EFFICIENT MULTIRESOLUTION ALGORITHMS FOR COMPUTING LIGHTNESS, SHAPE-FROM-SHADING, AND OPTICAL FLOW

Demetri Terzopoulos
MIT Artificial Intelligence Laboratory
545 Technology Square
Cambridge, MA 02139

Abstract

Problems in machine vision that are posed as variational principles or partial differential equations can often be solved by local, iterative, and parallel algorithms. A disadvantage of these algorithms is that they are inefficient at propagating constraints across large visual representations. Application of multigrid methods has overcome this drawback with regard to the computation of visible-surface representations. We argue that our multiresolution approach has wide applicability in vision. In particular, we describe efficient multiresolution iterative algorithms for computing lightness, shape-from-shading, and optical flow, and evaluate the performance of these algorithms using synthesized images.

1. INTRODUCTION

A number of computational tasks in low-level machine vision have been formulated as variational principles (minimization problems) or as (elliptic) partial differential equations (PDEs) (e.g., [1, 2, 8, 9, 10, 15, 17]). Under certain (self adjointness) conditions, PDE formulations can be linked to variational principles, as necessary conditions for minima, through the Euler–Lagrange equations of the calculus of variations [4]. An attractive feature of many variational principle and associated PDE formulations, once discretized, is that their solutions can be computed by iterative algorithms requiring only local computations which can be performed in parallel by many simple processors in locally-connected networks or grids. Such algorithmic structures are appealing, both in view of the apparent structure of biological vision systems and the imminent proliferation of massively parallel, locally connected VLSI processors for vision.

Visual representations usually possess certain essential global properties (consistency, smoothness, minimal energy, etc.) which the variational principle or PDE formulations aim to capture formally. Given only local processing capabilities, global properties must be satisfied indirectly, typically by propagating visual information across grids through iteration. Substantial computational inefficiency can result since the computational grids tend to become extremely large in machine vision applications. Convergence of the iterative process is often so slow as to nearly nullify the potential benefits of massive parallelism. A case in point is the local, iterative computation of visible-surface representations from scattered, local estimates of surface shape [14–16].

Multiresolution processing in hierarchical representations can be effective in counteracting the computational sluggishness of local, iterative solutions to vision problems posed as variational principles or PDEs. Multigrid methods [7], efficient techniques for solving PDEs numerically, have been adapted successfully in our previous work to the computation of visible-surface representations [14, 15]. An objective of this paper is to demonstrate that this methodology has broad applicability in vision (see also [14, 6]). After a brief overview of multigrid methods, we study, in turn, the iterative computation of lightness, shape-from-shading, and optical flow from images. We present empirical evidence that our multiresolution algorithms can be orders of magnitude more efficient than conventional single level versions.

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 N00014-75-C-0643, and the System Development Foundation. The author gratefully acknowledges the financial support of the Natural Sciences and Engineering Research Council of Canada and the Fonds F.C.A.C. Québec, Canada.

2. MULTIGRID METHODS

Progress has recently been made in applied numerical analysis with regards to multigrid methods (see, e.g., [3, 7]). We have drawn upon these techniques and the associated theory in our work on multiresolution computational vision to gain computational and representational leverage. Our adaptation of these methods provide an efficient means of computing consistent visual representations at multiple scales. In multigrid methods, a hierarchy of discrete problems is formulated and local, multilevel relaxation schemes are applied to accelerate convergence. Our algorithms have several components: (i) multiple visual representations over a range of spatial resolutions, (ii) local intralevel processes that iteratively 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 improve the accuracy of coarser ones, and (iv) adaptive (recursive) coordination strategies [3] that enable the hierarchy of representations and component processes to cooperate towards increasing efficiency (see [14, 15] for details).

Generally, the intralevel processes are familiar Gauss–Seidel or Jacobi relaxation [5], the prolongation processes are local Lagrange (polynomial) interpolations, and the restriction processes are local averaging operations [3]. The precise form of these processes is problem-dependent. The algorithms in this paper employ simple injection for the fine-to-coarse restrictions and bilinear interpolation for coarse-to-fine prolongation. Appropriate relaxation operations are derived by discretizing the continuous vision problems. The finite difference method [5] can be employed when a problem is posed as a PDE, whereas the finite element method [13], a more general and powerful discretization technique, can be applied directly to variational principle formulations [14–16].

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, assuming no precise knowledge about either reflectance or illuminance.

The retinex theory of lightness and color proposed by Land and McCann [12] 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 [9] 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 color.

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 $B(x,y) = S(x,y) \times R(x,y)$. Denoting the logarithms of the above functions as lowercase quantities, we have $b(x,y) = s(x,y) + r(x,y)$. Next, Horn employed the Laplacian operator Δ which gives $d(x,y) = \Delta b(x,y) = \Delta s(x,y) + \Delta r(x,y)$. In the Mondrian situation illuminance is assumed to vary smoothly so that $\Delta s(x,y)$ will be finite everywhere, while $\Delta r(x,y)$ will exhibit pulse doublets at intensity edges separating neighboring regions. A thresholding operator T can be applied to discard the finite part: $T[d(x,y)] = \Delta r(x,y) \equiv f(x,y)$. 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 PDE by convolution with the appropriate Green's function. We will instead pursue an iterative solution which is also local and parallel, hence apparently biologically feasible. The finite difference method can be applied directly. 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 $r_{xx} = (r^h_{i+1,j} - 2r^h_{i,j} + r^h_{i-1,j})/h^2$ 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 whose coefficient matrix is sparse and banded [5].

Rearranging, the Jacobi relaxation step is given by

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

Jacobi relaxation is suitable for parallel implementation, whereas Gauss-Seidel relaxation is better suited to a serial computer and, moreover, requires less storage.

The synthesized Mondrian images shown in Figure 1 were input to a four level lightness algorithm (with grid sizes 129×129, 65×65, 33×33, and 17×17). The grid function $f^h_{i,j}$ was computed by maintaining only the local peaks in the Laplacian of $r^h_{i,j}$ at each level. Zero boundary conditions were provided around the edges of the images, and the computation was started from the zero initial approximation $r^h_{i,j} = 0$. Figure 2 shows the reconstructed Mondrian, which lacks much of the illumination gradient. Reconstruction required 33.97 work units, where a work unit is the amount of computation required for an iteration on the finest grid. 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 algorithm required about 500 work units to obtain 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 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.

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 $B(x,y)$ is denoted by $R(p,q)$, where $p = u_x$ and $q = u_y$ are

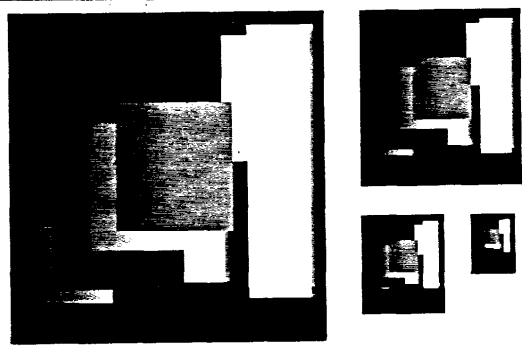

Figure 1. Synthetic Mondrian images containing patches of uniform reflectance and an illumination gradient which increases quadratically from left to right. The three smaller images are increasingly coarser sampled versions of the largest image which is 129×129 pixels, quantized to 256 irradiance levels.

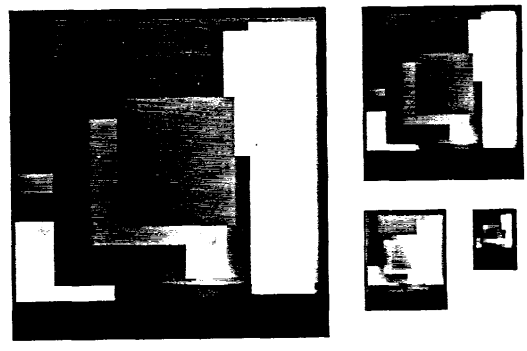

Figure 2. The reconstructed Mondrian computed after 33.97 work units by the four-level multiresolution lightness algorithm. Most of the illumination gradient in Figure 1 has been eliminated.

the first partial derivatives of the surface function at (x,y). The shape-from-shading problem can be posed as a nonlinear, first-order PDE in two unknowns, called the image-irradiance equation [11]: $B(x,y) - R(p,q) = 0$. Clearly, 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 one explicit constraint on surface orientation. Ikeuchi and Horn [11] employed an additional surface smoothness constraint. An appropriate set of boundary conditions is necessary to solve the problem, and they suggested the use of occluding boundaries of surfaces. Since the p–q parameterization of surface orientation becomes unbounded at occluding boundaries however, they reparameterized surface orientation in terms of the stereographic mapping: $f = 2p\alpha$, $g = 2q\alpha$, where $\alpha = (\sqrt{1+p^2+q^2} - 1)/(p^2+q^2)$.

The above considerations were formalized in a variational principle involving the minimization of the functional

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

The first integral incorporates the surface smoothness constraint. The second is a least-squares term which attempts to coerce the solution into satisfying the image irradiance equation, thus treating the image irradiance equation as a penalty constraint weighted by a factor λ. The Euler–Lagrange equations are given by the following system of coupled PDEs

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

Discretizing the above equations on a uniform grid with spacing h using the standard finite difference approximations, we obtain the Jacobi relaxation scheme

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

where $\Phi[f_{i,j}^h] = [f_{i-1,j}^h + f_{i+1,j}^h + f_{i,j-1}^h + f_{i,j+1}^h]/4$ and $\Phi[g_{i,j}^h] = [g_{i-1,j}^h + g_{i+1,j}^h + g_{i,j-1}^h + g_{i,j+1}^h]/4$ 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 $R_g = \partial R/\partial g$. We employ the Gauss-Seidel form of the relaxation in our multilevel algorithm. Appropriate boundary conditions may be obtained from occluding boundaries in the image (see [11] for a discussion).

A four level shape-from-shading algorithm (with grid sizes 129×129, 65×65, 33×33, and 17×17) was tested on the synthetically-generated Lambertian sphere images shown in Figure 3. Surface orientation was specified around the occluding boundary of the sphere, which was marked as a discontinuity, and the computation was started from the zero initial approximation, $f = g = 0$ within the sphere. The solution was obtained after 6.125 work units. 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. Unlike the lightness algorithm, however, the shape-from-shading algorithm employs shading information and the image irradiance equation to constrain the surface shape within the surface boundaries. For this reason, convergence is expected to be faster.

The surface normals computed by the shape-from-shading algorithm at the three coarsest resolutions are represented in Figure 4 as "needles." These needles are shown lying on a perspective view of the surface in depth. The depth representation was computed by a (four-level) multiresolution surface reconstruction algorithm [14–16] using the normals as surface orientation constraints. Nodes on the occluding boundary of the sphere were marked as depth discontinuities and the computation was started from the zero depth initial approximation. The surface reconstruction required 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 optical flow field and its discontinuities can be an important source of information about the arrangement and the motions of visible surfaces. The optical flow problem is to compute optical flow from a discrete series of images.

Horn and Schunck [10] suggested a technique for determining optical flow in the restricted case where the observed velocity of image irradiance patterns can be attributed directly to the movement of surfaces in the scene. Under these circumstances, the relation between the change in image irradiance at a point (x, y) in the image plane at time t and the motion of the irradiance pattern is given by the flow equation $B_x u + B_y v + B_t = 0$, where $B(x, y, t)$ is the image irradiance, and $u = dx/dt$ and $v = dy/dt$ are the optical flow components.

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. Thus, the velocity field will vary smoothly almost everywhere. Horn and Schunck formulated the optical flow problem as finding the flow functions $u(x, y)$ and $v(x, y)$ which minimize the functional

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

where α is a constant. The first term is the smoothness constraint, while the second term is a least-squares penalty functional which coerces the flow field into satisfying the flow equation as much as possible. The Euler–Lagrange equations for the above functional are given by [10]

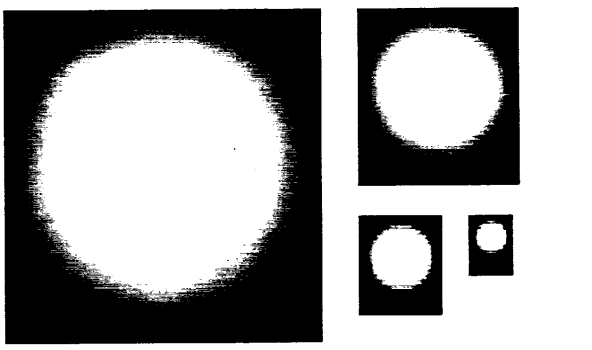

Figure 3. Synthetic images of a Lambertian sphere illuminated by a distant point source perpendicular to the image plane. The three smaller images are increasingly coarser sampled versions of the largest image which is 129×129 pixels, quantized to 256 irradiance levels.

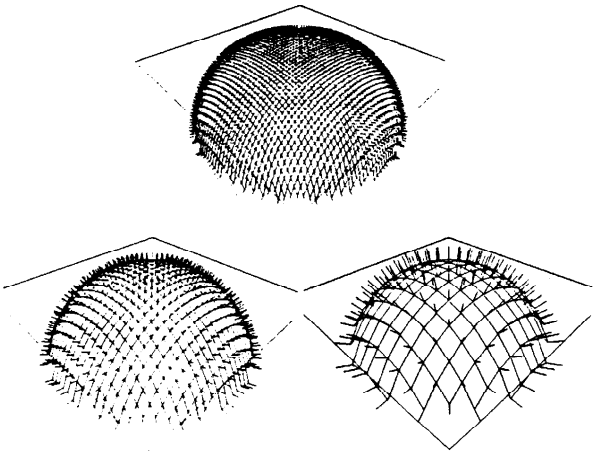

Figure 4. Surface normals which were computed after 6.125 work units by the four-level multiresolution shape-from-shading algorithm are shown as "needles" for the three coarsest levels (the finest resolution surface is too dense to illustrate as a 3-D plot). The surfaces were computed from the normals by a multiresolution surface reconstruction algorithm after 8.8 work units.

$$B_x^2 u + B_x B_y v - \alpha^2 \Delta u - B_x B_t,$$
$$B_x B_y u + B_y^2 v = \alpha^2 \Delta v - B_y B_t.$$

Assuming a cubical network of nodes with spacing h, where i, j, and k index nodes along the x, y, and t axes respectively, we use the following standard finite difference formulas to discretize the differential operators: $[B_x]_{i,j,k}^h = \frac{1}{2h}(B_{i+1,j,k}^h - B_{i-1,j,k}^h)$; $[B_y]_{i,j,k}^h = \frac{1}{2h}(B_{i,j+1,k}^h - B_{i,j-1,k}^h)$; $[B_t]_{i,j,k}^h = \frac{1}{h}(B_{i,j,k+1}^h - B_{i,j,k}^h)$; $\Delta^h u = \frac{1}{h^2}(\Phi[u_{i,j,k}^h] - u_{i,j,k}^h)$; $\Delta^h v = \frac{1}{h^2}(\Phi[v_{i,j,k}^h] - v_{i,j,k}^h)$; where $\Phi[u_{i,j,k}^h] = \frac{1}{4}(u_{i-1,j,k}^h + u_{i,j+1,k}^h + u_{i+1,j,k}^h + u_{i,j-1,k}^h)$ and $\Phi[v_{i,j,k}^h] = \frac{1}{4}(v_{i-1,j,k}^h + v_{i,j+1,k}^h + v_{i+1,j,k}^h + v_{i,j-1,k}^h)$. Other approximations are possible; for example, those suggested by Horn and Schunck [10] which, however, require over four times the computation per iteration. Given dynamic images over at least three frames, a symmetric central difference formula $[B_t]_{i,j,k}^h = \frac{1}{2h}(B_{i,j,k+1}^h - B_{i,j,k-1}^h)$ is preferable.

Substituting the above approximations into the Euler–Lagrange equations and solving for $u_{i,j,k}^h$ and $v_{i,j,k}^h$ yields the following Jacobi relaxation formula

$$u_{i,j,k}^{h\ (n+1)} = \Phi[u_{i,j,k}^h]^{(n)} - \frac{\nu_{i,j,k}^{h\ (n)}}{\mu_{i,j,k}^{h\ (n)}}[B_x]_{i,j,k}^{h\ (n)};$$
$$v_{i,j,k}^{h\ (n+1)} = \Phi[v_{i,j,k}^h]^{(n)} - \frac{\nu_{i,j,k}^{h\ (n)}}{\mu_{i,j,k}^{h\ (n)}}[B_y]_{i,j,k}^{h\ (n)},$$

where $\mu^h_{i,j,k} = \left([B_x]^h_{i,j,k}\right)^2 + \left([B_y]^h_{i,j,k}\right)^2 + \frac{1}{h^2}\alpha^2$ and $\nu^h_{i,j,k} = [B_x]^h_{i,j,k}\Phi[u^h_{i,j,k}] + [B_y]^h_{i,j,k}\Phi[v^h_{i,j,k}] + [B_t]^h_{i,j,k}$. Appropriate boundary conditions are the natural boundary conditions of zero normal derivative at the boundary of Ω. They can be enforced by copying values to boundary nodes from neighboring interior nodes.

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 expanding uniformly over two frames (Figure 5). The velocity field was specified around the occluding boundary of the sphere, and the computation was started from the zero initial approximation, $u = v = 0$ within the sphere. The occluding boundary itself was marked as a velocity field discontinuity. The solution computed on the three coarsest levels after 4.938 work units is shown in Figure 6 as velocity vectors in the x–y plane. 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. The comments about the convergence speed of the shape-from-shading algorithm apply here also. Employing the Horn-Schunck relaxation formulas, Glazer [6] also reports improvements in the convergence rate of a multilevel optical flow algorithm relative to a single level algorithm.

6. CONCLUSION

Once discretized, problems in machine vision posed as variational principles or partial differential equations are amenable to local support, parallel, and iterative solutions. Due to the locality of the iterative process, however, these computations are inherently inefficient at propagating constraints over the large representations typically encountered. Multiresolution processing can overcome this inefficiency by exploiting coarser representations which trade off resolution for direct interactions over larger distances. As was shown in our previous applications to the surface reconstruction problem [14–16] and, in this paper, to the lightness, shape-from-shading, and optical flow problems, dramatic increases in efficiency can result.

Using our approach, it is clearly possible to develop multiresolution iterative algorithms for other vision problems, including image registration [1], interpolating the motion field either along contours [8] or over regions, computing shape-from-contour [2], and for solving iteratively the structure-from-motion problem [17]. In fact, any iterative (relaxation) processes which seeks global consistency, but whose processors are restricted to simple, local interactions can benefit from the approach, most evidently when it is governed by a variational principle or partial differential equation.

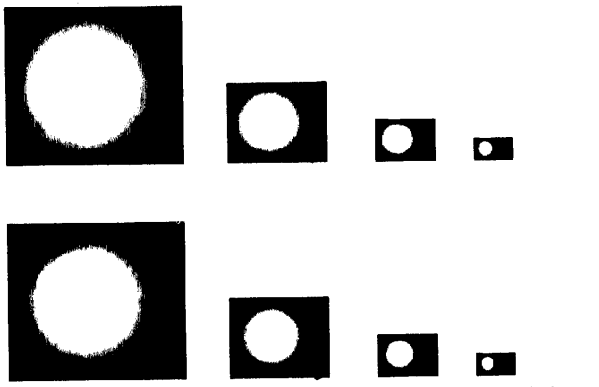

Figure 5. Synthetic images of a Lambertian sphere at four resolutions illuminated by a distant point source perpendicular to the image plane (top). The three smaller images are increasingly coarser sampled versions of the largest image which is 129×129 pixels, quantized to 256 irradiance levels. The frames for the second time instant (bottom) show an expanded sphere.

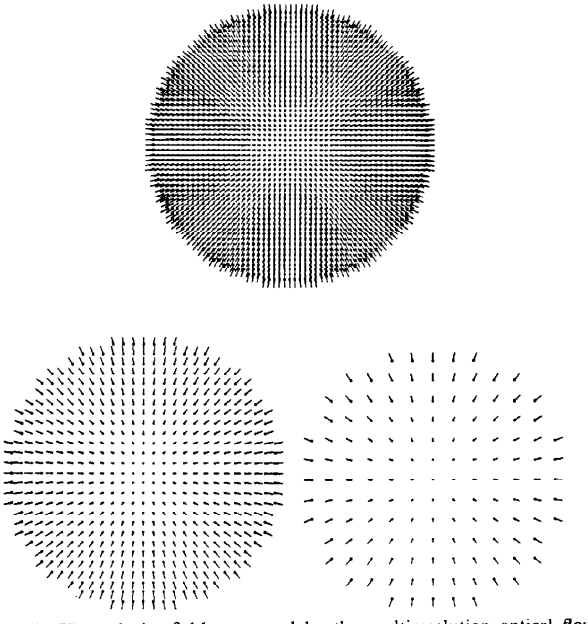

Figure 6. The velocity field computed by the multiresolution optical flow algorithm after 4.938 work units is shown at the three coarsest resolutions (the finest level solution is too dense to plot here).

References

1. **Bajcsy, R., and Broit, C.,** "Matching of deformed images," *Proc. Sixth Int. J. Conf. Pattern Recognition*, Munich, 1982, 351–353.
2. **Brady, J.M., and Yuille, A.,** "An extremum principle for shape from contour," *IEEE Trans. Pat. Anal. Mach. Intel.*, **PAMI-6**, 1984, 288–301.
3. **Brandt, A.,** "Multi-level adaptive solutions to boundary-value problems," *Math. Comp.*, **31**, 1977, 333–390.
4. **Courant, R., and Hilbert, D.,** *Methods of Mathematical Physics*, Vol. I, Interscience, London, 1953.
5. **Forsythe, G.E., and Wasow, W.R.,** *Finite Difference Methods for Partial Differential Equations*, Wiley, New York, 1960.
6. **Glazer, F.,** "Multilevel relaxation in low-level computer vision," *Multiresolution Image Processing and Analysis*, A. Rosenfeld (ed.), Springer-Verlag, New York, 1984, 312–330.
7. **Hackbusch, W., and Trottenberg, U., (ed.),** *Multigrid Methods*, Lecture Notes in Mathematics, Vol. 960, Springer-Verlag, New York, 1982.
8. **Hildreth, E.C.,** Computations underlying the measurement of visual motion, MIT A.I. Lab., Cambridge, MA, AI Memo No. 761, 1984.
9. **Horn, B.K.P.,** "Determining lightness from an image," *Computer Graphics and Image Processing*, **3**, 1974, 111–299.
10. **Horn, B.K.P., and Schunck, B.G.,** "Determining optical flow," *Artificial Intelligence*, **17**, 1981, 185–203.
11. **Ikeuchi, K., and Horn, B.K.P.,** "Numerical shape from shading and occluding boundaries," *Artificial Intelligence*, **17**, 1981, 141–184.
12. **Land, E.H., and McCann, J.J.,** "Lightness and retinex theory," *J. Opt. Soc. Amer.*, **61**, 1971, 1–11.
13. **Strang, G., and Fix, G.J.,** *An Analysis of the Finite Element Method*, Prentice-Hall, Englewoods Cliffs, NJ, 1973.
14. **Terzopoulos, D.,** Multilevel reconstruction of visual surfaces: Variational principles and finite element representations, MIT A.I. Lab., Cambridge, MA, 1982, AI Memo No. 671, reprinted in *Multiresolution Image Processing and Analysis*, A. Rosenfeld (ed.), Springer-Verlag, New York, 1984, 237–310.
15. **Terzopoulos, D.,** "Multilevel computational processes for visual surface reconstruction," *Computer Vision, Graphics, and Image Processing*, **24**, 1983a, 52–96.
16. **Terzopoulos, D.,** "The role of constraints and discontinuities in visible-surface reconstruction," *Proc. 8th Int. J. Conf. AI*, Karlsruhe, W. Germany, 1983b, 1073–1077.
17. **Ullman, S.,** *The Interpretation of Visual Motion*, MIT Press, Cambridge, MA, 1979a.