Surface Registration with a Helmholtz Reciprocity Image Pair

Peter Tu, Paulo R. S. Mendonça, James Ross and James Miller
GE Global Research Center
Niskayuna, NY 12065, USA
{tu,mendonca,ross,millerjv}@crd.ge.com

Abstract

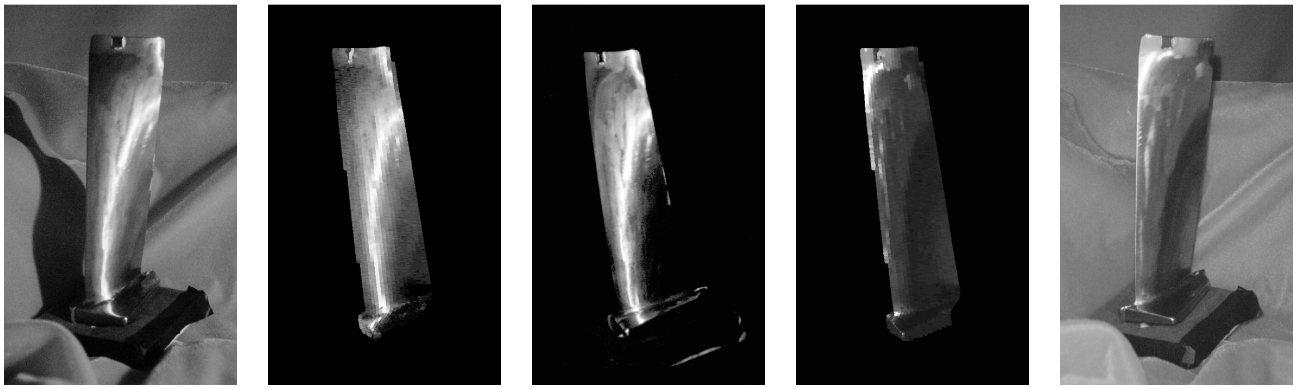
This paper introduces a novel surface model registration technique based on the Helmholtz reciprocity principle. Initially, a 3D surface model is generated for an object of interest. Two images of the object are then acquired under controlled lighting conditions. Given one of the images and an estimate of the surface pose, Helmholtz reciprocity is used to predict the appearance of the object surface as seen in the other image. This prediction is as accurate as one produced by a complete modeling of the BRDF of the surface, without requiring the BRDF to be explicitly measured. The position and orientation of the model are updated in order to minimize an appropriate metric of the dissimilarity between the predicted image and the observed second image. Experimental results are demonstrated for objects possessing a variety of surface properties.

1. Introduction

In many applications a 3D surface model of a known object needs to be aligned with the location and orientation of the object in space. For an object arbitrarily placed before a set of cameras, *3D to 2D registration* is the problem of estimating the pose of the the model that best aligns its projections with the images. An important scenario for the application of registration methods is an inspection system for industrial parts, in which a gauging system can be used to align a CAD model of a known part with images of a newly manufactured item. Model registration is also useful in the medical arena. For example, CT scans of patients could be used to generate sets of iso-surfaces that would then be aligned with images of the patients on an operating table enabling a system for image guided surgery.

If range images in the coordinate system of the cameras can be generated, algorithms such as ICP [1] can be used. Range images can be acquired using various strategies such as textured light reconstruction or dense stereo reconstruction. Textured light techniques such as laser striping can be confounded by various types of surface finishes such as polished metal and fiberglass composites. Dense stereo reconstruction algorithms usually rely on surface texture to establish image-to-image correspondences, and therefore face difficulties when dealing with textureless objects. Recent advances in dense stereo reconstruction are able to address some of these issues, but their implementation can be costly [4, 14, 10].

An alternative approach is to view the 3D surface structure as a generative model for the images. If an initial estimate of the object pose is reasonable, gradient descent can be used to improve the pose estimate by minimizing a measure of the difference between the generated and observed images. By making simple assumptions regarding the source of illumination and the properties of the surface material, a model of the surface can be rendered as seen from a given camera, and then compared to the actual image of the surface. An example would be the work of Horn and Bachman [3], who used a pre-computed model of surface topography and of terrain reflectance to align satellite images. Wildes *et al* [13] took a similar approach, substituting the modeling of terrain reflectance by a texture map of a topographical map of an area to be registered to aerial imagery. Viola and Wells [11] have shown that *mutual information* (MI) is a good metric for measuring the difference between misaligned images. In that work a stereo pair is used to generate the intensity values of one image based on the intensity values of the other. This is done by projecting each point on the surface of a given object into both images and reprojecting the intensity value found in the first image into the second image. Although this approach is usually more realistic than direct rendering of the surface, it assumes a Lambertian model for surface reflectance, which is a poor approximation for most materials, in particular for those that present specularities. In contrast to other techniques, our approach does not force us to settle for this approximation, and therefore we can apply our results to data sets that do not adhere to this model. This goal is achieved by using a generative model based on the *Helmholtz reciprocity principle*, which exploits the symme-



(a) (b) (c) (d) (e)

Figure 1. Prediction of pixel intensities with Helmholtz reciprocity and with Lambertian model. Images (a) and (c) are a Helmholtz pair, while images (e) and (c) are a Lambertian pair, i.e., the light source was fixed when acquiring these images. The Helmholtz prediction of (c) using (3) is shown in (b). The image in (d) displays the prediction of (c) using texture mapping assuming the Lambertian model. The error in the prediction from the Lambertian model is striking, and it is mainly due to its limitations in handling specularities. The Helmholtz prediction, on the other hand, is visually accurate.

try resulting from acquiring two images by swapping the positions of a camera and a light source. This setup allows for the generation of reprojected images which are superior to the Lambertian, Gouraud [2] or Phong [6] models, being in fact in full agreement with the *bidirectional reflectance distribution function* (BRDF) of the surface.

Helmholtz reciprocity has been introduced into computer vision in the context of dense reconstruction [4, 14]. Although dense reconstruction followed by an ICP algorithm can be used as a registration method, this paper shows that given one image of a Helmholtz stereo pair and the surface in the correct pose, the second image in the Helmholtz pair can be generated in accordance to the BRDF of the surface. Thus, reconstruction followed by ICP would be replaced by a search in image space, which is where the original measurements (the observed images) are made. Moreover, reconstruction of specular surfaces is a difficult problem, and although possible solutions to the problem have been found [10], they are computationally expensive. Finally, if the goal of the registration is to track the pose of an object, the computational cost of a full 3D reconstruction is prohibitive, that being the reason for why most model-based trackers rely on making measurements directly on image space.

Questions could be raised about the practicality of using a Helmholtz stereo rig for registration, since it requires carefully controlled lighting. However, by taking images with the light source of the Helmholtz grid off, a measure of ambient light is obtained, which then can be subtracted from the images acquired by the Helmholtz rig. In fact, for all the experimental results shown in this paper that was the procedure used, and there was no attempt to block ambient light.

In section 2 the Helmholtz reciprocity principle is reviewed. Section 3 gives a general description of our registration method, with emphasis on the prediction of one image of an object in a Helmholtz rig given an estimate of its pose and the other image. That section also discusses different metrics that can be employed to quantify discrepancies between the predicted and observed images. Experimental results are described in section 4, and conclusions are presented in section 5.

2. Theoretical Background

Recent works [4, 14, 10] have successfully made use of Helmholtz reciprocity in stereo reconstruction. This principle determines that the BRDF of a surface [5] is symmetric on the incoming and outgoing angles. The original statement by von Helmholtz [12] referred only to *specular reflections*. A more general result was stated by Lord Rayleigh [8]. It is interesting to note that, for generic surfaces, the derivation of the Helmholtz reciprocity principle from basic physics has been established only recently [9].

The BRDF of a point $\mathbf{p}$ on a surface is defined, for a light ray at an incoming direction $\mathbf{v}_1$, to be the ratio between the outgoing radiance at a direction $\mathbf{v}_2$ and the irradiance of the light ray, and it is denoted by $\rho(\mathbf{p}, \mathbf{v}_1, \mathbf{v}_2)$. Helmholtz reciprocity implies that $\rho(\mathbf{p}, \mathbf{v}_1, \mathbf{v}_2) = \rho(\mathbf{p}, \mathbf{v}_2, \mathbf{v}_1)$. Consider a camera and a point light source arbitrarily positioned. Let $\mathbf{v}_1$ be the unit vector pointing from $\mathbf{p}$ to the optical center $\mathbf{c}_1$ of the camera, and $\mathbf{v}_2$ the unit vector pointing from $\mathbf{p}$ to the location $\mathbf{c}_2$ of the light source. The radiance $I_{1,2}$ received

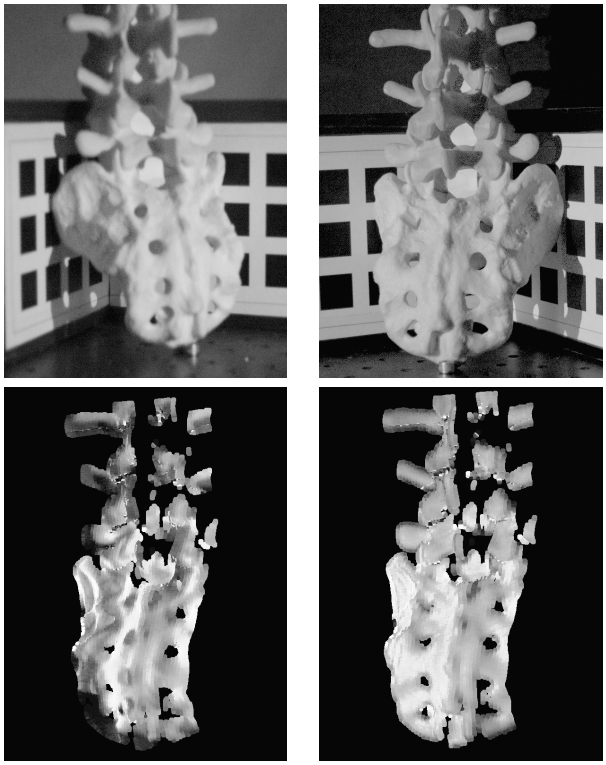


Figure 2. Example of Helmholtz prediction. The upper-left figure shows a coccyx as seen by a camera at $\mathbf{c}_1$. The upper-right displays the same object as seen by a camera at $\mathbf{c}_2$. The lower-left figure gives a predicted image given an incorrect pose estimate of the model, and the lower-right shows image prediction given proper model alignment.

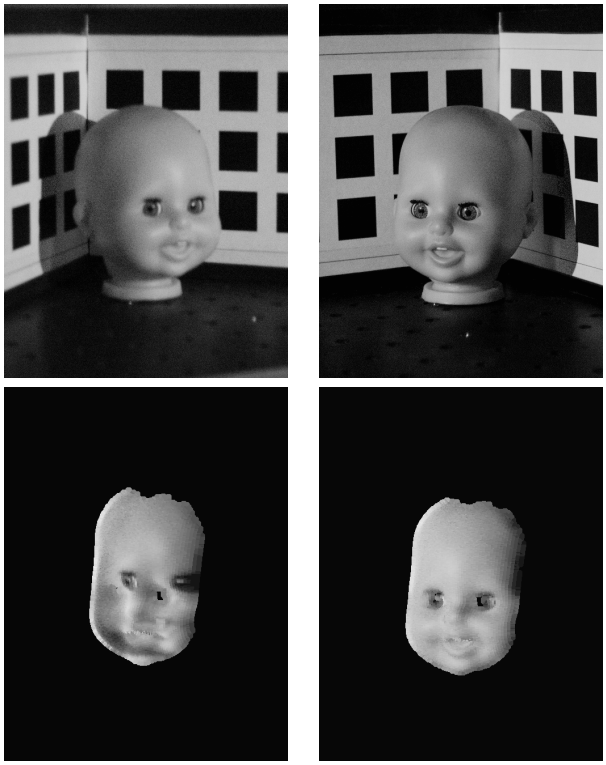


Figure 3. Another example of Helmholtz prediction, this time for a doll's head, following the scheme described in the caption of Fig. 2

by the camera from $\mathbf{p}$ will be given by [5]

$$I_{1,2}(\mathbf{p}) = \eta\rho(\mathbf{p}, \mathbf{v}_1, \mathbf{v}_2)\mathbf{n}\cdot\mathbf{v}_2\frac{1}{\|\mathbf{c}_2 - \mathbf{p}\|^2}, \qquad (1)$$

where $\mathbf{n}$ is the surface normal at $\mathbf{p}$ and η is a scale factor. If the positions of the camera and the light source are swapped, the new radiance $I_{2,1}$ received by the camera will be

$$I_{2,1}(\mathbf{p}) = \eta\rho(\mathbf{p}, \mathbf{v}_1, \mathbf{v}_2)\mathbf{n}\cdot\mathbf{v}_1\frac{1}{\|\mathbf{c}_1 - \mathbf{p}\|^2}. \qquad (2)$$

Substituting (1) in (2) one can compute, given $\mathbf{n}$ and the measured intensity $I_{2,1}$, an estimate of the intensity of the corresponding pixel value in the other image, given by [4]

$$\hat{I}_{1,2} = I_{2,1}\frac{\mathbf{n}\cdot\mathbf{v}_2\|\mathbf{c}_1 - \mathbf{p}\|^2}{\mathbf{n}\cdot\mathbf{v}_1\|\mathbf{c}_2 - \mathbf{p}\|^2}. \qquad (3)$$

From (3) it can be seen that by acquiring a pair of images in which the positions of the camera and the light source are swapped, the knowledge of the surface normals allows for any pixel intensity in one image to be predicted from the other image *regardless of the BRDF of the surface*, yielding more accurate results than the simple assumption of a Lambertian surface. This advantage is illustrated in Fig. 1. Fig. 1(a) shows one image of a Helmholtz pair, the other image shown in Fig. 1(c). The result in (3) was used to predict the image in 1(c) from the image in 1(a), and the result is presented in image 1(b). A similar prediction was done starting with the image in 1(e), but using constant illumination and a Lambertian model. The result of this new prediction is shown in image 1(d). The failure of the Lambertian model in this example is clear.

3. Algorithms for Registration

The high-level procedure to perform model registration can be summarized in three steps: prediction of the model appearance, comparison of the predicted appearance against observed images, and refinement of the model pose to optimize the match between the predicted and observed images.

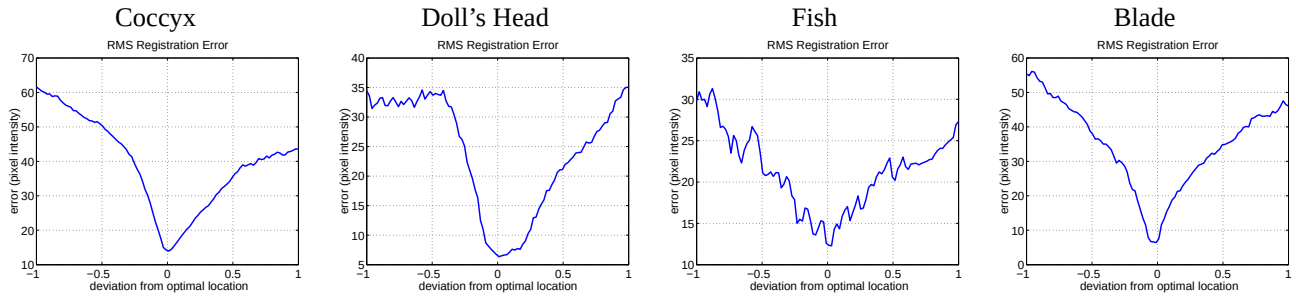


Figure 4. This figure shows the shape of the ϵ_{RMS} cost function for four objects in a neighborhood of optimal pose alignment.

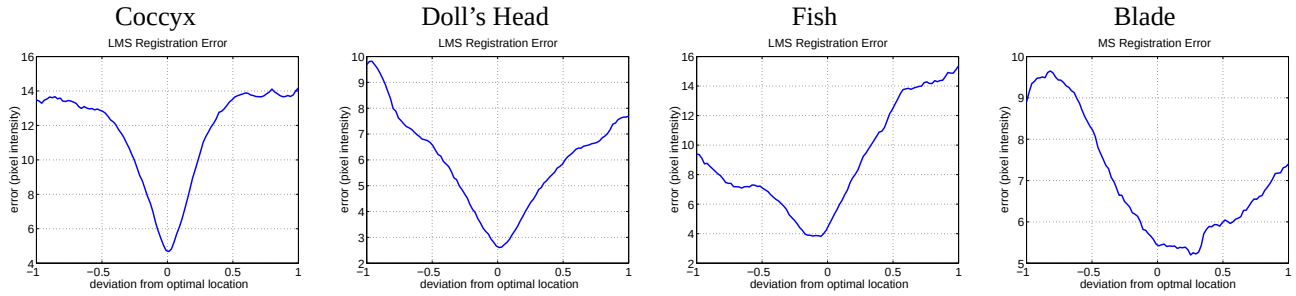


Figure 5. This figure shows the shape of the ϵ_{LMS} cost function for four objects in a neighborhood of the optimal pose alignment. The minimum cost for the blade corresponds to a misalignment.

3.1. Helmholtz Prediction

Initially, a light source is positioned at $\mathbf{c}_2$, and a camera at $\mathbf{c}_1$ captures an image of the object. The positions of the camera and light source are then swapped, and a second image is acquired. Next, assume that an estimate for the pose of the object model is available. For a given model point $\mathbf{p}$, the distances from the point to the (calibrated) camera centers $\mathbf{c}_1$ and $\mathbf{c}_2$ can be easily obtained. The pre-computed surface normal $\mathbf{n}$ at $\mathbf{p}$ is determined directly from the model. Once the viewing directions $\mathbf{v}_1$ and $\mathbf{v}_2$ associated to the camera centers $\mathbf{c}_1$ and $\mathbf{c}_2$ are computed, a ray is then cast from $\mathbf{c}_1$ to $\mathbf{p}$. The intensity of the pixel through which the ray passes is recorded as $I_{1,2}$, and using (3), $\hat{I}_{2,1}$ is computed. For a correct pose estimation for the model, this prediction corresponds to the intensity of the image of $\mathbf{p}$ as seen by the camera at $\mathbf{c}_2$. This procedure is repeated for each data point in the model, generating a prediction of the image as seen from $\mathbf{c}_2$. It is important to observe that this prediction is in agreement with a full modeling of the surface's BRDF, without requiring the BRDF to be explicitly measured.

3.2. Image Metrics

The next step is to determine an appropriate metric for the comparison of the measured and predicted images. The most direct way to measure image dissimilarity is by the root mean square of pixel differences, ϵ_{RMS}, given by:

$$\epsilon_{\mathrm{RMS}} = \sqrt{\frac{1}{N}\sum_{x}\sum_{y}(I_2(x,y) - \hat{I}_2(x,y))^2} \qquad (4)$$

where N is the number of pixels. Fig. 4 shows numerical values of this metric for four different objects, each of which having unique geometrical and textural properties. Let $\mathbf{t} = \begin{bmatrix}0 & 0 & 0\end{bmatrix}^{\mathrm{T}}$ in centimeters and $\boldsymbol{\theta} = \begin{bmatrix}0 & 0 & 0\end{bmatrix}^{\mathrm{T}}$ in degrees be vectors representing the correct pose of the objects. The plots in Fig. 4 show the result of (4) when the pose of the object is perturbed in arbitrary translational and rotational directions, denoted by $s\boldsymbol{\Delta}\mathbf{t}$ with $\|\boldsymbol{\Delta}\mathbf{t}\| = 8$ cm and $s\boldsymbol{\Delta}\boldsymbol{\theta}$ with $\|\boldsymbol{\Delta}\boldsymbol{\theta}\| = 14°$, respectively, for different values of the parameter s, which measures how big the perturbed pose deviates from the optimal one. This corresponds to a one-dimensional slice of the full six-dimensional SE(3) manifold in which the pose parameters lie, and, therefore, cannot offer a full picture of the optimization landscape. It is reassuring, however, that on this slice at least the correct pose corresponds to a minimum of

(4). The original coccyx image can be seen in Fig. 2, the doll's head in Fig. 3.

It is clear that any gradient based optimization algorithm would have difficulties converging to the true solution in the case of the fish data set. As shown in Fig. 6, this is a highly textured surface, with a background having the same material and color properties as the foreground. This problem calls for a different type of dissimilarity measure, such as the median of the square of the pixel differences, which should produce a metric ϵ_{LMS} more robust to image outliers

$$\epsilon_{\mathrm{LMS}} = \underset{(x,y)}{\mathrm{median}}\{(I_2(x,y) - \hat{I}_2(x,y))^2\}. \qquad (5)$$

Fig. 5 shows values of ϵ_{LMS} for the same four objects. Note that the cost curve for the blade has a minimum which is displaced from the optimal alignment position. On the other hand, the ϵ_{RMS} cost curve of the same object is quite smooth, and has a minimum very close to the position of optimal alignment. The model is perturbed from the optimal position in the same way as used to produce Fig. 4. The blade, shown in Fig. 7, is almost textureless, which could explain the shape of ϵ_{LMS} for this object. These results suggest that registration should be performed using ϵ_{RMS} when dealing with objects characterized by smooth, textureless surfaces, and ϵ_{LMS} is preferred when highly textured surfaces are concerned.

3.3. Pose Optimization

Once a prediction of the model appearance is compared against an actual image, the difference between the two can be used to drive an optimization algorithm that will refine the pose of the model. This can be carried out by optimizing any of the cost functions in (4) or (5), where $\hat{I}_2(x,y)$, for all pixel coordinates (x,y), is a function of the same parameters $\mathbf{R}$ and $\mathbf{t}$, corresponding to a rotation matrix and a translation vector with respect to the initial pose of the model. The dependency of $\hat{I}_2(x,y)$ on the orientation $\mathbf{R}$ and location $\mathbf{t}$ of the model is made explicit in (3), since

Algorithm 1 Model registration.

1: Capture image of object with camera at $\mathbf{c}_1$ and light source at $\mathbf{c}_2$
2: Capture image of object with camera at $\mathbf{c}_2$ and light source at $\mathbf{c}_1$
3: Initialize estimation of pose of object model
4: **while** Convergence has not been reached **do**
5: Estimate image seen by camera at $\mathbf{c}_2$
6: Compute distance metric between predicted image and actual image
7: Update estimate of model pose to optimize metric
8: **end while**

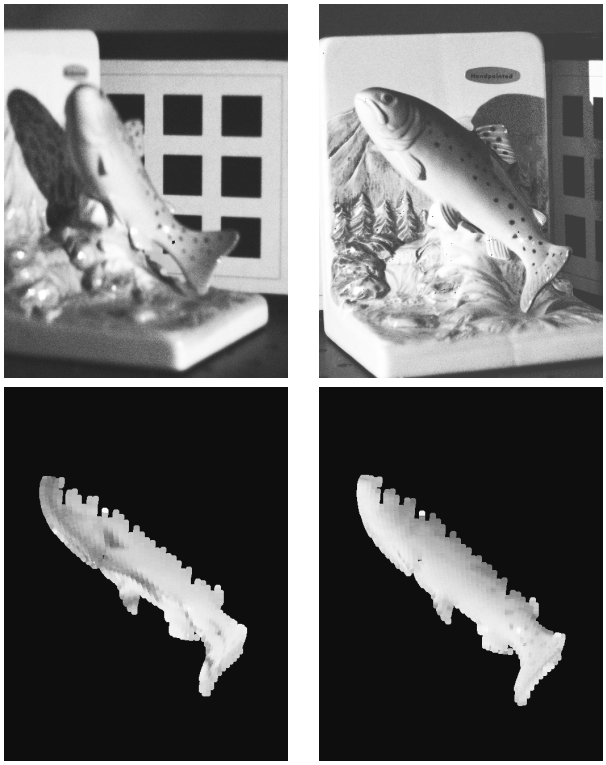


Figure 6. Example of Helmholtz prediction, following the scheme described in the caption of Fig. 2.

$\mathbf{p} = \mathbf{p}(\mathbf{R},\mathbf{t})$, $\mathbf{n} = \mathbf{n}(\mathbf{R})$, and (x,y) are the coordinates of the projection of the point $\mathbf{p}$, i.e., $(x,y) = \mathbf{x} = \mathbf{x}(\mathbf{p})$. Therefore the pose $(\hat{\mathbf{R}},\hat{\mathbf{t}})$ of the model can be obtained as

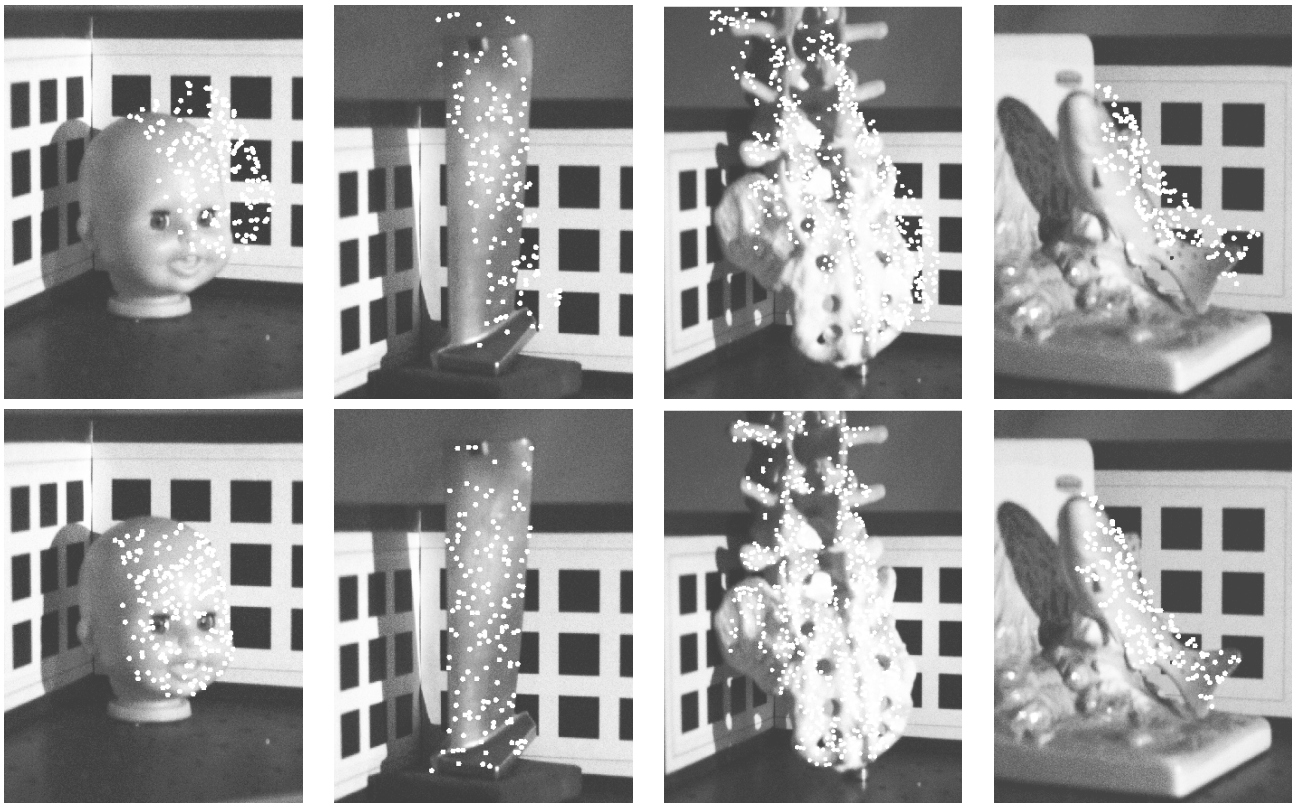
$$(\hat{\mathbf{R}},\hat{\mathbf{t}}) = \arg\max_{(\mathbf{R},\mathbf{t})} \sum_{\mathbf{p}} (I_2(\mathbf{x}(\mathbf{p}(\mathbf{R},\mathbf{t}))) - \hat{I}_2\mathbf{x}(\mathbf{p})(\mathbf{R},\mathbf{t}))^2 \qquad (6)$$

or

$$(\hat{\mathbf{R}},\hat{\mathbf{t}}) = \arg\max_{(\mathbf{R},\mathbf{t})} \underset{\mathbf{p}}{\mathrm{median}}\{(I_2(\mathbf{x}(\mathbf{p}(\mathbf{R},\mathbf{t}))) - \hat{I}_2(\mathbf{x}(\mathbf{p})(\mathbf{R},\mathbf{t})))^2\}. \qquad (7)$$

In the examples shown in this work the optimization method adopted to solve (6) or (7) was conjugate gradient [7], with derivatives computed via finite differences, although many other options are possible. A summary of the algorithm is shown in Alg. 1.

To begin the optimization process it is necessary to have an initial estimation of the pose that is close enough to the true position so that the optimization algorithm will converge. For the registration of industrial parts it is usually the case that a good initial guess is readily available. For



(a) (b) (c) (d)

Figure 8. The figures in the top row show the projection of the point cloud of a misaligned 3D model on its corresponding image. The bottom row shows the projection after Helmholtz registration. The initial position and orientation for the model in columns (a), (b) and (c) were $\mathbf{t} = [2\ 2\ 2]^{\mathrm{T}}$ cm and $\mathbf{w} = [10\ 10\ 10]^{\mathrm{T}}$ degrees. For (d), $\mathbf{t} = [1\ 1\ 1]^{\mathrm{T}}$ cm and $\mathbf{w} = [5\ 5\ 5]^{\mathrm{T}}$ degrees.

tracking applications it is customary to postpone the initialization problem, and at every iteration the current estimation of the pose provides an initial guess for the next iteration that should be close to the ground truth. Finally, the technique presented here is in no worse shape than methods such as registration by maximization of mutual information, which also require a good estimate of initial pose [11].

Because the Helmholtz reciprocity principle yields an exact generative model, there should be zero difference between the predicted and observed images given perfect alignment. Generally, there will be a discrepancy between the predicted image and the actual image as seen by the camera at $\mathbf{c}_2$. This is a result of model mis-alignment, which can be quantified using RMS (the root of mean squared differences), LMS (the median of squared differences), or MI (mutual information). With a properly chosen metric, conjugate gradient is used to update the model's transformation matrix. After the model is re-positioned in the scene, another predicted image is generated, and the cost of the current orientation is again computed. This series of steps is repeated until convergence is reached. Algorithm 1 summarizes these steps, and section 3.2 provides a discussion of the various image comparison metrics.

4. Experimental Results

In order to validate the technique introduced here, a series of experiments with real data was performed. A Helmholtz stereo pair was set up by placing point light sources as close as possible to the optical center of two identical cameras, but avoiding the lights from being occluded by the cameras themselves. Three images were acquired for the objects shown in fig. 8: one with the point sources off, to measure ambient light, and two images for alternate light and camera positions. The background image was then subtracted from each image in the Helmholtz pair. A 3D model for each object was obtained by sweeping the object with a laser stripe and performing stereo reconstruction. The 3D points of the model were then perturbed from their original position by a translation of 2.0 cm in each of the x, y and z directions, and by a rotation of 10° around each of

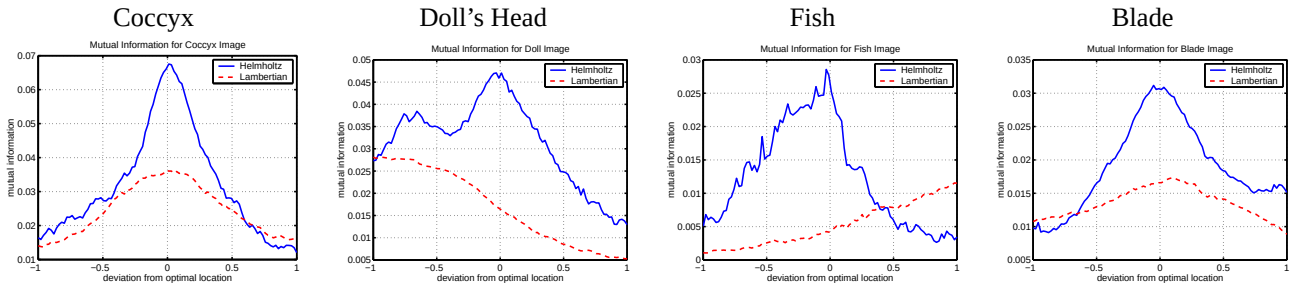


Figure 9. This figure shows the mutual information between a predicted and actual image for both the Helmholtz generative model and the Lambertian image approximation scheme. Note that for both the fish and the doll's head, the Lambertian approach fails to indicate the optimal model pose while the Helmholtz scheme succeeds in doing so for all four objects.

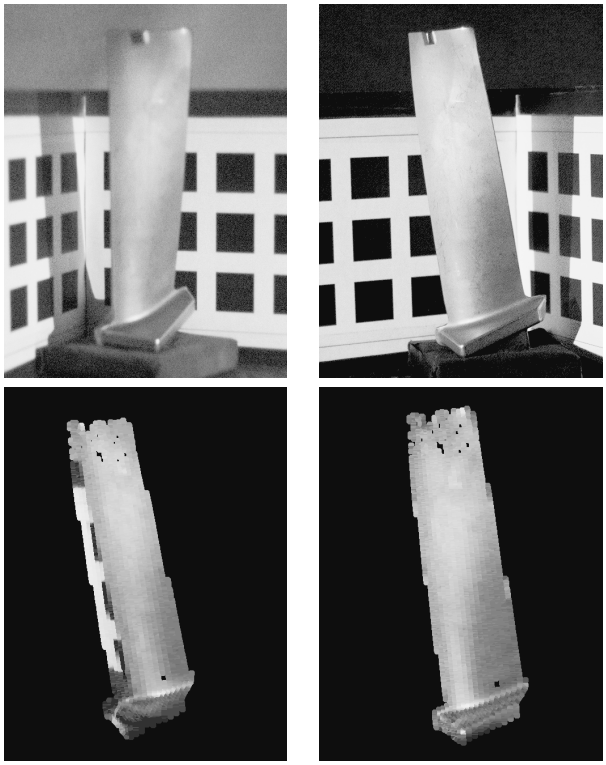


Figure 7. Example of Helmholtz prediction, following the scheme described in the caption of Fig. 2.

the x, y and z axes. Observe that this corresponds to a total translation of 6.9 cm and a rotation of $17.3°$. Since the cameras used in the 3D model reconstruction were the same used for the registration, optimal alignment is obtained with zero translation and rotation. The matrix $\mathbf{R}$ was represented through an exponential map, i.e., $\mathbf{R} = \exp([\mathbf{w}]_\times)$, where $[\mathbf{w}]_\times$ is the anti-symmetric matrix built from the entries of $\mathbf{w}$ such that $[\mathbf{w}]_\times \mathbf{x} = \mathbf{w} \times \mathbf{x} \forall \mathbf{x}$, $\mathbf{w}$ is the vector around which the rotation is performed, and the magnitude of $\mathbf{w}$ is the rotation angle. Using the algorithm described in section 3 with ϵ_{LMS} as the metric, good alignment (final translation < 2 mm and final rotation $< 1°$) was achieved for all data sets with the exception of the fish. This object had to be initialized with translations of 1.0 cm in x, y and z directions, as well as with rotations of $5°$ around the x, y and z axes. This corresponds to a total translation of 1.7 cm and a rotation of $8.8°$. As a quick experiment to verify the robustness of the registration to the initial pose of the model, the initial translations and rotations applied to the models were multiplied by -1, and the registration algorithm was rerun. Again, convergence within 2 mm and $1°$ was obtained. Fig. 8 shows initial and final poses for each object. The difficulty in convergence for the fish model can be attributed to two factors: the cluttered background (which has the same texture as the fish) and the small size of the object.

5. Conclusion

This paper introduces a technique for registering 3D models to 2D images based on Helmholtz reciprocity. By exploiting this principle the algorithm can predict the appearance of the back-projected model in agreement with its BRDF without having to explicitly know the BRDF. This is a great advantage over techniques which assume a Lambertian model, valid only for certain types of surfaces. In particular, such algorithms are not capable of handling shinny or specular surfaces, as demonstrated in Fig. 1. After the appearance of the model has been predicted, a suitable image metric is used to quantify the discrepancy between predicted and observed images. Since the predicted image should be in agreement with the BRDF of the object, this discrepancy can be attributed to misalignment of the object, and it can therefore drive a search for optimal registration parameters. The effectiveness of this algorithm was

demonstrated in a number of registration experiments with different objects.

An interesting point can be made about combining mutual information with the method introduced in this paper. The strength of mutual information comes from the fact that it does not perform a point to point comparison of predicted and measured images, but rather scores the overall similarity between the global distribution of intensities, according to

$$\epsilon_{\mathrm{MI}} = H[I_2] + H[\hat{I}_2] - H[I_2, \hat{I}_2], \tag{8}$$

where $H[x]$ is the entropy of the random variable x. In a preliminary experiment, the value of (8) was evaluated in a similar fashion to that used to produce Figs. 4 and 5, with images produced by a Helmholtz stereo rig. Additionally, the experiment was reproduced with a fixed light source, and assuming a Lambertian model. The results are shown in Fig. 9. In all cases, the technique developed here has shown to provide, in entropy units, more information than using a Lambertian model. Additionally, the Lambertian scheme failed to indentify the correct model pose for both the fish and the doll's head, whereas the optimization landscape for the Helmholtz method is much more well behaved.

References

[1] P. J. Besl and N. D. McKay. A method for registration of 3-D shapes. *IEEE Trans. Pattern Analysis and Machine Intell.*, 14(2):239–256, Feb. 1992.

[2] H. Gouraud. Continuous shading of curved surfaces. *IEEE Trans. on Comput.*, 20(6):623–629, June 1971.

[3] B. K. P. Horn and B. L. Bachman. Using synthetic images to register real images with surface models. Technical Report AIM-437, MIT, Massachusetts, USA, Aug. 1977.

[4] S. Magda, D. J. Kriegman, T. E. Zickler, and P. N. Belhumeur. Beyond lambert: Reconstructing surfaces with arbitrary BRDFs. In *Proc. 8th Int. Conf. on Computer Vision*, volume II, pages 391–398, Vancouver, Canada, July 2001.

[5] F. E. Nicodemus. *Radiometry*, volume IV of *Applied Optics and Optical Engineering*. Academic Press, New York, 1967.

[6] B. T. Phong. Illumination for computer-generated pictures. *Comm. of the ACM*, 18(6):311–317, June 1975.

[7] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, Cambridge, UK, 2^{nd} edition, 1992.

[8] J. W. S. B. Rayleigh. On the law of reciprocity in diffuse reflection. *Philosophical Magazine*, 49:324–325, 1900.

[9] W. C. Snyder. Structured surface BRDF reciprocity: Theory and counterexamples. *Applied Optics*, 41(21):4307–4313, July 2002.

[10] P. Tu and P. R. S. Mendonça. Surface reconstruction via Helmholtz reciprocity with a single image pair. In *Proc. Conf. Computer Vision and Pattern Recognition*, pages 541–547, Madison, USA, June 2003.

[11] P. Viola and W. M. Wells III. Alignment by maximization of mutual information. *Int. Journal of Computer Vision*, 24(2):137–154, Sept. 1997.

[12] H. von Helmholtz. *Treatise on Physiological Optics*. Dover, New York, 1925. Translated from the 1856 German edition.

[13] R. P. Wildes, D. J. Hirvonen, S. C. Hsu, R. Kumar, W. B. Lehmen, B. Matei, and W. Y. Zhao. Video georegistration: Algorithm and quantitative evaluation. In *Proc. 8th Int. Conf. on Computer Vision*, volume 2, pages 343–350, July 2001.

[14] T. Zickler, P. N. Belhumeur, and D. J. Kriegman. Helmholtz stereopsis: Exploiting reciprocity for surface reconstruction. In A. Heyden, S. G., M. Nielsen, and P. Johansen, editors, *Proc. 7th European Conf. on Computer Vision*, volume 3 of *Lecture Notes in Computer Science 2352*, pages 869–884, Copenhagen, Denmark, May 2002.