An Integral Formulation for Differential Photometric Stereo

James J. Clark and Holly Pekau
McGill University
Centre for Intelligent Machines
Montreal, Quebec
clark@cim.mcgill.ca

Abstract

In this paper we present an integral formulation of the active differential photometric stereo algorithm proposed by Clark [1] and by Iwahori et al *[3, 4]. The algorithm presented in this paper does not require measurement of derivatives of image quantities, but requires instead the computation of integrals of image quantities. Thus the algorithm is more robust to sensor noise and light source position errors than the Clark-Iwahori algorithm. We show that the algorithm presented in the paper can be efficiently implemented in practice with a planar distributed light source, and present experimental results demonstrating the efficacy of the algorithm.*

1. Introduction

Clark [1] and Iwahori *et al* [3, 4] independently proposed an active photometric technique for determining the absolute depth of surfaces. This algorithm worked for an arbitrary, non-zero, surface albedo, and assumed only a very general model for the reflectance properties of the surface. The algorithm uses small controlled motions of a point light source nearby the object and required only the solution of linear equations. This method built on three different shape from shading techniques: 1) the photometric stereo algorithm of Woodham [7], 2) the differential photometric stereo algorithm of Wolff [6] and 3) the nearby point light source photometric stereo algorithm of Iwahori *et al* [2].

The geometry of the imaging process for the Clark-Iwahori active differential photometric stereo algorithm is depicted in figure 1. The direction of the ray of light emanating from a point source, located at a position $\vec{t} = (t_x, t_y, t_z)^T$ with respect to a coordinate system centered on the focal point of the camera, that intersects a surface at the point $(X, Y, Z)^T = (-Zx/f, -Zy/f, Z)^T = -Z\vec{\chi}$ is given by

$$\hat{s} = \frac{Z\vec{\chi} + \vec{t}}{|Z\vec{\chi} + \vec{t}|} \tag{1}$$

The vector $\vec{\chi} = (x/f, y/f, -1)^T$ indicates the position of the image plane point $\vec{x} = (x, y)^T$ with respect to the world coordinate system centered on the camera focal point.

This algorithm takes as a general model for the relation between the measured image irradiance E and the point light source position the following:

$$E(\vec{x}) = K(\vec{x}) \left(\frac{R(\hat{s}(\vec{x}); \hat{n}(\vec{x}), \hat{\chi}(\vec{x}))}{|Z(\vec{x})\vec{\chi} + \vec{t}|^2} \right) \tag{2}$$

The $|Z\vec{\chi} + \vec{t}|^2$ term represents the reduction in light per unit area over the distance from the light source to the body. The surface albedo can be defined as the ratio of the reflectance of the surface to the reflectance of an ideal Lambertian reflector with the same shape. In general then, the albedo will depend on location on the surface as well as view angle, surface normal, and light source direction. The reflectance model given above requires that the albedo have the form $A(\vec{x}, \hat{\chi}, \hat{s}, \hat{n}) = K(\vec{x}) f(\hat{\chi}, \hat{s}, \hat{n})$. For the rest of the paper we will refer to K as the albedo and absorb the $f(\hat{\chi}, \hat{s}, \hat{n})$ into $R(\hat{\chi}, \hat{s}, \hat{n})$.

The reflectance will, in general, depend on the viewing direction, $\hat{\chi}$, and the surface normal, $\hat{n}$, but as we are concerned with only a particular point on the surface, the only quantity that the reflectance depends on that varies as the light source moves is the light source direction, $\hat{s}$. Hence, in what follows, we will omit the dependence of $R()$ on $\hat{\chi}$ and $\hat{n}$, unless required by the mathematical analysis. Note that we are assuming that the reflectance does not depend on the distance of the light source to the surface point, and that the dependence of the image irradiance on the distance of the light source to the surface is handled by the $|Z\vec{\chi} + \vec{t}|^2$ term in the equation above. This assumption is not very restrictive and is valid for most of the shading models currently used in the Computer Vision community.

0-7695-0149-4/99 $10.00

Using this geometry and the rather weak assumptions on the object's reflectance map, it was shown in [1, 3] that the Z component of the distance (what we will call the "depth") from the camera lens centre (or pinhole) to the object is given by

$$Z = -\left(\frac{((\nabla_t E)^T \vec{t} + 2E)}{(\nabla_t E)^T \vec{\chi}}\right) \quad (3)$$

where $\nabla_t E$ refers to the gradient of the image intensity with respect to the light source position.

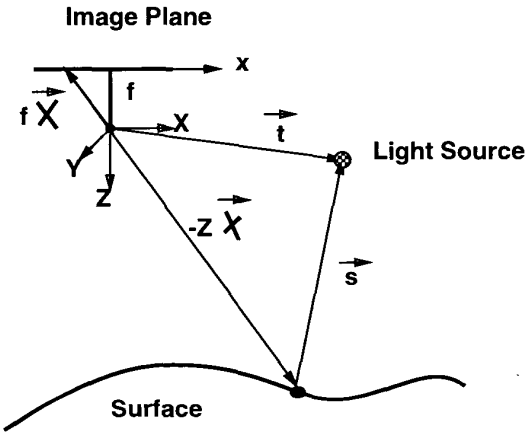

Figure 1. Geometry of the active photometric stereo scheme.

2. Integral Formulation

It can be seen that the Clark-Iwahori algorithm requires the computation of the gradient of the image intensity with respect to the light source position. As such, the solution provided by the algorithm will be extremely sensitive to noise in the image measurements and errors in the positioning of the light source. In this paper we will show that the above differential algorithm can be transformed, via application of the divergence theorem, into an algorithm that requires knowledge of the integral of image intensity over various surfaces and volumes in the light source position space. No derivatives need be computed. This integral algorithm promises to be much more robust in the face of image noise and position errors.

To understand how the differential algorithm (equation (3)) can be converted into an integral formulation, first consider the following relationship:

$$\nabla_t^T (E\vec{t}) = E\nabla_t^T \vec{t} + (\nabla_t E)^T \vec{t} = 3E + (\nabla_t E)^T \vec{t} \quad (4)$$

where we have used the identity $\nabla_t^T \vec{t} = 3$. From this we see that

$$(\nabla_t E)^T \vec{t} = \nabla_t^T (E\vec{t}) - 3E \quad (5)$$

Similarly, we have that

$$(\nabla_t E)^T \vec{\chi} = \nabla_t^T (E\vec{\chi}) \quad (6)$$

(where we have made use of the fact that $\vec{\chi}$ does not depend on the light source position $\vec{t}$).

Combining these two results with the differential relation of equation (3) gives

$$(\nabla_t E)^T (Z\vec{\chi} + \vec{t}) = \nabla_t^T (EZ\vec{\chi}) + \nabla_t^T (E\vec{t}) - 3E = -2E \quad (7)$$

Hence,

$$\nabla_t^T (E[Z\vec{\chi} + \vec{t}]) = E \quad (8)$$

This equation states that the image intensity $E = E(\vec{t}, \vec{\chi})$ is the divergence of the vector field $\vec{F} = E[Z\vec{\chi} + \vec{t}]$.

The well-known divergence theorem relates the volume integral of the divergence of a vector field (obeying certain continuity requirements) to the surface integral of the dot product of that vector field with the unit vector normal to the surface. In our case, the divergence theorem states that

$$\int\int\int_V \nabla_t^T \vec{F} dV = \int\int_S \vec{F}^T \hat{N} dS \quad (9)$$

where S is the surface of a volume V in light source position space, and $\hat{N}$ is the unit vector normal to the surface. Using equation (8) above, this can be expressed as

$$\int\int\int_V E(\vec{t}) dV = \int\int_S E[Z\vec{\chi} + \vec{t}]^T \hat{N} dS \quad (10)$$

We can solve this equation for the depth Z,

$$Z = Z(\vec{\chi}) = \frac{\int\int\int_V E dV - \int\int_S E[\vec{t}^T \hat{N}] dS}{\int\int_S E[\vec{\chi}^T \hat{N}] dS} \quad (11)$$

Notice the similarity between this equation and equation (3). The main difference is that this equation requires the image intensity to be integrated, while equation (3) requires the image intensity to be differentiated. The integral formulation should be much less sensitive to image noise than the differential formulation. The cost of this increased robustness is, however, the increase in the number of images that must be acquired in order to compute the integrals. It may be possible, in some cases, to reduce the number of images required by using an extended light source rather than a point light

source. In this way, the image sensors perform the integration by accumulating the reflected light coming from different parts of the extended source. This is only possible (or at least straightforward) with a light source position volume that is polyhedral. In the case of a polyhedral volume we can write the surface integrals as sums of surface integrals over the polyhedral faces. On each of these planar faces the quantity $\vec{t}^T\hat{N}$ is constant, as is the quantity $\vec{\chi}^T\hat{N}$. Thus we can write,

$$Z = Z(\vec{\chi}) = \frac{\int\int\int_V EdV - \sum_{i\in faces}(\vec{t}^T\hat{N}_i)\int\int_{S_i} EdS}{\sum_{i\in faces}(\vec{\chi}^T\hat{N}_i)\int\int_{S_i} EdS} \tag{12}$$

Note that, in the above equation, the integrands only involve the image intensity. The integrals can therefore be computed simply by using an extended light source corresponding to the polyhedral surfaces, and measuring the total light reflected by the object. For example, if one used a cubic light source position volume, with a square, planar, distributed light source, then only six images would be needed to obtain the required surface integrals. The volume integral could conceivably be obtained by using some sort of volumetric distributed light source, but this would likely be impractical. The volume integral would more practically be obtained by scanning a point light source through the cubic volume in some fashion. The number of images required for computation of the volume integral will depend on the smoothness of the image intensity as a function of light source position and on the particular integration method being used.

3. Experiments Using Real Imagery

To demonstrate the performance of our algorithm we performed a set of experiments. The first order of business was to implement the distributed light source to be used in computing the neccessary surface and volume integrals. A planar light source with a square light emitting area was constructed from a 2mm wide ring illluminator (diameter = 55mm) attached to a flexible fiber optic cable (diameter = 10mm) which fitted into an incandescent light source (Intralux 5000, 185 Watts). The fiber optic ring emitted light which was converted into an approximately uniform square planar form by an attachment to the ring. The side of the attachment that fit onto the light emitting side of the ring fiber optic was a reflective cone (length = 100mm, base diameter = 70mm). The cone reflected light inward through the hollow center of the fiber optic ring. The attachment that fit on the other side of the ring was a panel with a square hole covered by a diffusing medium (frosted glass) to shape and distribute the emitted light into a square planar area. The dimensions of the square area were 50mm x 50mm. Images of the light source are shown in figures 2 and 3.

The planar light source was moved to positions where it represented the faces of a 50mm x 50mm cube by a robotic arm with six degrees of freedom: x-coordinate, y-coordinate, z-coordinate, yaw, pitch, and roll. An industrial robot with a 3 degree of freedom cartesian arm and a 3 degree of freedom spherical wrist was used. The geometry of the experimental setup is illustrated in Figure 4. A calibration procedure was performed to obtain the relative transformations between the camera coordinate systems, the light source position, the robot coordinate systems and an object reference coordinate systems. Details of the calibration process can be found in [5]. The normal vectors of the illumination cube faces were aligned with the axes of the robotic arm's cartesian joints to simplify the computations and the calibration process.

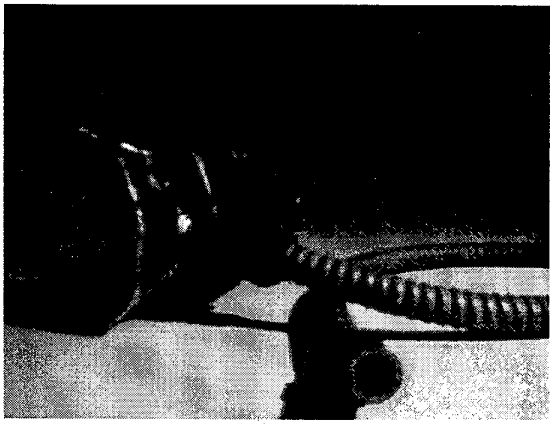

Figure 2. Planar light source (off)

The distributed light source that we constructed is intended to model a planar distribution of isotropic point light sources. Thus, light should be emitted from both sides of the plane. This is difficult to arrange in practice, so in our experiments we always positioned the planar light source such that it was always completely to one side of the object being illuminated. In this way only the rays from one side of the light source fall on the object, so we do not have to worry about illuminating the back side of the planar light source. We could handle the situations which would require a two

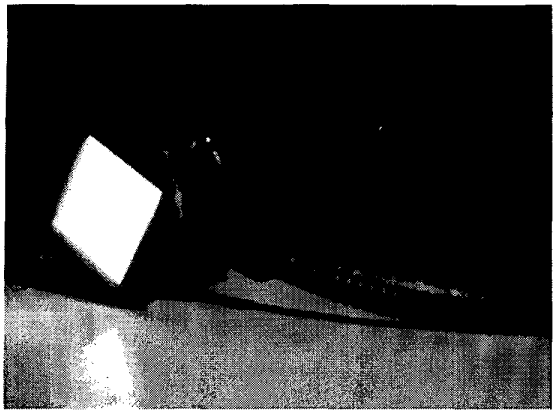

Figure 3. Planar light source (on)

sided planar light source by taking twice as many images, one set with the light source facing in one direction, and the other with the light source facing in the other direction.

The planar source was translated using the robot manipulator, perpendicular to the plane of illumination, to obtain the volume integral used in the calculation of scene depths. Ten such slices were used, spaced 5mm apart. The question of whether this spacing is adequate to represent point sources throughout the volume of the cube arises.

Two different objects were used to test our algorithm. The objects used were all painted a matte white to approximate Lambertian surfaces. Further testing of the algorithm could involve objects with surfaces that were not Lambertian as the algorithm to recover depth does not require knowledge of the reflectance map of the object. The first object tested was a truncated pyramid. The second object to be test was a doorknob. The doorknob object exhibited several different curved surfaces as well as one flat rectangular surface.

Figures 5 and 7 show views of the objects as illuminated by the planar light source, oriented with the surface normal parallel to the background plane ("X-direction" in figure 4).

3.1. Discussion of Experimental Results

The depth maps computed by our algorithm for the two test objects are shown in figures 6 and 8. It should be noted that in generating these figures no smoothing was done on the results computed using equation (12). The results presented in [1] using the Clark-Iwahori algorithm used a great deal of median filtering in an effort to combat the high noise sensitivity of that algorithm.

Even though our algorithm is more robust than that of the Clark-Iwahori algorithm, it is not without its problems.

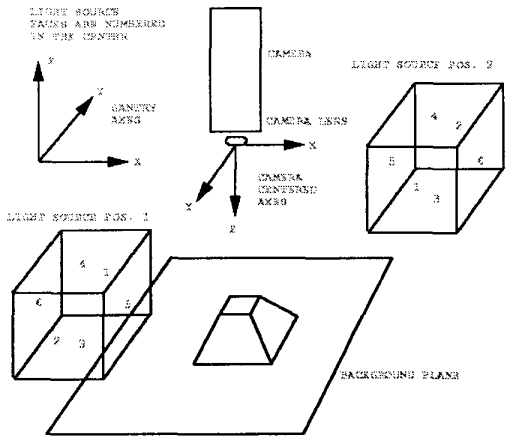

Figure 4. Geometry for the experimental system

It can be seen from the figures that reliable depth information is obtained only in certain regions of the image. This is due to a number of reasons. One of the primary reasons is that the uncertainty in the result can be shown to be inversely proportional to the image brightness (for the same reasons as described in [1]). Thus in dark areas of the scene, the depth uncertainty is very high. Problems also occur in areas of the scene that are too bright, as the image sensing elements will saturate. Thus our algorithm is limited to the areas of the scene that are not too dark and are not too bright. We can improve performance by increasing the illumination level when viewing areas with low reflectance, and likewise improve performance in viewing highly reflective areas by reducing the illumination level. This leads to an active illumination scheme, acting to extend the dynamic range of our algorithm. We have not implemented or tested such a scheme, however.

Another source of problems in applying our method results from the fact that the planar light source that we developed deviates from ideality in a number of ways. First, the edge of the square mask and the diffusing surface had a lip which obscured light emitted from the edges of the square planar source in directions close to parallel to the source plane. This made the image irradiance less than it would

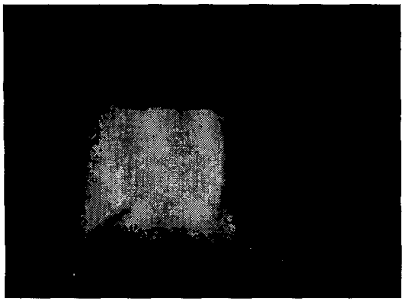

Figure 5. One View of the Pyramid Object

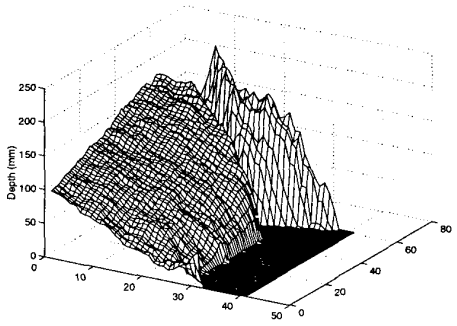

Figure 6. Depth Map of the Pyramid Object

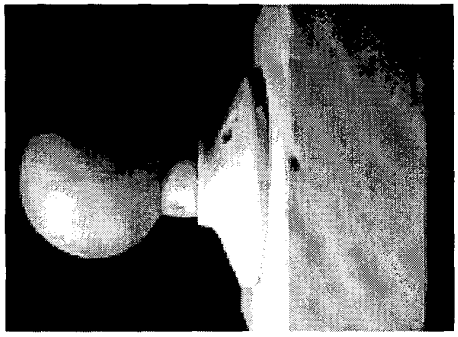

Figure 7. One View of the Door Knob Object

be given an ideal planar source at points corresponding to scene locations that would have fallen in the path of the obscured rays of light at the edge of the source. The plane of light produced by the light source was fairly uniform but the points closer to the center of the planar source were brighter than those at the edge. If the dependence of the brightness of a point in the source plane on the location of the point is known exactly it may be possible to incorporate this dependency into the depth equation. In this case, however, the dependence was unknown. The dependence was also not completely regular since the diffusing element was not completely uniform in the amount of incident light it transmitted at different locations.

Our algorithm assumes that we can accurately compute the volume integral in equation 12. In practice we do this by using a simple quadrature approximation based on slices of the volume obtained by shifting the planar light source through the volume at discrete locations. The number of slices required for a given level of accuracy in this approximation depends on the smoothness of the image intensity as a function of light source position. In addition, the light emitted in the directions perpendicular to the direction of translation of the plane source over all "slices" in the translation should be equal to the light emitted by the plane source when placed at the faces of the cube with normals pointing in these perpendicular directions. In other words, for an ideal source and an ideal translation increment, the volume integral over all the slices should be the same no matter what axis the planar source is translated along. The volume integral was also prone to lip or edge effects of the source plane. The points on the edge of the volume were therefore not well approximated by the translation of the planar source. In addition, the relatively large spacing between slices contributed to the discrepancy at the edges of the cube that were close to faces parallel to the direction of translation. This effect was minimized by choosing to translate the planar source along a direction that would minimize the amount of light originating from these ill- represented light source edge points that could be incident on the object under the geometric constraints of the scene. This direction of translation was that with the surface normal of the source plane closest to the direction of the object being illuminated (the z-axis). It allowed the ratio of incident light originating from the central points of the source cube to the incident light originating from badly modelled edge points of the source to be maximised.

The analysis underlying the Clark-Iwahori algorithm, and hence of the algorithm presented in this paper, assumes that all of the light falling on the object comes from the point

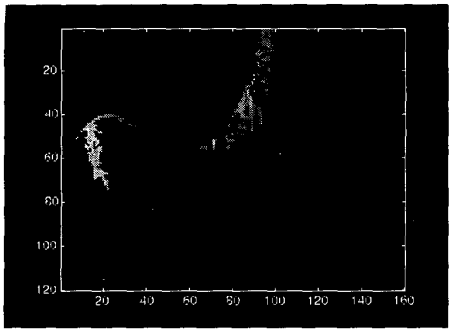

Figure 8. Depth Map of the Door Knob Object

source. In general, however, light reflecting from one part of the object may fall on other part of the object. These "inter-reflections" are not modeled by the algorithms and will produce errors in the resulting depth maps. For the pyramid object there was no self-inter-reflections, but there were some reflections from the background. Ideally the background should be as non-reflective as possible to minimise the amount of light incident on the object from scene inter-reflections. The background used was matte black construction paper with several white perpendicular grid lines used in the calibration process. This background is not perfectly nonreflective and the calibration lines proved to be sources of background noise in some images.

4. Conclusions

We have presented a reformulation of the Clark-Iwahori [1, 3, 4] active differential photometric stereo algorithm which replaces the measurement of differential quantities with measurements of integral quantites. The resulting algorithm is therefore more robust to image noise and light source positioning errors than the Clark-Iwahori algorithm. Our algorithm requires the computation of volume and surface integrals, which would normally require a much larger number of images to be captured than with the Clark-Iwahori algorithm. We showed how we can reduce the number of images required by using a single planar distributed light source. We demonstrated the operation of the algorithm using real video imagery of Lambertian objects illuminated by a robot controlled planar distributed light source.

Acknowledgements

This work was supported by the Natural Sciences and Engineering Research Council of Canada.

References

[1] Clark, J.J. 1992, Active photometric stereo, In *Proceedings of the Computer Vision and Pattern Recognition conference*, Champaign, Illinois, pp 29-34.

[2] Iwahori, Y., Sugie, H., and Ishii, N. 1990, Reconstructing shape from shading images under point light source illumination, In *Proceedings of the 10th International Conference on Pattern Recognition*, Atlantic City, pp 83-87.

[3] Iwahori, Y., Woodham, R.J., and Ishii, N. 1992, Shape from shading with a nearby moving point light source, in *Proceedings of the 2nd International Conference on Automation, Robotics and Computer Vision*, pp CV.5.5.1-5.5.5, Singapore.

[4] Iwahori, Y., Woodham, R.J., Tanaka, H., and Ishii, N. 1994, Moving point light source photometric stereo, in *IEICE Transactions on Information and Systems*, Vol. E77-D, No. 8, pp 925-929

[5] Pekau, H., 1997, Active differential photometric stereo, Undergraduate honours thesis, Dept. of Electrical Engineering, McGill University, Montreal, Canada.

[6] Wolff, L.B. 1989, Shape understanding from Lambertian photometric flow fields, In *Proceedings of the 1989 Computer Vision and Pattern Recognition Conference*, pp 46-53.

[7] Woodham, R.J. 1980, Photometric method for determining surface orientation from multiple images, *Optical Engineering*, Vol. 19, No. 1, pp 139-144.