MULTISENSOR IMAGE REGISTRATION

Ravi K. Sharma and Misha Pavel

Oregon Graduate Institute, Portland, OR

Abstract

We describe an approach for registration of images from different sensors that may generate local polarity reversals and inconsistent features. The developed technique is based on modifying the objective function by transforming the images into a representation that is less dependent on the local image intensity, and using an iterative process that reduces the effect of inconsistent features.

1. Introduction

Image registration is a critical component of multisensor systems that typically use different types of sensors. For instance, image fusion systems which aide in vehicular navigation [10] combine image information from TV, IR and radar. Another example is medical imaging. Several techniques have been developed in the past to register identical and similar types of images [3]. Optical flow recovery techniques such as gradient-based approaches [7, 9], correlation-based techniques [1], and spatio-temporal filtering, perform registration by estimating motion. Feature-based techniques register by finding a mapping between corresponding features such as corners, points and contour lines in the images [11, 8].

Images captured through different types of sensors pose additional difficulties to the traditional registration approaches. The grayscale characteristics of such images may differ locally. For example, the polarity of contrast changes (i.e. polarity reversals) can be arbitrary between TV and IR images; thus degrading the performance of the optical flow techniques which assume luminance constancy. The feature extraction task can be impaired because features in one image may take on a different geometric form in the other image. Finally, features present in one image may not be present in the other image (i.e. inconsistent or complementary features), e.g. TV and radar imagery. In this paper we present an approach for registration of multisensor images that addresses the above problems.

2. Multisensor Registration

We define registration as a mapping between images [6], both spatially and with respect to intensity,

$$G_2[I_2](x, y) = G_1[I_1](x + u(x, y, \mathbf{p}), y + v(x, y, \mathbf{p})) \qquad (1)$$

where I_1 and I_2 are images captured through different sensors, (x, y) denotes spatial location, G_1 and G_2 are non-invertible transformations that account for grayscale variations, (u, v) denotes the spatial motion between the images that is specified by a parametric motion model with parameters $\mathbf{p}$ (see Section 4.)

We use a hierarchical [2] or coarse-to-fine registration strategy based on the multiresolution Laplacian pyramid scheme [5], which represents the image as successive levels of bandpassed, sub-sampled images I^k, where k denotes the level. The registration parameters at each level are estimated by minimizing an objective function (error function) that embodies the definition of correspondence among the images. The objective function is generally expressed in terms of differences between the images.

3. Reducing Graylevel Disparities

If the images to be registered contain polarity reversals, an error function such as the sum of squared differences, may not necessarily be minimum at the point where the images are

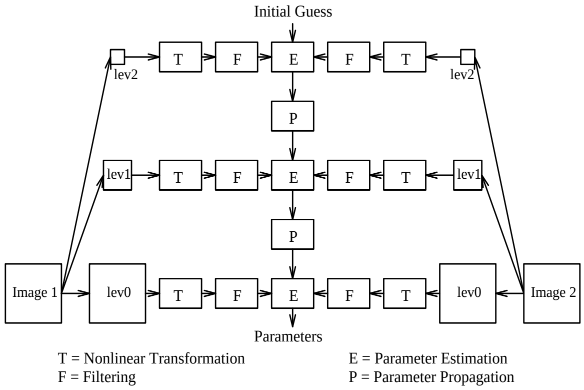


Figure 1: Block diagram.

perfectly registered. Figure 2(a), shows the error E, for a pair of images, as a function of translation. In this example, the point of correct registration is actually the maximum point on the error surface.

To address this problem, we must make the images as similar as possible before registration. We transform each image using transformations G_1 and G_2, to a representation that is invariant to the polarity reversals. The particular transformation used for this paper is the absolute of the pixel values at each resolution of the Laplacian pyramid. This transformation results in identical representations for signals which are exact opposites of each other.

We express the objective function in terms of the nonlinear pyramids as,

$$E^k(\mathbf{p}^k) = \sum_x \sum_y (G_1[I_1^k](x + u(x,y,\mathbf{p}^k), \\ y + v(x,y,\mathbf{p}^k)) - G_2[I_2^k](x,y))^2 \quad (2)$$

Figure 2(b) shows the error surface after the nonlinear transformation. Since the disparity between the graylevels is reduced, the minimum point on the error surface yields the correct registration.

Thus, applying G_1 and G_2 results in a new error function whose minimum is at the point of correct registration. We note that well-chosen transformation functions will allow our technique to benefit from both approaches – feature and motion based registration.

4. Motion models

Motion can be represented either locally or globally. In the local motion model, the motion of each pixel is described in the horizontal and vertical directions. In the global motion model, the motion over the entire image is described by a set of parameters. For our experiments we assumed a global motion model.

If the world scene is assumed to be a planar surface, the motion of the scene can be approximated by a projective model or perspective transformation since the image is formed by projection onto the camera plane.

$$\begin{aligned} u(x,y) &= p_1 + p_3 x + p_5 y + p_7 x^2 + p_8 xy \\ v(x,y) &= p_2 + p_4 x + p_6 y + p_7 xy + p_8 y^2 \end{aligned} \quad (3)$$

When the distance between the scene and the camera is large, the motion can be approximated by an affine model, with $p_7 = p_8 = 0$ in Equation 3. Translation is a special case of the projective model when $p_3 = p_6 = 1$ and $p_4 = p_5 = p_7 = p_8 = 0$ in Equation 3.

5. Parameter Estimation

The model parameters at each resolution are estimated by minimizing the error function in Equation 2. A standard numerical approach for the minimization is the Gauss-Newton method [2]. Parameter estimation proceeds from lower resolutions to higher resolutions - at each step the coarse solution obtained at a lower level is used to improve the registration accuracy at a higher level. The block diagram illustrating the different steps is shown in Figure 1.

The minimization process is sensitive to local minima in the error surface. A smoothing of the error surface reduces the local minima. We smooth the error surface by filtering the nonlinear pyramid representation with a lowpass filter. This operation causes a spatial spread of the features in the pyramid levels. Figure 2(c) shows the error surface after filtering. The error surface is smoother and the basin containing the global minimum is widened, leading to improvement in performance of the minimization procedure.

6. Complementary features

Complementary features (i.e. inconsistent features generated by sensors such as FLIR and radar) can cause errors in registration if their impact on the error function is significantly large. We have developed an iterative procedure to reduce the errors caused by the complementary features. We first register the images as well as is possible, using the approach described in the previous sections. We then compute a discrepancy image W that represents the error due to complementary features and the error due to the remaining mis-registration. To compute W, we estimate a local (pixel by pixel) graylevel affine transformation that accounts for local polarity reversals between the two images.

$$\begin{aligned} I_2(x,y) = \beta(x,y) I_1(x + u(x,y,\mathbf{p}), \\ y + v(x,y,\mathbf{p})) + \alpha(x,y). \end{aligned} \qquad (4)$$

W is the squared difference after the transformation,

$$\begin{aligned} W(x,y) = (I_2(x,y) - \beta(x,y) I_1(x + \\ u(x,y,\mathbf{p}), y + v(x,y,\mathbf{p})) - \alpha(x,y))^2, \end{aligned} \qquad (5)$$

which we represent by its Laplacian pyramid. We identify the highest ten percent of values in the two highest resolutions of the discrepancy pyramid and remove them from the error function computation. We assume that these are the regions containing the complementary features that cannot be matched. We then repeat the registration with the new error function.

7. Conclusion

Figures 3, 4, and 5 show the registration results of our experiments using images from several different sensors. We conclude that registration of different sensors is possible by modifying the objective function using a nonlinear transformation to reduce the disparities caused by polarity reversals, by smoothing of the objective function for improved parameter estimation, and by using an iterative procedure that reduces the effect of complementary features.

8. Acknowledgments

This work was supported by a grant to the Oregon Graduate Institute from NASA Ames Research Center, NCC2-811. We are grateful to Yacov Hel-Or for providing us with the source code of his gradient-based registration package, and to Todd Leen for useful discussions.

9. References

[1] P. Anandan. A computational framework and an algorithm for the measurement of visual motion. *Intl. Jou. of Comp. Vision*, 2(3):283–310, 1989.

[2] J. R. Bergen, P. Anandan, K. J. Hanna, and R. Hingorani. Heirarchical model based motion estimation. In *Second European Conf. on Comp. Vision (ECCV'92)*, pages 237–252, 1992.

[3] L. G. Brown. A survey of image registration techniques. *Computing Surveys*, 24(4):325–376, 1992.

[4] M. A. Burgess, T. Chang, D. E. Dunford, R. H. Hoh, W. F. Home, and R. F. Tucker. Synthetic vision technology demonstration executive summary. Technical Report DOT/FAA/RD-93/40,I, Research and Development Service, Washington, D.C., 1993.

[5] P. J. Burt and E. H. Adelson. The Laplacian pyramid as a compact image code. *IEEE Trans. on Communications*, Com-31(4):532–540, 1983.

[6] Y. Hel-Or. Studies in gradient-based registration. Technical report, NASA Ames Research Center.

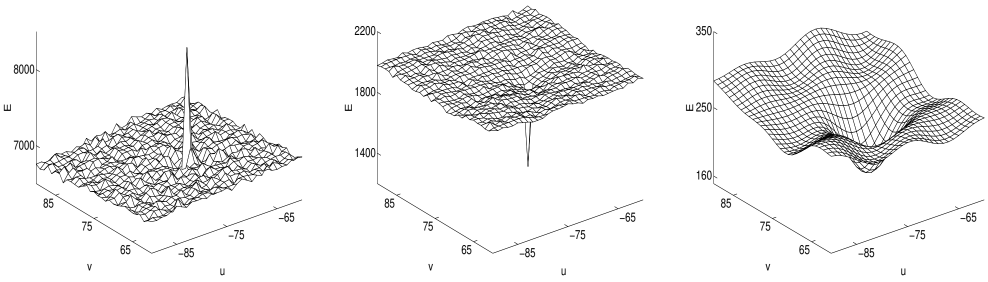
[7] B. K. P. Horn and B. G. Schunk. Determining optical flow. *Artificial Intell.*, 17:185–203, 1981.

[8] H. Li and Y. Zhou. Automatic visual/IR image registration. *Opt. Eng.*, 35(2):391–400, 1996.

[9] B. D. Lucas and T. Kanade. An iterative image registration technique with an application in stereo vision. In *Seventh Intl. Joint Conf. on Artificial Intelligence*, pages 674–679, 1981.

[10] R. K. Sharma and M. Pavel. Adaptive and statistical image fusion. *SID Digest of Technical Papers*, XXVII:969–972, 1996.

[11] Q. Zheng and R. Chellappa. A computational vision approach to image registration. In *Proc. of the Intl. Conf. on Pattern Recognition*, pages 193–197, The Hague, Netherlands, 1992.



(a) simple squared error (b) after transformation (c) after filtering

Figure 2: Error surfaces for Laplacian level 0.

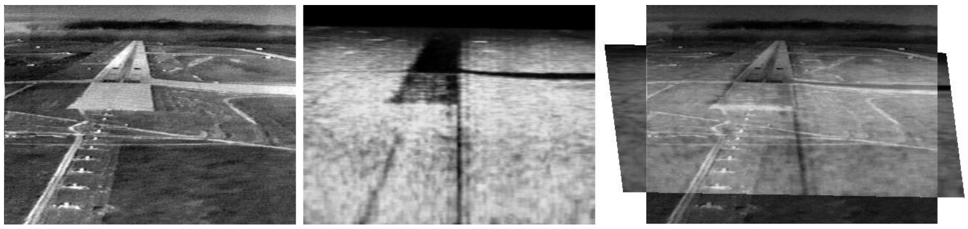


Figure 3: Registration of FLIR and radar images (Data from SVTD [4]).

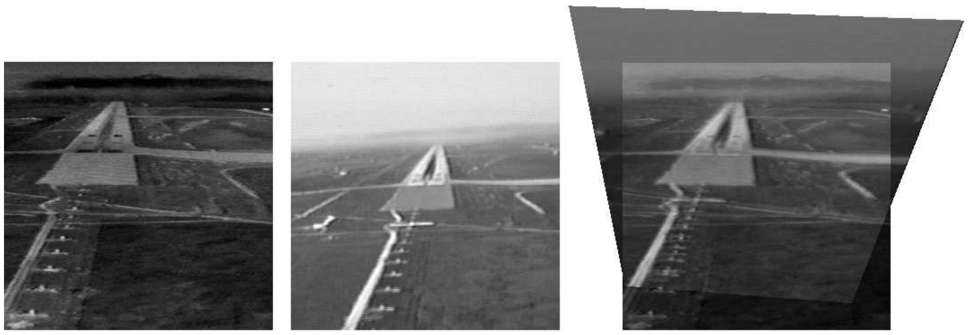


Figure 4: Registration of FLIR and visual images using projective model (Data from SVTD[4]).

Figure 5: Registration using affine model (Data from AMPS).