REGISTRATION AND MATCHING OF PERSPECTIVE SURFACE NORMAL MAPS

Craig Fancourt[1]

Sarnoff Corp.
Princeton, New Jersey, USA

ABSTRACT

We present a method for the registration and matching of perspective surface normal maps. Registration of two maps consists of optimally aligning their normals through a 2-D warping in the image plane in conjunction with a 3-D rotation of the normals. Once aligned, the average dot-product then serves as a match metric for automatic target recognition (ATR). We conduct an ATR experiment using synthesized views of 25 commercial vehicles, and obtain perfect recognition results when the test azimuth is within [-6°,+10°] of the reference pose, even when the normals are corrupted by up to 20° uniform random noise. The results suggest that needle maps are a rich yet compact representation of an object, which may be useful for exploiting information from stereo images, shape from shading algorithms, or sensors which obtain the normals from polarization information.

Index Terms **—** needle maps, surface normal maps, registration, automatic target recognition.

1. INTRODUCTION

We define a normal or needle map as a type of perspective image where each pixel represents a 3-D surface normal. It can be compactly represented by the notation $\mathbf{n}(\mathbf{r})$, where $\mathbf{n} \in \Re^3$ is the 3-D surface normal, and $\mathbf{r} \ni \Re^2$ is the 2-D pixel location in the image plane. Typically, the coordinate system for representing the normals is such that the $\hat{x}$ and $\hat{y}$ directions coincide with the image plane, while the $\hat{z}$ direction is perpendicular to the image plane.

Such maps can arise from a photometric stereo analysis of multiple images [1], a shape-from-shading analysis of a single image [2-3], or an analysis of the polarization of light reflected from surfaces [4-5].

1.1. Related work

To the best of our knowledge, no one has proposed a general method for registering needle maps, while there are several proposed methods for matching needle maps.

Automatic target recognition (ATR) with needle maps has tended to focus on two tracks (see [2] for a survey). One approach is to calculate local features from the normals, such as those related to the shape operator, and then match local features to known model features. However, the shape operator is calculated from the derivatives of the normals, and is thus extremely sensitive to noise. In addition, the shape operator, when estimated from normals, is not guaranteed to be symmetric, and thus may not always represent a proper integrable surface.

Another approach has been to integrate over the normals in an attempt to recover the underlying 3-D surface, and then match the recovered 3-D shape to known models. However, this procedure is computationally intensive and has difficulty with discontinuous surfaces, such as those due to edges or occlusions.

Here we show that accurate ATR can be achieved with needle maps directly. This is achieved through a registration procedure which optimally aligns the normals between a reference and test image through a 2-D warping in the image plane, in conjunction with a 3-D rotation of the normals. The average dot-product after alignment is then used as a matching criteria for ATR.

2. REGISTRATION OF SURFACE NORMAL MAPS

Registration of two images, which ostensibly capture the same scene, but at different times, from different viewpoints, or with different sensors, is the process of aligning them so that they coincide according to some well-defined criteria.

For 2-D intensity images, the standard optimization criteria is the mean-squared difference between the intensities of corresponding pixels in the two images. For needle maps, the analogous criteria is the average L2-norm of the difference between corresponding surface normals. However, if the two needle maps were obtained from differing viewpoints, then the surface normals of corresponding pixels in one image will be rotated with respect to the other. Therefore, the registration must include a rotation matrix operating on one of the needle maps.

Like intensity image registration, in order to optimally align the pixels in the image plane, the 2-D position of the surface normals of one map must be transformed or warped in the image plane.

All this suggests the following optimization criteria

$$J = \tfrac{1}{2} \sum_{\mathbf{r}_1} \left\| \mathbf{n}_2\left(f(\mathbf{r}_1;\mathbf{w})\right) - \mathbf{R}\,\mathbf{n}_1(\mathbf{r}_1) \right\|^2 \qquad (1)$$

where $\mathbf{n}_1(\mathbf{r}_1)$ is one needle map as a function of planar image coordinates $\mathbf{r}_1$, and $\mathbf{n}_2(\mathbf{r}_2)$ is a second needle map as a function of planar image coordinates $\mathbf{r}_2 = f(\mathbf{r}_1;\mathbf{w})$, parameterized by the planar transform weights $\mathbf{w}$, and $\mathbf{R}$ is a 3-D rotation matrix. Note that the rotation matrix could alternatively operate on $\mathbf{n}_2$, but the resulting equations are simpler for the present formulation.

Registration consists of optimizing the criteria with respect to the rotation matrix, $\mathbf{R}$, and the warping parameters, $\mathbf{w}$.

[1] Now at Merck Research Laboratories, Rahway, New Jersey, USA.

2.1. Rotation matrix

For a given planar transformation, there is an analytical solution for the 3-D rotation matrix [6]

$$\frac{\partial J}{\partial \mathbf{R}} = 0 \Rightarrow \mathbf{R} = \mathbf{U}\mathbf{V}^T \tag{2}$$

where the orthogonal matrices **U** and **V** are derived from the singular value decomposition of the outer product of the normals:

$$\mathbf{U}\mathbf{S}\mathbf{V}^T = SVD\left(\sum_{\mathbf{r}_1} \mathbf{n}_2\left(f(\mathbf{r}_1;\mathbf{w})\right)\mathbf{n}_1^T(\mathbf{r}_1)\right) \tag{3}$$

The orthogonality of **U** and **V** guarantees that **R** is also orthogonal and thus represents a pure rotation.

2.2. Image plane transformation

There is no analytical solution for the 2-D planar transform weights, and thus the solution must be found iteratively. We use the Gauss-Newton method

$$\Delta\mathbf{w} = -\left(\frac{\partial^2 J}{\partial \mathbf{w}\partial \mathbf{w}^T}\right)^{-1}\frac{\partial J}{\partial \mathbf{w}} \tag{4}$$

where the gradient is given by

$$\frac{\partial J}{\partial \mathbf{w}} = \sum_{\mathbf{r}_1}\frac{\partial \mathbf{r}_2^T}{\partial \mathbf{w}}\frac{\partial \mathbf{n}_2^T}{\partial \mathbf{r}_2}\left[\mathbf{n}_2(\mathbf{r}_2) - \mathbf{R}\,\mathbf{n}_1(\mathbf{r}_1)\right] \tag{5}$$

and the Hessian is approximated by the outer product

$$\frac{\partial^2 J}{\partial \mathbf{w}\partial \mathbf{w}^T} \approx \sum_{\mathbf{r}_1}\left[\frac{\partial \mathbf{r}_2^T}{\partial \mathbf{w}}\frac{\partial \mathbf{n}_2^T}{\partial \mathbf{r}_2}\right]\left[\frac{\partial \mathbf{r}_2^T}{\partial \mathbf{w}}\frac{\partial \mathbf{n}_2^T}{\partial \mathbf{r}_2}\right]^T \tag{6}$$

For the special case of an affine transform:

$$\mathbf{r}_2 = \mathbf{P}_1^T\mathbf{w}\,,\quad \frac{\partial \mathbf{r}_2^T}{\partial \mathbf{w}} = \mathbf{P}_1 \tag{7}$$

$$\mathbf{P}_1^T = \begin{bmatrix} 1 & x_1 & y_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & x_1 & y_1 \end{bmatrix} \tag{8}$$

$$\mathbf{w}^T = \begin{bmatrix} w_1 & w_2 & w_3 & w_4 & w_5 & w_6 \end{bmatrix} \tag{9}$$

The affine transform is capable of several types of geometric warpings, including the identity, translation, scaling, rotation, reflection, and shear. However, it does not in general preserve lengths and angles. Note that these results apply to 2-D images, and may or may not have corollaries for needle maps.

Evaluation of $\mathbf{n}_2(f(\mathbf{r}_1;\mathbf{w}))$ requires interpolation of the warped needle map at possibly non-integer planar coordinates. For intensity image registration, the most common interpolation methods are nearest neighbor, bi-linear, and higher-order non-linear methods. The nearest neighbor method is equally applicable to needle maps without modification. However, the other methods require modification because the normal components are not independent. This can be accounted for by either: (1) independent interpolation of the three components followed by renormalization; or (2) fitting the normals to a local model of the underlying surface, and then deriving the normals from the model. In this paper, we used bilinear interpolation with the first method.

2.2. Joint optimization

The 3-D rotation matrix, **R**, and 2-D planar transform weights, **w**, are implicitly coupled, related to each other through the hidden parameters of an unknown perspective. However, here we treat them as independent and optimize them as such. Thus, optimization proceeds by alternating between calculating **R**, and iterative optimization of **w**.

For a given rotation matrix, the 2-D planar transform is optimized until no further improvement in the criteria is observed. At this point, the rotation matrix is recalculated using the latest alignment between the two images, and then the optimization of the planar transform is begun anew.

The final stopping criteria is when no further improvement in the criteria is observed between successive recalculations of the rotation matrix. The algorithm is initialized with affine transform weights that correspond to the identity transform, and a rotation matrix calculated from the original normals images.

2.3. Examples

The following two examples shows the registration of two vehicles at the same elevation (30°), but at different azimuths (reference = 46°, test = 58°). The surface normal maps were generated using IRMA facet models with ray tracing. For display purposes, the normals directions are mapped into a color space. Fig. 1 shows the registration of an Accord with itself. The dot-product map shows that most errors occur at the boundaries between planes, as well as the side view mirror and wheel wells. Figure 2 shows the registration of an Accord with a Volvo. Here, the front grill area of the cars is also salient. After registration, the average dot-product of the Accord with itself was 0.9682, while for the Accord with the Volvo it was 0.9586, demonstrating the potential of registration as a matching method, even for two very similar vehicles.

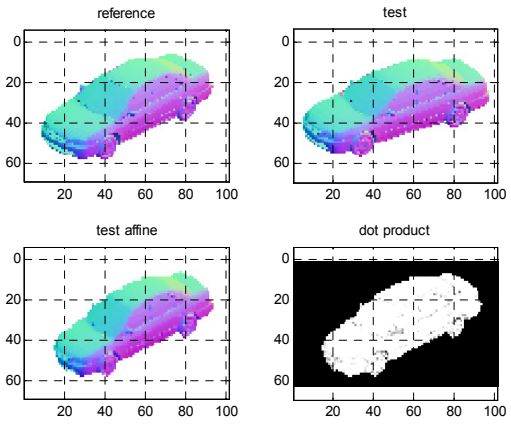

Fig. 1. Registration of two needle maps of the same vehicle (Accord) in different poses: (top left) reference (az=46°); (top right) test (az=58°); (bottom left) warped test; (bottom right) dot-product between reference and warped test.

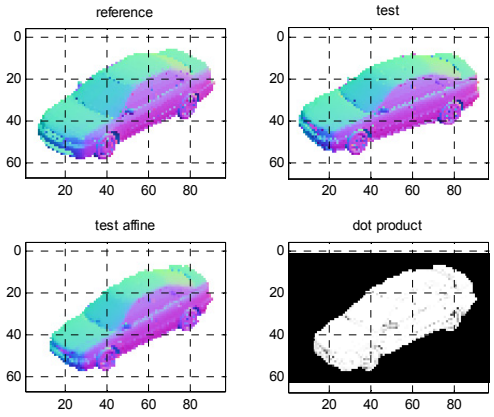

Fig. 2. Registration of two needle maps of different vehicles in different poses: (top left) reference (Accord, az=46°); (top right) test (Volvo, az=58°); (bottom left) warped test; (bottom right) dot-product between reference and warped test.

3. MATCHING OF SURFACE NORMAL MAPS

In one scenario, matching an unknown but segmented query needle map against a model database can be accomplished by the following steps: 1. Pose estimation of the unknown object; 2. Retrieval of database model images for the nearest available pose. Alternatively, the database model images can be generated on-the-fly at the estimated pose; 3. Registration of the query image to all database models; 4. Calculating a match score between the unknown image and each of the database models.

The match score may be derived from the minimum value of the registration criteria, or it may be more complicated features compared using the point correspondences obtained from the registration. We use the registration criteria itself.

Note that if the normals have unity norm and **R** is a true rotation matrix, then the criteria can be re-written as

$$J = \sum_{\mathbf{r}_1} 1 - \mathbf{n}_2^T(\mathbf{r}_2)\,\mathbf{R}\,\mathbf{n}_1(\mathbf{r}_1) \qquad (11)$$

which is one minus the dot-product, summed over all pixels. The match score is the mean dot-product after registration

$$S = 1 - \frac{J}{N} = \frac{1}{N}\sum_{\mathbf{r}_1} \mathbf{n}_2^T(\mathbf{r}_2)\,\mathbf{R}\,\mathbf{n}_1(\mathbf{r}_1) \qquad (12)$$

where N is the number of registered pixels. It will tend to have a value of one for identical objects, and zero for randomly unrelated objects. For matching, the query object is registered against all the reference models in a database, and the model with the largest average dot-product is chosen as the best match.

4. ATR EXPERIMENT

We conducted an automatic target recognition (ATR) experiment to determine the robustness of perspective surface normal map registration and matching to pose estimation errors and surface normal noise. To this end, we generated perspective normal maps of 25 vehicles using IRMA facet models with ray tracing. The vehicles consisted of cars, pickups, suv's, and vans, as follows:

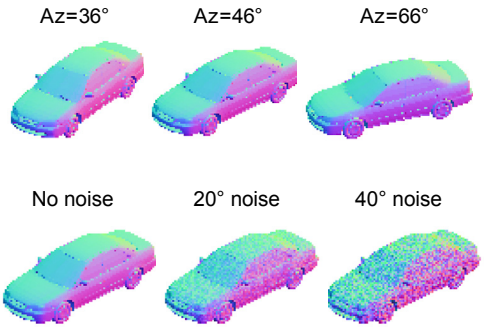

Fig. 3. Range of azimuths (top row) and noise levels (bottom row) of the test set in the ATR experiment, shown for the Accord.

cars: accord, beetle, bmw, lexus, markvii, mercedes, neon, ptcruiser, volvo.

pickups: chevy, dakota, elcamino, frontier, s10, tacoma.

suv's: 4runner, blazer, escalade, jimmy, landrover, mercedessuv, pathfinder, suburban.

vans: aerostar, astro.

The reference set consisted of the normals images of the 25 vehicles at 50 m distance, 30° elevation, 46° azimuth, and no noise. The test set consisted of the normals images of the same 25 vehicles at the same 50 m distance and 30° elevation, but with varying azimuth (-10° to +20° offset from reference), and uniform normals noise (0°, 20°, and 40°). The pixel resolution corresponding to these specifications was approximately 5 cm.

Figure 3 shows the range of azimuths and noise levels employed, while Figure 5 shows the complete reference vehicle set. Again, the normals directions were mapped to a color space.

Each test image was warped into each reference image, using the registration method discussed previously, and one minus the average dot-product after alignment served as the match metric between the two vehicles. The results are summarized in Figure 4.

For the case of no noise, the ATR results are perfect for pose offsets of –6° to +12°. For 20° noise, the results are similar, with perfect ATR results for pose offsets of –6° to +10°. In general, results are better for positive test azimuth differences, which exposes more of the side of a vehicle, compared to negative test azimuth differences, which exposes more of the front of a vehicle.

The results are surprisingly good for the 40° noise case, with an average recognition rate of 0.96 for pose offsets of –6° to +12°. This is significant because 45° noise is the point at which right angle planes become rounded, in effect losing their shape, and becoming indistinguishable from a continuously curved surface. Thus, needle map matching is very robust to noisy normals.

5. CONCLUSION

We presented a method for the registration of perspective surface normal maps, and a way of matching them using the registration criteria itself. Through experiments, we showed that needle maps are a compact, rich, and robust description of an object.

In particular, we showed that perfect ATR results can be obtained from a single view of 25 vehicles (if pose is estimated to within -6° to +10°), even for noisy (20° uniform) normals. For box-like structures such as vehicles, the errors in the normals can be as large as 40° and the primary surfaces are still distinct. This is

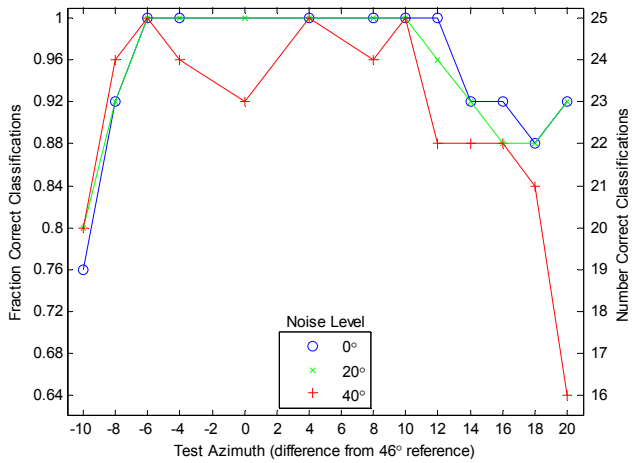

Fig. 4. ATR performance as a function of azimuth difference between test and reference poses for normal noise levels of 0°, 20°, and 40°.

particularly significant, given that 45° noise is the critical point where box-like structures start to appear as rounded surfaces.

This suggests that the method presented here may be applicable to the noisy normal estimates obtained from stereo images, shape-from-shading algorithms, or algorithms that obtain surface normal information from polarization data.

In addition, many of the improvements to standard image registration are also applicable to needle map registration. These include coarse-to-fine registration at multiple image scales, more realistic 2-D transforms such as the perspective transform, symmetric registration, and optical flow. We intend to purse all these improvements.

6. REFERENCES

[1] Woodham R.J., Photometric method for determining surface orientations from multiple images, *Optical Engineering*, vol. 19, pp. 139-144, 1980.

[2] Zhang R., Tsai P.S., Cryer J., and Shah M., Shape from Shading: A Survey, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 21, no. 8, pp 690-706, 1999.

[3] Worthington P.L. and Hancock E.R., Object recognition using shape-from-shading, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 23, no. 5, pp. 535-542, 2001.

[4] Wolff L.B., Surface orientation from polarization images, in *Proc. Optics Illumination and Image Sensing for Machine Vision II*, vol. 850, pp. 110-121, 1987.

[5] Sadjadi F.A. and Chun C.S.L., Passive polarimetric IR target classification, *IEEE Trans. on Aerospace and Electonics Systems*, vol. 37, no. 2, pp. 740-751, 2001.

[6] Umeyama S., Least-squares estimation of transformation parameters between two point patterns, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 13, no. 4, pp. 376–380, 1991.

7. ACKNOWLEDGMENTS

This work was supported by the Raytheon Corp. under contract #S7-6BW317X6272. Special thanks to Dave Doria and Eran Marcus for their guidance and for supplying the normal maps.

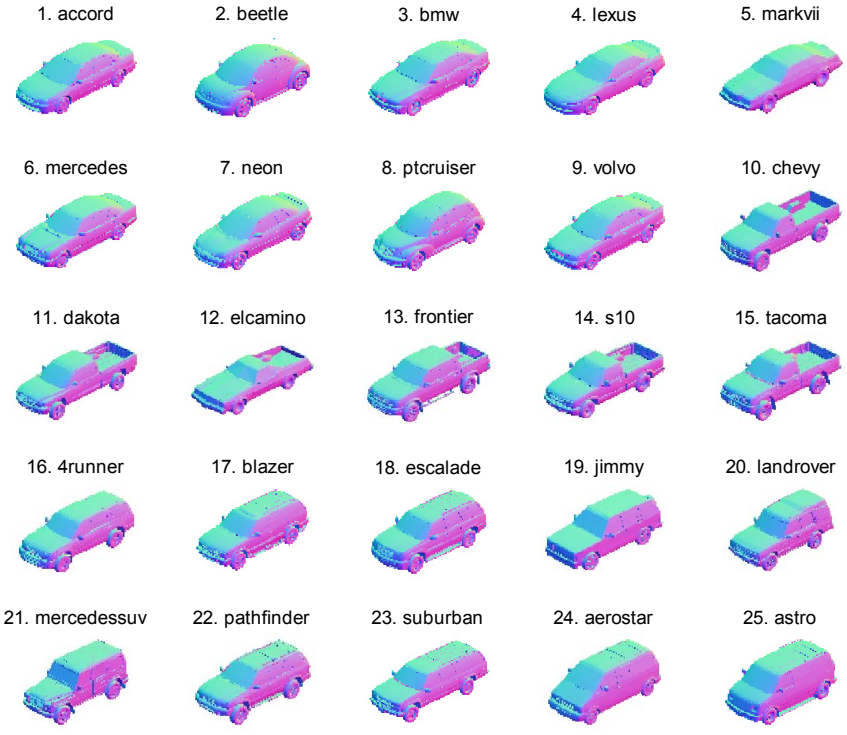

Fig. 5. Reference set of 25 vehicles at 30° elevation, 46° azimuth, and no normals noise.