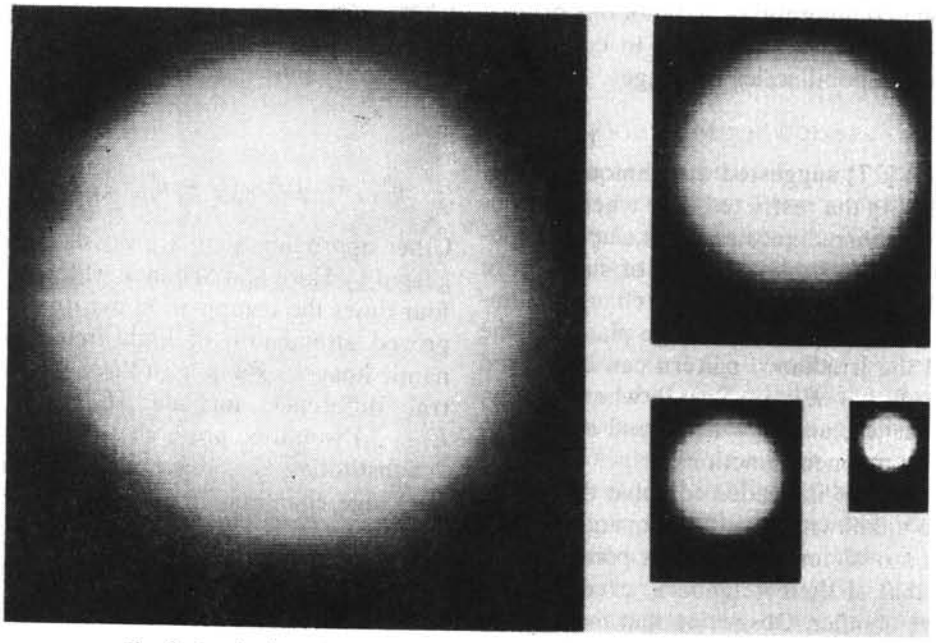

Fig. 6. Lambertian sphere images. These synthetic images input to the algorithm show a Lambertian sphere distantly illuminated from the viewing direction. The three smaller images are increasingly coarser sampled versions of the largest image which is 129 × 129 pixels, quantized to 256 irradiance levels.

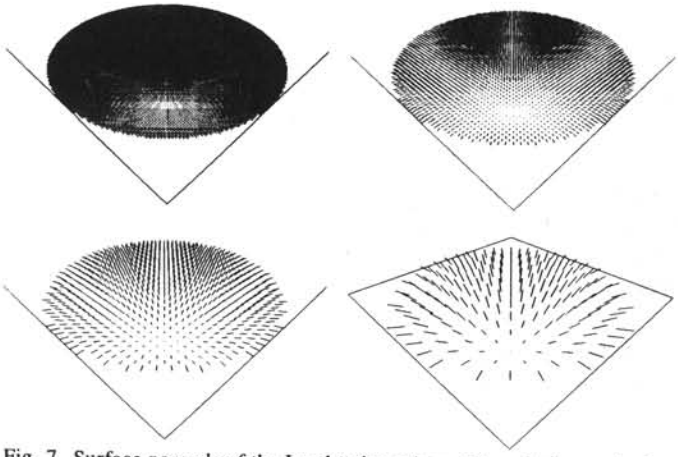

Fig. 7. Surface normals of the Lambertian sphere. The solution at the four resolutions that was obtained after 6.125 work units is shown.

case, the multilevel shape-from-shading algorithm is again much more efficient because it enables information to propagate quickly at the coaser scales.

To obtain a representation of the surface in depth, the surface normals in Fig. 7 were introduced as orientation constraints to a four-level surface reconstruction algorithm with identical grid sizes [43]. The normal vectors were first transformed from the f–g stereographic parameterization used in the shape-from-shading algorithm to the p–q gradient space parameterization used in the surface reconstruction algorithm using the formulas $p = -4f/(f^2 + g^2 - 4)$ and $q = -4g/(f^2 + g^2 - 4)$. Nodes outside the occluding contour of the sphere were treated as depth discontinuities. Fig. 8 shows the surfaces generated by the algorithm at the three coarsest resolutions. The reconstruction required an additional 8.8 work units.

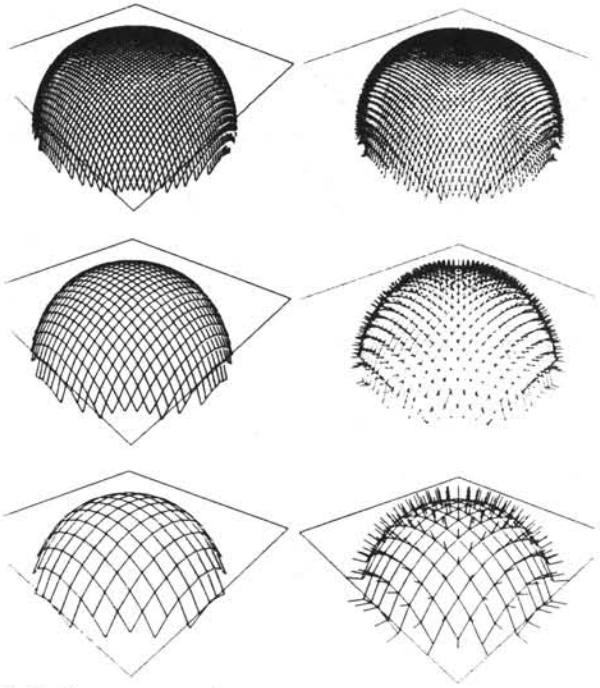

Fig. 8. Surface representations of the Lambertian sphere. The depth representations on the left were generated by a four-level surface reconstruction algorithm in 8.8 work units using the normal vectors in Fig. 7 as orientation constraints. On the right, the orientation constraints are depicted as "needles" on the reconstructed surfaces. Only the three coarsest levels are shown, since the finest resolution surface is too dense to render as a 3-D perspective plot.

V. The Optical Flow Problem

Optical flow is the distribution of apparent velocities of irradiance patterns in the dynamic image. The velocity field and its discontinuities can be an important source of

information about the configurations and motions of visible surfaces. The optical flow problem is to compute a velocity field from a temporal series of images.

A. Analysis

Horn and Schunck [27] suggested a technique for determining optical flow in the restricted case where the observed velocity of image irradiance patterns can be attributed directly to small interframe motions of surfaces in the scene. Under these circumstances, the change in image irradiance at a point (x, y) in the image plane at time t and the motion of the irradiance pattern can be related by the flow equation $E_x u + E_y v + E_t = 0$, where $E(x, y, t)$ is the image irradiance, and $u = dx/dt$ and $v = dy/dt$ are the optical flow component functions.

An additional constraint is needed to solve this linear equation for the two unknowns u and v. If opaque objects undergo rigid motion or deformation, most points have a velocity similar to that of their neighbors, except where surfaces occlude one another. Observing that the velocity field varies smoothly almost everywhere, optical flow can be determined by finding the flow functions $u(x, y)$ and $v(x, y)$ which minimize the functional

$$\mathcal{E}(u, v) = \alpha^2 \iint_\Omega (u_x^2 + u_y^2) + (v_x^2 + v_y^2)\, dx\, dy + \iint_\Omega (E_x u + E_y v + E_t)^2\, dx\, dy,$$

where α is a constant. The first term is the smoothness constraint, while the second is a least-squares penalty expression which coerces the flow field into satisfying the flow equation. Related variational formulations of the optical flow problem have been suggested (e.g., [13], [33]).

The Euler–Lagrange equations for the functional $\mathcal{E}$ are given by [27]

$$E_x^2 u + E_x E_y v = \alpha^2 \Delta u - E_x E_t,$$
$$E_x E_y u + E_y^2 v = \alpha^2 \Delta v - E_y E_t.$$

Assuming a cubical network of nodes with spacing h, where i, j, and k index nodes along the x, y, and t axes, respectively, the following finite difference formulas may be used to discretize the differential operators:

$$[E_x]^h_{i,j,k} = \frac{1}{2h}(E^h_{i+1,j,k} - E^h_{i-1,j,k}),$$

$$[E_y]^h_{i,j,k} = \frac{1}{2h}(E^h_{i,j+1,k} - E^h_{i,j-1,k}),$$

$$[E_t]^h_{i,j,k} = \frac{1}{h}(E^h_{i,j,k+1} - E^h_{i,j,k}),$$

$$\Delta^h u = \frac{4}{h^2}(\bar{u}^h_{i,j,k} - u^h_{i,j,k}),$$

$$\Delta^h v = \frac{4}{h^2}(\bar{v}^h_{i,j,k} - v^h_{i,j,k}),$$

where

$$\bar{u}^h_{i,j,k} = \tfrac{1}{4}(u^h_{i-1,j,k} + u^h_{i,j+1,k} + u^h_{i+1,j,k} + u^h_{i,j-1,k})$$

and

$$\bar{v}^h_{i,j,k} = \tfrac{1}{4}(v^h_{i-1,j,k} + v^h_{i,j+1,k} + v^h_{i+1,j,k} + v^h_{i,j-1,k}).$$

Other approximations are possible, including those suggested by Horn and Schunck which, however, require over four times the computation per iteration to gain some improved attenuation of high frequency error. Given dynamic images over at least three frames, a symmetric central difference formula $[E_t]^h_{i,j,k} = (1/2h)(E^h_{i,j,k+1} - E^h_{i,j,k-1})$ would be preferable, provided it is stable.

Substituting the above approximation into the Euler–Lagrange equations and solving for $u^h_{i,j,k}$ and $v^h_{i,j,k}$ yields the following Jacobi relaxation formulas

$$u^{h\,(n+1)}_{i,j,k} = \bar{u}^{h\,(n)}_{i,j,k} - \frac{\nu^{h\,(n)}_{i,j,k}}{\mu^{h\,(n)}_{i,j,k}} [E_x]^{h\,(n)}_{i,j,k},$$

$$v^{h\,(n+1)}_{i,j,k} = \bar{v}^{h\,(n)}_{i,j,k} - \frac{\nu^{h\,(n)}_{i,j,k}}{\mu^{h\,(n)}_{i,j,k}} [E_y]^{h\,(n)}_{i,j,k},$$

where

$$\mu^h_{i,j,k} = ([E_x]^h_{i,j,k})^2 + ([E_y]^h_{i,j,k})^2 + \frac{4}{h^2}\alpha^2$$

and

$$\nu^h_{i,j,k} = [E_x]^h_{i,j,k}\bar{u}^h_{i,j,k} + [E_y]^h_{i,j,k}\bar{v}^h_{i,j,k} + [E_t]^h_{i,j,k}.$$

The natural boundary conditions of the zero normal derivative are appropriate on the boundaries of surfaces. They can be enforced by copying values to boundary nodes from neighboring interior nodes.

B. Results

A four-level optical flow algorithm (with grid sizes 129 × 129, 65 × 65, 33 × 33, and 17 × 17) was tested on a synthetically generated image of a Lambertian sphere distantly illuminated from the viewing direction by a point source. The sphere expanded uniformly over two frames. The first frame, which is 129 × 129 pixels in size, and three coarser-sampled versions are shown in the left half of Fig. 9. The next frame, in which the sphere has expanded is shown in the right half of the figure. All images are quantized to 256 irradiance levels. The velocity field was specified around the occluding contour of the sphere, and by treating the contour as a possible flow field discontinuity, u and v were allowed to make discontinuous transitions across it. The computation was started from the zero initial approximation $u = v = 0$.

The solution computed on the three coarsest levels after 4.938 work units is shown in Fig. 10 as velocity vectors in xy-space. The total number of iterations performed on each level from coarsest to finest, respectively, is 40, 5, 4, and 3. In comparison, a single-level algorithm required 37 work units to obtain a solution of the same accuracy at the

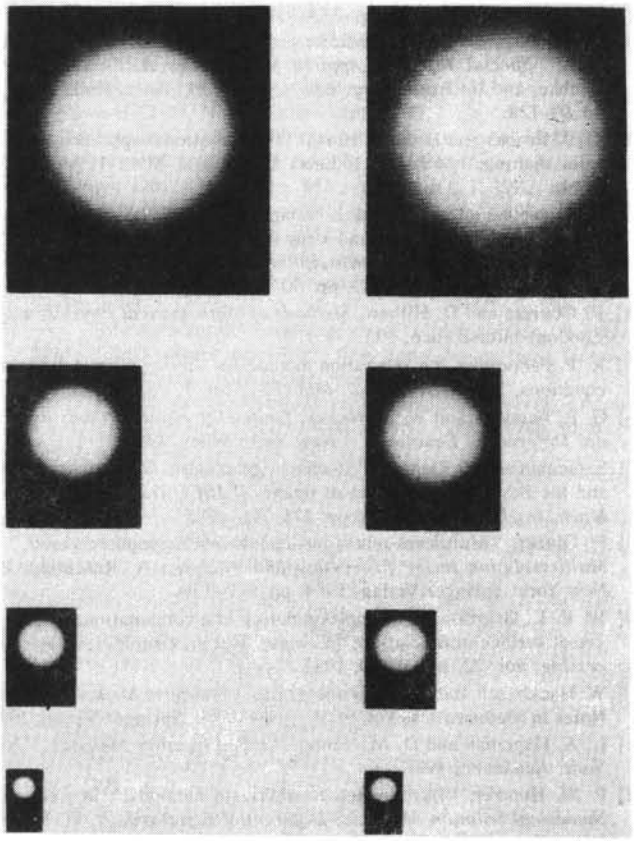

Fig. 9. Lambertian sphere images. These synthetic images input to the algorithms at four resolutions depict a uniformly expanding Lambertian sphere, distantly illuminated from the viewing direction. Frames for the first time instant are shown to the left of frames for the second time instant.

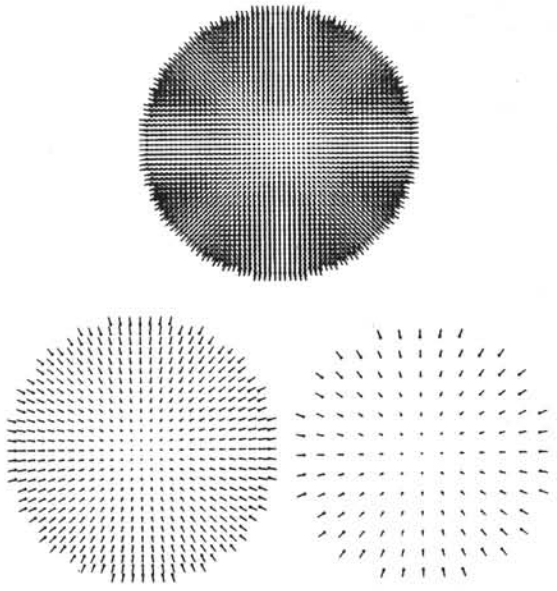

Fig. 10. Velocity vectors for the expanding Lambertian sphere. The solution at the three coarsest resolutions that was obtained after 4.938 work units is shown (the finest-level solution is too dense to depict).

finest level in isolation. Again, the multilevel algorithm is more efficient because it propagates information quickly at the coarser scales. Glazer [18] also reports improvements consistent with the above with regard to the convergence rate of a multilevel optical flow algorithm relative to a single-level algorithm. He employed the Horn–Schunck relaxation formulas in his implementation.

VI. A Methodological Perspective

A primary purpose of image analysis for computer vision is to reconstruct relevant physical characteristics of 3-D scenes from their images. This paper has developed efficient algorithms for three visual reconstruction problems—the computation of lightness (a 2-D, static problem), shape-from-shading (a 3-D, static problem), and optical flow (a 2-D, dynamic problem). However, the approach can be viewed in broader theoretical perspective.

As inverse mathematical problems, visual reconstruction problems tend to be ill-posed in that existence, uniqueness, and stability of their solutions cannot be guaranteed *a priori* [36]. Among the techniques that have been developed to come to grips with ill-posed problems is the method of *regularization* [45]. A major attraction of regularization analysis, from the point of view of the work in the present paper, is that it leads systematically to well-posed variational principles which are amenable to efficient solution by multigrid relaxation methods. As a systematic design strategy for visual algorithms, regularization analysis and multigrid methodology, in conjunction, promise to impact on a broader spectrum of visual reconstruction problems, including image restoration [17], stereopsis [32], registration [1], motion analysis [46], and so on.

Regularization achieves its goal by imposing stabilizing conditions on possible solutions. The most straightforward conditions are global smoothness constraints. Note that identical stabilizing functionals impose the smoothness constraint in both the shape-from-shading and optical flow formulations. The inevitable occurrence of visual discontinuities raises a crucial computational issue. It is important to realize that the iterative algorithms developed in this paper preserve discontinuities. The discontinuities appropriately restrict the propagation of smoothness constraints. The theoretical underpinnings of this technique are based on a generalized "controlled continuity" constraint which is developed in [43] and [44].

Another issue of concern is that the regularization of visual reconstruction problems cannot always be expected to lead to convex variational principles having a unique absolute extremum (and no other relative extrema). Unfortunately, classical relaxation or gradient descent methods are not directly applicable to nonconvex variational principles, since they often get trapped in relative extrema. Stochastic relaxation algorithms (such as simulated annealing) do not suffer this disadvantage, in principle [30]. However, since stochastic relaxation searches for absolute extrema with processors that are restricted to local interactions, it too suffers serious inefficiencies in propagating constraints. Moreover, the nondeterminism of the local computations exacerbates the slow convergence rates. The inefficiency may be ameliorated by developing stochastic multigrid methods.

VII. Conclusion

Many problems in early vision have been formulated as variational principles or as partial differential equations. Such formulations result naturally from the regularization analysis of ill-posed visual reconstruction problems. Once discretized, variational and differential formulations are amenable to numerical solution by iterative relaxation methods, which readily map into massively parallel computer architectures. Distributed local-support computations, however, are inherently inefficient at propagating constraints over the large network or grid representations prevalent in computer vision applications.

In prior work on surface reconstruction algorithms, we showed that multiresolution relaxation techniques can overcome this inefficiency without sacrificing the local-interconnect nature of the computations. This has been corroborated in the present paper by applying multigrid methods to the well-known problems of computing lightness, shape-from-shading, and optical flow from images. For each problem, the novel multiresolution algorithms are substantially more efficient than the traditional single-level versions.

Beyond its success as a (local) convergence acceleration strategy, multigrid methodology leads to iterative algithms that compute mutually consistent visual representations *over a range of spatial scales*. Such multiresolution representations appear to be crucial in effectively interfacing early visual processing to subsequent tasks such as recognition, manipulation, and navigation.

Acknowledgment

This paper derives from thesis research supervised by M. Brady and S. Ullman. M. Brady, M. Brooks, and B. Horn provided insightful comments on a draft.

References

[1] R. Bajcsy and C. Broit, "Matching of deformed images," in *Proc. 6th Int. Joint Conf. Pattern Recognition*, Munich, 1982, pp. 351–353.

[2] N. S. Bakhvalov, "Convergence of a relaxation method with natural constraints on an elliptic operator," *Zh. Vychisl. Mat. Mat. Fiz.*, vol. 6, pp. 861–885, 1966.

[3] D. H. Ballard, G. E. Hinton, and T. J. Sejnowski, "Parallel visual computation," *Nature*, vol. 306, no. 5938, pp. 21–26, 1983.

[4] K. E. Batcher, "Design of a massively parallel processor," *IEEE Trans. Comput.*, vol. C-26, 1980.

[5] A. Blake, "On lightness computation in Mondrian world," *Central and Peripheral Mechanisms of Color Vision*. New York: Macmillan, 1984.

[6] O. Braddick, F. W. Campbell, and J. Atkinson, "Channels in vision: Basic aspects," in *Handbook of Sensory Physiology: Perception*, vol. 8, R. Held, H. W. Leibowitz, and H. L. Teuber, Eds. Berlin: Springer, 1978, pp. 3–38.

[7] J. M. Brady and A. Yuille, "An extremum principle for shape from contour," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, pp. 288–301, 1984.

[8] K. Brand, "Multigrid bibliography," in *Multigrid Methods* (Lecture Notes in Mathematics, Vol. 960), W. Hackbusch and U. Trottenberg, Eds. New York: Springer-Verlag, 1982, pp. 631–650.

[9] A. Brandt, "Multi-level adaptive technique (MLAT) for fast numerical solution of boundary value problems," in *Proc. 3rd Int. Conf. Num. Meth. Fluid Mechanics*, Paris, 1972, *Lecture Notes in Physics, Vol. 18*. Berlin: Springer-Verlag, 1973.

[10] ——, "Multi-level adaptive solutions to boundary-value problems," *Math. Comput.*, vol. 31, pp. 333–390, 1977.

[11] ——, "Multi-level adaptive finite element methods: Variational problems," *Special Topics of Applied Mathematics*, J. Frehse, D. Pallaschke, and U. Trottenberg, Eds. New York: North-Holland, 1980, pp. 91–128.

[12] M. J. Brooks and B. K. P. Horn, "The variational approach to shape from shading," M.I.T. A.I. Lab., Cambridge, MA, AI Memo 813, 1984.

[13] N. Cornelius and T. Kanade, "Adapting optical flow to measure object motion in reflectance and x-ray image sequences," in *Proc. ACM SIGGRAPH/SIGGART Interdisciplinary Workshop Motion: Representation and Perception*, 1983, pp. 50–58.

[14] R. Courant and D. Hilbert, *Methods of Mathematical Physics*, vol. I. London: Interscience, 1953.

[15] R. P. Fedorenko, "A relaxation method for solving elliptic difference equations," *Zh Vychisl. Mat. Mat. Fiz.*, vol. 1, pp. 922–927, 1961.

[16] G. E. Forsythe and W. R. Wasow, *Finite Difference Methods for Partial Differential Equations*. New York: Wiley, 1960.

[17] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-7, pp. 721–741, 1985.

[18] F. Glazer, "Multilevel relaxation in low level computer vision," in *Multiresolution Image Processing and Analysis*, A. Rosenfeld, Ed. New York: Springer-Verlag, 1984, pp. 312–330.

[19] W. E. L. Grimson, "An implementation of a computational theory of visual surface interpolation," *Comput. Vision, Graphics, Image Processing*, vol. 22, pp. 39–69, 1983.

[20] W. Hackbusch and U. Trottenberg, Eds., *Multigrid Methods* (Lecture Notes in Mathematics, Vol. 960). New York: Springer-Verlag, 1982.

[21] L. A. Hageman and D. M. Young, *Applied Iterative Methods*. New York: Academic, 1981.

[22] P. W. Hemker, "Introduction to multigrid methods," in *Colloquim Numerical Solution of Partial Differential Equations*, V. G. Verwer, Ed. Amsterdam, The Netherlands: Dept. of Numerical Mathematics, Mathematical Center, 1980, pp. 59–97.

[23] E. C. Hildreth, "Computations underlying the measurement of visual motion," *Artificial Intell.*, vol. 23, pp. 309–354, 1984.

[24] W. D. Hillis, "The connection machine," Ph.D dissertation, Dep. Elec. Eng. and Comput. Sci., Massachusetts Inst. Technol., Cambridge, 1985.

[25] B. K. P. Horn, "Determining lightness from an image," *Comput. Graphics Image Processing*, vol. 3, pp. 111–299, 1974.

[26] —, "Obtaining shape from shading information," in *The Psychology of Computer Vision*, P. H. Winston, Ed. New York: McGraw-Hill, 1975, pp. 115–155.

[27] B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artificial Intell.*, vol. 17, pp. 185–203, 1981.

[28] R. A. Hummel and S. W. Zucker, "On the foundations of relaxation labeling processes," *IEEE Trans. Pattern Anal. and Mach. Intell.*, vol. PAMI-5, pp. 267–287, 1983.

[29] K. Ikeuchi and B. K. P. Horn, "Numerical shape from shading and occluding boundaries," *Artificial Intell.*, vol. 17, pp. 141–184, 1981.

[30] S. Kirkpatrick, C. D. Gelatt, Jr., and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, pp. 671–680, 1983.

[31] E. H. Land and J. J. McCann, "Lightness and retinex theory," *J. Opt. Soc. Amer.*, vol. 61, pp. 1–11, 1971.

[32] D. Marr and T. Poggio, "Cooperative computation of stereo disparity," *Science*, vol. 195, pp. 283–287, 1977.

[33] H.-H. Nagel, "Constraints for the estimation of displacement vector fields from image sequences," in *Proc. 8th Int. J. Conf. AI*, Karlsruhe, West Germany, 1983, pp. 945–951.

[34] K. A. Narayanan, D. P. O'leary, and A. Rosenfeld, "Image smoothing and segmentation by cost minimization," *IEEE Trans. Syst., Man, Cybernetics*, vol. SMC-12, pp. 91–96, 1982.

[35] R. A. Nicolaides, "On the l^2 convergence of an algorithm for solving finite element equations," *Math. Comput.*, vol. 31, pp. 892–906, 1977.

[36] T. Poggio and V. Torre, "Ill-posed problems and regularization analysis in early vision," in *Proc. DARPA Imge Understanding Workshop*, New Orleans, LA, L. S. Baumann, Ed., 1984, pp. 257–263.

[37] A. Rosenfeld, Ed., *Multiresolution Image Processing and Analysis*. New York: Springer-Verlag, 1984.

[38] A. Rosenfeld, R. Hummel, and S. W. Zucker, "Scene labeling by relaxation operations," *IEEE Trans. Syst., Man, and Cybernetics*, vol. 6, pp. 420–433, 1976.

[39] G. Smith, "The recovery of surface orientation from image irradiance," in *Proc. DARPA Image Understanding Workshop*, Palo Alto, CA, L. S. Baumann, Ed., 1982, pp. 132–141.
[40] G. Strang and G. J. Fix, *An Analysis of the Finite Element Method*. Englewood Cliffs, NY: Prentice-Hall, 1973.
[41] D. Terzopoulos, "Multilevel reconstruction of visual surfaces: Variational principles and finite element representations," M.I.T. Artificial Intell. Lab., Cambridge, MA, AI Memo 671, 1982; reprinted in *Multiresolution Image Processing and Analysis*, A. Rosenfeld, Ed. New York: Springer-Verlag, 1984.
[42] ——, "Multilevel computational processes for visual surface reconstruction," *Comput. Vision, Graphics, Image Processing*, vol. 24, pp. 52–96, 1983.
[43] ——, "Multiresolution computation of visible-surface representations," Ph.D. dissertation, Elec. Eng. and Comput. Sci., Massachusetts Inst. Technol., Cambridge, 1984.
[44] ——, "Regularization of inverse visual problems involving discontinuities," *IEEE Trans. Pattern Anal. Mach. Intell.*, 1986, to be published.
[45] A. N. Tikhonov and V. A. Arsenin, *Solution of Ill-Posed Problems*. Washington, DC: Winston, 1977.
[46] S. Ullman, "Relaxation and constrained optimization by local processes," *Comput. Graphics Image Processing*, vol. 10, pp. 115–125, 1979.