# BRDF Invariant Stereo using Light Transport Constancy

Liang Wang, Ruigang Yang[*] James Davis[†]

December 5, 2006

### Abstract

Nearly all existing methods for stereo reconstruction assume that scene reflectance is Lambertian and make use of brightness constancy as a matching invariant. We introduce a new invariant for stereo reconstruction called *Light Transport Constancy*, which allows completely arbitrary scene reflectance (BRDFs). This invariant can be used to formulate a rank constraint on multiview stereo matching when the scene is observed by several lighting configurations, in which only the lighting intensity varies. In addition, we show that this multiview constraint can be used with as few as two cameras and two lighting configurations. Unlike previous methods for BRDF invariant stereo, Light Transport Constancy does not require precisely configured or calibrated light sources, or calibration objects in the scene. Importantly, the new constraint can be used to provide BRDF invariance to any existing stereo method, whenever appropriate lighting variation is available.

**Keywords**: Stereo, BRDF, Rank Constraint, Light Transport Constancy, Non-Lambertian.

---

[*] Liang Wang and Ruigang Yang are with the University of Kentucky.

[†] James Davis is with University of California, Santa Cruz.

# 1 Introduction

Stereo reconstruction of scene depth is a well studied and important topic in computer vision. Most existing stereo methods rely on the assumption that objects in the scene reflect light equally in all directions. This assumption on surface reflectance is commonly referred to both as a Lambertian BRDF and as “brightness constancy.” Unfortunately, this assumption is violated for nearly all real world objects, leading to incorrect depth estimates.

Several methods for overcoming this limitation have been proposed, but all require some combination of calibrated light sources, calibration objects in the scene, or smoothness assumptions on the surface reflectance. This paper introduces *light transport constancy* as a constraint on stereo matching. Light transport constancy is based on the observation that surface reflectance properties are usually static. That is, for a static scene, the percentage of light reflected by a particular surface patch remains constant for a given viewing direction. This constraint has not been previously exploited and allows stereo correspondence to be correctly determined for surfaces with an arbitrarily complex BRDF and does not require calibrated light sources or objects.

As an intuitive introduction to this constraint, consider the scene configuration in Figure 1. The scene is illuminated by a single point light source, $L$. A particular point in the scene, $x_i$, will reflect light to each of cameras $C_1$ and $C_2$ according to:

$$E_{C_j}(x_i) = L(x_i)R(x_i, \theta_L, \theta_{C_j}) \tag{1}$$

where $E_{C_j}(x_i)$ is the radiance in the direction of $C_j$ from the point $x_i$, $L(x_i)$ is the observed irradiance of point $x_i$, and $R(x_i, \theta_L, \theta_{C_j})$ is the BRDF at point $x_i$, indexed by the vectors in the direction of $L$ and $C_j$. Throughout the text, direction vectors are written as single variables for notational simplicity (e.g. $\theta_L, \theta_{C_j}$) despite the fact they represent 2D quantities. Also for the sake of simplicity, we do not include the dependency of wavelength in this exemplary scenario.

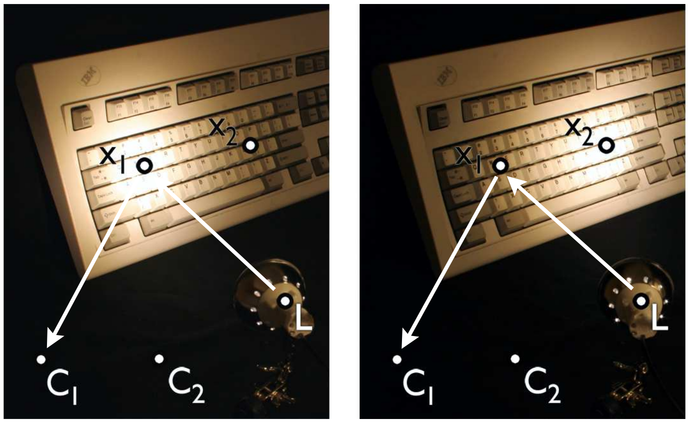

Figure 1: (Left) The BRDF at $x_1$ determines the percentage of light reflected from light source $L$ toward each of cameras $C_1$ and $C_2$. (Right) The spatial position of all components is the same, but the light distribution has been altered by rotating the light about its light bulb (i.e., steering the light beam to a different place). Although the incident intensity at $x_1$ has changed, the percentage of light reflected remains constant.

The traditional Lambertian assumption is that the reflectance (BRDF) is equal in the directions of $C_1$ and $C_2$, i.e.,

$$R(x_i, \theta_L, \theta_{C_1}) = R(x_i, \theta_L, \theta_{C_2}) \tag{2}$$

Thus we legitimately have $E_{C_1}(x_i) = E_{C_2}(x_i)$. However, this relation will not in general hold true for arbitrary BRDFs.

Light transport constancy assumes that the surface BRDF, $R(x_i, \theta_L, \theta_{C_j})$, remains constant under variable illumination. If we vary the lighting conditions so that the irradiance varies by a factor of $k(x_i)$, then the observed reflected radiance, $E'_{C_j}(x_i)$, will also vary by a factor of $k(x_i)$.

$$E'_{C_j}(x_i) = k(x_i)L(x_i)R(x_i, \theta_L, \theta_{C_j}) \tag{3}$$

Note that in general neither the irradiance nor the change in irradiance will be equal at different scene points. That is, $L(x_1) \neq L(x_2)$ and $k(x_1) \neq k(x_2)$. This is in contrast to the assumption made in many vision algorithms that the light source is a precisely isotropic emitter. Consider the two scene variants in Figure 1. The configuration of components is identical, but the emitted light intensity field has been changed by rotating the flashlight. The emitted light is not uniform in all directions, and thus $L(x_1) \neq L(x_2)$ and $k(x_1) \neq k(x_2)$.

One thing distinctly worth noticing is that light sources mentioned in this paper are geometrically static, i.e. stationary during image acquisition. Illumination variations simply come from variable radiant intensity distributions, instead of any spatial position variation of light sources.

Redefining our observation, $E''_{C_j}(x_i)$, as the ratio of two different lighting conditions, gives:

$$E''_{C_j}(x_i) = \frac{E'_{C_j}(x_i)}{E_{C_j}(x_i)} = \frac{k(x_i) \cdot L(x_i) \cdot R(x_i, \theta_L, \theta_{C_j})}{L(x_i) \cdot R(x_i, \theta_L, \theta_{C_j})} = k(x_i) \tag{4}$$

Note that the observations are invariant to camera viewpoint and $E''_{C_1}(x_i) = E''_{C_2}(x_i)$ regardless of the surface BRDF.

The simplified formulation just given is sufficient to design a practical stereo system which uses two cameras and a single uncalibrated light source. Practically, this design is easier to implement than existing methods for BRDF invariant stereo, because it requires fewer known or precisely calibrated scene components.

More important from a theoretical standpoint, the introductory formulation can be extended to handle incident lighting for which a single constant $k_i$ can not explain the lighting variation. By factoring the incident light field into a number of basis functions which vary independently, a series of linear equations which relate observations to lighting and reflectance can be derived. We can then use light transport constancy to formulate a rank constraint on multiview stereo matching, providing a relation between observations, lighting complexity, and BRDF complexity. One implication of this relation is that stereo matching can be performed precisely even when scenes contain arbitrary BRDFs.

This paper makes several contributions: the derivation of a rank constraint for stereo using light transport constancy which allows correspondence of arbitrary surface BRDFs, a practical implementation which is easier to reproduce than existing methods for BRDF invariant stereo, and an evaluation of our method on several real scenes to show that it is both practical and effective.

The rest of the paper is organized as follows. We first review the literature in Section 2, in particular stereo matching techniques for non-Lambertian surfaces. Then we develop our light transport constancy and discuss its variations with different lighting and BRDFs in Section 3. Experimental results are presented in Section 4, using several images captured from scenes with arbitrary BRDFs. Finally we conclude in Section 5.

## 2 Related Work

All stereo depth recovery methods make explicit or implicit assumptions about which image features are held constant. The primary differences arise from the choice of invariant. A number of possible invariants that allow stereo matching have been explored.

Stereo matching of specular surfaces has most commonly been approached by treating specularities as outliers to the brightness constancy invariant, which should be detected and either removed or avoided [2, 3, 1, 14, 15]. An alternate approach treats surfaces as diffuse-plus-specular and formulates a multiview constraint that all observations must lie on a line in color space [26]. Unfortunately, all of these methods limit the range of surface BRDFs to those which can be represented as a simple combination of diffuse and specular terms. The light transport constancy invariant presented in this work allows stereo matching of surfaces with completely arbitrary BRDF.

Jin et al. show that a multiview rank constraint on reflectance complexity is implied by a diffuse-plus-specular surface model and use this constraint to reconstruct non-Lambertian surfaces [11]. Although our work also formulates a rank constraint, we rely on a different matching invariant and allow for truly arbitrary surface BRDFs at each scene point.

Helmholtz stereopsis [16, 28, 29, 30] allows matching of arbitrary BRDFs using reciprocity. That is, $R(x_i, \theta_A, \theta_B) = R(x_i, \theta_B, \theta_A)$. By collocating point light sources with each camera it is possible to record reciprocal pairs using two different lighting conditions, such that image A is illuminated by light B, and image B is illuminated by light A. Due to reciprocity, the reflected light to cameras A and B will be equal. Unfortunately, this method requires the light sources be collocated with the optical center of each camera. Although acceptable results are possible by simply placing the light nearby, a proper implementation requires calibrated optics to ensure collocation. The method presented in this paper makes use of a different property and does not require the position of light sources to be precisely calibrated or even known.

Orientation constancy has been used to allow reconstruction of scenes with arbitrary BRDF in both photometric stereo and multiview stereo configurations [24, 9]. Although very accurate results are possible, these methods require a known calibration object with BRDF similar to the unknown scene, as well as distant cameras and light sources. In contrast, this work does not require a known object and allows for arbitrarily located light and camera positions.

Unlike many previous approaches that make use of *geometric* illumination changes, our formulation requires *radiometric* illumination variations, i.e., rather than changing position, the light sources in our work change only their intensities. Prior approaches using radiometric variations include structured light (e.g. [10, 23]), and the more general space-time stereo framework [27, 6]. Image intensity ratios are also a well studied method for recovering depth which are often formulated as using radiometric variation [4, 19].

It is argued in [25] that image ratios are only applicable to diffuse surfaces. Our method is fundamentally different from that work in that we assume radiometric variations only while the derivation in [25] is based on an illumination distribution which includes geometric variation. Our experimental results demonstrate that using radiometric variation, scenes with arbitrary surface BRDFs can be effectively reconstructed using image ratios.

The invariant proposed by this paper, light transport constancy, has not previously been explored for stereo matching. However, in the case of laser scanning, it was explicitly identified and articulated by Curless and Levoy [5]. In addition, it has implicitly been used in other domains. Magda et al. capture hundreds of images illuminated by precisely calibrated light source positions on two concentric spheres surrounding an object. The two sampled representations of the incoming illumination field can then be aligned to find the depth of a given scene point [16].

# 3 Light Transport Constancy

Light transport constancy (LTC) can be used to formulate a general constraint on multi-baseline stereo matching regardless of the surface BRDF complexity, provided that sufficient illumination variations and viewpoints are available. A point to emphasize here is that when we discuss illumination variation in the scope of LTC, we mean *radiometric* variations of the light source, i.e., changes in the radiant intensity. This is fundamentally different from *geometric* lighting variations, i.e., moving the light source around, as required in many photometric stereo methods.

This section first presents the rank constraint in the context of multiple point light sources, each of which varies independently. We then show how this can be applied to arbitrary lighting by replacing point lights with arbitrary lighting basis functions. Finally, we expand the formulation to include the concept of BRDF complexity and show that simple BRDFs also provide a rank constraint.

## 3.1 LTC as a rank constraint

The simplified introduction given in section 1 assumes that the irradiance is due to a single light source and varies by a single multiplier, $k_i$. We now formally introduce our radiometric model.

For a single point $x_i$ on the display surface, it is illuminated by a point light source and observed by several cameras. For the sake of simplicity, let us for now assume the camera have just one channel (e.g., a gray-scale camera).

The irradiance at $x_i$ is denoted as $D(x_i, \lambda)$ where $\lambda$ is the wavelength. Let $R(x_i, \lambda, \theta_L, \theta_{C_j})$ be the spectral reflectance (i.e., BRDF) of $x_i$ indexed by the incident direction $\theta_L$ and viewing direction $\theta_{C_j}$. If $t(\lambda)$ is the spectral response for the camera, then the irradiance detected by the camera sensor is:

$$I_{C_j}(x_i) = \int_{\Lambda} D(x_i, \lambda) \cdot R(x_i, \lambda, \theta_L, \theta_{C_j}) \cdot t(\lambda) d\lambda, \tag{5}$$

where $\Lambda$ is the camera's spectrum. Note that strictly speaking, the integration should include a cosine term to account for the fore-shortening effect. Since we are dealing with a static scene, it is a per-point scale factor and we consolidate it in the BRDF. Finally, the measured irradiance $I_{C_j}(x_i)$ is converted to a pixel value via a camera response function. For the scope of this paper, we assume that the camera has a linear response, in other words, the camera is measuring relative irradiance directly. To deal with cameras with non-linear responses, standard radiometric calibration procedures (e.g. [18, 7]) should be applied to correct the pixel values.

If we change only the intensity of $D(x_i, \lambda)$ by a scale factor $k(x_i)$ and keep everything else fixed, $I_{C_j}(x_i)$ will be modulated by $k(x_i)$ according to equation 5 and our assumption of linear camera response. This concurs with our intuitive introduction in section 1. We now expand to derive a series of linear equations that can accommodate an arbitrary number of light sources. These equations are the basis for a rank constraint on stereo matching.

Figure 2 shows a scene observed from multiple cameras and illuminated by multiple light sources. We can explain the perceived irradiance from a particular scene point, $x_i$, in the direction of a particular camera, $C_j$, as a combination of the reflected light from each individual source, $Light_1..Light_M$.

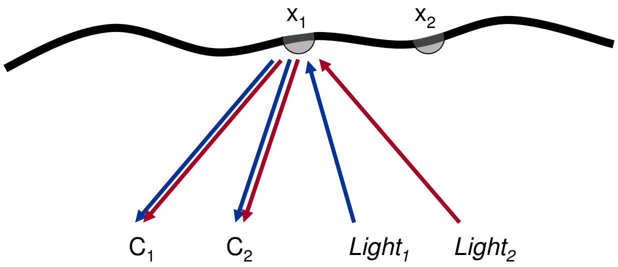

Figure 2: Light reflected toward camera $C_1$ can be explained as a combination of reflected light from each of $\mathit{Light}_1$ and $\mathit{Light}_2$.

$$I_{C_j}(x_i) = \int\limits_{\Lambda} D_1(x_i, \lambda) \cdot R(x_i, \lambda, \theta_{L_1}, \theta_{C_j}) \cdot t(\lambda)\, d\lambda + \int\limits_{\Lambda} D_2(x_i, \lambda) \cdot R(x_i, \lambda, \theta_{L_2}, \theta_{C_j}) \cdot t(\lambda)\, d\lambda + \ldots \tag{6}$$

For notational convenience we will hereafter drop the indexing for scene location, $x_i$, since it is understood that each scene location is considered separately. Further, we denote integration constants for particular pairs of light-camera directions as

$$R_{C_1L_1} = \int\limits_{\Lambda} D_1(\lambda) \cdot R(\lambda, \theta_{L_1}, \theta_{C_1}) \cdot t(\lambda)\, d\lambda \tag{7}$$

Equation 6 can be rewritten using the new notation as:

$$I_{C_j} = R_{C_jL_1} + R_{C_jL_2} + R_{C_jL_3} + \ldots \tag{8}$$

We can include the notion of lighting variation in which $D_i(\lambda)$ is modulated by a scalar $L_{iV_j}$. Let $I_{C_1V_1}$ be the observed irradiance at camera $C_1$ under the illumination variation $V_1$, we can write a sequence of bilinear equations relating the observations from each camera, $C_1..C_J$, under illumination conditions, $V_1..V_N$:

$$\begin{aligned} I_{C_1V_1} &= L_{1V_1}R_{C_1L_1} + L_{2V_1}R_{C_1L_2} + \dots \\ I_{C_2V_1} &= L_{1V_1}R_{C_2L_1} + L_{2V_1}R_{C_2L_2} + \dots \\ &\dots \\ I_{C_1V_2} &= L_{1V_2}R_{C_1L_1} + L_{2V_2}R_{C_1L_2} + \dots \\ I_{C_2V_2} &= L_{1V_2}R_{C_2L_1} + L_{2V_2}R_{C_2L_2} + \dots \\ &\dots \end{aligned} \tag{9}$$

Note that light transport constancy holds that $R_{C_jL_m}$ is constant for a given pair of light source and camera position regardless of how we vary the illumination conditions. In addition, the illumination variation for a given light source, $L_{mV_n}$, does not depend on either the BRDF or the camera viewpoint.

This set of linear equations can be rewritten in matrix form as:

$$\begin{array}{cc} & \text{\# of cameras} \\ \text{\# of variations} & \begin{bmatrix} I_{C_1V_1} & I_{C_2V_1} & \cdots \\ I_{C_1V_2} & I_{C_2V_2} & \\ I_{C_1V_3} & I_{C_2V_3} & \\ \vdots & & \end{bmatrix} \end{array} = \begin{array}{cc} & \text{\# of lights} \\ \text{\# of variations} & \begin{bmatrix} L_{1V_1} & L_{2V_1} & \cdots \\ L_{1V_2} & L_{2V_2} & \\ L_{1V_3} & L_{2V_3} & \\ \vdots & & \end{bmatrix} \end{array} \begin{array}{cc} & \text{\# of cameras} \\ \text{\# of lights} & \begin{bmatrix} R_{C_1L_1} & R_{C_2L_1} & \cdots \\ R_{C_1L_2} & R_{C_2L_2} & \\ \vdots & & \end{bmatrix} \end{array} \tag{10}$$

Let us denote the matrix on the left side **I**, and the two matrices on the right **L** (lighting modulation matrix) and **R** (reflectance matrix). From the factorization, we can see that there is a rank constraint on matrix **I**. When the number of light sources, $M$, is less than both the number of lighting variations and the number of cameras, matrix **I** has rank of at most $M$. This constraint allows stereo correspondence to be determined.

### 3.2 Rank constraint with multiple color channels

In the case of color cameras, irradiance $I_{C_j}$ is typically represented as a triple of three intensity values, each representing a distinct color channel in red, green or blue. Let us denote them as $\{I^r_{C_j}, I^g_{C_j}$, and $I^b_{C_j}\}$. Similarly we further decompose the spectral response of the light source $Light_i$ into three separate channels: $\{D_i^r(\lambda), D_i^g(\lambda), D_i^b(\lambda)\}$ (imagine that we have a 3-color light projector).

By plugging in different camera/light spectral responses $t(\lambda)$ and $D(\lambda)$ for each color channel in equation 5, $\{I_{C_j}^r, I_{C_j}^g$, and $I_{C_j}^b\}$ can be obtained as

$$\begin{aligned} I_{C_j}^r &= R_{C_jL_1^r}^r + R_{C_jL_1^g}^r + R_{C_jL_1^b}^r \\ I_{C_j}^g &= R_{C_jL_1^r}^g + R_{C_jL_1^g}^g + R_{C_jL_1^b}^g \\ I_{C_j}^b &= R_{C_jL_1^r}^b + R_{C_jL_1^g}^b + R_{C_jL_1^b}^b \end{aligned} \tag{11}$$

where $R_{C_jL_1^m}^l = \int D_1^m(\lambda) \cdot R(\lambda, \theta_{L_1}, \theta_{C_1}) \cdot t^l(\lambda) d\lambda$ and $l, m \in \{r, g, b\}$.

With multiple views and multiple lights, we can rewrite the matrix in equation 10 for color input as:

$$\begin{array}{cc} & \text{(3}\times\text{) \# of cameras} \\ \text{\# of variations} & \begin{bmatrix} I_{C_1V_1}^r & I_{C_1V_1}^g & I_{C_1V_1}^b & \cdots \\ I_{C_1V_2}^r & I_{C_1V_2}^g & I_{C_1V_2}^b & \\ I_{C_1V_3}^r & I_{C_1V_3}^g & I_{C_1V_3}^b & \\ \vdots & & & \end{bmatrix} \end{array} = \begin{array}{cc} & \text{(3}\times\text{) \# of lights} \\ \text{\# of variations} & \begin{bmatrix} L_{1V_1}^r & L_{1V_1}^g & L_{1V_1}^b & \cdots \\ L_{1V_2}^r & L_{1V_2}^g & L_{1V_2}^b & \\ L_{1V_3}^r & L_{1V_3}^g & L_{1V_3}^b & \\ \vdots & & & \end{bmatrix} \end{array} \begin{array}{cc} & \text{(3}\times\text{) \# of cameras} \\ \text{(3}\times\text{) \# of lights} & \begin{bmatrix} R_{C_1L_1^r}^r & R_{C_1L_1^r}^g & R_{C_1L_1^r}^b & \cdots \\ R_{C_1L_1^g}^r & R_{C_1L_1^g}^g & R_{C_1L_1^g}^b & \\ R_{C_1L_1^b}^r & R_{C_1L_1^b}^g & R_{C_1L_1^b}^b & \\ \vdots & & & \end{bmatrix} \end{array} \tag{12}$$

The reflectance matrix $\mathbf{R}$ consists of many $3 \times 3$ sub-matrices. Each matrix, typically called as a *color mixing* matrix, records the interaction of the spectral responses of the light source and camera. Typically, the responses of cameras and projectors are wide band and have large overlaps [8]. Thus each sub-matrix has a general form shown above. Nevertheless, as far as stereo matching is concerned, we are only interested in the rank of matrix $\mathbf{I}$ on the left side, not the actual decomposition. Therefore the rank constraint we have developed for gray-scale images can be simply extended. That is, matrix $\mathbf{I}$ has rank of at most $3 \times M$. Note that although $\mathbf{I}$ now has a higher rank, it has $3\times$ the columns as well, so that on balance we expect little change in the outcome.

There are two special cases we can consider. First, with a white light captured by color cameras, the number of columns in the lighting matrix $\mathbf{L}$ reduces by a factor of three.

Therefore the rank constraint on $\mathbf{I}$ is at most $M$, i.e., same as using gray-scale cameras. Since we are measuring 3× the data but have only 1× the rank constraint, we expect white lighting and color cameras to be a desirable measurement configuration. Second, with color lighting captured by gray-scale cameras, we have two subcases. If the three color channels scale independently, $\mathbf{I}$ has only $J$ columns where $J$ is the number of cameras, but its rank constraint remains $3 \times M$. Naturally this is undesirable since more cameras will be required to ensure that $\mathbf{I}$ has a sufficient number of columns. On the other hand, if the three channels scale in the same way, $\mathbf{I}$'s rank remains $M$, which is the same as the gray-scale case.

Dealing with color images is a direct extension from the gray-scale case. Because the notation is cumbersome, we will resume the assumption of a gray-scale world in our remaining discussion.

## 3.3 Arbitrary lighting basis functions

Light transport constancy applies even when light sources are not simple point light sources. Each light in the preceding analysis can be replaced with a lighting basis function, each of which might have broad spatial support.

In general, the irradiance value from a scene point, $x_i$, in the direction of camera $C_j$ can be written as an integral over all incoming light directions. Therefore equation 5 can be modified as the following for a more general lighting setup.

$$I_{C_j} = \iint\limits_{\Phi\,\Lambda} D(\lambda, \phi) \cdot R(\lambda, \phi, \theta_{C_j}) \cdot t(\lambda) d\lambda d\phi \tag{13}$$

where $D(\lambda, \phi)$ is the incident light irradiance function indexed by incoming angle $\phi$, and $\Phi$ ranges over a hemisphere.

The irradiance field $D$ can be decomposed into a linear combination of basis vectors:

$$D(\lambda, \phi) = k_{L_1} D_1(\lambda, \phi) + k_{L_2} D_2(\lambda, \phi) + \ldots + k_{L_M} D_M(\lambda, \phi) \tag{14}$$

It is conceptually helpful to think of each basis as a separate light source. We previously discussed individual point lights as the basis, however area lights represented as a piecewise constant basis, or a wavelet decomposition of the incident illumination field would work equally well. By truncating the wavelet expansion after a sufficient amount of variation has been accounted for, very general lighting can be modeled using a finite set of coefficients. The graphics community has in fact used such an expansion to represent incident illumination fields [20].

We can now rewrite equation 13, taking into account the lighting bases and indexed by illumination condition.

$$I_{C_j V_n} = k_{L_1 V_n} \iint\limits_{\Phi\,\Lambda} D_1(\lambda,\phi)\cdot R(\lambda,\phi,\theta_{C_j})\cdot t(\lambda) d\lambda d\phi + k_{L_2 V_n} \iint\limits_{\Phi\,\Lambda} D_2(\lambda,\phi)\cdot R(\lambda,\phi,\theta_{C_j})\cdot t(\lambda) d\lambda d\phi.. \tag{15}$$

That is, the observation from camera $C_j$ under illumination condition $V_n$, is a summation over the individual lighting bases, each modified by their own variation multiplier, $k_{L_m V_n}$.

Notice that each integral term is constant because it relies only on the lighting basis and the surface BRDF. Just as was true in the case of discrete point light sources, lighting variation will induce a set of bilinear equations. These equations can be written identically to equation 10 by redefining variables in terms of the new continuous formulation.

$$\begin{aligned} L_{mV_n} &= k_{L_m V_n} \\ R_{C_j L_m} &= \iint\limits_{\Phi\,\Lambda} D_m(\lambda,\phi)\cdot R(\lambda,\phi,\theta_{C_j})\cdot t(\lambda) d\lambda d\phi \end{aligned} \tag{16}$$

## 3.4 Limited BRDF complexity

So far we have formulated the problem assuming completely arbitrary surface reflectance. However, most real world BRDFs are not arbitrary, and it is unlikely that the reflectance is truly independent in every camera direction. In this case we can further factor the reflectance matrix, **R**, into a set of reflectance bases, **B**, and a mixing matrix **M**.

$$\begin{array}{cc} & \text{\# of cameras} \\ \text{\# of lights} & \begin{bmatrix} \quad R \quad \end{bmatrix} \end{array} = \begin{array}{cc} & \text{\# of BRDF bases} \\ \text{\# of lights} & \begin{bmatrix} \quad B \quad \end{bmatrix} \end{array} \begin{array}{cc} & \text{\# of cameras} \\ \text{BRDF bases} & \begin{bmatrix} \quad M \quad \end{bmatrix} \end{array} \tag{17}$$

We now have a trilinear equation **I=LBM**, which has a rank constraint on **I** if either **L** or **B** has a small number of columns. For example, if the surface is Lambertian, then a single BRDF basis describes the outgoing light in all camera directions, and **B** has a single column. Thus we have a rank constraint if either the illumination or the BRDF is sufficiently "simple." In this work we address completely arbitrary BRDFs and have not evaluated the expected complexity of real world BRDFs.

## 3.5 Stereo matching

It is not necessary to find an actual factorization of the observation matrix **I** in order to evaluate stereo correspondence. It is sufficient to calculate the singular values of matrix **I** and select the disparity which results in a matrix of minimum rank.

Because the matrix will be corrupted with noise, it is impossible to calculate rank exactly. Conceptually, we prefer matrices which have most of their energy in the first few principal components rather than those with evenly distributed energy. Thus, we use moments to approximate the notion of minimum rank and select the disparity with minimum score. If the singular values of I are encoded in $w_1..w_n$, then we choose the disparity which minimizes $\Re$.

$$\Re = \sum_i (i \cdot w_i^2) / \sum_i w_i^2 \tag{18}$$

When a single light source and only two cameras are used, simply minimizing the second singular value is equivalent to equation 18. However, in general it is impossible to use the second (or any particular) singular value as a matching metric, because the expected rank of the matrix is not known a priori.

The introductory matching metric which uses image ratios given in equation 4, is also equivalent to equation 18. A proof of this equivalence is provided in the Appendix. When only two cameras are used, this simpler matching metric is quite convenient, because it allows existing stereo implementations to be used without modification.

Scharstein and Szeliski have introduced a taxonomy of stereo algorithms which includes matching cost, aggregation, and disparity selection [22]. Light transport constancy and the implied rank constraint are local operators and replace only the matching cost in existing stereo algorithms. Aggregation, disparity selection, and any global regularization are all orthogonal issues, and the new invariant introduced in this work can be used in conjunction with a wide variety of existing algorithms.

## 4 Experiments

To facilitate the evaluation of our technique, we captured several stereo data sets under varying illumination conditions. Our data acquisition setup includes up to four synchronized VGA ($640 \times 480$) cameras and two light projectors, as shown in Figure 3. The cameras are calibrated with respect to each other, but the projectors are completely uncalibrated. Note that much simpler light sources could be substituted–for example, the flashlight shown in Figure 1. We use projectors only because they allow the light distribution to be controlled remotely rather than by physically manipulating the light source. The actual light output of the projector is unknown to our algorithm. We used several types of patterns for lighting variation (shown in Figure 4), attempting to verify that our results work for both low and high frequency variation. The first is a smooth ramp that is used in the minimum configuration of two lighting variations. The second is a randomly moving Gaussian blob that exhibits low-frequency brightness variation. The third is a pattern acquired from a real flashlight. And the last is a stripe pattern with random intensity values which exhibits high frequency variation. Unless noted otherwise, all the experiments were carried out with the low-frequency (blob) pattern since we expected this to most closely mimic a spotlight which is brighter in the center of its field, similar to the motivational example shown in Figure 1.

Another practical issue to mention is the dynamic range. Saturated pixels (e.g., from specular highlights) will violate the rank constraint we have developed. In our experiments we carefully control the exposure to avoid saturation. It is also possible to combine images taken with multiple exposures to generate a high-dynamic-range (HDR) image (e.g. [17, 7]).

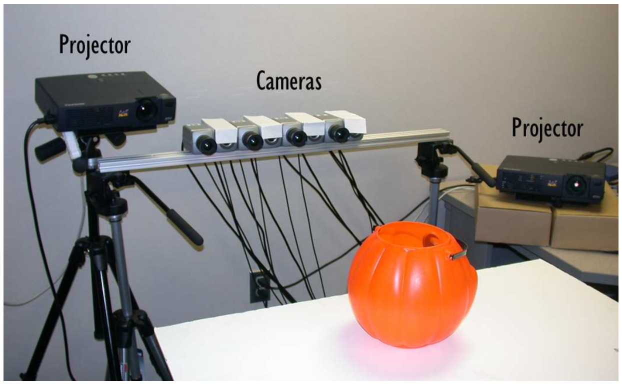

Figure 3: Our experimental setup with four cameras and two variable light sources.

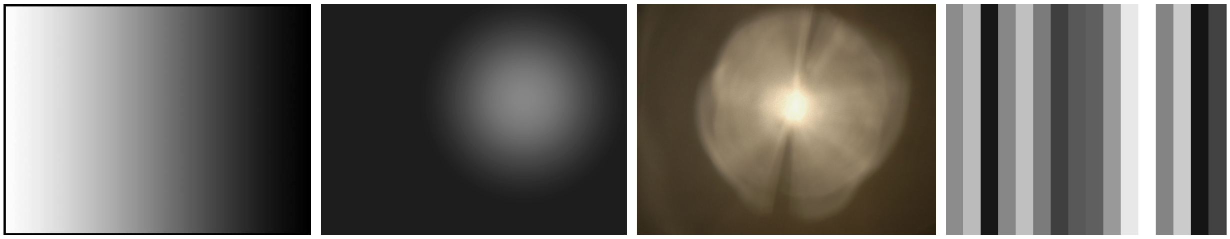

Figure 4: Patterns used for lighting variation. From left to right: *ramp* lighting (boxed for illustration purpose), *blob* lighting, *flashlight*, *stripe* lighting.

Two-view stereo is the dominant method by which stereo algorithms are evaluated. Although our method is inherently multiview, we defer to tradition and first evaluate our method in the arrangement we believe will be most commonly implemented. Following these evaluations we provide some analysis of the rank constraint when multiple cameras and lights are present. Finally we show some quantitative evaluations with a ground-truth data set.

### 4.1 Two-view with one light source

In this setup, we used two cameras and a single light source position. We experimented with gray-scale images to evaluate our method against traditional stereo.

**Minimal configuration.** We captured gray-scale images from each of two cameras under two different lighting variations. Figure 5 shows the two lighting variations from the viewpoint of one of the cameras. The first lighting pattern is a flat gray-field and the second is the ramp in Figure 4. Brightness constancy (i.e., traditional intensity difference based on lambertian surfaces) was evaluated using one of the two lighting configurations. Light transport constancy was evaluated by first computing a new image as the ratio of the two illumination conditions, as given in equation 4. This process is mathematically equivalent to evaluating the rank constraint. The resulting ratio image is shown in Figure 6. Note that neither the specular highlights nor any other view-dependant effect are visible in the ratio image.

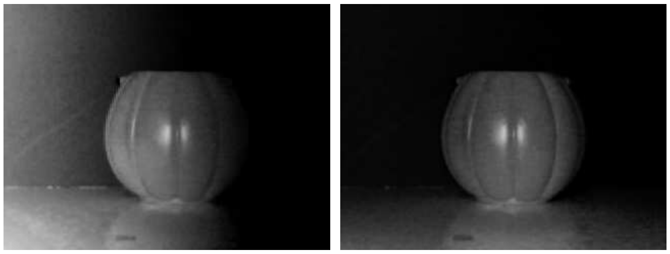

Figure 5: A plastic pumpkin illuminated by a single light source under two different lighting conditions.

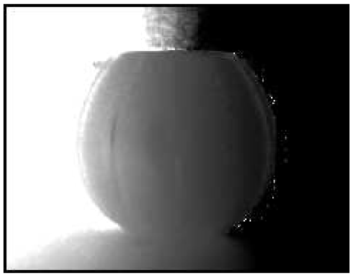

Figure 6: The ratio of images taken under two lighting conditions.

Standard stereo matching was applied to the stereo pairs arising from both brightness constancy and light transport constancy using a Sum-of-Absolute-Differences (SAD) metric. Because we are interested in the performance of a local matching operator, we use a winner-takes-all approach and simply accept the minimum SAD disparity as correct rather than applying a global regularization method.

Figure 7 shows the stereo results from each method. The left column is derived from brightness constancy, and the right column is from light transport constancy. The first row shows the disparity map computed by each method. Depth is coded such that white pixels indicate depths closer to the camera. The second row shows the same data along a single scan-line as scaled disparity values. In both visualizations, it is clear that our new method has superior results. Note the garbled depth values in the case of brightness constancy. In the third row of Figure 7, we investigate the reason that our method performs well by plotting the matching profile for a single pixel. Note that brightness constancy has no clear global minimum, whereas our method has a very clear minimum at the correct disparity. This presumably leads to much better depth estimates.

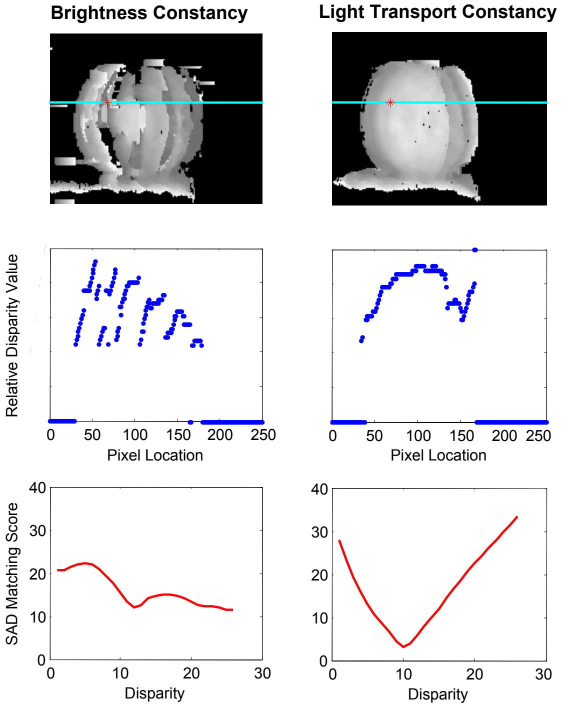

Figure 7: Results from using brightness constancy (left column) and light transport constancy (right column). (Row 1) Disparity maps computed by stereo matching using each invariant. (Row 2) Scaled disparity estimates along a single scan line. (Row 3) Matching profile for the pixel marked with a red cross.

**Together with existing stereo methods.** In order to validate that existing stereo methods can be adapted to handle non-Lambertian objects, we tested the same two sets of gray-scale stereo pairs with a stereo implementation available on the web [13]. This implementation happens to be based on graph cuts [12], which allowed us to further verify that no undesirable artifacts are caused by integration with a global regularization method.

Since we have computed a ratio image to use for matching, absolutely no modification to the existing code was required. The computed disparity maps are shown in Figure 8. Similar to the winner-takes-all example above, the disparity map computed using light transport constancy shows much better results.

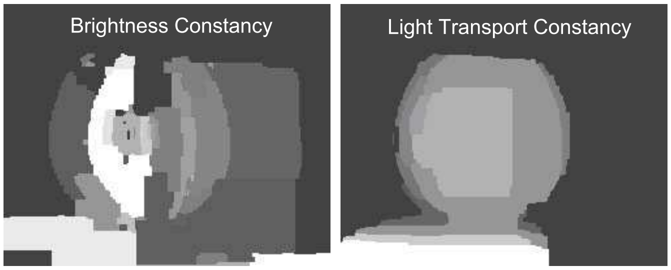

Figure 8: Disparity maps computed using an unmodified graph-cut stereo algorithm with brightness constancy (left) and our new invariant (right).

**Increased lighting variation.** It is possible that our improved results come merely because by imposing lighting variation more information is available when computing disparity, rather than because our new invariant actually performs better. To evaluate whether this is true, we computed disparity using a data set with six lighting variations, as shown in Figure 9. Brightness constancy was evaluated as the Sum-of-Absolute-Differences over the vector of all six image pairs. Light transport constancy was evaluated as a rank constraint over the same input images. Although it is clear that additional lighting variations improve the result from brightness constancy, the result from light transport constancy also improves. We conclude that additional lighting variations improve the results from either constraint but that our new invariant performs better on objects such as the pumpkin, which exhibit non-Lambertian effects.

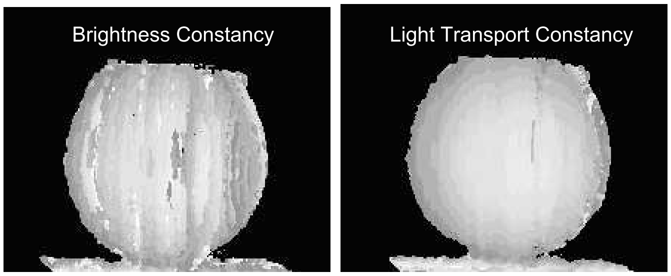

Figure 9: Disparity maps computed from a data set with six illumination variants. Left is from brightness constancy; right is from light transport constancy.

**Using a simulated flashlight.** We captured a lighting pattern of a regular flashlight (shown as one pattern in Figure 4). To facilitate automatic data acquisition, we use a projector to display five variations of the flashlight pattern with shift or rotation, simulating the scenario described in Figure 1. Good results can be obtained as shown in Figure 10.

Figure 10: Reconstructed depth map using a simulated flashlight with five lighting variations.

**Complex reflectance.** We further experimented with scenes containing more complex surface material properties. We first captured a piece of silk glued onto a slightly curved surface. The view dependent reflectance of the silk is very obvious in the stereo pair, as shown in Figure 11. Using seven lighting variations, we evaluate brightness constancy against our new invariant and find that light transport constancy is better able to deal with this highly non-Lambertian scene. The improvement is particularly obvious in the plot of disparity along a scan-line, shown in the bottom row of Figure 12. Brightness constancy results in many incorrect disparity estimates, whereas light transport constancy results in a smooth curve.

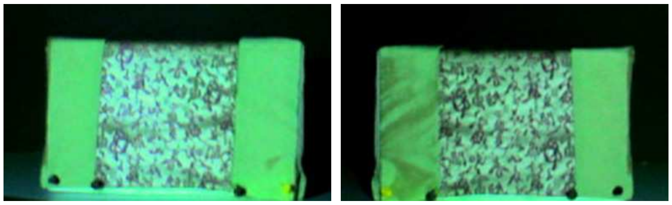

Figure 11: Silk cloth from two different viewpoints. Note the non-Lambertian reflectance.

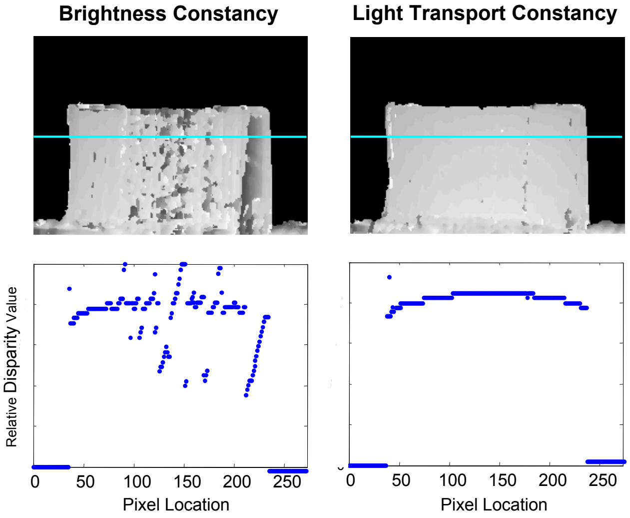

Figure 12: (Top) Disparity maps computed using brightness constancy and light transport constancy (LTC). (Bottom) Scaled disparity values along a single scan-line. Note how much more robustly LTC estimates depth.

**Multi-channel color.** The advantage of light transport constancy over brightness constancy is further demonstrated in Figure 13. We captured full color images of a lady's purse made from materials with a complex anisotropic BRDF. Note the surface color changes in the stereo image pair: the right side of the purse appears to be blue in one image and pink in the other image. With a white light source, we captured just two lighting variations in full color. We use as few lighting variations as possible to illustrate the effectiveness of our approach. All color images were used to compute each of light transport constancy and brightness constancy. The reconstructed depth maps are shown in the bottom row of Figure 13. We would not expect brightness constancy to perform well under these conditions, and indeed we see that the computed object depth is erroneous in the region exhibiting color change. In contrast, light transport constancy is able to evaluate depth accurately.

Note that it is not required to use the multichannel color formulation to compute disparity on colored objects such as this. We converted the input images to gray-scale to experiment with the formulation given in equation 10, and found the result to be qualitatively similar to that in Figure 13, which was computed using the full 3-channel color formulation given in equation 11, with the caveat that the rank on the matrix $I$ is expected to be one because we use a gray-scale light source. As we have discussed in Section 3.2, this is a more favorable configuration for matching.

**Complex geometry.** Our next data set is a live tree with substantial specular highlights. This scene would be challenging for traditional stereo algorithms due to the non-Lambertian effects and because there are many depth discontinuities. For this setup, we used the high-frequency (stripe) pattern with 30 variations to calculate the disparity map shown in Figure 14. With such a large number of lighting conditions, we would anticipate good performance. As expected, the results are of high quality. Individual leaves are well represented by clean boundaries and smooth estimates of depth, despite the fact that no global regularization method was applied.

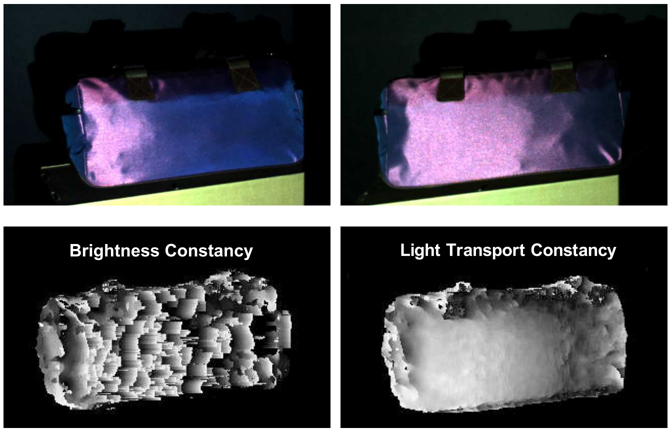

Figure 13: Stereo reconstruction of a lady's purse with anisotropic BRDFs. (Top row) the left and right images under one lighting condition; note the color changes in two images. (Bottom left) reconstructed depth map using brightness constancy. (Bottom right) reconstructed depth map using light transport constancy.

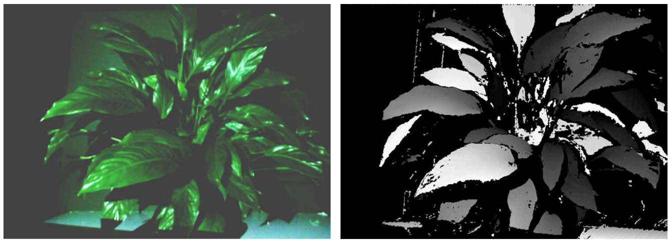

Figure 14: (Left) Tree with non-Lambertian reflectance properties and many depth discontinuities. (Right) Disparity map computed from thirty lighting variations.

## 4.2 Multi-view with two light sources

To evaluate the behavior of the rank constraint under multiview conditions, we computed disparity on the pumpkin scene using four cameras, two light sources, and thirty lighting variations. The resulting disparity map can be seen in Figure 15. As a whole, the results are very good, with smooth estimates of depth across the surface of the pumpkin. There is an error in the lower left corner which we believe is caused by occlusion from some camera viewpoints. Accounting for partial occlusion is typically handled during the aggregation stage of stereo processing, and, as mentioned earlier, we focus on the matching cost in this work.

**Analysis of singular values.** When two light sources are used, the rank of the observation

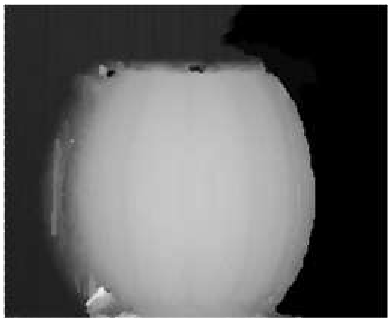

Figure 15: Disparity map for the pumpkin calculated from multiple cameras and multiple light sources.

matrix is limited to 2 for surfaces with arbitrary BRDF. In this case, we expect the third singular value to be minimized at the correct disparity. However, if the complexity of the surface reflectance is limited, the rank may be lower. This could happen either if the surface was actually Lambertian or merely because it appears Lambertian from the limited set of viewpoints available.

To provide some insight into the behavior of our rank constraint, we plotted the $2^{nd}$, $3^{rd}$, and $4^{th}$ singular values as a function of disparity for two different scene points, drawn from the multiview example above. For the scene point in the top plot of Figure 16, we see that the $2^{nd}$ singular value has an obvious minimum and that the combined metric $\Re$ is minimized at this same disparity. However, in the case of the scene point in the bottom plot, $\Re$ is minimized at the same disparity as the $3^{rd}$ singular value. Although the $4^{th}$ singular value is not precisely zero as would be expected in an ideal environment without noise, we can see that $\Re$ has an easily locatable global minimum which confirms that our approximation of "minimum rank" is performing as expected.

## 4.3 Quantitative Evaluation

While the Middlebury stereo evaluation web site [21] has become the gold standard to evaluate performance of stereo algorithms, the datasets there do not include lighting variations therefore cannot be used for our approach. In order to generate our own "ground-truth"

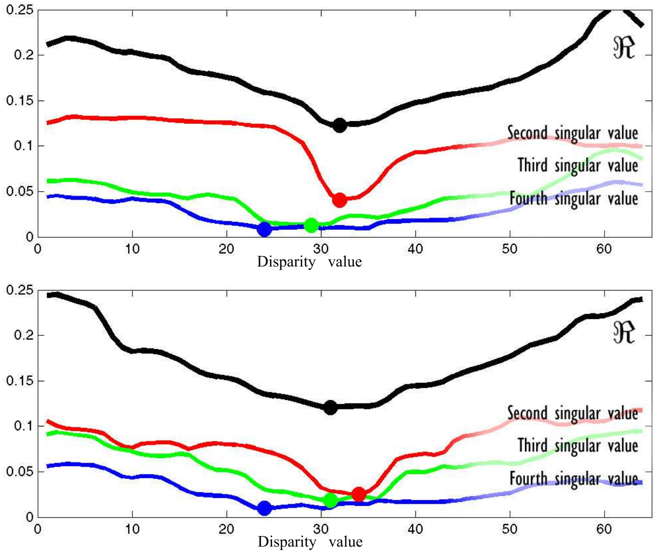

Figure 16: Normalized singular values for two particular scene points. The x-axis represents the disparity. Dots indicate the minimum on each curve. The moment has been scaled to fit on the same graph together with the singular values. Note that the moment is minimized together with a different singular value in each case.

data, we project a single vertical strip pattern from the light projector and calculate the depth along this strip using traditional stereo. The strip is swept across the scene simulating a laser triangulation-based scanner. Since only the strip is illuminated, disparity can be calculated unambiguously. The advantage of this approach as opposed to using a real laser range scanner is that the ground-truth data is automatically registered in the stereo cameras' coordinate frame.

Figure 17 shows a data set we captured. Note that we have not implemented sub-pixel disparity interpolation, so some of the groves on the surface are not visually noticeable in the depth map. Bad pixels around the silhouette (due to occlusions) are manually removed.

We generated depth results under varying patterns and compared the depth maps with the ground-truth data. If a pixel's disparity differs more than one pixel from the ground truth, we label it as a bad match. The error rates from two methods, one using the brightness constancy (BC) and the other using LTC, are summarized in Table 1 and Figure 18. In general, the error rates for both methods reduce as the number of lighting variations increases. This is not surprising because there are more data to work with. With just a

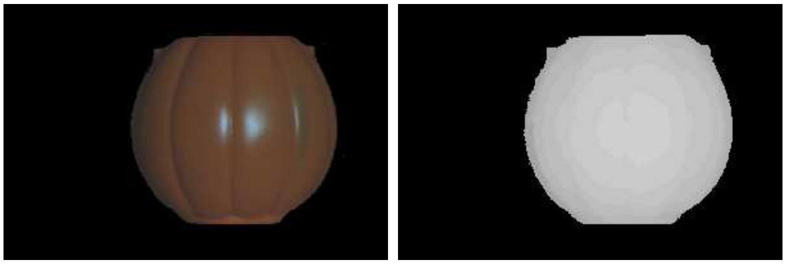

Figure 17: The ground truth dataset.Left is one color image and right is its corresponding depth map. Bad pixels due to occlusions are manually removed.

few low-frequency lighting variations, the error rate from BC is very large and changes quite arbitrarily. Results from LTC are much better. On the other hand, with high frequency lighting, both BC and LTC can generate much more accurate results and the difference in error rates is much smaller. These rapid lighting variations in fact "mask" the surface reflectance properties. This is similar to the fact that regular structured light scanners using binary-coded patterns can get decent results from shiny objects. Nevertheless, LTC always outperform BC in all testing cases.

| # of lighting variations | Low-frequency lighting (blob) | | | | | | High-frequency lighting (stripe) | | |
|---|---|---|---|---|---|---|---|---|---|
| | 2 | 3 | 4 | 8 | 16 | 32 | 8 | 16 | 32 |
| BC Error (%) | 54.5 | 65.0 | 61.1 | 40.7 | 44.1 | 21.3 | 3.44 | 3.95 | 3.25 |
| LTC Error (%) | 13.6 | 9.27 | 5.9 | 5.28 | 3.9 | 3.10 | 3.33 | 3.13 | 2.36 |

Table 1: Error rate of depth maps computed with brightness constancy (BC) and light transport constancy (LTC). Different lighting patterns (shown in Figure 4) are used.

## 5 Conclusion

Light transport constancy is a new invariant for multiview stereo matching which allows the depth of surfaces with arbitrary BRDF to be computed. We introduce a rank constraint based on this invariant which allows stereo algorithms to combine observations of non-Lambertian surfaces from different viewpoints in a theoretically principled way.

Our rank constraint can be applied with as few as two cameras and two lighting configurations. In addition, unlike existing methods for non-Lambertian stereo matching, our method does not require that light sources be precisely calibrated nor does it require known

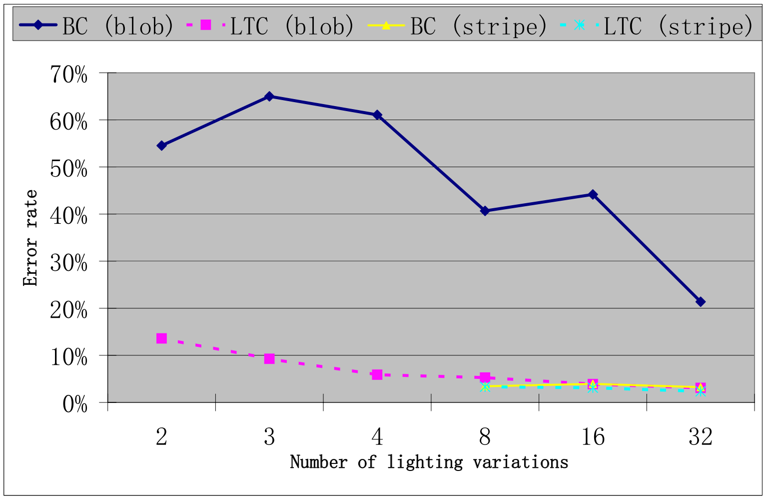

Figure 18: A plot of the error rates using data from Table 1.

calibration objects in the scene.

The rank constraint implied by light transport constancy can easily be employed as a replacement to brightness constancy. Thus, whenever sufficient lighting variation is available, any existing stereo algorithms can be enhanced to allow matching of non-Lambertian surfaces.

We have verified experimentally that stereo matching is possible using our rank constraint. In addition, we show that it performs better than brightness constancy on a variety of scenes.

A few aspects of our work may limit the conditions under which light transport constancy can be used. The rank constraint requires multiple illumination conditions to be available. All previously existing methods for arbitrary BRDF stereo also require illumination variation [15, 24], and it is interesting to wonder if this is a fundamental requirement. In addition, we do not consider the issue of inter-reflection in our formulation. In scenes with strong inter-reflection (e.g. concave and shiny objects), some points may have a higher rank than the rest (consider inter-reflections as additional light sources). Experiments are needed to see if inter-reflection can be treated as a secondary effect or noise. Finally, the rank constraint is a multiview constraint, and we do theoretically require more camera viewpoints than light source positions when the surface BRDF is truly arbitrary. However, the BRDF of most real surfaces is not arbitrary, and we have shown that BRDF complexity can be traded for lighting complexity. Thus an interesting avenue for future work would be to characterize the actual matrix rank, and thus the actual number of viewpoints required, for a wide class of naturally occurring scenes and lighting.

**Acknowledgment** Helpful early discussions were held with Gaurav Garg, Jeffrey Ho, Diego Nehab, Szymon Rusinkiewicz, and Vaibhav Vaish. This work is supported in part by University of Kentucky Research Foundation, U.S. Department of Homeland Security, and U.S. National Science Foundation grant IIS-0448185.

# 6 Appendix

We show in this appendix that the multiview rank constraint is equivalent to the absolute difference of ratio images when only two viewpoints and two illumination conditions are present. That is, given image intensities from illumination conditions a and $k_1 a$ from camera A, and $b$ and $k_2 b$ from camera B, we want to show that the observation matrix's second singular value has a minimum at the same disparity as $|k_1 - k_2|$.

The observation matrix can be written as: $\begin{bmatrix} a & b \\ k_1 a & k_2 b \end{bmatrix}$

Its second singular value $s_2$ can be calculated as:

$$s_2 = \frac{1}{2}\left[ 2(a^2 + k_1^2 + k_2^2 b^2 + b^2) - 2\sqrt{\begin{matrix} a^4(k_1^2+1)^2 + b^4(k_2^2+1)^2 + \\ 2a^2b^2(k_1k_2+1)^2 - 2a^2b^2(k_1-k_2)^2 \end{matrix}} \right]^{\frac{1}{2}} \tag{19}$$

Let us define $d = k_1 - k_2$ such that it is positive, reversing the role of $k_1$ and $k_2$ if necessary. Note also that $a$, $b$, $k_1$, $k_2$ are all positive due to physical constraints.

It can be shown that $s_2 = 0$ if and only if $k_1 = k_2$, (given non-zero $a$ and $b$). Similarly, it is obvious that $|k_1 - k_2| = 0$ only when $k_1 = k_2$. It remains to be shown that $s_2$ is related to $d$ by a monotonic relationship, such that an increase in $s_2$ always implies an increase in $d$.

Now if we replace $k_2$ with $k_1 - d$ in equation 19 and take the derivative of $s_2$ with respect to $d$, we have:

$$\begin{aligned}\frac{\partial(s_2)}{\partial(d)} =& \frac{1}{4\sqrt{s_2}}[4(k_1+d)b^2 - \frac{1}{G}(4b^4((k_1+d)^2+1)(k_1+d)+ \\ & 4a^2b^2(k_1(k_1+d)+1)k_1 - 4a^2b^2d],\end{aligned} \tag{20}$$

where

$$G = \sqrt{\begin{aligned}& a^4(k_1^2+1)^2 + b^4((k_1+d)^2+1)^2 + \\ & 2a^2b^2(k_1(k_1+d)+1)^2 - 2a^2b^2d^2\end{aligned}} \tag{21}$$

We need to show that this derivative is always positive, $\frac{\partial(s_2)}{\partial(d)} \geq 0$, which is the same as showing:

$$\begin{aligned}4(k_1+d)b^2 >& \frac{1}{G}(4b^4((k_1+d)^2+1)(k_1+d)+ \\ & 4a^2b^2(k_1(k_1+d)+1)k_1 - 4a^2b^2d)\end{aligned} \tag{22}$$

Taking square of both sides and simplifying results in:

$$\begin{aligned}& 64a^2b^6d^2 + 128a^2b^6k_1^2d^2 + 64a^2b^6k_1^3d + 64a^2b^6k_1d^3 + \\ & 64a^4b^4k_1^3d + 64a^2b^6k_1d + 64a^4b^4k_1d + 64a^4b^4k_1^2d^2 > 0\end{aligned} \tag{23}$$

The above inequality holds true because all variables are positive, and thus the rank constraint is equivalent to using the absolute difference of the ratio images.

# References

[1] D. Bhat and S. Nayar. Stereo in the presence of specular reflection. In *Proceedings of International Conference on Computer Vision (ICCV)*, pages 1086–1092, 1995.

[2] A. Blake. Specular Stereo. In *Proc. 9th Int. Joint Conf. Artificial Intell. (IJCAI)*, volume 2, pages 973–976, 1985.

[3] G. Brelstaff and A. Blake. Detecting Specular Reflections using Lambertian Constraints. In *Proceedings of International Conference on Computer Vision (ICCV)*, pages 297–302, 1988.

[4] B. Carrihill and R. Hummel. Experiments with the intensity ratio depth sensor. *Computer Vision, Graphics, and Image Processing*, 32:337–358, 1985.

[5] B. Curless and M. Levoy. Better optical triangulation through spacetime analysis. In *Proceedings of International Conference on Computer Vision (ICCV)*, pages 987–994, 1995.

[6] J. Davis, D. Nehab, R. Ramamoorthi, and S. Rusinkiewicz. Spacetime stereo: A unifying framework for depth from triangulation. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 27(2):296–302, 2005.

[7] P. E. Debevec and J. Malik. Recovering High Dynamic Range Radiance Maps from Photographs. *Proceedings of ACM Siggraph*, pages 369–378, 1997.

[8] K. Fujii, M.D. Grossberg, and S.K. Nayar. A Projector-Camera System with Real-Time Photometric Adaptation for Dynamic Environments. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 814–821, 2005.

[9] Aaron Hertzmann and Steven M. Seitz. Example-Based Photometric Stereo: Shape Reconstruction with General, Varying BRDFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 27(8):1254–1264, 2005.

[10] J. A. Jalkio, R. C. Kim, and S.K. Case. Three dimensional inspection using multistripe structured light. *Optical Engineering*, 24(6):966–974, 1985.

[11] H. Jin, S. Soatto, and A. Yezzi. Multi-view Stereo beyond Lambert. In *Proceedings of Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 171–178, June 2003.

[12] V. Kolmogorov and R. Zabih. Computing Visual Correspondence with Occlusions Using Graph Cuts. In *Proceedings of International Conference on Computer Vision (ICCV)*, pages 508–515, 2001.

[13] Vladimir Kolmogorov. An implementation of graph-cuts stereo. http://www.cs.cornell.edu/People/vnk/software.html, 2003.

[14] Y. Li, S. Lin, H. Lu, S.B. Kang, and H-Y Shum. Multibaseline Stereo in the Presence of Specular Reflections. In *International Conference on Pattern Recognition*, pages 573–576, 2002.

[15] Stephen Lin, Yuanzhen Li, Sing Bing Kang, Xin Tong, and Heung-Yeung Shum. Diffuse-Specular Separation and Depth Recovery from Image Sequences. In *Proceedings of European Conference on Computer Vision (ECCV)*, pages 210–224, 2002.

[16] S. Magda, T. Zickler, D. Kriegman, and P. Belhumeur. Beyond Lambert: Reconstructing Surfaces with Arbitrary BRDFs. In *Proceedings of International Conference on Computer Vision (ICCV)*, pages 297–302, 2001.

[17] S. Mann and R. Picard. Being Undigital with Digital Cameras:Extending Dynamic Range by Combining Differently Exposed Pictures. In *Proc. of IST s 48th Annual Conference*, pages 442–448, 1995.

[18] T. Mitsunaga and S. K. Nayar. Radiometric Self Calibration. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 380–387, 1999.

[19] T. Miyasaka and K. Araki. Development of real time 3-d measurement system using intensity ratio method. In *Proceedings of Machine Vision and Three Dimensional Imageing System for Inspection and Metrology II, Intelligent System and Advanced Manufacturing*, 2001.

[20] R. Ng, R. Ramamoorthi, and P. Hanrahan. All-frequency shadows using non-linear wavelet lighting approximation. *ACM Transactions on Graphics (SIGGRAPH)*, 22:376–381, 2003.

[21] D. Scharstein and R. Szeliski. www.middlebury.edu/stereo.

[22] D. Scharstein and R. Szeliski. A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms. *International Journal of Computer Vision*, 47(1):7–42, May 2002.

[23] D. Scharstein and R. Szeliski. High-accuracy stereo depth maps using structured light. In *Proceedings of Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 195–202, 2003.

[24] Adrien Treuille, Aaron Hertzmann, and Steven M. Seitz. Example-Based Stereo with General BRDFs. In *Proceedings of European Conference on Computer Vision (ECCV)*, pages 457–469, 2004.

[25] L. Wolff and E. Angelopoulou. Three-dimensional stereo by photometric ratios. *Journal of the Optical Society of America*, 11(11):3069–3078, 1994.

[26] R. Yang, M. Pollefeys, and G. Welch. Dealing with Textureless Regions and Specular Highlights—A Progressive Space Carving Scheme Using a Novel Photo-consistency Measure. In *Proceedings of International Conference on Computer Vision (ICCV)*, pages 576–584, 2003.

[27] Li Zhang, Brian Curless, and Steven M. Seitz. Spacetime Stereo: Shape Recovery for Dynamic Scenes. In *Proceedings of Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 367–374, 2003.

[28] T. Zickler, P. N. Belhumeur, and D. J. Kriegman. Helmholtz stereopsis: exploiting reciprocity for surface reconstruction. In *Proceedings of European Conference on Computer Vision (ECCV)*, pages 869–884, 2002.

[29] T. Zickler, P. N. Belhumeur, and D. J. Kriegman. Toward a Stratification of Helmholtz Stereopsis. In *Proceedings of Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 548–554, 2003.

[30] T. Zickler, J. Ho, D. J. Kriegman, J. Ponce, and P. N. Belhumeur. Binocular Helmholtz Stereopsis. In *Proceedings of International Conference on Computer Vision (ICCV)*, pages 1411–1417, 2003.

[]Liang Wang received his BS degree from the School of Computer Science, Beijing University of Aeronautics and Astronautics in 2004. He is currently a Ph.D student in the Computer Science Department at the University of Kentucky. His research interests lie in computer vision, especially in 3D reconstruction and stereo matching.

[]Ruigang Yang is an Assistant Professor in the Computer Science Department at the University of Kentucky. He received his Ph.D. degree in Computer Science from University of North Carolina at Chapel Hill in 2003. Prior to coming to UNC-Chapel Hill, he earned a M.S. degree in Computer Science from Columbia University in 1998. Dr. Yang's research interests include computer graphics, computer vision, and multimedia. He is a recipient of U.S. NSF CAREER award in 2004, and a member of the IEEE Computer Society and ACM.

[] James E. Davis is currently an Assistant Professor at University of California at Santa Cruz, with research focused primarily on acquiring digital models of the real world for use in computer graphics. He also serves as a technical advisor for VSeeLab, a startup developing video based collaboration tools. After obtaining his Ph.D. from Stanford

University in 2002, he spent two years at Honda Research Institute USA, working on real-time range sensing for humanoid robotics applications. Dr. Davis has received awards for both his research and teaching, including the IEEE ICRA Best Vision Paper Award in 2003, and Excellence in Teaching recognition given to 2% of USSC faculty in 2006.