社団法人　情報処理学会　研究報告
IPSJ SIG Technical Report

2006－CVIM－153（41）
2006／3／17

Revisiting Geotensity in the Presence of Specular Reflection

Atsuto MAKI
Graduate School of Informatics, Kyoto University　Kyoto 606-8501 Japan
E-mail: maki@i.kyoto-u.ac.jp

Abstract The geotensity constraint governs the relationship between four or more images of a moving object in spite of the illumination variance due to object motion. It applies to 3D surface reconstruction as well as to generate illumination image basis for Lambertian surface under a distant point light source. This paper presents a simple method for extending the availability of the constraint to the case that specularities are also present. The key idea is to utilise the fact that highlights shift on the surface due to object motion. That is, we employ five or more images as inputs, and interchangeably utilise a certain intensity subset consisting of four projected intensities which is the least influenced by the specular component. Further, we show an algorithm to directly linearise the image basis by factoring out the specular components as residuals with respect to a linear combination of the intensity subset.

Key words 3D surface reconstruction, illumination image basis, geotensity, specular reflection

1 Introduction

Given a set of images, in each of which an object is viewed from a different direction, the fundamental issue in extracting 3D information of the object out of 2D images is to match corresponding points in those images so that these points are the projections of an identical point on the surface of it. For the point correspondence, typically exploited is the constraint that the corresponding parts of the images have equivalent image intensities, regarding the variation in illumination as noise. It has been successfully applied to stereo (see for example [5, 7]) where two images are taken simultaneously as the lighting of the object is identical in each image. However, when we consider replacing the stereo camera with a single camera observing an object in motion, unfortunately the constraint is nearly always invalid as non-uniform lighting causes the intensity at a specific location on the surface of an object to change as the object moves. Among the few efforts for this issue, whereas *photometric motion* [17] treated the illumination variance due to object motion in terms of optical flow, *geotensity* constraint [11] has been recently derived to overcome the problem with respect to camera geometry, and to replace the constant intensity constraint. Based on the notion of linear intensity subspaces [18], the geotensity constraint governs the relationship between four (or more) images of a moving object, and it can be computed and applied automatically to the task of 3D surface reconstruction.

The algorithm for surface reconstruction using geotensity constraint proceeds basically in two stages. The first stage is to derive the parameters of the geotensity constraint by analysing coordinates and image intensities of some sample points on the object in motion. That is, computing structure from motion obtains the geometric parameters of the situation, whereas computing the linear image subspace obtains the lighting parameters of the situation. By combining both sets of parameters we arrive at the geotensity constraint. Using the same set of images, the second stage is to take each pixel in an arbitrary reference image in turn and search for the depth along the ray from the optical centre of the camera passing through the pixel. The depth is evaluated by measuring the agreement of the entire set of projected intensities of a point on the object surface with the geotensity constraint.

Although the availability of the constraint was limited in principle to Lambertian surface as is the case with some other sophisticated approaches [19, 21], the thrust of this paper is to extend it to the situation that the object surface partly takes on specular reflection. In the case of stereo, in the presence of specular reflection, it has been shown possible to determine trinocular configurations such that at least one stereo pair can provide correct depth estimate at each scene point visible to all cameras [3]. In contrast, given a single static camera, we propose to employ five or more images as inputs, and interchangeably utilise a certain *intensity subset* consisting of four projected image intensities that is the least influenced by the specular component of surface reflection. The strategy is motivated by the fact that the specularities shift on the surface thanks to the object motion.

The other objective of this paper is to construct an illumination image basis from images of an object in motion, as introduced in [15] but even when specularities are present. Given the correspondence

between input images is correctly computed despite the specularities, we propose to linearise the pixel intensities simultaneously as we re-align the input images because otherwise non-linear artifact will naturally appear in the resulting basis as long as specular reflections are present in the input images. In practice, we present a simple algorithm that enables us to factor out the specular components as residuals with respect to a linear combination of an intensity subset so that we can generate a linearised image basis directly as the output.

2 Geotensity

The term geotensity constraint[1] accounts for a constraint between four or more images of an object from different views under static lighting conditions. As the background for correspondence search we first consider some issues respecting geometry and image intensity that form the basic of the constraint. Although a thorough description of the geotensity constraint can be found in [11], we briefly review the constraint while reformulating it so as to prepare for the extension for the case with specularities which we will introduce in Section 3.

2.1 Preliminaries

We initially need to find some number of corresponding sample points by an independent mechanism as seen for example in [1, 22]. Given point correspondence for some sample points, we can derive a constraint on geometry by the coordinates, and also a photometric constraint by observing the intensities on these points.

Solving for geometry. For simplicity, we will concern ourselves with the *affine* and *scaled-orthographic* camera models for projection. Consider the i^{th} world point $\mathbf{X}_i = (X_i, Y_i, Z_i)^\top$ on the surface of an object projected to image point $\mathbf{x}_i(j) = (x_i(j), y_i(j))^\top$ in the j^{th} frame. The affine camera model defines this projection to be $\mathbf{x}_i(j) = \mathbf{M}(j)\mathbf{X}_i + \mathbf{t}(j)$, where $\mathbf{M}(j)$, an arbitrary 2×3 matrix, and $\mathbf{t}(j)$, an arbitrary 2 vector, encode the motion parameters of the object. The solution to the structure from motion problem using singular value decomposition is well known for this case; given at least four point trajectories $\mathbf{x}_i(j)$ observed through at least two frames the set $\mathbf{M}(j)$, $\mathbf{X}_i$ and $\mathbf{t}(j)$ can be uniquely recovered up to an arbitrary affine ambiguity [20]. The result is affine structure.

Given the solution to structure from motion using the affine camera model, the Euclidean structure and motion parameters fitting the weak perspective camera model can be recovered. A result of choosing the first frame to be canonical [8] is that the structure vectors have the form, $\mathbf{X}_i = (\mathbf{x}_i^\top(1), Z)^\top$, and we can derive the relationship, which effectively describes the *epipolar constraint* between two images:

$$\mathbf{x}_i(j) = \mathbf{M}(j) \begin{pmatrix} \mathbf{x}_i(1) \\ Z \end{pmatrix} + \mathbf{t}(j) \ . \tag{1}$$

Solving for image intensity. Assuming a static camera and a single point light source, we consider the intensity $I_i(j)$ of the i^{th} point on the surface of a moving object projected into the j^{th} image. For Lambertian surface, since the intensity can be equivalently represented as if it were generated on the surface of a static object under inversely moving point light source, we can then express $I_i(j)$ in terms of the image formation equation process so that

$$I_i(j) = \max(\mathbf{b}_i{}^\top \mathbf{s}(j), 0) \ . \tag{2}$$

The 3-vector $\mathbf{b}_i$ is defined at the first frame to be the product of the albedo with the inward facing unit normal for the i^{th} point whereas the 3-vector $\mathbf{s}(j)$ is the product of the strength of the light source with the unit vector for its direction. Note that $\mathbf{s}(j) = \mathbf{R}^\top(j)\ \mathbf{s}(1)$, where the 3×3 matrix, $\mathbf{R}(j)$, is the rotation of the object from the first frame, which is taken to be the reference, to the j^{th} frame. Multiplication of $\mathbf{R}^\top(j)$ represents virtually inverse rotation of the light source. The rotation matrix is directly computed from the 2×3 matrix $\mathbf{M}(j)$ that is given above by solving for the structure from motion problem. Although the maximum operator zeroes negative components [9] which correspond to the shadowed surface points, we assume that there are no shadows in the subsequent formulations.

Given the correspondence for n feature points through m images, we record the corresponding pixel intensities, $I_i(j)$, in an $n \times m$ matrix as $\mathbf{I} = (I_i(j))$, which we call the *illumination matrix*. Then, we can form the matrix equation, $\mathbf{I} = \mathbf{BS}$, where $\mathbf{B}$ is an $n \times 3$ matrix containing the rows $\mathbf{b}_i^\top$, and $\mathbf{S}$ is a $3 \times m$ matrix containing the columns $\mathbf{s}(j)$. Then, the familiar form for solution by singular value decomposition to obtain a rank 3 approximation to the matrix $\mathbf{I}$ is such that

$$\mathbf{I} = \breve{\mathbf{B}}\breve{\mathbf{S}} = (\breve{\mathbf{B}}\mathbf{A}^{-1})(\mathbf{A}\breve{\mathbf{S}}) \ . \tag{3}$$

As is well known, the solution is unique up to an arbitrary invertible 3×3 transformation $\mathbf{A}$ which transforms $\breve{\mathbf{S}}$ into $\mathbf{S}$, or $\breve{\mathbf{s}}(j)$ into $\mathbf{s}(j)$, for each column by

$$\mathbf{S} = \mathbf{A}\breve{\mathbf{S}} \ , \quad \mathbf{s}(j) = \mathbf{A}\breve{\mathbf{s}}(j) \ , \tag{4}$$

where $\breve{\mathbf{s}}(j)$ denotes each column of $\breve{\mathbf{S}}$. In the following we call matrix $\mathbf{S}$ (and $\breve{\mathbf{S}}$ as well) *light source matrix* for convenience of explanation.

[1] *Geotensity* stands for "geometrically corresponding pixel intensity."

2.2 The geotensity constraint

The basic conditions for applying the geotensity constraint to surface reconstruction are as follows: (i) The scene consists of a single moving object that is convex (therefore no self-shadowing). (ii) There is a single distant light source[2]. (iii) The object has Lambertian surface properties while the surface may or may not be textured. However, the condition (iii) will be relaxed in Section 3.

Evaluating the set of intensities for correspondence. At each pixel, $\mathbf{x}$, in the first image, to search for the depth Z we can recall equation 1 for the geometric constraint imposed on a sequence of images so that

$$I(j;\mathbf{x},Z) = I[\mathbf{x}(j)] = I[\mathbf{M}(j)\begin{pmatrix}\mathbf{x}(1)\\ Z\end{pmatrix} + \mathbf{t}(j)] \ . \tag{5}$$

$I(j;\mathbf{x},Z)(j = 1,...,m; m \geq 4)$ indicates the set of image intensities in the j^{th} frame at the coordinates determined by $\mathbf{x}$ in the first image, guess of depth Z, and the motion parameters $\mathbf{M}(j)$ and $\mathbf{t}(j)$. The task is now to evaluate the set of intensities $I(j;\mathbf{x},Z)$. If full Euclidean lighting conditions have been recovered in advance so that $\mathbf{s}(j)$ is known without ambiguity, we define $\hat{\mathbf{b}}^\top$ as

$$\hat{\mathbf{b}}^\top = [I(1),...,I(m)]\,\mathbf{S}^\top(\mathbf{S}\mathbf{S}^\top)^{-1} \tag{6}$$

where $\mathbf{S}$ is a $3\times m$ light source matrix containing the columns $\mathbf{s}(j)(j = 1,...,m)$. For a single light source with all images taken with the light source in the so-called *bright cell* [3], the estimated values of the intensities would then be

$$\hat{I}(j;\mathbf{x},Z) = \hat{\mathbf{b}}^\top\mathbf{s}(j) \ . \tag{7}$$

It should be noted that exactly the same estimation of $\hat{I}(j;\mathbf{x},Z)$ is available also in the case that the light source direction is determined only up to the ambiguity. This is easily confirmed by substituting equation 4 to equation 6, and then to equation 7, where matrix $\mathbf{A}$ turns out to be cancelled. Thus, the set of equations 6 and 7 can be represented equivalently as

$$\begin{aligned}\hat{I}(j;\mathbf{x},Z) &= \hat{\mathbf{b}}^\top\breve{\mathbf{s}}(j) \ , \\ \hat{\mathbf{b}}^\top &= [I(1),...,I(m)]\,\breve{\mathbf{S}}^\top(\breve{\mathbf{S}}\breve{\mathbf{S}}^\top)^{-1} \ .\end{aligned} \tag{8}$$

Estimating $\hat{I}(j;\mathbf{x},Z)$ by equation 8, we can define the error function to evaluate the set of intensities $I(j;\mathbf{x},Z)$ as

$$E(\mathbf{x},Z) = \sum_{j=1}^{m}(I(j;\mathbf{x},Z) - \hat{I}(j;\mathbf{x},Z))^2 \ . \tag{9}$$

[2]However, an extension to the case of a few distant light sources has been proposed in [12].

[3]The cell of light source directions that illuminate all points on the object [2].

Computing the correspondence. At each pixel, $\mathbf{x}$, in the first reference image we measure the error, E, in the geotensity constraint at regular small intervals of depth, Z. When the depth is correct we expect the error to approach zero and when it is incorrect we expect the error to be large. The geotensity constraint can be therefore stated simply as $E(\mathbf{x},Z) = 0$. It is clear that as the depth parameter is varied the location of the corresponding points in each image will trace out the corresponding epipolar line in each image. We then choose such depth Z that minimises the error $E(\mathbf{x},Z)$ as the depth estimate.

3 Dealing with specularities

The technique of evaluating the set of projected intensities discussed in the previous section is deficient in dealing with specular reflection. This is because the notion of linear image basis which plays the essential role is only valid for Lambertian surface. In order to cope with specular reflection within the same framework, we propose to employ five or more images as inputs, and interchangeably utilise a certain subset consisting of four intensities, which are required at the minimum for the geotensity constraint to be applied.

Since we consider a moving object as the target, each surface point may or may not have specularity depending on the object's pose as a specular lobe in general has some certain limited directional range. Evidence suggests that away from specularities real surfaces do not deviate from the Lambertian model by more than a factor of three or so [4]. We may thus expect that a certain subset of projected intensities exists that satisfies the geotensity constraint to a reasonable extent. At least it is very unlikely that some specular reflection is added to an identical surface point throughout all the input image frames (except the case the object consists of a particular kind of material such as metal). The problem to cope with the situation is then attributed to that of finding an *intensity subset* which is the least influenced by specular component of surface reflection.

3.1 Search with intensity subset

As the simplest case, let us consider that five images are given as input so that $m = 5$, and assume that specular reflection occurs mostly in either of them due to object motion. Given an $n \times m$ illumination matrix as $\mathbf{I} = (I_i(j))$, we define m different $n \times (m-1)$ *illumination submatrices*, $\mathbf{I}_{\bar{k}}(k = 1,...,m)$, by skipping the k^{th} column, $\mathbf{I}(k)$, such as

$$\mathbf{I}_{\bar{k}} = [\mathbf{I}(1) \ ... \ \mathbf{I}(k-1) \ \mathbf{I}(k+1) \ ... \ \mathbf{I}(m)] \ . \tag{10}$$

We then recall equation 3 and compute a rank 3 approximation to each $\mathbf{I}_{\bar{k}}$,

$$\mathbf{I}_{\bar{k}} = \breve{\mathbf{B}}_{\bar{k}}\breve{\mathbf{S}}_{\bar{k}} = (\breve{\mathbf{B}}_{\bar{k}}\mathbf{A}^{-1})(\mathbf{A}\breve{\mathbf{S}}_{\bar{k}}) \ , \tag{11}$$

where $\breve{\mathbf{S}}_{\bar{k}}(k = 1, ..., m)$ in this case is a $3 \times (m-1)$ matrix containing the columns $\breve{\mathbf{s}}(j)(j = 1, ..., m-1)$. Importantly, for computing each $\breve{\mathbf{S}}_{\bar{k}}$ we employ RANSAC, a robust random sampling and consensus technique, to ensure that artifacts which are caused by specularities (or some intensities not fulfilling the assumed conditions, e.g. self-shadowing) do not distort the correct solution.

At the correct depth in correspondence search, an intensity subset, $I(j; \mathbf{x}, Z)(j = 1, ..., m-1)$, excluding the k^{th} element should be properly validated with corresponding light source matrix $\breve{\mathbf{S}}_{\bar{k}}$ if specular reflection occurs in the k^{th} frame. Hence, with each $\breve{\mathbf{S}}_{\bar{k}}$, we can exploit equation 8 to estimate $\hat{I}(j; \mathbf{x}, Z)$ $(j = 1, ..., m-1)$ and then evaluate the subset of intensities by

$$E_{\bar{k}}(\mathbf{x}, Z) = \sum_{j=1, j\neq k}^{m} (I(j; \mathbf{x}, Z) - \hat{I}(j; \mathbf{x}, Z))^2 \ . \tag{12}$$

Namely, we obtain m different $E_{\bar{k}}(\mathbf{x}, Z)(k = 1, ..., m)$ due to m different candidates of intensity subset (together with corresponding $\breve{\mathbf{S}}_{\bar{k}}$) and one of them must be appropriate according to the above assumption.

In order to search for the depth Z, we need to judge the correct error among those defined by equation 12. Although it is not possible to know which subset of projected intensities is the least influenced by specularity in advance, since the error $E_{\bar{k}}(\mathbf{x}, Z)$ generally becomes larger with specularity, it is sensible to choose the smallest one so that

$$E(\mathbf{x}, Z) = \min_{k} E_{\bar{k}}(\mathbf{x}, Z) \ . \tag{13}$$

Just as in the case without specular reflection, we expect the error to approach zero when the depth is correct, and choose such depth Z that minimises the error $E(\mathbf{x}, Z)$ as the depth estimate.

3.2 Algorithm summary

Our algorithm for estimating the depth in the presence of specularities can be summarised as following:

1° Decompose each illumination submatrix $\mathbf{I}_{\bar{k}}$ using SVD and yield $\breve{\mathbf{S}}_{\bar{k}}$.
2° At point $\mathbf{x}$, measure $I(j; \mathbf{x}, Z)$ by equation 5 for a guess of depth Z.
3° Estimate $\hat{I}(j; \mathbf{x}, Z)$ by equation 8, for all $\breve{\mathbf{S}}_{\bar{k}}$ using $I(j; \mathbf{x}, Z)(j \neq k)$.
4° Compute $E_{\bar{k}}(\mathbf{x}, Z)$ and then $E(\mathbf{x}, Z)$ by equations 12 and 13.
5° Choose such depth Z that minimises $E(\mathbf{x}, Z)$ as the depth estimate.

Alternatively, we may simplify step 3° by estimating $\hat{I}(j; \mathbf{x}, Z)$ just in one way, rather than examining it for all the possible candidates of $\breve{\mathbf{S}}_{\bar{k}}$. That is, we simply determine k such that the k^{th} element has the highest value among the set of projected intensities, and compute $E_{\bar{k}}(\mathbf{x}, Z)$ accordingly. It is based on the observation that setting the element of highest intensity aside must be the choice for the solution to be the least influenced by specular reflection at the correct depth. As we can then skip estimating different $E_{\bar{k}}(\mathbf{x}, Z)$ in step 4°, it is also less expensive in the computational viewpoint.

Finally, the basic principle in the proposition applies also to a larger number of input images, $m > 5$, and we then need choose the size of the intensity subset. Although a large size implies higher tolerance against noise as long as specular reflection appears just in one frame as assumed above, a crude observation tells that it is rather optimal to keep the size minimum, i.e. four elements, to allow a higher probability of including no specularities. However, the trade-off analysis is also motion dependent and it is beyond the scope of this paper.

4 Linearising image basis

Based on the correspondence between input images of an object in motion, which is computed pixelwise, we can generate images of the object as if they were captured in an identical pose as seen in the reference frame but under varying lighting. As has been described in [14, 15], this is done essentially by re-aligning the pixels in the input images. We can either employ a set of three images from the generated images directly as an image basis, or apply the PCA to the re-aligned images and form a basis by the three eigenvectors corresponding to the highest eigenvalues. In either case, however, non-linear artifact will naturally appear in the resulting basis as long as specular reflections are present in the input images even though the correspondence is correctly computed despite the specularities. Since the linear functionality of the basis will then deteriorate, we wish to replace the specularities with some values that linearly agree to each other, i.e., linearise the appearance of image basis.

Let us consider the case that m images are given as input and that for simplicity we should typically replace an intensity, $I(h)$, that is most likely to be specular among a corresponding intensity set, $I(j)(j = 1, ..., m)$. For each corresponding set of pixels in a set of re-aligned images we simply determine h such that the h^{th} element has the highest value and accordingly generate an intensity subset,

$$\begin{aligned} & I(j)(j = 1, ..., m-1) \\ = \ & [I(1) \ ... \ I(h-1) \ I(h+1) \ ... \ I(m)] \ . \end{aligned} \tag{14}$$

Since the intensity subset, $I(j)(j = 1, ..., m-1)$, excluding the h^{th} element should be consistent with corresponding $3 \times (m-1)$ light source matrix, $\breve{\mathbf{S}}_{\bar{h}}$, containing the columns $\breve{\mathbf{s}}(j)(j = 1, ..., m-1)$ (excluding $\breve{\mathbf{s}}(h)$), we can estimate a linearised value of $I(h)$ by

$$\hat{I}(h) = \hat{\mathbf{b}}_{\bar{h}}^{\top} \breve{\mathbf{s}}(h) \ , \tag{15}$$

where we compute $\hat{\mathbf{b}}_{\bar{h}}^{\top}$ as

$$\hat{\mathbf{b}}_{\bar{h}}^{\top} = [I(1), ..., I(m-1)] \, \breve{\mathbf{S}}_{\bar{h}}^{\top} (\breve{\mathbf{S}}_{\bar{h}} \breve{\mathbf{S}}_{\bar{h}}^{\top})^{-1} \ , \tag{16}$$

in an analogous way to equation 6 but without using $I(h)$. We can then factor out the specular component of $I(h)$ as a residual with respect to the linear estimates by

$$I(h) - \hat{I}(h) \ . \tag{17}$$

The residual should be zero when no specularity is present, given the re-alignment is accurate. If some residual is observed, we can linearise the corresponding set of pixels by directly replacing $I(h)$ with $\hat{I}(h)$. This operation at each pixel in a set of re-aligned images allows us to linearise the entire image basis.

5 Experiments

We illustrate the performance of the proposed scheme using an object whose surface is moderately glossy. Figure 1 shows three input images among five that were used. They were taken while the object pose varies under a point light source. As can be observed, specularities appear on different part of the surface according to the object pose relative to the light source. Since the light source is placed beside the camera, in each image, parts of the surface where the orientation is close to the direction to the camera tend to take on specular reflection.

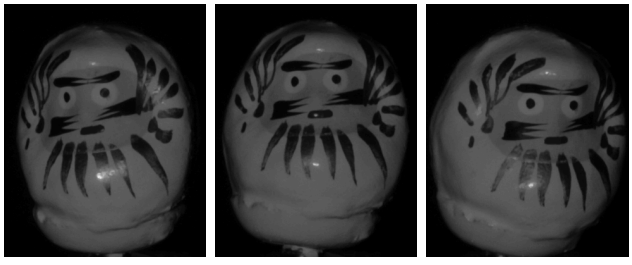

Figure 1: Example input images of a dharma doll. Captured in different poses under a point light source placed beside the camera (Nikon D70). Image size: 320×400 pixels. It is observed that the specularities shift on the surface thanks to the object motion.

Figure 2 shows the resulting depth maps computed for the midmost input image in Figure 1. For comparison, we carried out the depth search by three different algorithms within the common framework, i.e., differencing, the geotensity constraint, and the proposed – the geotensity constraint using the intensity subset. In each case we used a 15×15 template for the search to suppress the error arising from image noise. It turned out that the proposed method allows an improved estimation whereas the other two tend to fail more severely in the region where the surface is either with little texture, or with specularity such as on the top of the *belly*, although the erroneous estimates in the bottom (conceivably due to complicated inter-reflection) are of an issue for all of them.

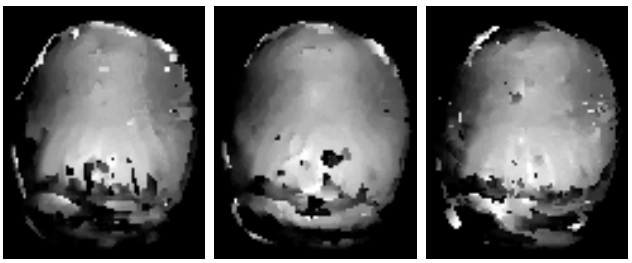

Figure 2: Three depth maps estimated referring the midmost image in Figure 2 (the lighter, the closer). From the left: differencing, the geotensity, and the geotensity with *intensity subset*. The profile of the dharma doll is also shown for reference.

Figure 3 shows re-aligned images of dharma doll (top), each of which is also separated to specular components and linearised images (bottom) after removing the residuals. Since the motion of the object is in such a way that it spans the 3D space, the residuals were removed in different parts of the surface in different images. Comparing the top and bottom rows of the figure, in particular, it can be noticed that the strong specularity that is present due to the round shape of the object at locally closest points (viewed from the camera) has been removed in the linearised images.

As a whole, however, it should be stressed that we have presented early results only demonstrating the principle of the theoretical proposition. One could for instance generate image bases by applying the PCA to the re-aligned images and compare the cases with and without the linearisation.

6 Summary and discussion

For the problem of 3D object surface reconstruction by a single static camera, we have proposed to relax the condition for geotensity constraint to be applicable to the case that specular reflection is added to the object surface. The key idea is to utilise such a certain subset of intensities that is the least influenced by specular reflection. To our knowledge it is one of the few trials of developing correspondence scheme in the presence of specularity as well as under varying image intensity due to object motion.

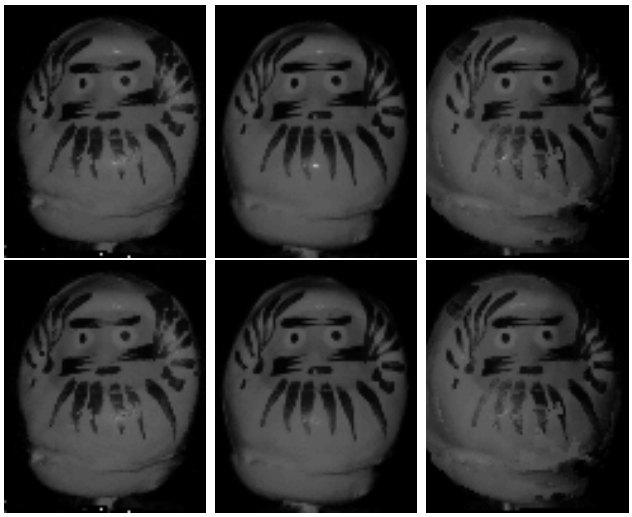

Figure 3: Top: Re-aligned images generated from the dharma images. Bottom: Re-aligned and then linearised images after removing the specular components.

On the basis of the proposed scheme for correspondence, we have also presented an algorithm to directly linearise the image basis, by factoring out the specular components as residuals with respect to a linear combination of an intensity subset.

Although we have illustrated our scheme only for the case of using five input images, the principle also applies to the case that a larger number of images are available. However, we further need some investigations as to how the size of the intensity subset in such cases can be optimally determined, which is one of the issues for future work. Another related extension will be to deal with shadows as well as specularities as have been suggested in other different contexts such as photometric linearisation [13] or quotient image [16]. It will be also interesting to further investigate the issues of estimating illumination [10] and of recovering reflectance properties of the object surface [6] in this framework.

Acknowledgments. This work is supported by MEXT, Japan, under a Grant-in-Aid for Scientific Research (No.16680010) and in part by national project on "Development of high fidelity digitisation software for large-scale and intangible cultural assets".

References

[1] P.A. Beardsley, P. Torr, and A. Zisserman. 3D model acquisition from extended image sequences. In *4rd ECCV*, pages 683–695, 1996.

[2] P.N. Belhumeur and D.J. Kriegman. What is the set of images of an object under all possible illumination conditions? *IJCV*, 28:3:245–260, 1998.

[3] D. N. Bhat and S. Nayar. Stereo in the presence of specular reflection. In *5th ICCV*, pages 1086–1092, 1995.

[4] G. Brelstaff and A. Blake. Detecting specular reflections using lambertian constraints. In *2nd ICCV*, pages 297–302, 1988.

[5] F. Devernay and O. Faugeras. Computing differential properties of 3D shapes from stereoscopic images without 3D models. In *CVPR*, pages 208–213, 1994.

[6] F. Du, T. Okabe, Y. Sato, and A. Sugimoto. Reflectance estimation from motion under complex illumination. In *17th ICPR*, pages 218–222, 2004.

[7] P. Fua. Object-centered surface reconstruction: combining multi-image stereo and shading. *IJCV*, 16:35–56, 1995.

[8] R.I. Hartley. Euclidean reconstruction from uncalibrated views. In *Proc. Applications of Invariance in computer vision*, pages 238–256. Springer-Verlag, 1994.

[9] B.K.P. Horn. *Robot Vision*. The MIT Press, 1992.

[10] A. Maki. Estimation of illuminant direction and surface reconstruction by geotensity constraint. *Pattern Recognition Letters*, 21:13-14:1115–1123, 2000.

[11] A. Maki, M. Watanabe, and C.S. Wiles. Geotensity: Combining motion and lighting for 3d surface reconstruction. *IJCV*, 48:2:75–90, 2002.

[12] A. Maki and C. Wiles. Geotensity constraint for 3d surface reconstruction under multiple light sources. In *6th ECCV*, pages 725–741, 2000.

[13] Y. Mukaigawa, H. Miyaki, S. Mihashi, and T. Shakunaga. Photometric image-based rendering for image generation in arbitrary illumination. In *8th ICCV*, pages II:652–659, 2001.

[14] A. Nakashima and A. Maki. Synthesizing pose and lighting variation from object motion. In *10th ICIP*, pages III:725–728, 2003.

[15] A. Nakashima, A. Maki, and K. Fukui. Constructing illumination image basis from object motion. In *7th ECCV*, pages III:195–209, 2002.

[16] M. Nishiyama and O. Yamaguchi. Face recognition using the classified appearance-based quotient image. In *7th FG*, 2006. To appear.

[17] A. Pentland. Photometric motion. *IEEE-PAMI*, 13:9:879–890, 1991.

[18] A. Shashua. *Geometry and photometry in 3D visual recognition*. PhD thesis, Dept. Brain and Cognitive Science, MIT, 1992.

[19] D. Simakov, D. Srolova, and R. Basri. Dense shape reconstruction of a moving object under arbitrary, unknown lighting. In *9th ICCV*, pages 1202–1209, 2003.

[20] C. Tomasi and T. Kanade. Shape and motion from image streams under orthography: a factorization method. *IJCV*, 9:2:137–154, 1992.

[21] M. Weber, A. Blake, and R. Cipolla. Towards a complete dense geometric and photometric reconstruction under varying pose and illumination. In *BMVC*, pages 83–92, 2002.

[22] C.S. Wiles, A. Maki, and N. Matsuda. Hyperpatches for 3d model acquisition and tracking. *IEEE-PAMI*, 23:12:1391–1403, 2001.