# BRDF Invariant Stereo using Light Transport Constancy

James E. Davis
*UC Santa Cruz*
*davis@cs.ucsc.edu*

Ruigang Yang
*U. of Kentucky*
*ryang@cs.uky.edu*

Liang Wang
*U. of Kentucky*
*lwangd@cs.uky.edu*

## Abstract

*Nearly all existing methods for stereo reconstruction assume that scene reflectance is Lambertian, and make use of color 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 in several lighting configurations. In addition, we show that this multiview constraint can be used with as few as two cameras and two lighting configurations. Unlikely previous methods for BRDF invariant stereo, Light Transport Constancy does not require precisely configured or calibrated light sources, nor 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.*

## 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 "color 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 simply asserts that the percentage of light reflected by a particular surface patch (the BRDF) 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:

$$I_{Cj}(x_i) = L(x_i)\, R(x_i, \theta_L, \theta_{Cj}) \quad (1)$$

where $I_{Cj}(x_i)$ is the reflected intensity in the direction of $C_j$ from the point $x_i$, $L(x_i)$ is the incident light intensity at point $x_i$, and $R(x_i, \theta_L, \theta_{Cj})$ is the reflectance function or BRDF at point $x_i$, indexed by the vectors in the direction of L and $C_j$.

The traditional Lambertian assumption is that the reflected light is equal in the directions of $C_1$ and $C_2$.

$$R(x_i, \theta_L, \theta_{C1}) = R(x_i, \theta_L, \theta_{C2}) \quad (2)$$

Thus we legitimately have $I_{C1}(x_i) = I_{C2}(x_i)$. However this relation will not in general hold for arbitrary BRDFs.

Light transport constancy assumes that the surface reflectance function, $R(x_i, \theta_L, \theta_{Cj})$, remains constant under variable illumination. If we vary the lighting conditions, so that the incident illumination varies by a factor of $k(x_i)$, then the observed reflected light, $I'_{Cj}(x_i)$, will also vary by a factor of $k(x_i)$.

$$I'_{Cj}(x_i) = k(x_i)\, L(x_i)\, R(x_i, \theta_L, \theta C_j) \quad (3)$$

Note that in general neither the incident light intensity, nor the change in intensity, 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

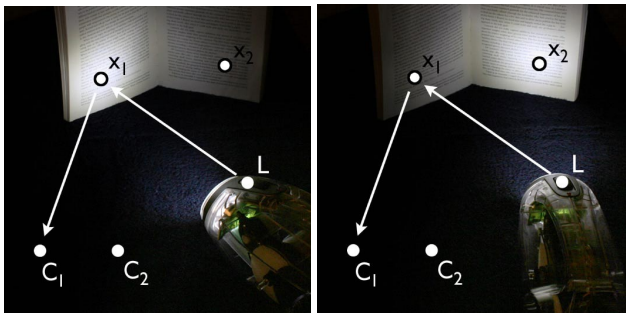


**Figure 1 (left) The reflectance function at $x_1$ determines the percentage of light reflected from light source L towards each of cameras $C_1$ and $C_2$. (right) The spatial position of all components is the same, but the light distribution has been altered. Although the incident intensity at $x_1$ has changed, the percentage of light reflected remains constant.**

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 $k(x_1) \neq k(x_2)$.

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

$$\begin{aligned} I''_{Cj}(x_i) &= \frac{I'_{Cj}(x_i)}{I_{Cj}(x_i)} \\ &= \frac{k(x_i)\cdot L(x_i)\cdot R(x_i,\theta_L,\theta_{Cj})}{L(x_i)\cdot R(x_i,\theta_L,\theta_{Cj})} \\ &= k(x_i) \end{aligned} \tag{4}$$

Note that the observations are invariant to camera viewpoint and $I''_{C1}(x_i) = I''_{C2}(x_i)$ 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. This design is practically easier to implement than existing methods for BRDF invariant stereo, because it requires fewer known or precisely calibrated scene components.

More importantly 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 showing that it is both practical and effective.

## 2. Related Work

All stereo depth recovery methods make an explicit or implicit assumption 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 color constancy invariant, which should be detected and either removed or avoided [1, 2, 3, 11, 12]. 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 [19]. Unfortunately all of these methods limit the range of surface reflectance functions 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 [9]. 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 allows matching of arbitrary BRDFs and uses reflectance function reciprocity as an invariant. That is, $R(x_i, \theta_A, \theta_B) = R(x_i, \theta_B, \theta_A)$ [13, 21, 22, 23]. 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 will require calibrated optics to ensure collocation. The method presented in this paper makes use of a different invariant 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 [8, 17]. 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.

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 [13].

This paper makes use of illumination variation to formulate a constraint on stereo matching. Prior authors have also made use of illumination change as a correspondence aid. For example, spacetime stereo formulates multiview stereo matching in the presence of illumination variation and achieves good results [6, 20]. Image intensity ratios are also a well studied method for recovering depth that utilize illumination variation [4, 14]. Surprisingly none of these prior authors have discussed BRDF invariance. In fact some have gone so far as to argue that image ratios are only applicable

to diffuse surfaces [18]. We believe that some of these prior methods implicitly support arbitrary scene BRDFs, and an important contribution of this work is introducing and characterizing a new invariant which describes precisely when this is true.

## 3. Light Transport Constancy

Light transport constancy can be used to formulate a general constraint on multi-baseline stereo matching regardless of the surface BRDF complexity, provided that sufficient lighting variation and viewpoints are available.

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 reflectance functions also provide a rank constraint.

### 1.1. LTC as a rank constraint

The simplified introduction given in section 1 assumes that the incident lighting field is due to a single light source and varies by a single multiplier $k_i$. This section derives 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 reflected illumination 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, $L_1..L_M$.

$$I_{Cj}(x_i) = L_1(x_1)R(x_i,\theta_{L1},\theta_{Cj})+L_2(x_1)R(x_i,\theta_{L2},\theta_{Cj})+\ldots \quad (5)$$

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 will simplify reflectance function notation by defining subscripted scalar values which correspond to the BRDF for particular pairs of light-camera directions.

$$R_{C1L1} = R(x_i, \theta_{L1}, \theta_{C1}) \quad (6)$$

Equation (5) can be rewritten using the new notation as:

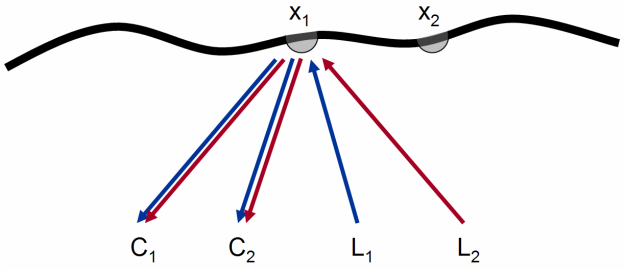


**Figure 2 Light reflected towards camera $C_1$ can be explained as a combination of reflected light from each of $L_1$ and $L_2$.**

$$I_{Cj} = L_1R_{CjL1} + L_2R_{CjL2} + L_3R_{CjL3} + \ldots \quad (7)$$

Light transport constancy makes use of the fact that the reflectance function, $R_{CjLm}$, remains constant when the illumination conditions change. In general the illumination intensity from each light source varies arbitrarily in each illumination condition. We can include the notion of lighting variation by writing $L_{1V1}$ for the illumination from source $L_1$ under the first illumination *variation*, and $I_{C1V1}$ for the observed intensity 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, $V_1..V_N$, to the incident illumination from each light source.

$$\begin{aligned}
&I_{C1V1} = L_{1V1}R_{C1L1} + L_{2V1}R_{C1L2} + L_{3V1}R_{C1L3} + \ldots \\
&I_{C2V1} = L_{1V1}R_{C2L1} + L_{2V1}R_{C2L2} + L_{3V1}R_{C2L3} + \ldots \\
&\ldots \\
&I_{C1V2} = L_{1V2}R_{C1L1} + L_{2V2}R_{C1L2} + L_{3V2}R_{C1L3} + \ldots \\
&I_{C2V2} = L_{1V2}R_{C2L1} + L_{2V2}R_{C2L2} + L_{3V2}R_{C2L3} + \ldots \\
&\ldots
\end{aligned} \quad (8)$$

Note that, light transport constancy says that $R_{CjLm}$ 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, $L_{mVn}$, is related only to the light source and does not depend on either the reflectance function or the camera viewpoint.

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

$$\begin{matrix} \#\text{ variations} \\ \end{matrix}\begin{bmatrix} I_{C1V1} & I_{C2V1} & \cdots \\ I_{C1V2} & I_{C2V2} & \\ I_{C1V3} & I_{C2V3} & \\ \vdots & & \end{bmatrix} = \begin{matrix} \#\text{ variations} \\ \end{matrix}\begin{bmatrix} L_{1V1} & L_{2V1} \\ L_{1V2} & L_{2V2} \\ L_{1V3} & L_{2V3} \\ L_{1V4} & L_{2V4} \\ \vdots & \end{bmatrix} \begin{matrix} \#\text{ lights} \\ \end{matrix}\begin{bmatrix} R_{C1L1} & R_{C2L1} & R_{C3L1} \\ R_{C1L2} & R_{C2L2} & R_{C3L2} \end{bmatrix} \quad (9)$$

(# cameras over the first matrix; # lights over the second matrix; # cameras over the third matrix)

Note 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.

### 1.2. 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 reflected light from a scene point, $x_i$, in the direction of camera $C_j$ can be written as an integral over all incoming light directions.

$$I_{Cj} = \int_{\phi} L(\phi) \cdot R(\phi, \theta_{Cj}) \, \partial\phi \tag{10}$$

where $I_{Cj}$ is the reflected intensity, $L(\phi)$ is the incident light intensity function indexed by incoming angle $\phi$. As before, $R(\phi, \theta_{Cj})$ is the reflectance function at $x_i$.

In order to accommodate more complex variation, the incident illumination field L can be decomposed into a set of basis vectors.

$$L(\phi) = L_1(\phi) + L_2(\phi) + \ldots + L_M(\phi) \tag{11}$$

It is conceptually helpful to think of each basis as a separate light source. Each light has compact angular support, varies its output independently and resides in a disjoint segment of the direction hemisphere ($\phi$). This allows for area light sources, rather than the simple point light sources discussed previously. The only requirement is that from the perspective of a particular scene point, the incident illumination from this area source varies by a uniform factor.

We can now rewrite equation (10) taking into account the lighting bases and indexed by illumination condition.

$$I_{CjVn} = k_{L1Vn} \int_{\phi} L_1(\phi) R(\phi, \theta_{Cj}) \partial\phi + k_{L2Vn} \int_{\phi} L_2(\phi) R(\phi, \theta_{Cj}) \partial\phi + \ldots \tag{12}$$

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, $k_{LmVn}$.

The key thing to notice is that each integral term is constant since 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 set (8) by redefining variables in terms of the new continuous formulation.

$$\begin{aligned} L_{mVn} &= k_{LmVn} \\ R_{CjLm} &= \int_{\phi} L_m(\phi) R(\phi, \theta_{Cj}) \, \partial\phi \end{aligned} \tag{13}$$

Although most readers will find the mental model of disjoint lighting bases helpful, there is no need for the bases to be this restricted. For example, 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, completely arbitrary 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 [15].

## 1.3. 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**.

$$\underset{\#\text{ lights}}{\overset{\#\text{ cameras}}{\begin{bmatrix} \mathbf{R} \end{bmatrix}}} = \underset{\#\text{ lights}}{\overset{\#\text{ BRDF bases}}{\begin{bmatrix} \mathbf{B} \end{bmatrix}}} \underset{\#\text{ BRDF bases}}{\overset{\#\text{ cameras}}{\begin{bmatrix} \mathbf{M} \end{bmatrix}}} \tag{14}$$

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 the illumination is sufficiently "simple", or if the surface reflectance is sufficiently "simple". In this work we allow arbitrary BRDFs so require a limited number of lighting bases.

## 1.4. 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.

Since 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 $\mathfrak{M}$.

$$\mathfrak{M} = \sum_i i \cdot w_i^2 \Big/ \sum_i w_i^2 \tag{15}$$

When a single light source and only two cameras are used, then simply minimizing the second singular value is equivalent to equation (15). 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 provably equivalent to equation (15). Conveniently, it allows existing stereo implementations to be used without modification, simply by matching ratio images rather than raw camera images.

Scharstein and Szeliski have introduced a taxonomy of stereo algorithms which includes matching cost, aggregation, and disparity selection [16]. 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 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. It should be noted 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.

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, two cameras and a single light source. Following these evaluations we provide some analysis of the rank constraint when multiple cameras and lights are present.

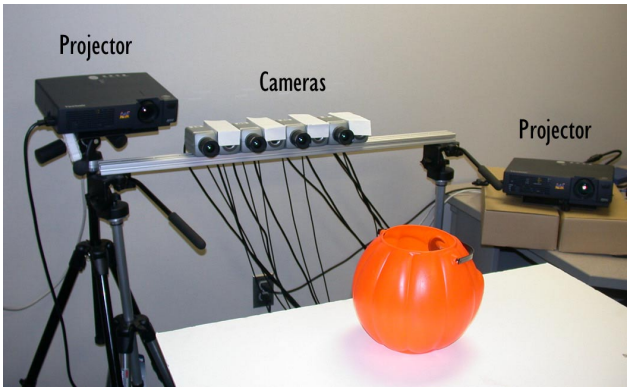


**Figure 3 Our experimental setup with four cameras and two variable light sources.**

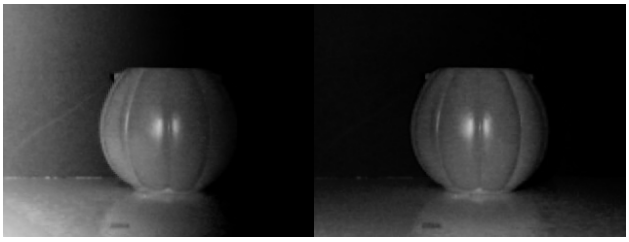

**Figure 4 A plastic pumpkin illuminated by a single light source, under two different lighting conditions.**

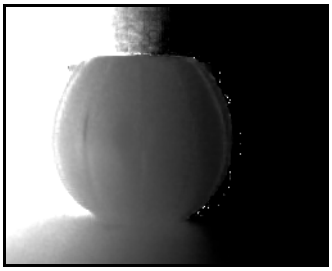

**Figure 5 The ratio of images taken under two lighting conditions. This ratio can be used to compute stereo in place of the raw images, allowing BRDF invariance.**

We first experimented to evaluate our method against traditional stereo. We captured images from each of two cameras under two different lighting configurations. Figure 4 shows the two lighting variations from the viewpoint of one of the cameras. Color constancy 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 is mathematically equivalent to evaluating the rank constraint. The resulting ratio image is shown in Figure 5. Note that neither the specular highlights nor any other view dependant effect are visible in the ratio image.

Standard stereo matching was applied to the stereo pairs arising from both color constancy and light transport constancy using a Sum-of-Absolute-Differences (SAD) metric. Since 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 6 shows the stereo results from each method. The left column is derived from color constancy, while 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 color constancy.

In the third row of Figure 6, we investigate the reason that our method performs well by plotting the matching profile for a single pixel. Note that color constancy has no clear global minimum, while our method has a very clear minimum at the correct disparity. This presumably leads to much better depth estimates.

In order to validate that existing stereo methods can be adapted to handle non-Lambertian objects, we tested the same two sets of stereo pairs with a stereo implementation available on the web [7]. This implementation happens to be based on graph cuts [10], allowing 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 7. Similar to the winner-takes-all example above, the disparity map computed using light transport constancy shows much better results.

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. In order to evaluate whether this is true, we computed disparity using a data set with six lighting variations, as shown in Figure 8. Color constancy was evaluated as the sum-of-absolute-difference 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 color constancy, the result from light transport constancy also improves. We conclude that additional lighting variations will improve the results from either constraint, but that our new invariant performs better on objects such as the pumpkin which exhibit non-Lambertian effects.

Our second test scene is 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 9. Using seven lighting variations, we evaluate color 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 10. Color constancy results in many incorrect disparity estimates, while light transport constancy results in a smooth curve.

To evaluate our method on a more complex scene, we chose a live tree with substantial specular highlights. This scene would be challenging for traditional stereo algorithms both due to the non-Lambertian effects and because there are many depth discontinuities. We use thirty lighting variations to calculate the disparity map shown in Figure 11. 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.

To evaluate the behavior of the rank constraint under multi-view conditions, we computed disparity on the pumpkin scene using four cameras, two light sources and thirty lighting variations. The resulting disparity map can be scene in Figure 12. 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.

When two light sources are used, the rank of the observation 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.

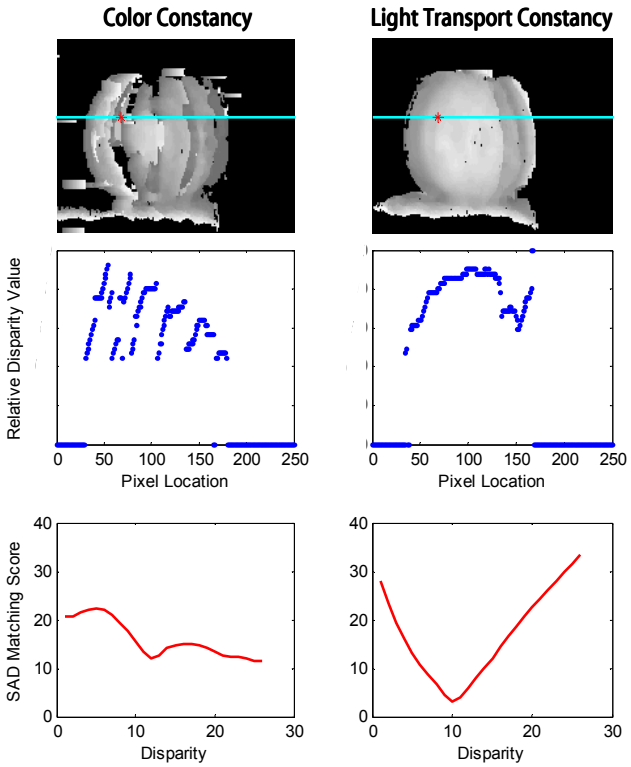


**Figure 6 (row 1) Disparity maps computed by stereo matching using each invariant. (row 2) Scaled disparity estimates along a single scan line. (row 3) Matching profile for the pixel marked with a red cross.**

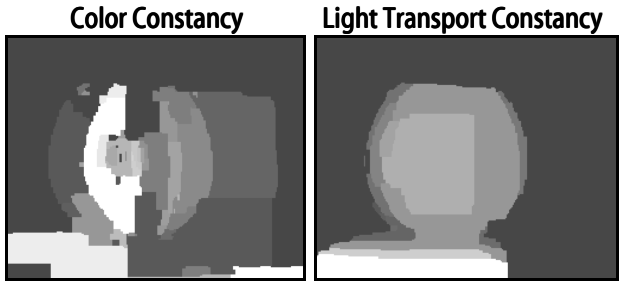


**Figure 7 Disparity maps computed using an unmodified graph-cut stereo algorithm together with our new invariant.**

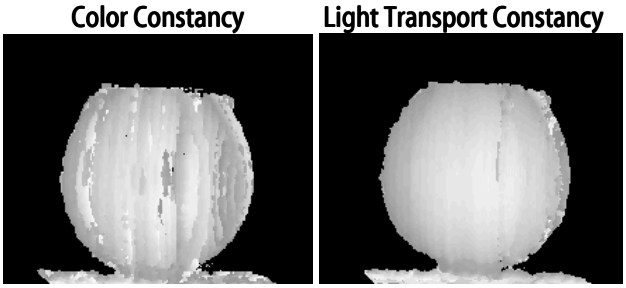


**Figure 8 Disparity maps computed using six illumination variants. Our new invariant performs better than traditional stereo in this case as well.**

To provide some insight into the behavior of our rank constraint, we plotted the 2nd, 3rd, and 4th 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 13, we see that the 2nd singular value has an obvious minimum and that the combined metric $\mathfrak{M}$ is minimized at this same disparity. However, in the case of the scene point in the bottom plot, $\mathfrak{M}$ is minimized at the same disparity as the 3rd singular value. Although the 4th singular value is not precisely zero as would be expected in as ideal environment without noise, we can see that $\mathfrak{M}$ has an easily locatable global minimum which confirms that our approximation of "minimum rank" is performing as expected.

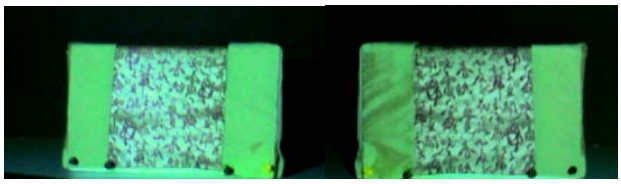

**Figure 9 Silk cloth from two different viewpoints. Note the non-Lambertian reflectance.**

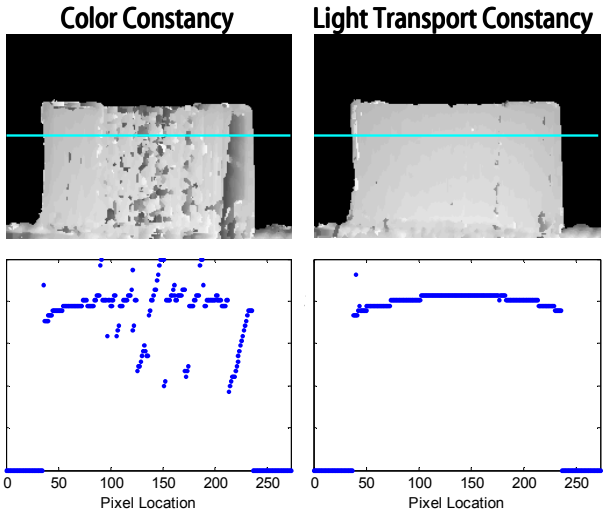


**Figure 10 (top) Disparity maps computed using color constancy and light transport constancy (LTC). (bottom) Scaled disparity values along a single scan-line. Note how much more robustly LTC estimates depth.**

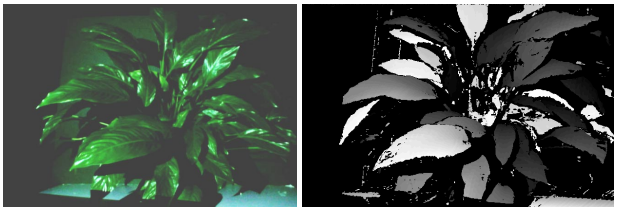

**Figure 11 (left) Tree with non-Lambertian reflectance properties and many depth discontinuities. (right) Disparity map computed from thirty lighting variations.**

## 5. Conclusions

*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, we do not require that light sources be precisely calibrated, nor do we require known calibration objects in the scene.

The rank constraint implied by light transport constancy can easily be employed as a replacement to color 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 color constancy on a variety of scenes.

There are a few aspects of our work which 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 [13, 17], and it is interesting to wonder if this is a fundamental requirement. In addition, 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 actual number of viewpoints required, for a wide class of naturally occurring scenes and lighting.

## Acknowledgements

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 and NSF award IIS-0448185.

## References

[1] D. Bhat and S. Nayar, "Stereo in the presence of specular reflections," presented at Int. Conf. on Computer Vision (ICCV), 1995.

[2] A. Blake, "Specular Stereo," presented at Int. J. Conf. on Artificial Intelligence, 1985.

[3] G. Brelstaff and A. Blake, "Detecting specular reflections using Lambertian constraints," presented at Int. Conf. on Computer Vision, 1988.

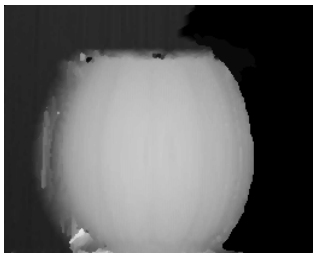

**Figure 12 Disparity map for the pumpkin calculated from multiple cameras and multiple light sources.**

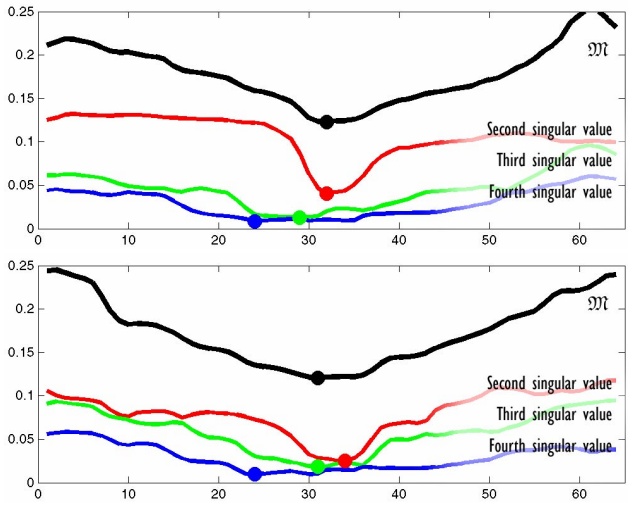


**Figure 13 Normalized singular values for two different scene points. Dots indicate the minimum on each curve. The metric $\mathfrak{M}$ has been scaled to fit on the same graph together with the singular values. Note that $\mathfrak{M}$ is minimized together with a different singular value in each case.**


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

[5] B. Curless and M. Levoy, "Better optical triangulation through spacetime analysis," presented at IEEE International Conference on Computer Vision (ICCV95), 1995.

[6] J. Davis, D. Nehab, R. Ramamoorthi, and S. Rusinkiewicz, "Spacetime Stereo : A Unifying Framework for Depth from Triangulation," *Trans. on Pattern Analysis and Machine Intelligence (PAMI)*, vol. 27, 2005.

[7] http://www.cs.cornell.edu/People/vnk/software.html

[8] A. Hertzmann and S. M. Seitz, "Shape and materials by example: a photometric stereo approach," *IEEE Computer Vision and Pattern Recognition (CVPR)*, 2003.

[9] H. Jin, S. Soatto, and A. J. Yezzi, "Multi-view stereo beyond Lambert," *Computer Vision and Pattern Recognition (CVPR)*, pp. 171-8, 2003.

[10] V. Kolmogorov, R. Zabih, and S. Gortler, "Graphcut: Multi-camera Scene Reconstruction via Graph Cuts," presented at European Conf. on Computer Vision, 2002.

[11] Y. Li, S. Lin, H. Lu, S. Kang, and H.-Y. Shum, "Multibaseline Stereo in the Presence of Specular Reflections," presented at Int. Conf on Pattern Recognition, 2002.

[12] S. Lin, Y. Li, S. Kang, X. Tong, and H.-Y. Shum, "Diffuse-Specular Seperation and Depth Recovery from Image Sequences," presented at European Conference on Computer Vision (ECCV), 2002.

[13] S. Magda, D. J. Kriegman, T. Zickler, and P. N. Belhumeur, "Beyond Lambert: Reconstructing surfaces with arbitrary BRDFs," *Proceedings of the IEEE International Conference on Computer Vision*, vol. 2, pp. 391, 2001.

[14] T. Miyasaka and K. Araki, "Development of Real Time 3-D Measurement System Using Intensity Ratio Method," presented at Machine Vision and Three Dimensional Imageing System for Inspection and Metrology II, Intelligent System and Advanced Manufacturing, 2001.

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

[16] D. Scharstein and R. Szeliski, "A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms," *International Journal of Computer Vision*, vol. 47, pp. 7-42, 2002.

[17] A. Treuille, A. Hertzmann, and S. M. Seitz, "Example-based stereo with general BRDFs," presented at European Conference on Computer Vision (ECCV), 2004.

[18] L. Wolff and E. Angelopoulou, "Three-dimensional stereo by photometric ratios," *J. Opt. Soc. Am. A*, vol. 11, pp. 3069-3078, 1994.

[19] 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," *Proceedings of the IEEE International Conference on Computer Vision*, vol. 1, pp. 576, 2003.

[20] L. Zhang, B. Curless, and S. M. Seitz, "Spacetime stereo: Shape recovery for dynamic scenes," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. II/367-II/374, 2003.

[21] T. Zickler, P. N. Belhumeur, and D. J. Kriegman, "Toward a stratification of Helmholtz stereopsis," presented at IEEE Computer Vision and Pattern Recognition, 2003.

[22] T. E. Zickler, P. N. Belhumeur, and D. J. Kriegman, "Helmholtz stereopsis: exploiting reciprocity for surface reconstruction," *International Journal of Computer Vision*, vol. 49, pp. 215, 2002.

[23] T. E. Zickler, J. Ho, D. J. Kriegman, J. Ponce, and P. N. Belhumeur, "Binocular Helmholtz stereopsis," *Proceedings of the IEEE Int. Conf. on Computer Vision*, vol. 2, pp. 1411, 2003.