Extracting Affine Deformations from Image Patches - I: Finding Scale and Rotation

R. Manmatha and John Oliensis *
Computer Science Department
University of Massachusetts at Amherst, Amherst, MA-01002
manmatha@cs.umass.edu

Abstract

Image deformations due to relative motion between an observer and an object may be used to infer 3-D structure. Up to first order these deformations can be written in terms of an affine transform. Here, a novel approach is adopted to measuring affine transforms which correctly handles the problem of corresponding deformed patches. The patches are filtered using gaussians and derivatives of gaussians and the filters deformed according to the affine transform. The problem of finding the affine transform is therefore reduced to that of finding the appropriate deformed filter to use. In the special case where the affine transform can be written as a scale change and an in-plane rotation, the gaussian and first derivative equations have been solved for the scale. This case arises frequently in robot navigation. The robustness of the method has been demonstrated experimentally.

1 Introduction

Changes in the relative orientation of a surface with respect to a camera cause deformations in the image of the surface. Deformations can be used to infer local surface geometry and depth from motion. Since a repeating texture pattern can be thought of as a pattern in motion, shape from texture can also be derived from deformations [1].

Up to first order, there are four kinds of image deformations - a scale change, rotation in the image plane and shear along the x and y axes. The position of the surface's projection in the image also changes with motion. To first order, this image translation together with the deformation can be described using a six parameter affine transformation **A**where

$$\mathbf{Ar} = \begin{vmatrix} u_0 \\ v_0 \end{vmatrix} + \begin{vmatrix} u_x & u_y \\ v_x & v_y \end{vmatrix} \begin{vmatrix} r_x \\ r_y \end{vmatrix} \quad (1)$$

r_x and r_y are the image coordinates and u_0 and v_0 the image translation.

*This research was supported by the National Science Foundation under grants CDA-8922572 and IRI-9113690.

Consider the case where an image patch F_1 is deformed into a patch F_2 (either in the same image or in another image) by an unknown affine transform. The problem of measuring the affine transform is to first find the corresponding patch F_2 given F_1 and second to recover the affine parameters from the two patches. Even if the centroids of the image patches are matched, the precise size and shape of F_2 is difficult to determine since it is a function of the unknown deformation. If this correspondence is not precisely done, the affine parameters will be determined incorrectly. Thus the problem is more difficult than in standard correspondence problems e.g. the determination of optical flow. Existing methods usually ignore this problem.

1.1 Previous Work

Affine transforms between image patches have been recovered in three different ways (for more details see [1]):

1. By linearizing the image intensities or the filter outputs *with respect to the image coordinates.* Linearization limits these algorithms to cases when the affine transforms are small.

2. Matching image intensities by searching through the space of all affine parameters. This approach adopts a brute force search strategy which is potentially slow.

3. Line based methods which match the closed boundaries of corresponding regions.

Of these approaches, only the third deals with the problem of corresponding deformed patches. However, it is limited to homogenous regions with closed boundaries.

The rest of this paper assumes that the translation is already known or has been recovered (say using a traditional flow algorithm). For another approach to recovering translation, see [1].

2 Deformation of Filters

By applying deformed filters, the problem of corresponding deformed image patches is converted to the simpler problem of finding the appropriate deformed filter to use. For simplicity, it is assumed in this paper that the affine transform is given by $\mathbf{A} = s\mathbf{R}(\theta)$ i.e. a scale change s and a rotation θ. Suppose two image patches F_1 and F_2 are related by an affine deformation

$$F_1(\mathbf{r}) = F_2(\mathbf{A}\mathbf{r}). \quad (2)$$

Then the output of F_1 filtered with a gaussian is equal to the output of F_2 filtered with a gaussian deformed by the affine transform, i.e.

$$F_1 * G(., \sigma) = F_2 * G(., s\sigma), \quad (3)$$

where $G(., \sigma)$ denotes a gaussian with standard deviation σ, and * denotes a convolution at the origin (for a proof of this statement and a straightforward extension to arbitrary affine transforms see [1]). Intuitively, this equation implies that the gaussian-weighted brightnesses of the two patches must be equal.

Note that in equation (3) only the scale s appears since the gaussian is rotationally invariant. In applications such as robot navigation, the depth is directly related to the scale and may therefore be conveniently recovered. The equation is applicable to arbitrary dimensions.

Similar equations may be written for other derivatives of gaussians. For example the first and second derivative equations obtained by differentiating (3) are

$$F_1 * dG(., \sigma)/d\mathbf{r} = s\mathbf{R}F_2 * dG(., s\sigma)/d\mathbf{r} \quad (4)$$

and

$$F_1 * d^2G(., \sigma)/d(\mathbf{r}\mathbf{r}^T) = s^2\mathbf{R}^T F_2 * d^2G(., s\sigma)/d(\mathbf{r}\mathbf{r}^T)\mathbf{R} \quad (5)$$

All these equations are applicable even for large values of the affine transform. Also note that the first derivative equation is a vector equation while the second derivative equation is a matrix equation.

Equations (3-5) may be used to solve for the parameters of the affine transform.

3 Solution of the Gaussian Equation for Scale

To solve equation (3), for example, requires finding a gaussian of the appropriate scale $s\sigma$ given σ. One strategy is to search through the space of scale changes. This is a brute force approach and computationally expensive. A more elegant solution is to linearize the problem with respect to σ. This gives an equation linear in the unknown α

$$F_1 * G(., \sigma) \approx F_2 * G(., \sigma) + \alpha\sigma^2 F_2 * \nabla G(., \sigma) \quad (6)$$

where $s = 1 + \alpha$.

Equation (6) is not very stable if solved at a single scale. By using gaussians at several different scales σ_i an overconstrained linear system is obtained which can be solved using singular value decomposition (for details see [1]). It should be emphasized that the equations at several scales are solved simultaneously rather than in a coarse to fine approach.

The method differs from other linearization methods in that *it linearizes with respect to* σ *rather than the image coordinates* $\mathbf{r}$. The advantage of doing so is that the linearization point can be shifted (i.e the right-hand side of (6) can be linearized with respect to a σ different from the one on the left-hand side - see [1] for specifics). The correct linearization point is recovered by consistency checking. This allows affine deformations of arbitrary size to be dealt with.

3.1 Experiments

In implementations, filters at 5 different scales were used. The filter widths were 10,14,20,28 and 40 pixels (the σ was given by width = 8 * σ). Experiments were carried out on sine wave gratings as well as random dot images. These images were scaled and also rotated - the rotations being arbitrary. The results were within 5% for scale changes up to 1.8. The experiments were repeated with gaussian as well as uniform noise and the performance did not degrade significantly. An experiment on a real image was also carried out and the scale used to recover depth. The depths recovered were within a few percent.

4 Solving the First Derivative of Gaussian Filters for Scale and Rotation

The first derivative equations (4) can also be solved for scale by first taking the magnitude of both sides and then linearizing around σ. The resulting equations are no longer linear in α and are solved using the Gauss-Newton technique [1]. Having solved for scale, the rotation is then recovered. Again, the method gives good results.

References

[1] R. Manmatha and J. Oliensis. Measuring the affine transform – I: Scale and Rotation. Technical Report 92–74, University of Massachusetts at Amherst, MA, 1992. Also to be published in *Proc. of the Darpa Image Understanding Workshop 1993.*