A Characterization of Inherent Stereo Ambiguities

Simon Baker, Terence Sim, and Takeo Kanade

The Robotics Institute
Carnegie Mellon University
Pittsburgh, PA 15213

Abstract

The complete set of measurements that could ever be used by a stereo algorithm is the plenoptic function or light-field. We give a concise characterization of when the light-field of a Lambertian scene uniquely determines its shape, and, conversely, when stereo is inherently ambiguous. We show that stereo computed from the complete light-field is ambiguous if and only if the scene is radiating light of a constant intensity (and color) over an extended region.

1. Introduction

Although stereo has been studied extensively, it has never been precisely delineated when the process has a unique solution and when it is inherently ambiguous. Kutulakos and Seitz [10] recently pointed out that stereo is not always unique and proposed the concept of the *photo-hull* to quantify any ambiguity. They provided very little insight, however, into when stereo is actually ambiguous and when the solution is in fact theoretically unique.

Stereo has traditionally been thought of as primarily a 2-camera problem. More recently it has increasingly been treated as an n-camera problem where $n \geq 2$. As pointed out in [5, 15] (for example) the complete set of visual information that could ever be made available to a stereo algorithm is the *plenoptic function* [1] or *light-field* [12]; i.e. the radiance of light in free space given as a function of 3D position, 2D direction, wavelength, polarization, and time.

In this paper we analyze stereo computed from the entire light-field. In particular we present a characterization of when the light-field of a Lambertian scene uniquely determines its shape, and conversely when stereo is inherently ambiguous. (Since the light-field is the set of all passive visual measurements that could ever be made, this question is the same as asking whether the shape of the scene can ever be uniquely determined without actively changing the illumination conditions.) It turns out that our characterization is particularly concise and intuitively very natural; stereo is

unique (given the light-field) *if and only if* there is no extended region on the surface of an object in the scene that is radiating a constant intensity, (color, and polarization.)

Our analysis shows that constant intensity regions are the only inherent ambiguities in stereo; i.e. the only ambiguities that cannot be resolved by using more visual measurements. It also shows that constant intensity regions are *always* ambiguous for stereo, for any number and arrangement of cameras. (Constant intensity regions are not always ambiguous for *shape-from-silhouette* and so part of our analysis consists of formalizing the difference between these two processes.) Our analysis unfortunately does not prove anything about the uniqueness of n-camera stereo. The most we can say is that, in the absence of constant intensity regions and given enough cameras, n-camera stereo is *unlikely* to have significant inherent ambiguities. Complete absence of ambiguity is not guaranteed though (at least by this paper.)

Proving these claims is quite involved and requires that a number of technical details be addressed. To avoid burdening the reader with all of the details at once, we organize this paper as follows. In Section 2 we introduce the question of whether the shape of the scene is unique given the light-field. We proceed to informally outline the major arguments in the proof of our claims. In Section 3 we formally state our claims as a theorem and discuss its implications. We sketch the proof of the theorem in Appendix A. A complete proof will be provided in a future technical report [2].

2. Stereo Ambiguities and Uniqueness

2.1. The Plenoptic Function or Light-Field

The *plenoptic function* [1] or *light-field* [12] is a function which specifies the radiance of light in free space. It is typically assumed to be a 5D function of position (3D) and orientation (2D). In addition, it is also sometimes modeled as a function of wavelength, polarization, and time, depending on the application. We ignore these effects in this paper because: (1) our results can easily be generalized to the case that light can be distinguished based on its wavelength
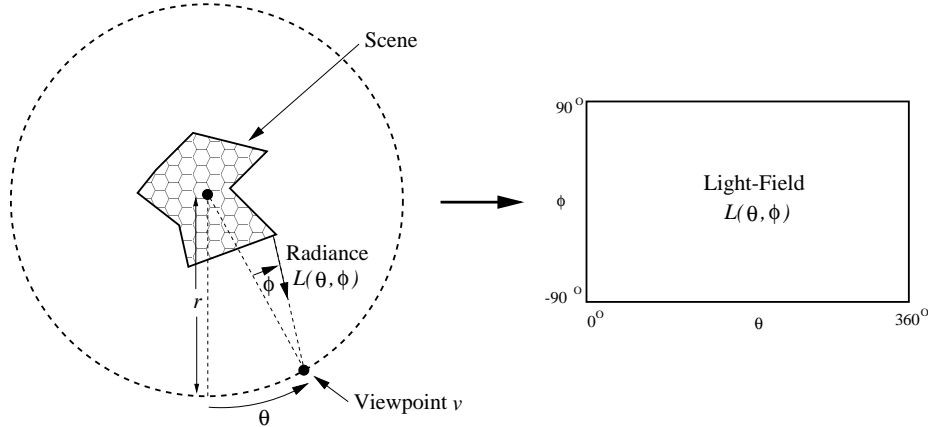
Figure 1. An illustration of the 2D light-field of a 2D scene [12]. The scene is conceptually placed within a circle, radius r. The angle to the viewpoint v around the circle is measured by the angle θ, and the direction that the viewing ray makes with the radius of the circle is denoted ϕ. For each pair of angles θ and ϕ, the radiance of light reaching the viewpoint is denoted $L(\theta, \phi)$, the *light-field*. Although the light-field of a 3D scene is actually 4D, we will continue to use the 2D notation of this figure in this paper for ease of explanation. Everything derived here also holds for the 4D light-fields of 3D scenes. The circle can also be replaced with an arbitrary piecewise smooth curve (surface in 3D) in more complex scenes (and need not even be connected.)

or polarization, and (2) there is an implicit assumption in stereo that the images are captured at the same time, or equivalently that the scene and illumination do not change.

Assuming there is no absorption, scattering, or emission of light through the air [14], the light-field is only a 4D function, a function of 2D direction defined on a 2D surface [7, 12]. (Similarly, the light-field of a 2D scene is 2D rather than 3D, as illustrated in Figure 1.) We make the "no absorption" assumption in this paper, and also, for ease of explanation, assume that the surface that the light-field is defined on is a sphere (a circle in 2D.) At no point is this spherical surface property required in our analysis. All of our results do generalize to the case that the light-field is defined on an arbitrary piecewise smooth surface (consisting of a finite number of connected components.)

2.2. Uniqueness from Lambertian Light-Fields?

There are three components in the formation of a light-field: (1) the shape of the scene, (2) its reflectance properties, and (3) the illumination conditions. We make piecewise smoothness assumptions about the shape of the scene, piecewise smoothness assumptions about the illumination, and the Lambertian assumption (made by all brightness constancy algorithms) about the reflectance properties.

Light-fields of Lambertian scenes contain shape information in that any two rays which image the same point in the scene must necessarily have the same intensity. Is this information enough to constrain the shape of the scene uniquely? Kutulakos and Seitz studied this question (in the n-camera stereo case) and showed that *if there is any ambiguity* there is one special solution (the *photo hull*) which is the union of all of the solutions and therefore contains them

all [10]. But, is there ever any ambiguity in the solution?

Kutulakos and Seitz did give one example where there is ambiguity (albeit ignoring inter-reflections. See Figure 3 in their paper.) Is this generally the case, or is the solution unique most of the time? In this paper, we attempt to answer this question by characterizing when the light-field of a scene uniquely determines its shape, and conversely when there are multiple scenes that generate the same light-field.

We formulate our results in terms of the light radiating outwards from the surface of the scene (i.e. the light-field) rather than in terms of the shape, reflectance, and incoming illumination. This leads to a much more concise characterization than would otherwise be possible. The light-field is also the information that a stereo algorithm has to work with. Our characterization will therefore be more useful to a stereo algorithm that is trying to determine whether there is a unique solution or not (rather than a characterization in terms of quantities that are unknown to the algorithm, such as, shape, reflectance properties, and illumination.)

2.3. The Ambiguous Cases

There are two simple scenarios in which the shape of the scene cannot be uniquely determined from the light-field. The first such case, illustrated in Figure 2(a), occurs when there are points in the scene that are not visible in the light-field. (The shape of the scene is, of course, completely unconstrained anywhere it is not visible.) This trivial case is really just an artifact of our parameterization of the light-field by points on a sphere (or by a circle in 2D). Since the light-field can always be defined on a piecewise smooth surface, possibly disconnected, so that every point in the scene is visible somewhere, this scenario can be ignored.
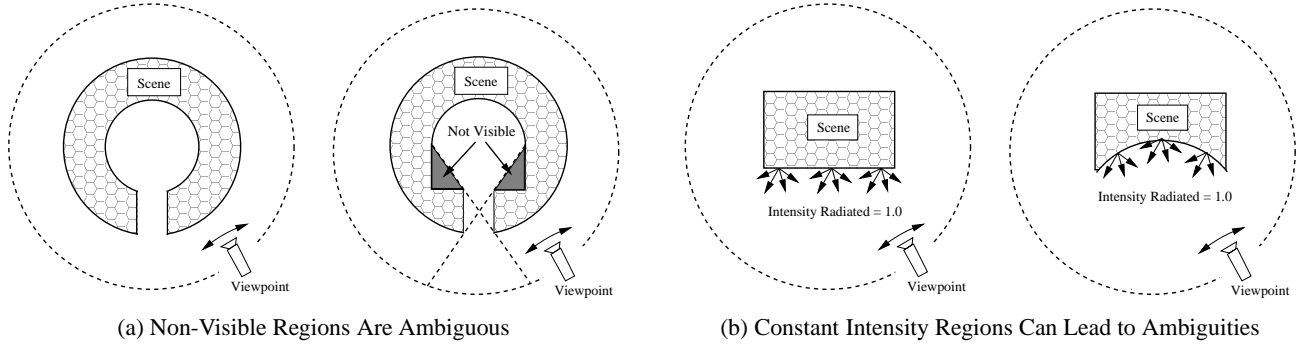
Figure 2. Examples of situations in which the shape of the scene is ambiguous given the light-field: (a) if there are parts of the scene that are not visible in any part of the light-field, the shape of those regions cannot be uniquely determined, and (b) if the light radiated from the scene is constant over an extended region, in this case the entire bottom face of the rectangle on the left and the entire curved bottom of the "carved-out" rectangle on the right, the shape of the two scenes can sometimes be in-distinguishable; i.e. their light-fields can be set up to be exactly the same with an appropriate choice for the albedo variations. The curved region in case (b) does not have uniform albedo, but has an albedo distribution that leads to the same amount of light being radiated from each point after inter-reflections are taken into account. (See the body of the text for a derivation of what the albedo variation must be.)

The second ambiguous case is more significant. It occurs when the intensity of the light radiated from the scene is constant over an extended area. See Figure 2(b) for an example of such a scenario. In this case, the rectangle and the modified "carved-out" rectangle with the concavity both have the same *visual hull* [11]. If the albedo variation in the concavity is set up in a way that the intensity of the light radiated outwards is constant across the concavity (and is the same as that for the rectangle), the two scenes have *exactly the same light-field.* The shape of the scene is therefore ambiguous, even given the entire light-field. (The ambiguous example presented in [10] also has constant intensity regions, as mentioned above, ignoring inter-reflections.)

To complete the proof that this is an ambiguous case we must show that it is actually possible to configure the albedo variation to achieve a constant radiance over an extended region. Suppose x is a point in the region that we wish to make radiate a constant intensity. (See Figure 3 for an illustration.) Suppose x has albedo $alb(x)$. The light falling on x can be divided into two components, that inter-reflected from the constant intensity region, and that coming from the rest of the scene (which may be either direct illumination or inter-reflected from some other part of the scene.) Suppose that the radiance of the constant intensity region is $inter$. Denote the foreshortened solid angle (see [8]) subtended by the non-constant "rest" of the scene by $fssa(x)$. The foreshortened solid angle subtended by the constant intensity region is therefore $\pi - fssa(x)$. Finally, denote the foreshorten-weighted average incoming illumination radiance from the rest of the scene $illum(x)$; i.e. the total incoming irradiance from the rest of the scene divided by $fssa(x)$. The total incoming irradiance is therefore:

$$fssa(x) \times illum(x) + [\pi - fssa(x)] \times inter. \quad (1)$$

Multiplying this expression by the albedo gives the amount of light radiated by point x. The constraint to be solved to obtain a constant intensity region is therefore:

$$inter = \quad (2)$$

$$alb(x)\left(fssa(x) \times illum(x) + [\pi - fssa(x)] \times inter\right).$$

This equation can always be solved for $alb(x)$. We just have to make sure that the resulting albedo is valid (the surface does not radiate more light than it received); i.e. we need to
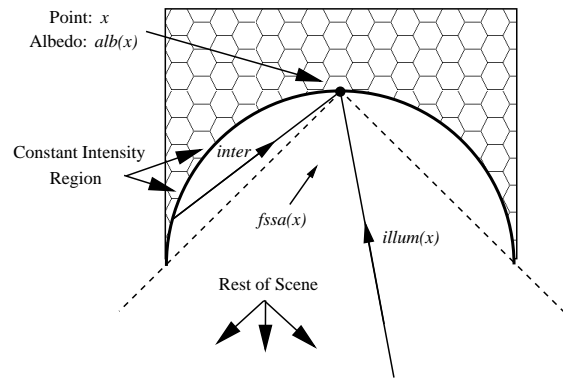

Figure 3. A derivation of the albedo variation required to create a constant radiance region. Consider a point x with albedo $alb(x)$. Suppose that the required constant radiance is $inter$, the foreshortened solid angle of the scene that does not have the constant radiance is $fssa(x)$, and that the foreshorten-weighted average incoming radiance from that part (i.e. the rest) of the scene is $illum(x)$.

check that $alb(x) \leq 1/\pi$ [8]. Substituting this inequality into Equation (2) and simplifying gives:

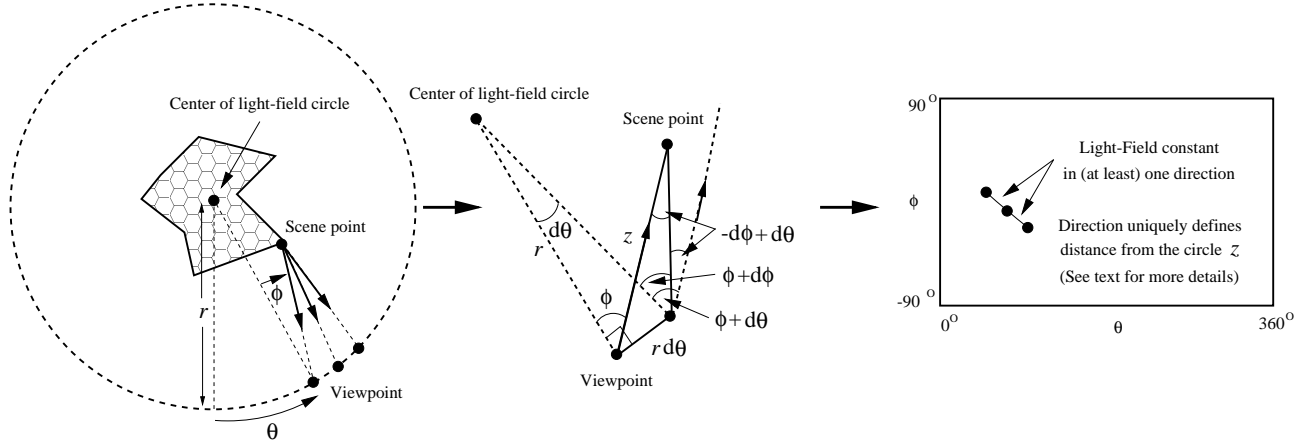$$inter \leq illum(x). \quad (3)$$

3

Figure 4. If the light-field is differentiable at (θ, ϕ), it is locally constant in at least one direction (the direction orthogonal to the gradient.) If the gradient is non-zero, the light-field is locally constant in only one direction and that direction *uniquely* defines the distance to the point in the scene (i.e. the depth z.) The depth is therefore uniquely defined wherever the gradient exists and is non-zero. The depth (scene shape) is also globally unique if the light-field is not constant in an extended region (i.e. one containing an open subset) since under this assumption (and appropriate smoothness assumptions about the scene shape), there can only be "isolated points and curves" where $L(\theta, \phi)$ is either not differentiable or has zero gradient. The depth at these points is then uniquely determined because of the assumed continuity of the scene. This argument can be extended to the 4D light-fields of 3D scenes where there is a 3D hyper-plane in which the light-field is locally constant and which uniquely defines the scene depth.

Equation (3) means that we can always find an albedo variation to create a constant intensity region so long as the constant radiance desired is less than or equal to the minimum (computed over the constant radiance region) foreshorten-weighted average incoming radiance $illum(x)$ (the total incoming irradiance divided by the foreshortened solid angle subtended by the rest of the scene.) The scenario in Figure 2(b) is therefore realizable in practice. It is a valid ambiguous case where more than one differently shaped scenes share exactly the same light-field.

2.4. Uniqueness of Stereo from Light-Fields

We have just shown that there are ambiguous cases where two differently shaped scenes generate the same light-field. In the example that we exhibited, there are extended regions that are radiating a constant intensity. We now argue that if there are no such regions, the light-field uniquely determines the shape of the scene; i.e. all ambiguous cases include an extended constant intensity region.

The main step in the (2D) proof is illustrated in Figure 4. Suppose that θ and ϕ are a pair of angles defining a point in the light-field $L(\theta, \phi)$. If θ changes to $\theta + d\theta$ and ϕ changes to $\phi + d\phi$ in a way that the same point in the scene is imaged, the geometry is as sketched in the middle of Figure 4. Applying the sine rule to the solid triangle gives:

$$\frac{\sin\left[180° - (-d\phi + d\theta) - (90° - \phi)\right]}{z} = \frac{\sin\left[-d\phi + d\theta\right]}{r \cdot d\theta} \tag{4}$$

where z is the distance to the scene point (measured from the viewpoint circle.) Rearranging this expression gives:

$$\frac{d\phi}{d\theta} = 1 - \frac{r}{z} \cdot \cos\phi. \tag{5}$$

This equation means that there is a *one-to-one* relationship between the direction $\frac{d\phi}{d\theta}$ in the light-field and the distance z to the point in the scene. The 2D light-field of a Lambertian scene will therefore be locally constant in the direction defined by Equation (5). If the light-field is differentiable and has non-zero gradient, this direction will be the direction orthogonal to the gradient. The (direction of the) gradient of $L(\theta, \phi)$, if it exists and is non-zero, therefore uniquely determines the distance to the scene. (Note that this argument is similar to the analysis of the extraction of depth from *epipolar-plane images* or EPIs [3] and the extraction of depth from calibrated normal flow.)

In general, of course, there will be many points at which the light-field $L(\theta, \phi)$ is either not differentiable or has zero gradient. The argument above therefore does not apply at these points. However, assuming that the light-field is not constant in an extended region (i.e. one containing an open subset or equivalently a small disk with non-zero radius) the gradient can be zero (or undefined) only on a set of "isolated points and curves." (See Appendix A.1 for more details.) The depth will then also be uniquely defined at these isolated points and curves using the assumed piecewise continuity of the shape of the scene. So, assuming that every point in the scene is visible somewhere in the light-field, the light-field uniquely defines the scene shape *so long as there are no extended constant intensity regions.*
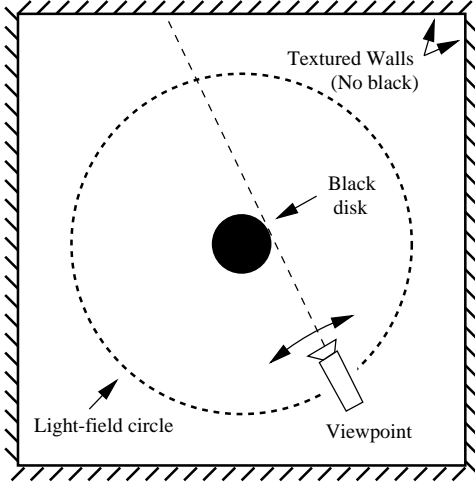
4

Figure 5. An example of a scene which contains an extended constant intensity region, and yet which has a unique light-field. The uniqueness can be deduced using a *shape-from-silhouette* argument [6, 13]. (See the body of the text for more details.) To distinguish stereo from shape-from silhouette, we exclude the use of any tangent rays (or rays infinitesimally close to tangent rays) from stereo. Once such rays are excluded, the scene shape is then *always ambiguous* if there is an extended constant intensity region.

2.5. Stereo Versus Shape-From-Silhouette

In Section 2.3 we showed that there are ambiguous cases where the light-field does not uniquely define the shape of the scene. In Section 2.4 we argued that all ambiguous cases contain an extended region that is radiating a constant intensity. Are all cases that contain such a region ambiguous?

Perhaps surprisingly, the answer is no. For example, Figure 5 illustrates a scene which contains an extended constant intensity region (the black disk, or sphere in 3D, in the middle of the scene) and which has a unique light-field.

The fact that the light-field of this scene cannot be generated by any other scene can be deduced using a similar argument to that used in *shape-from-silhouette* algorithms [6, 13]. Since the walls are textured, the argument of the previous section implies that the volume outside the black disk (the region to the right of the dashed ray in the figure) is empty space. We can then deduce that there is a black object in the middle of the room, the *visual hull* [11] (defined by an infinite number of cameras) of which is a disk (a sphere in 3D). Since the only shape which has a disk (sphere) as its visual hull is the disk (sphere) itself, the object in the middle of the room must be a disk (sphere.) The shape of the scene is therefore unique. (Note that if the disk is replaced with a rectangle, the situation is the same as Figure 2(b), and the scene shape is then ambiguous instead.)

As pointed out, the uniqueness of the scene shape in this example is deduced using a shape-from-silhouette like ar-

gument. We would like to distinguish stereo algorithms and shape-from-silhouette algorithms. The key distinguishing aspect of shape-from-silhouette is that it uses rays that are tangent to the surfaces of the objects in the scene. Stereo, on the other hand, primarily only "matches" rays that are radiated from the body of the objects in the scene.

It turns out that if we exclude tangent rays from the light-field, and thereby remove the source of information used by shape-from-silhouette algorithms, the situation is changed completely. The shape of the scene is then *always ambiguous* if there is an extended constant intensity region. The proof of this fact is constructive and operates by "carving" out a concave hole in the constant intensity region *without changing the visual hull.* The albedo is then modified in the carved region to restore the constant intensity region. (The fact that this can always be done was shown in Section 2.3.) The details of the proof are included in Appendix A.2.

3. Formalization as a Theorem

So far our arguments have been presented informally in order to provide the reader with a brief overview of our results. We now formally state our claims as a theorem.

Theorem: *Consider the set of Lambertian scenes consisting of an occupied volume (a 3D sub-manifold of $[0,1]^3$ with boundary [4]) bounded by a finite collection of continuously differentiable surface patches (each a connected 2D manifold with possibly overlapping 1D boundaries.) Each surface patch has a piecewise continuously differentiable albedo variation on it (the 2D manifold can be partitioned into a finite collection of 2D sub-manifolds with possibly overlapping 1D boundaries where the albedo variation is continuously differentiable on each sub-manifold in isolation.) Suppose that some of these surfaces are also area light-sources and are radiating (finite valued) light in an isotropic manner, piecewise differentially across their surfaces. Suppose that S_1 is a scene in this set and L_1 a light-field of that scene with the following properties: (1) L_1 is defined on a finite collection of continuously differentiable closed surface patches within $(0,1)^3$ and is a closed 4D manifold (with or without boundary), (2) every point (in free space and) on the surface of S_1 lies in an open subset of (free space or) the surface of S_1 that is everywhere visible in a 4D open subset of L_1, and (3) no ray in L_1 is tangent to S_1. Then, there is another scene S_2 (that has a different occupied volume) in the set (and for which point (2) above holds) that also generates L_1 if and only if L_1 is constant in a 4D open subset.*

Although this formulation is quite lengthy, the only real assumptions are that the scene contains: (1) a collection of Lambertian objects with piecewise smooth surfaces and
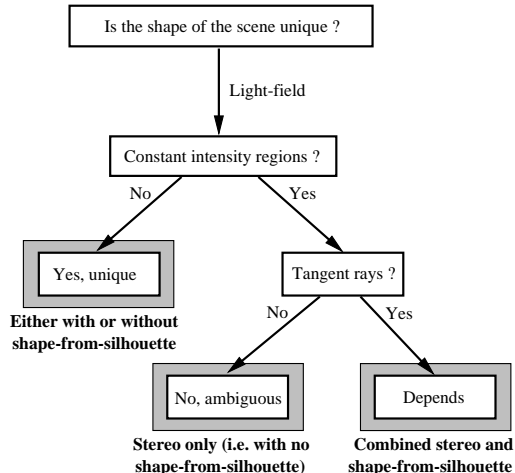
```
            Is the shape of the scene unique ?
                          │
                     Light-field
                          ▼
             Constant intensity regions ?
                    No ╱     ╲ Yes
                      ╱        ╲
              ┌─────────────┐    ▼
              │ Yes, unique │  Tangent rays ?
              └─────────────┘   No ╱   ╲ Yes
            Either with or without  ╱     ╲
            shape-from-silhouette ▼         ▼
                          ┌──────────────┐ ┌─────────┐
                          │ No, ambiguous │ │ Depends │
                          └──────────────┘ └─────────┘
                          Stereo only (i.e. with no  Combined stereo and
                          shape-from-silhouette)     shape-from-silhouette
```

Figure 6. Our characterization of when the light-field of a Lambertian scene uniquely determines its shape. If the light-field contains no extended constant intensity regions, the shape of the scene is *always unique.* If the light-field does contain any such regions, but does not contain any tangent rays, the scene shape is *always ambiguous.* If the light-field also contains tangent rays, the shape may or may not be unique. (See the body of the text for more details.)

piecewise smooth albedo variations, and (2) a collection of extended isotropic light sources. Point light-sources in particular are not allowed. In the real world, there are, of course, no such things. Moreover, it is always possible to approximate a point light-source by a very small source. (It is also possible to approximate a "laser-beam" by placing a light source behind a "wall" with a small hole in it.)

We also assume that the scene is contained in $[0,1]^3$. This is to avoid technical problems with rays of light that go off into infinity and do not intersect the scene. The "walls" of $[0,1]^3$ can reflect or absorb any light that reaches them, contain light-sources, or can even be textured like in Figure 5. These scenarios are all compatible with the theorem.

3.1. Interpretation of the Theorem

The theorem has two parts, the *if* and the *only if*. The *if* part says that the shape of the scene is *always ambiguous* whenever there is an extended region in the scene that is radiating a constant intensity, *even if the entire light-field is given.* This result also clearly holds for any collection of images, and under weaker assumptions, such as in the absence of the Lambertian assumption. Constant intensity regions are therefore *inherent ambiguities* and in the sense that they cannot be resolved by adding more visual measurements. The importance of this result is that it means that as soon as a stereo algorithm encounters a constant intensity region, it can immediately deduce that it will need to use *a priori* assumptions in order to resolve the inherent ambiguity.

This result applies to stereo algorithms that do not take

advantage of the shape-from-silhouette information provided by tangent rays. When this information is added, the shape of the scene may or may not be unique. In particular, the example in Figure 5, in which the scene shape is unique, can easily be turned into an ambiguous case by simply adding a pair of black hemispheres to the scene and placing them against the walls, diagonally across from each other. Although the shape of the two hemispheres can be uniquely determined, the silhouette of the black sphere in the center of the scene can no longer be determined and the overall scene shape is ambiguous. A complete characterization of when the combination of stereo and shape-from-silhouette is unique *given the full light-field including tangent rays* appears to be much more complex than for stereo or shape-from-silhouette alone and is beyond the scope of this paper. (This question is probably much harder than characterizing the information in silhouettes themselves because it is the light-field that is given, not the silhouettes.)

The *only if* part of the theorem says that the shape of the scene is *always unique*, so long as the light-field is not constant anywhere in an extended region. (This result holds whether the stereo algorithm uses the shape-from-silhouette information or not.) It does of course require that the entire light-field be given and that the Lambertian assumption holds. Although this second result does not apply to the n-camera stereo problem, it does at least indicate that there are unlikely to be stereo ambiguities in practice if every point in the scene is imaged by enough cameras (and assuming that there are no extended constant intensity regions.)

Stereo ambiguities can therefore be categorized into two types: (1) those caused by constant intensity regions, and (2) those caused by the fact that only a finite number of cameras are available rather than the entire light-field. One way to interpret the analysis in this paper is that ambiguities of the first kind are *inherent* and will always be a problem, whereas the ambiguities of the second kind will be unlikely to occur if enough cameras are used. Figure 6 contains a schematic diagram illustrating the results of this paper.

Acknowledgments

References

[1]   E. Adelson and J. Bergen. The plenoptic function and elements of early vision. In Landy and Movshon, editors, *Computational Models of Visual Processing*. MIT Press, 1991.

[2] S. Baker, T. Sim, and T. Kanade. A fundamental theorem of stereo? Technical Report CMU-RI-TR-01-07, The Robotics Institute, Carnegie Mellon University, (In Preparation) 2001.

[3] R. Bolles, H. Baker, and D. Marimont. Epipolar-plane image analysis: An approach to determining structure from motion. *Intl. Journal of Computer Vision*, 1(1):7–56, 1987.

[4] T. Bröcker and K. Jänich. *Introduction to differential topology*. Cambridge University Press, 1982.

[5] J.-X. Chai, S.-C. Chan, H.-H. Shum, and X. Tong. Plenoptic sampling. In *Computer Graphics Proceedings, Annual Conference Series (SIGGRAPH)*, pages 307–318, 2000.

[6] P. Giblin and R. Weiss. Reconstruction of surfaces from profiles. In *Proceedings of the 1st IEEE International Conference on Computer Vision*, pages 136–144, 1987.

[7] S. Gortler, R. Grzeszczuk, R. Szeliski, and M. Cohen. The lumigraph. In *Computer Graphics Proceedings, Annual Conference Series (SIGGRAPH)*, pages 43–54, 1996.

[8] B. Horn. *Robot Vision*. McGraw Hill, 1996.

[9] J. Koenderink. *Solid Shape*. MIT Press, 1990.

[10] K. Kutulakos and S. Seitz. A theory of shape by shape carving. *Intl. Journal of Computer Vision*, 38(3):197–216, 2000.

[11] A. Laurentini. The visual hull concept for silhouette-based image understanding. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(2):150–162, 1994.

[12] M. Levoy and M. Hanrahan. Light field rendering. In *Computer Graphics Proceedings, Annual Conference Series (SIGGRAPH)*, pages 31–42, 1996.

[13] W. Martin and J. Aggarwal. Volumetric description of objects from multiple views. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 5(2):150–158, 1983.

[14] S. Nayar and S. Narasimhan. Vision in bad weather. In *Proc. of the 7th IEEE Intl. Conf. on Computer Vision*, 1999.

[15] H.-H. Shum, A. Kalai, and S. Seitz. Omnivergent stereo. In *Proc. of the 7th IEEE Intl. Conf. on Computer Vision*, 1999.

A. Sketch of the Proof of the Theorem

We now sketch a proof the theorem. A more complete proof will be provided in a future technical report [2].

A.1. No Constant Regions Implies Uniqueness

The first step in the proof that the occupied volume of the scene is unique if there are no extended constant intensity regions is to argue that if the distance to the surface of the scene is uniquely defined everywhere in the light-field then the occupied volume of the scene is unique. This follows from the assumption that all of the scene, and all of the free space, is visible from somewhere in the light-field; i.e. it follows from point (2) in the theorem. The formal proof of this step proceeds by assuming that there is a point that is occupied in one of S_1 and S_2 and not in the other. Since the unoccupied point is visible in the light-field there is a ray that sees it. The distance to the scene along this ray is therefore different in the two scenes. The contra-positive of the required result has therefore been proven.
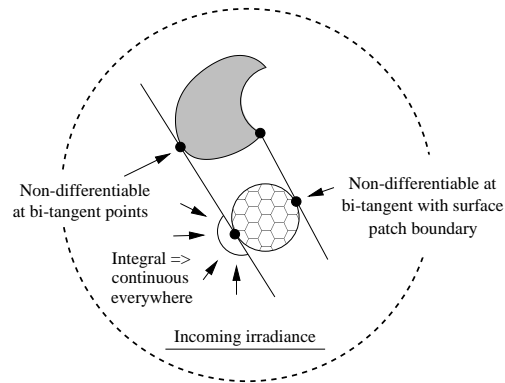

Figure 7. The incoming irradiance is continuous since it is an integral (except at the boundaries of the surface patches where the surface normal is discontinuous.) The radiated light is therefore also continuous except at the albedo and illumination source discontinuities. The incoming radiance is therefore differentiable since it is the integral of a piecewise continuous function (except at the discontinuities; i.e. where the visibility changes at bi-tangent points with the surface patches, their boundary curves, and the albedo and illumination source discontinuities.)

The second step is to argue that if the distance to the surface of the scene is unique at at least one point in every open subset of the surface, the occupied volume of the scene is still unique. The proof of this is similar, but just adds in some continuity arguments. If there is a point that is occupied in one of S_1 and S_2, but not in the other, then there is an open subset of 3D space that is occupied in one and not the other because the occupied volumes are 3D manifolds. It is then possible to find an open subset on one of the surfaces, which is in that 3D open subset, and for which the distance to every point in it is different in the two scenes. Again, this is the contra-positive of the required result.

In Section 2.4 we presented an argument that the distance to the surface of the scene is unique: (1) for points in the light-field where the gradient is well defined (i.e. where the light-field is differentiable), and (2) for points where the gradient is non-zero. (We just showed this for 2D scenes.) So, to complete the uniqueness proof, we need to show that the light-field is differentiable (and non-zero) at some point (and an open subset of that point) in every open subset. The non-zero gradient part of this is trivial; it follows directly from the assumption that there are no constant intensity open subsets. We now show that the light-field is differentiable at at least one point in every open subset (and an open subset of that point.) Afterwards, we extend the proof of Section 2.4 to 3D scenes. The proof will then be complete.

The incoming radiance at any point on a surface in the scene is a (foreshorten-weighted) integral of the incoming scene radiance over the hemisphere of incoming directions. See Figure 7 for an illustration. Since it is an integral of a real valued function, it is continuous (except at the bound-

ary of the surface patches where the surface normal is discontinuous.) The light radiated from the surface patches is therefore also a continuous function, except, in addition, where the albedo variation is discontinuous and at the illumination source boundaries.

The incoming irradiance is therefore a differentiable function (since it is the integral of a continuous function, the scene radiance) except where the "visibility" of surface patches changes. This only occurs at points of bi-tangency between surface patches, and at points of bi-tangency between surface patches and the 1D boundaries of surface patches, and the 1D boundaries where the albedo variation or illumination source is discontinuous. The radiance of light is differentiable at the same points that the incoming irradiance is differentiable. The radiated light is therefore differentiable on the surfaces in the scene, except: (1) at the surface patch boundaries, (2) at the albedo and illumination source discontinuities, (3) at points of bi-tangency, and (4) at points of bi-tangency with the surface patch boundaries and the albedo and illumination source discontinuities. All of these types of points are 1D manifolds at most. The first two follow by definition. The set of bi-tangency points of a pair of surfaces and the set of bi-tangency points of a surface and a curve are also 1D [9].

Since the sets of non-differentiable points are all (contained in) 1D manifolds, there is always a point in any open subset of the surface of the scene that is differentiable (and is contained in a differentiable open subset.) Finally, the differentiability of the radiated light across the surfaces in the scene can then easily be transferred to the differentiability of the light-field on the surfaces that it is defined upon.

The extension of the argument in Section 2.4 to 3D scenes is messy. We therefore just sketch an outline of the complete argument. In 3D the light-field is a 4D manifold. If the gradient of the light-field is non-zero, the light-field must be locally constant in a 3D subspace. This 3D subspace can be divided into a 1D subspace that is in the same direction as the gradient in the 3D scene, and a 2D subspace orthogonal to it. The same argument as in Section 2.4 can then be performed in the 2D slice through the 3D scene defined by the ray in the light-field being considered and the direction in the 3D scene corresponding to the light-field gradient. The 1D subspace then takes the role of the direction orthogonal to the light-field gradient in the 2D proof.

A.2. Constant Regions Imply Ambiguity

The proof that constant intensity regions are always ambiguous is a combination of the arguments in Sections 2.3 and 2.5. Suppose that there is an constant open subset in the light-field. We can then find a point on the surface of S_1 that is viewed by a point in the open subset of the light-field. (If the constant intensity region did not originate from the
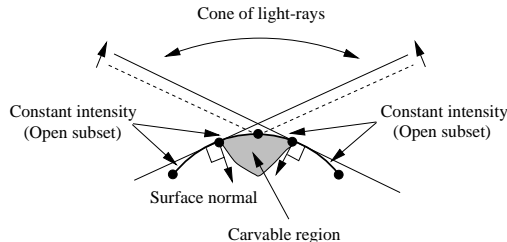

Figure 8. The proof that constant regions are always ambiguous is constructive and operates by finding a point on the surface of S_1 in an constant intensity open region where the surface can be "carved" away and made more concave. To be able to do this without the light-field being changed requires the assumption that there are no tangent rays in the light-field. The incoming rays for an entire open region therefore all lie in a cone (bounded away from the tangent plane). The occupied volume of S_1 can be carved away beneath the open region without the visual hull being changed.

surface of the scene, it came from the boundary $[0, 1]^3$. In this case, it is possible to add a small black volume to the scene close to boundary because the light-field is defined on a closed surface in $(0, 1)^3$. The scene is ambiguous in this case and so w.l.o.g. we can assume the the constant intensity region is on the surface of the scene.) We can also choose this point so that is lies in an open region of the surface that is all viewed by the constant open subset of the light-field.

We can then find a an open subset of the surface and a cone in which all of the incoming light-rays lie, everywhere in that open region. (This step requires the assumption that there are no tangent rays in the light-field, combined with the fact that the light-field is defined on a closed surface to bound the rays away from the tangent plane over the entire open subset. The open subset may have to be made smaller in this process.) The open subset is then further reduced until the boundary of the cone is tangent to the surface at the boundary of the open region. A region is then carved out of the occupied volume of S_1 beneath this open region of the surface and the albedo changed there as in Section 2.3. (Carving means removing the open region from the surface patch and some of the scene volume beneath it and adding in an additional surface patch to form the new surface.) As is illustrated in Figure 8, neither the visual hull nor the light-field is changed in this carving step; i.e. we have modified S_1 to create a different scene S_2 with the same light-field.

There is a minor technical detail to deal with here in that changing the shape of the surface changes $illum(x)$ and may make it smaller than $inter$ and violate Equation (3). This problem can be removed by noting: (1) there is a global scale ambiguity between the illumination and the albedo and so we can always find a new stereo solution with $illum(x)$ multiplied by an arbitrary constant, and (2) it is always possible to change the shape by a finite amount without $illum(x)$ dropping to zero because of the continuity of the light-field. (See Figure 7 and Section A.1 above.)