MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A.I. Memo No. 803 October, 1984

MULTIGRID RELAXATION METHODS AND THE ANALYSIS OF LIGHTNESS, SHADING, AND FLOW

Demetri Terzopoulos

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 necessary for low-level vision. Application of multigrid methods can overcome this drawback, as we established in previous work on 3-D surface reconstruction. In this paper, we develop efficient multiresolution iterative algorithms for computing lightness, shape-from-shading, and optical flow, and we evaluate the performance of these algorithms on synthetic images. The multigrid methodology that we describe is broadly applicable in low-level 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 visual reconstruction problems.

This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's Artificial Intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-80-C-0505 and 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., Québéc, Canada.

1. Introduction

Variational principles and partial differential equations have played a significant role in the mathematical formulation of low-level visual information processing problems (representative examples include [Horn, 1974, 1975; Ullman, 1979; Horn & Schunck, 1981; Ikeuchi & Horn; 1981; Narayanan *et al.*, 1982; Bajcsy & Broit, 1982; Hummel & Zucker, 1983; Grimson, 1983; Terzopoulos, 1982, 1983; Nagel, 1983; Hildreth, 1984; Brady & Yuille, 1984]). An attractive feature of variational and differential formulations (once discretized) is the possibility of computing the desired solutions by a popular class of numerical relaxation algorithms. These iterative algorithms require only local computations which can usually be performed in parallel by many locally communicating processors distributed in computational networks or grids.

Local, parallel algorithms are appealing in the context of low-level vision [Rosenfeld *et al.*, 1976; Ullman, 1979; Ballard *et al.*, 1983]. At a certain level of abstraction they do not appear incompatible with the apparent structure of advanced biological vision systems. Moreover, they are ideally suited to implementation on massively parallel computers with numerous simple, locally interconnected processing elements. Such potentially powerful architectures will certainly proliferate, pending imminent advances in VLSI technology [Batcher, 1980; Hillis, 1981].

The desired solutions to many visual problems appear to possess certain global properties (consistency, smoothness, minimal energy, etc.), which are expressed formally by the variational principle or associated partial differential equation formulations.[1] Given only local communication capabilities among processing elements, however, global properties can only be satisfied *indirectly*, typically by iteratively propagating visual constraints across the grid network. Indirect propagation can result in substantial computational inefficiency, since the computational grids necessary for low-level vision applications tend to be extremely large. Convergence of the iterative process is often so slow as to nearly neutralize the computational power offered by massive parallelism. Indeed, for fine discretizations on large grids, excruciatingly slow convergence rates have been observed in iterative algorithms for computing lightness [Blake, 1984; see also Horn, 1974], shape-from-shading [Ikeuchi & Horn, 1981; Smith 1982], optical flow [Horn & Schunck, 1981; Nagel, 1983], 3-D surfaces [Grimson, 1983; Terzopoulos, 1982, 1983], 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 *multigrid relaxation methods* which are gaining popularity in applied numerical analysis [Hackbusch & Trottenberg, 1982]. The computational structure of

[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 [Courant & Hilbert, 1953].

multigrid methods bears an interesting analogy to the multiresolution nature of spatial frequency channels in the human early visual system [Braddick, *et al.*, 1978]. The methods are also related to certain multiresolution image processing structures that have been proposed, notably pyramids [Rosenfeld, 1984].

In earlier work, we developed an efficient surface reconstruction algorithm based on multigrid relaxation methods [Terzopoulos, 1982, 1983] and we suggested, as has Glazer [1984], 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. We develop novel multiresolution algorithms for each problem. Our empirical results indicate that these algorithms offer order-of-magnitude gains in efficiency over their conventional single level counterparts.

2. Multigrid Methodology

Pioneering investigations into multigrid methodology include the work of [Fedorenko, 1961], [Bakhvalov, 1966], [Brandt, 1973, 1977], and [Nicolaides, 1977]. It has been applied to many boundary value problems (see [Brand, 1982] for an extensive bibliography) and there has also been some development in the context of variational problems [Nicolaides, 1977; Brandt, 1980].

2.1. 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 off 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 [Brandt, 1977]. 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

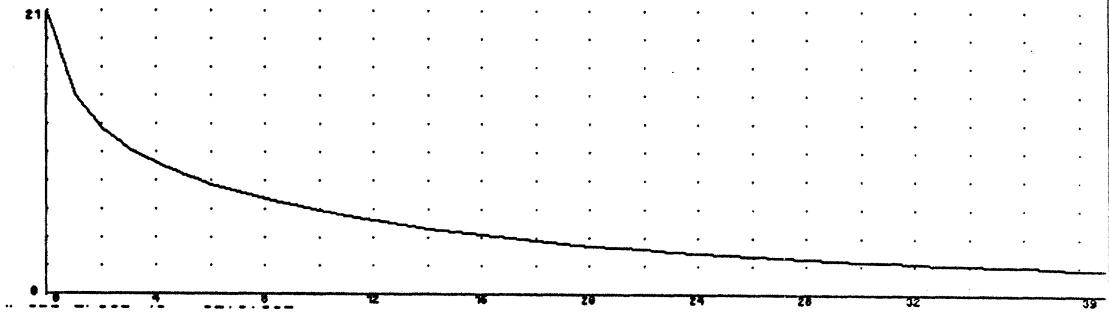

Figure 1. Asymptotic error reduction by relaxation. The mean square (dynamic residual) error is plotted as a function of the iteration number for a sequence of (Gauss–Seidel) relaxation iterations of a surface reconstruction algorithm. The curve exhibits a typical behavior of local iterative methods: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

nodes in the grid [Brandt, 1977]. This can be compared to typical complexities of $O(N^3)$ operations for the solution of model problems by standard (single level) relaxation. As a consequence, multigrid methods potentially offer dramatic increases in efficiency over standard relaxation methods in low-level vision applications, since N tends to be very large (order 10^4 to 10^6, or more). 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.

Our adaptation of multigrid methods to visual processing has a number of features: (i) multiple visual representations covering a range of spatial resolutions, (ii) local, iterative relaxation processes that propagate constraints within each representational level, (iii) local coarse-to-fine *prolongation* processes that allow coarser representations to constrain finer ones, (iv) fine-to-coarse *restriction* processes that allow finer representations to constrain and improve the accuracy of coarser ones, and (iv) (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 prolongation processes are local Lagrange (polynomial) interpolations, and the restriction processes are local averaging operations. The exact form of these operations is problem-dependent.

2.2. Discretization

Appropriate relaxation processes can be derived by local discretization of the continuous

problems. The *finite element method* [Strang & Fix, 1973], a general and powerful local discretization technique, can be applied directly to variational principle formulations of visual problems [Terzopoulos, 1982]. When the visual problem is posed as a partial differential equation, local discretization may be carried out using the *finite difference method* [Forsythe & Wasow, 1960].

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 large number of small subdomains or elements E whose dimensions depend on a fundamental size h. The approximation within elements depends on a small number of parameters — the values of the solution, and/or some of its derivatives, at a set of nodes associated with each element. The power of the method stems from the fact the local approximations can be based on low-order polynomials. This makes it relatively easy to express the continuous functional as a discrete summation over all the element contributions. If the variational principle is quadratic, the resulting discrete problem takes the form of a large system of linear equations $\mathbf{A}^h\mathbf{u}^h = \mathbf{f}^h$, where $\mathbf{u}^h$ is the vector of nodal variables. The finite element method can also be characterized as a systematic procedure for generating finite element approximating spaces whose local-support basis functions make $\mathbf{A}^h$ sparse (i.e., most of its elements are zero).

The finite difference method is applied differently. 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. If the differential operator is linear (as are the Euler–Lagrange equations of quadratic variational principles) and a linear finite difference approximation is employed, the discrete system is again a linear system $\mathbf{A}^h\mathbf{u}^h = \mathbf{f}^h$. Although the total number of nodes N is generally large, each finite difference equation involves only a few nodal variables. Therefore, the linear system is again sparse.

While the finite difference method is generally easier to apply, the finite element method offers a much sounder convergence theory, as well as a flexibility that allows the spatially nonuniform discretization of domains having complicated shapes. Nonetheless, both discretization techniques yield large, sparse systems of linear equations in a wide range of visual applications. A great deal of effort in numerical analysis has been directed to the solution of such systems, which turn out to be especially well suited for solution by local, parallel, iterative methods, particularly the relaxation methods that we have been discussing.

2.3. Multigrid Structure and Coordination

Our spatially uniform discretizations of the continuous visual problems in this paper will yield uniform grids at each level of the multigrid hierarchy. Application of multigrid methods can be simplified substantially given a 2:1 decrease in grid resolution from any level to the next

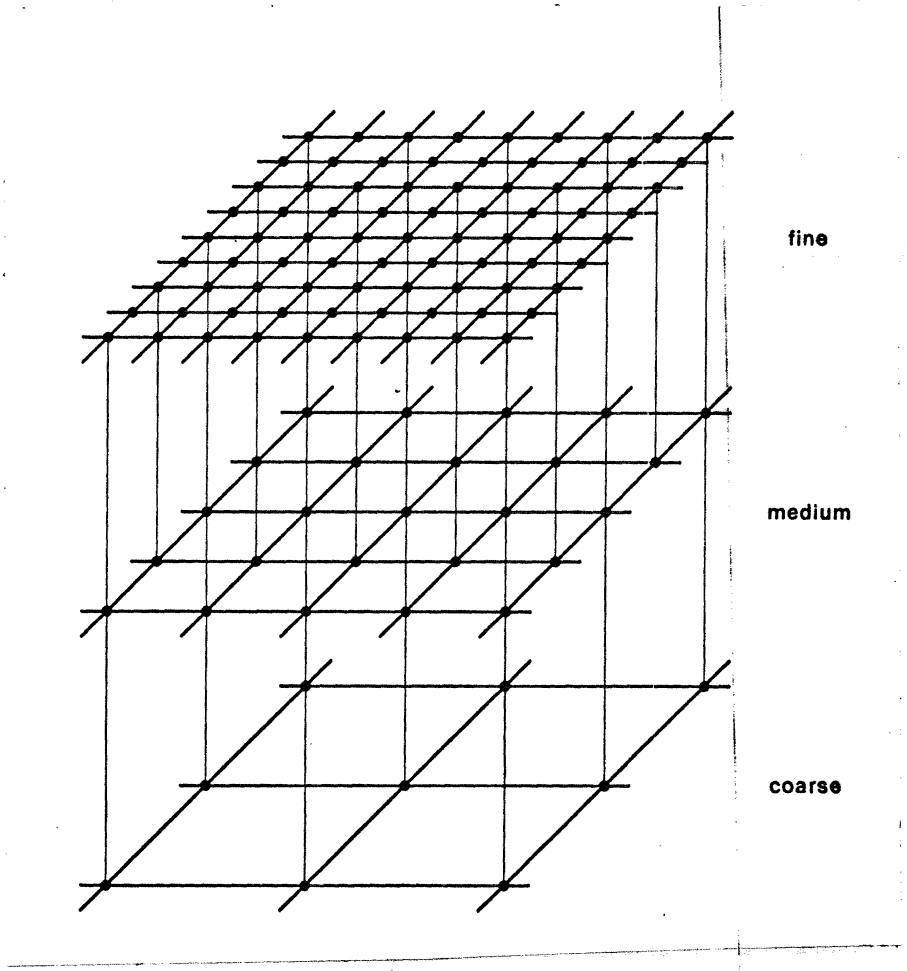

Figure 2. Possible grid organization of a multiresolution algorithm. A small portion of three levels of the 2:1 multigrid hierarchy is shown. Only nearest-neighbor interprocessor connections are included.

coarser level. Fortunately, this resolution ratio appears to be near optimal with regard to multigrid convergence rates [Brandt, 1977]. Fig. 2 illustrates a portion of three grids of a 2:1 multigrid hierarchy. In a serial implementation the central processor operates at each grid node sequentially, whereas in a fully parallel implementation, each node represents a separate processing element within a distributed local-interconnect architecture (see Fig. 2).

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 prolongation, and an *adaptive* multigrid coordination scheme which was employed successfully in our surface reconstruction algorithm (see [Terzopoulos, 1982, 1983] for details). The general coordination scheme first performs a sufficient number of relaxation iterations to solve the coarsest level discrete system $\mathbf{A}^{h_1}\mathbf{u}^{h_1} = \mathbf{f}^{h_1}$ to desired accuracy (procedure SOLVE), and then proceeds to the finest level $l = L$ according to

```
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 a prolongation 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 λ it calls RELAX a total of $n_2^{\lambda-l}(n_1 + n_3)$ times on level $l \neq 1$ and it calls SOLVE $n_2^{\lambda-1}$ times on level 1. In general, most of the relaxation iterations are performed on the coarser levels [Hemker, 1980].

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 [Brandt, 1977]. 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.

3. 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.

3.1. Analysis

The retinex theory of lightness and color proposed by Land and McCann [1971] is based on the observation that illuminance and reflectance patterns differ in their spatial properties. Illuminance changes are usually gradual and, therefore, typically give rise to smooth illumination gradients, while reflectance changes tend to be sharp, since they often originate from abrupt pigmentation changes and surface occlusions. Horn [1974] 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 corresponding 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 appropriate Green's function. We instead pursue a local, iterative solution based on the finite difference method. Suppose that Ω is covered by a uniform square grid with spacing h. We can approximate $\Delta r = r_{xx} + r_{yy}$ using the order h^2 approximations $r_{xx} = (\mathrm{r}^h_{i+1,j} - 2\mathrm{r}^h_{i,j} + \mathrm{r}^h_{i-1,j})/h^2$ and $r_{yy} = (\mathrm{r}^h_{i,j+1} - 2\mathrm{r}^h_{i,j} + \mathrm{r}^h_{i,j-1})/h^2$ to obtain a standard discrete version of Poisson's equation $(\mathrm{r}^h_{i+1,j} + \mathrm{r}^h_{i-1,j} + \mathrm{r}^h_{i,j+1} + \mathrm{r}^h_{i,j-1} - 4\mathrm{r}^h_{i,j})/h^2 = \mathrm{f}^h_{i,j}$. This denotes a system of linear equations with sparse coefficient matrix.

Rearranging, the Jacobi relaxation step is given by

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

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

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

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

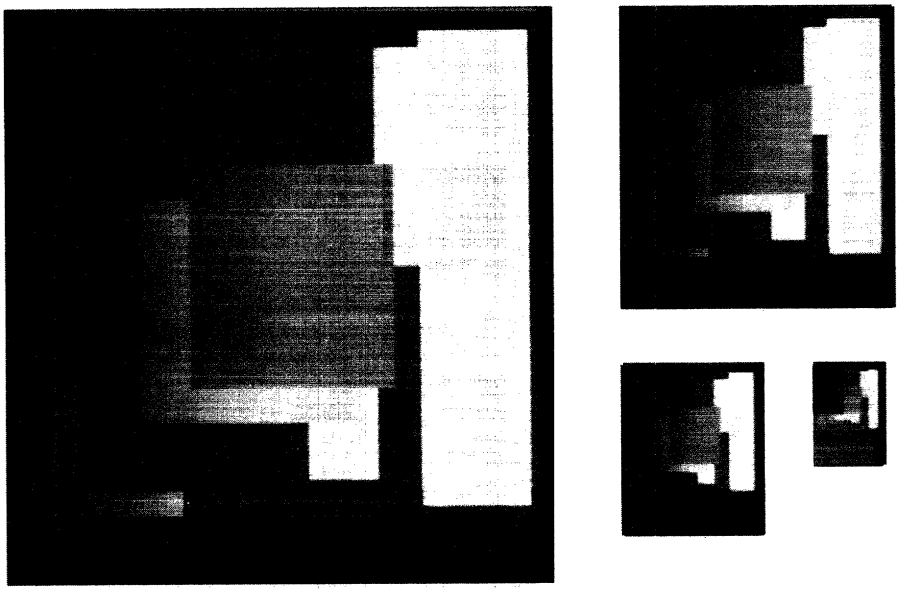

Figure 3. Synthesized Mondrian images. These images, input to the algorithm, contain patches of uniform reflectance and a left-to-right illumination gradient. The three smaller images are increasingly coarser sampled versions of the largest image which is 129×129 pixels, quantized to 256 irradiance levels.

We note in passing that 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 $\mathcal{E}(v) = \int\int_\Omega \frac{1}{2}(v_x^2 + v_y^2) - fv\,dx\,dy$ [Courant & Hilbert, 1953]. Blake [1984] 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.

3.2. 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 $\mathrm{f}_{i,j}^h$, shown in Fig. 4, was computed by maintaining only the peaks in the Laplacian of $\mathrm{r}_{i,j}^h$. Zero boundary conditions were provided around the edges of the images, and the computation was started from the zero initial approximation $\mathrm{r}_{i,j}^h = 0$.

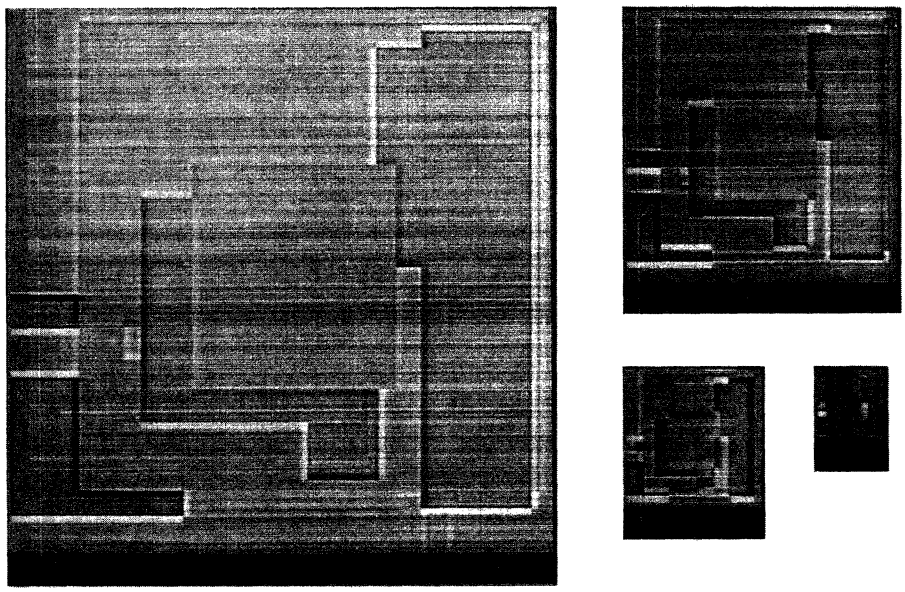

Figure 4. The grid function $f^h_{i,j}$ on each level. These functions were obtained by maintaining only the peaks in the Laplacian of $r^h_{i,j}$ at each level.

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. 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.

4. 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.

4.1. Analysis

Let $u(x, y)$ be a surface patch with constant albedo defined over a bounded planar region Ω.

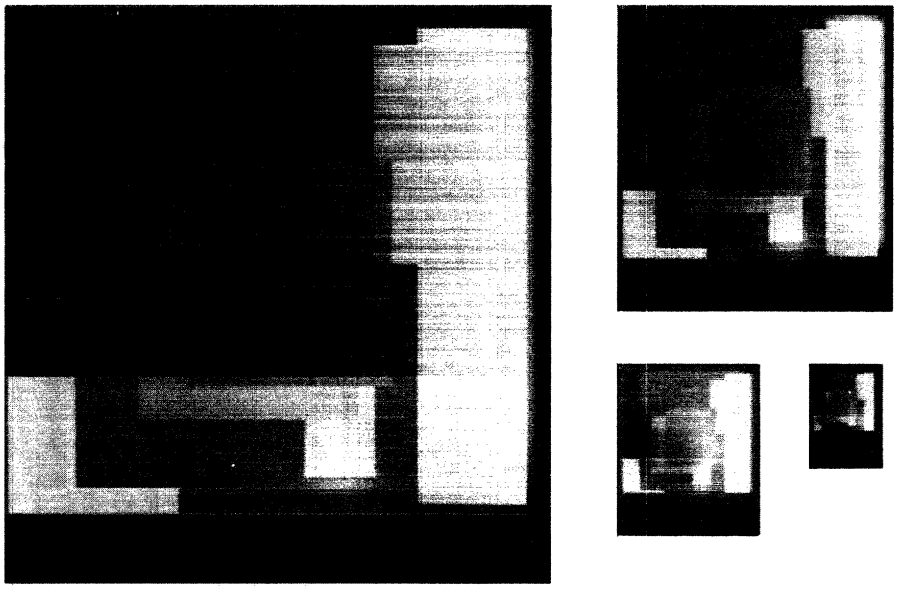

Figure 5. The reconstructed Mondrian. This is the solution computed after 33.97 work units by the four-level lightness algorithm. Most of the illumination gradient in Fig. 3 has been eliminated.

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: $E(x, y) - R(p, q) = 0$ [Horn, 1975]. 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 [1981] 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 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

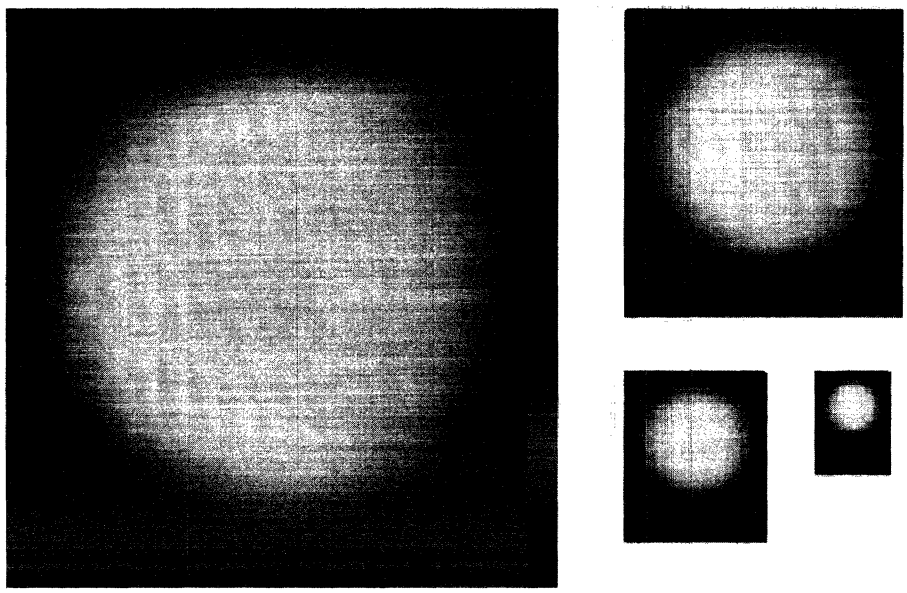

Figure 6. Lambertian sphere images. These synthetic images input to the algorithm show a Lambertian sphere distantly illuminated from the viewing direction. The three smaller images are increasingly coarser sampled versions of the largest image which is 129×129 pixels, quantized to 256 irradiance levels.

a penalty constraint weighted by a factor λ. Other variational formulations for shape-from-shading have been suggested, e.g., [Brooks & Horn, 1984].

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

$$\mathrm{f}_{i,j}^{h\ (n+1)} = \Phi[\mathrm{f}_{i,j}^h]^{(n)} + \lambda[E_{i,j} - R(\mathrm{f}_{i,j}^{h\ (n)}, \mathrm{g}_{i,j}^{h\ (n)})][R_f]_{i,j}^{(n)},$$
$$\mathrm{g}_{i,j}^{h\ (n+1)} = \Phi[\mathrm{g}_{i,j}^h]^{(n)} + \lambda[E_{i,j} - R(\mathrm{f}_{i,j}^{h\ (n)}, \mathrm{g}_{i,j}^{h\ (n)})][R_g]_{i,j}^{(n)},$$

where $\Phi[\mathrm{f}_{i,j}^h] = [\mathrm{f}_{i-1,j}^h + \mathrm{f}_{i+1,j}^h + \mathrm{f}_{i,j-1}^h + \mathrm{f}_{i,j+1}^h]/4$ and $\Phi[\mathrm{g}_{i,j}^h] = [\mathrm{g}_{i-1,j}^h + \mathrm{g}_{i+1,j}^h + \mathrm{g}_{i,j-1}^h + \mathrm{g}_{i,j+1}^h]/4$ 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, we prefer to use the analogous Gauss-Seidel relaxation in our 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.

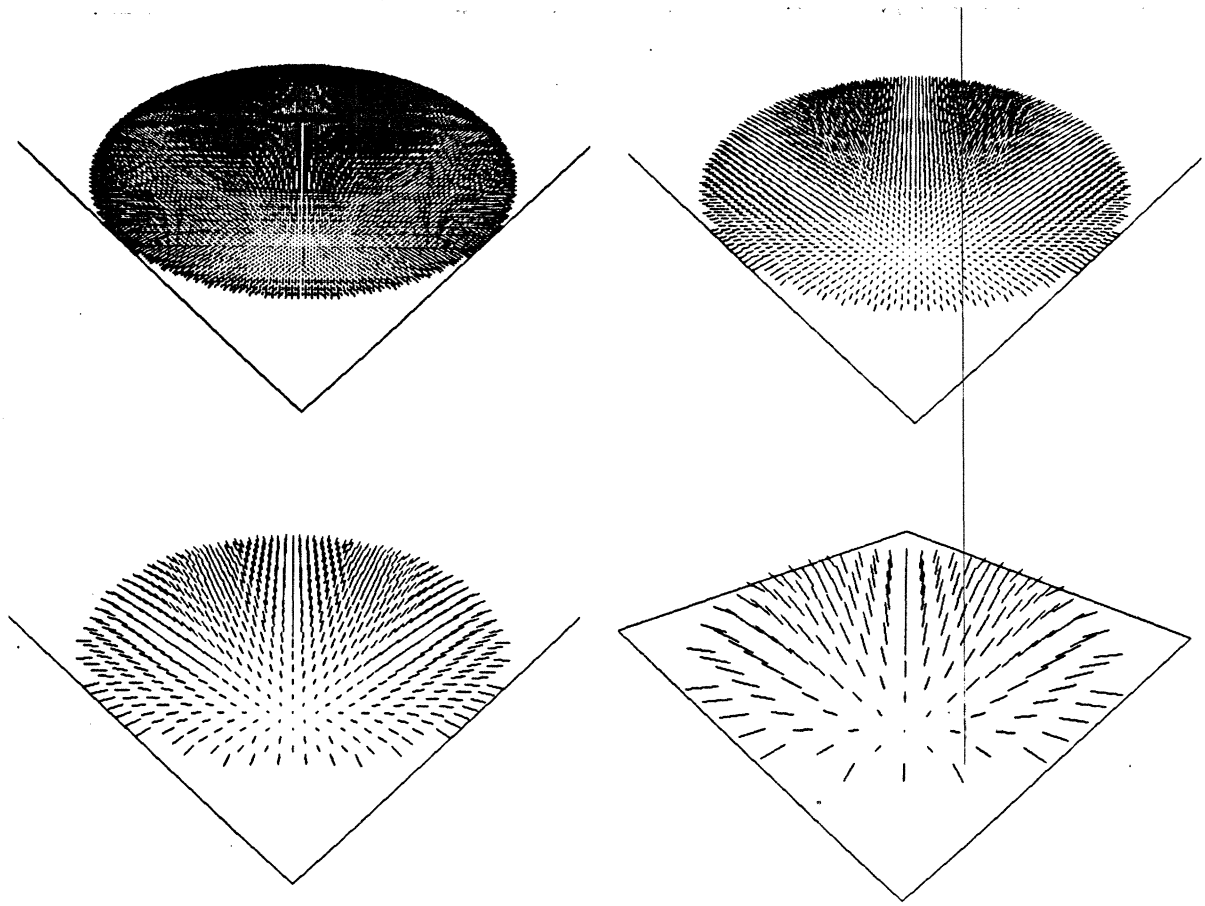

Figure 7. Surface normals of the Lambertian sphere. The solution at the four resolutions that were obtained after 6.125 work units are shown.

4.2. 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, we employed the expression $R(f,g) = \max[0, \cos i]$, 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 [Ikeuchi & Horn, 1981], 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 $f = g = 0$.

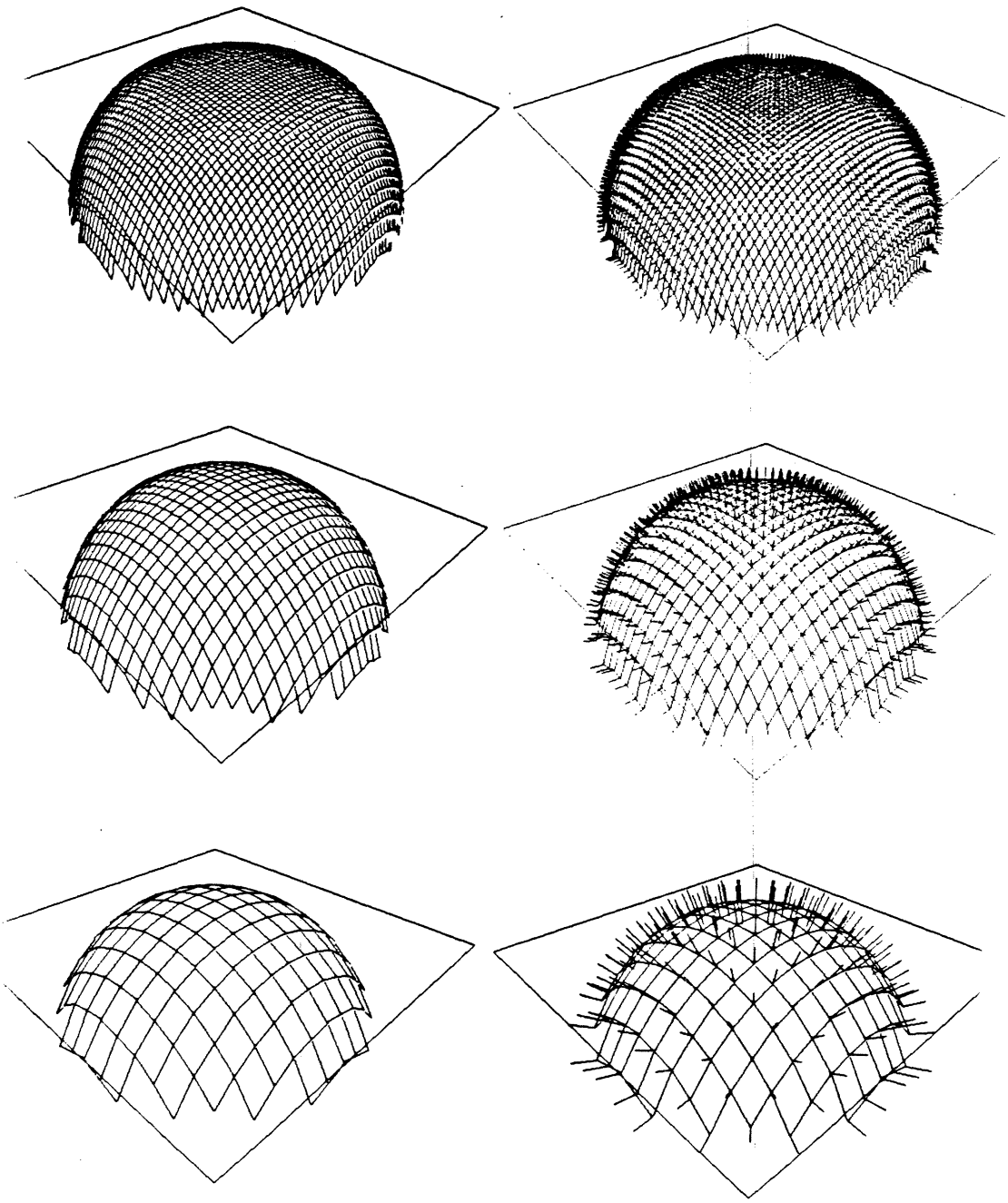

Figure 8. Surface representations of the Lambertian sphere. The depth representations on the left were generated by a four-level surface reconstruction algorithm in 8.8 work units using the normal vectors in Fig. 7 as orientation constraints. On the right, the orientation constraints are depicted as "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.

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 case, the multilevel shape-from-shading algorithm is again much more efficient because it enables information to propagate quickly at the coarser 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 [Terzopoulos, 1984a]. 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.

5. 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.

5.1. Analysis

Horn and Schunck [1981] 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

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

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., [Nagel, 1983], [Cornelius and Kanade, 1983]).

The Euler–Lagrange equations for the functional $\mathcal{E}$ are given by [Horn and Schunck, 1981]

$$\begin{aligned} E_x^2 u + E_x E_y v &= \alpha^2 \Delta u - E_x E_t, \\ E_x E_y u + E_y^2 v &= \alpha^2 \Delta v - E_y E_t. \end{aligned}$$

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

$$\begin{aligned} [E_x]^h_{i,j,k} &= \frac{1}{2h}\left(E^h_{i+1,j,k} - E^h_{i-1,j,k}\right), \\ [E_y]^h_{i,j,k} &= \frac{1}{2h}\left(E^h_{i,j+1,k} - E^h_{i,j-1,k}\right), \\ [E_t]^h_{i,j,k} &= \frac{1}{h}\left(E^h_{i,j,k+1} - E^h_{i,j,k}\right), \\ \Delta^h u &= \frac{4}{h^2}\left(\Phi[\mathrm{u}^h_{i,j,k}] - \mathrm{u}^h_{i,j,k}\right), \\ \Delta^h v &= \frac{4}{h^2}\left(\Phi[\mathrm{v}^h_{i,j,k}] - \mathrm{v}^h_{i,j,k}\right), \end{aligned}$$

where $\Phi[\mathrm{u}^h_{i,j,k}] = \frac{1}{4}\left(\mathrm{u}^h_{i-1,j,k} + \mathrm{u}^h_{i,j+1,k} + \mathrm{u}^h_{i+1,j,k} + \mathrm{u}^h_{i,j-1,k}\right)$ and $\Phi[\mathrm{v}^h_{i,j,k}] = \frac{1}{4}\left(\mathrm{v}^h_{i-1,j,k} + \mathrm{v}^h_{i,j+1,k} + \mathrm{v}^h_{i+1,j,k} + \mathrm{v}^h_{i,j-1,k}\right)$. 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 $[E_t]^h_{i,j,k} = \frac{1}{2h}\left(E^h_{i,j,k+1} - E^h_{i,j,k-1}\right)$ would be preferable, provided it is stable.

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

$$\begin{aligned} \mathrm{u}^h_{i,j,k}{}^{(n+1)} &= \Phi[\mathrm{u}^h_{i,j,k}]^{(n)} - \frac{\nu^h_{i,j,k}{}^{(n)}}{\mu^h_{i,j,k}{}^{(n)}}[E_x]^h_{i,j,k}{}^{(n)}, \\ \mathrm{v}^h_{i,j,k}{}^{(n+1)} &= \Phi[\mathrm{v}^h_{i,j,k}]^{(n)} - \frac{\nu^h_{i,j,k}{}^{(n)}}{\mu^h_{i,j,k}{}^{(n)}}[E_y]^h_{i,j,k}{}^{(n)}, \end{aligned}$$

where $\mu^h_{i,j,k} = \left([E_x]^h_{i,j,k}\right)^2 + \left([E_y]^h_{i,j,k}\right)^2 + \frac{4}{h^2}\alpha^2$ and $\nu^h_{i,j,k} = [E_x]^h_{i,j,k}\Phi[\mathrm{u}^h_{i,j,k}] + [E_y]^h_{i,j,k}\Phi[\mathrm{v}^h_{i,j,k}] + [E_t]^h_{i,j,k}$. 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.

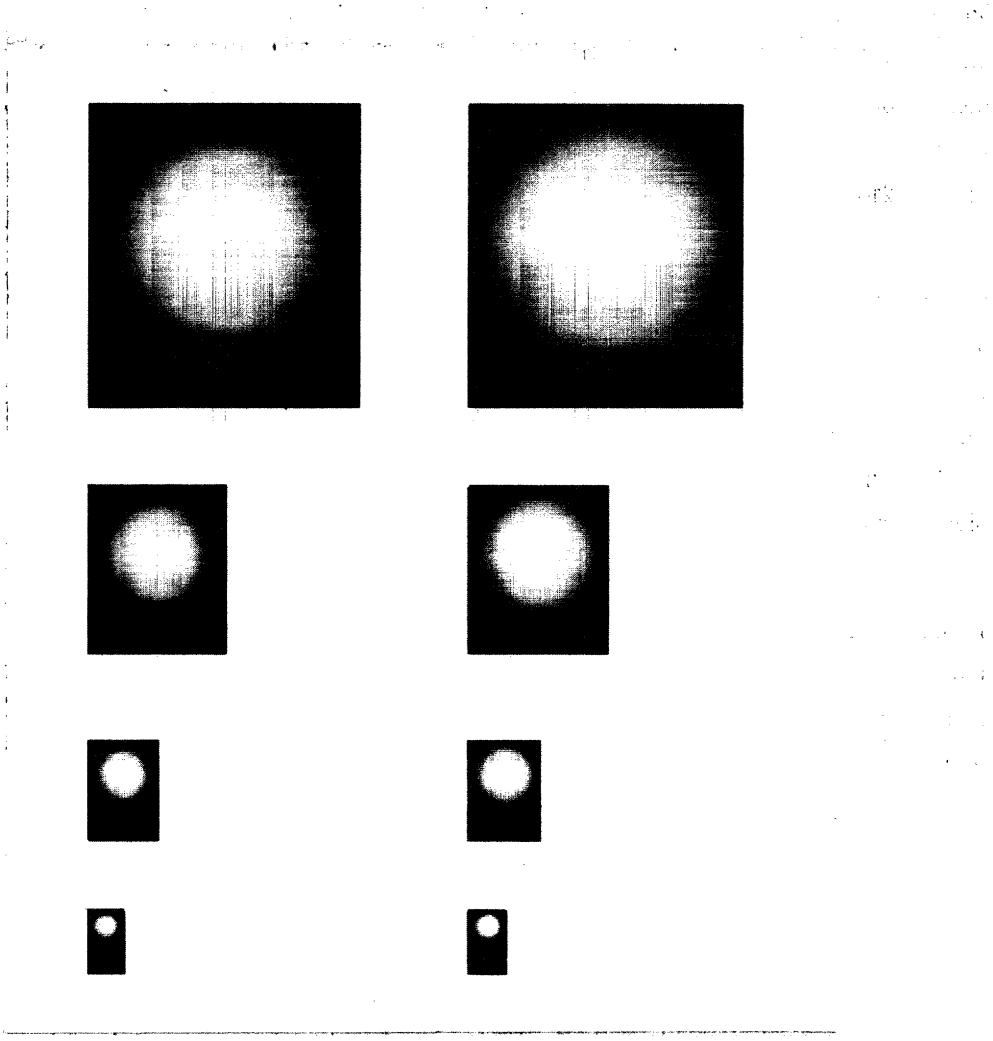

Figure 9. Lambertian sphere images. These synthetic images input to the algorithm at four resolutions depict a uniformly expanding Lambertian sphere, distantly illuminated from the viewing direction. Frames for the first time instant are shown to the left of frames for the second time instant.

5.2. 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 are shown in Fig.

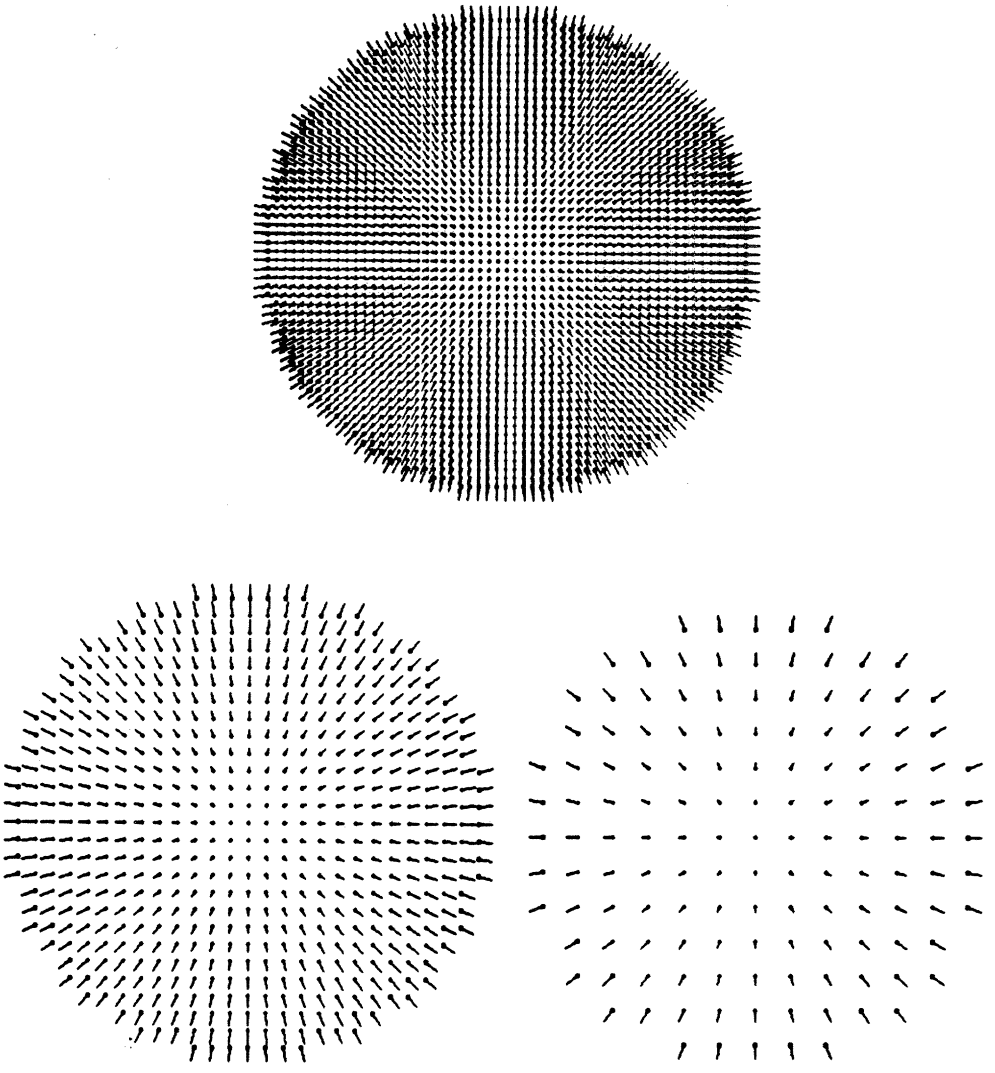

Figure 10. Velocity vectors for the expanding Lambertian sphere. The solution at the three coarsest resolutions that were obtained after 4.938 work units are shown (the finest-level solution is too dense to depict).

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 finest level in isolation. Again, the multilevel algorithm is more efficient because it propagates information quickly at the coarser scales. Glazer [1984] also reports improvements consistent with ours 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 for his implementation.

6. Multigrid Methods, Regularization, and Stochastic Relaxation

A primary purpose of low-level visual processing is to reconstruct relevant physical characteristics of 3-D scenes from their images. We have considered in this paper three different visual reconstruction problems — the computation of lightness from an image (a 2-D, static reconstruction problem), shape-from-shading (a 3-D, static problem), and optical flow (a 2-D, dynamic problem). It was possible to apply multigrid methods because each of these problems was formulated as a variational principle or associated partial differential equation.

As inverse mathematical problems, visual reconstruction problems tend to be mathematically ill-posed, in that existence, uniqueness, and stability of their solutions cannot be guaranteed *a priori* [Poggio and Torre, 1984]. Among the systematic techniques that have been developed to tackle ill-posed problems is the method of *regularization* [Tikhonov and Arsenin, 1977]. Through regularization analysis, ill-posed visual problems can be restated as well-posed variational principles by restricting the possible solutions with appropriate stabilizing functionals. In general, the smoothness properties of stabilizers must be controlled near discontinuities [Terzopoulos, 1984b]. Interestingly, the same stabilizer was used to impose the smoothness constraint in both the shape-from-shading and optical flow problems.

A major attraction of regularization analysis is that it leads systematically to variational principles which permit advantageous use of multigrid relaxation methods. As a visual algorithm design strategy, regularization analysis applied in conjunction with multigrid methodology promises to impact on a broader spectrum of visual reconstruction problems, including image reconstruction and discontinuity detection [Geman and Geman, 1984], stereopsis [Marr and Poggio, 1977], registration [Bajcsy & Broit, 1982], motion field interpolation [Hildreth, 1984], shape-from-contour [Brady & Yuille, 1983], and structure-from-motion [Ullman, 1979].

An 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, without 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 [Kirkpatrick *et al.*, 1983; Hinton & Sejnowski, 1983]. Nonetheless, since stochastic relaxation searches for absolute extrema with processors that are restricted to local interactions, it too suffers serious inefficiencies in propagating constraints. The inherently slow convergence rates are further aggravated by the nondeterministic nature of the local computations. Multigrid methods may ameliorate these problems by facilitating constraint propagation through the use of coarser scales.

7. Conclusion

Many important problems in low-level computer vision can be formulated as variational

principles or as partial differential equations. A particular source of such formulations is 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. However, distributed local-support computations are inherently inefficient at propagating constraints over the large network or grid representations that are encountered in computer vision applications.

In our previous work on surface reconstruction algorithms, we established 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 successfully applying multigrid methods to the well-known problems of computing lightness, shape-from-shading, and optical flow from images. The novel multiresolution algorithms that we designed in the context of each of these problems were shown to be substantially more efficient than the published single level versions.

Beyond its effectiveness as a (local) convergence acceleration strategy, our adaptation of multigrid methodology also leads to iterative algorithms that compute mutually consistent visual representations *over a range of spatial scales.* Multiresolution representations appear to be crucial in interfacing low-level visual processing to subsequent tasks such as recognition, manipulation, and navigation.

Acknowledgements

This paper derives from thesis research supervised by Michael Brady and Shimon Ullman. Insightful comments on a draft were provided by Michael Brady, Michael Brooks, and Berthold Horn.

References

Bajcsy, R., and Broit, C., [1982], "Matching of deformed images," *Proc. Sixth Int. J. Conf. Pattern Recognition*, Munich, 351–353.

Bakhvalov, N.S., [1966], "Convergence of a relaxation method with natural constraints on an elliptic operator," *Zh. vych. Mat. mat. Fiz. (USSR Comp. Math. and Math. Phys.)*, 6, 861–885.

Ballard, D.H, Hinton, G.E., and Sejnowski, T.J., [1983], "Parallel visual computation," *Nature*, **306**, 5938, 21–26.

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

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

Braddick, O., Campbell, F.W., and Atkinson, J., [1978], "Channels in vision: basic aspects," *Handbook of Sensory Physiology: Perception*, Vol. 8, R. Held, H.W. Leibowitz, and H.L. Teuber (ed.), Springer, Berlin, 3–38.

Brady, J.M., and Yuille, A., [1984], "An extremum principle for shape from contour," *IEEE Trans. Pattern Analysis and Machine Intelligence*, PAMI-6, 288–301.

Brand, K., [1982], "Multigrid Bibliography," *Multigrid Methods*, Lecture Notes in Mathematics, Vol. 960, Hackbusch, W., and Trottenberg, U. (ed.), Springer–Verlag, New York, 631–650.

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

Brandt, A., [1977], "Multi-level adaptive solutions to boundary-value problems," *Math. Comp.*, **31**, 333–390.

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

Brooks, M.J., and Horn, B.K.P., [1984], A variational approach to shape from shading, MIT A.I. Lab., Cambridge, MA, AI Memo No. 813.

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

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

Fedorenko, R.P., [1961], "A relaxation method for solving elliptic difference equations," *Zh. vych. Mat. mat. Fiz. (USSR Comp. Math. and Math. Phys.)*, **1**, 922–927.

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

Geman, S., and Geman, D., [1984], "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Analysis & Machine Intelligence*, to appear.

Glazer, F., [1984], "Multilevel relaxation in low level computer vision," *Multiresolution Image Processing and Analysis*, A. Rosenfeld (ed.), Springer–Verlag, New York, 312–330.

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

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

Hemker, P.W., [1980], "Introduction to multigrid methods," *Colloqium Numerical Solution of Partial Differential Equations*, V.G. Verwer (ed.), Dept. of Numerical Mathematics, Mathematical Center, Amsterdam, 59–97.

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

Hillis, W.D., [1981], The connection machine, MIT A.I. Lab., Cambridge, MA, AI Memo No. 646.

Hinton, G.E., and Sejnowski, T.J., [1983], "Analyzing cooperative computation," *Proc. 5th Annual Conf. of the Cognitive Society of America*, Rochester, NY.

Horn, B.K.P., [1974], "Determining lightness from an image," *Computer Graphics and Image Processing*, 3, 111–299.

Horn, B.K.P., [1975], "Obtaining shape from shading information," *The Psychology of Computer Vision*, P.H. Winston (ed.), McGraw–Hill, New York, 115–155.

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

Hummel, R.A., and Zucker, S.W., [1983], "On the foundations of relaxation labeling processes," *IEEE Trans. Pattern Analysis and Machine Intelligence*, PAMI-5, 267–287.

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

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

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

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

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

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

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

Poggio, T., and Torre, V., [1984], Ill-posed problems and regularization analysis in early vision, MIT A.I. Lab., Cambridge, MA, AI Memo No. 773.

Rosenfeld, A. (ed.), [1984], *Multiresolution Image Processing and Analysis*, Springer–Verlag, New York.

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

Smith, G., [1982], "The recovery of surface orientation from image irradiance," *Proc. DARPA Image Understanding Workshop*, Palo Alto, CA, 132–141.

Strang, G., and Fix, G.J., [1973], *An Analysis of the Finite Element Method*, Prentice–Hall, Englewoods Cliffs, NJ.

Terzopoulos, D., [1982], Multilevel reconstruction of visual surfaces: Variational principles and finite element representations, MIT A.I. Lab., Cambridge, MA, AI Memo No. 671, reprinted in *Multiresolution Image Processing and Analysis*, A. Rosenfeld (ed.), Springer–Verlag, New York, 1984.

Terzopoulos, D., [1983], "Multilevel computational processes for visual surface reconstruction," *Computer Vision, Graphics, and Image Processing*, 24, 52–96.

Terzopoulos, D., [1984a], Multiresolution computation of visible-surface representations, Ph.D. thesis, Department of Electrical Engineering and Computer Science, MIT, Cambridge, MA.

Terzopoulos, D., [1984b], "Controlled-smoothness stabilizers for the regularization of ill-posed visual problems involving discontinuities," *Proc. DARPA Image Understanding Workshop*, New Orleans, November, 225–229.

Tikhonov, A.N., and Arsenin, V.A., [1977], *Solutions of Ill-Posed Problems*, Winston & Sons, Washington, DC.

Ullman, S., [1979], "Relaxation and constrained optimization by local processes," *Computer Graphics and Image Processing*, 10, 115–125.