Predictive Interpolation for Registration

D.G. Bailey

Institute of Information Sciences and Technology,
Massey University, Private bag 11222, Palmerston North
D.G.Bailey@massey.ac.nz

Abstract

Predictive interpolation is a new registration method that uses the bilinear interpolation equation to predict the pixel values of one image from those in another. The method requires that the images already be registered to the nearest pixel, which is not a serious limitation since hierarchical methods may be used to perform the initial registration. Registration accuracy of better than 5% of the width of a pixel can be expected between pairs of images, and this may be improved significantly if multiple images are to be registered to one another.

Keywords: interpolation, registration, linear prediction.

1. Introduction

Image fusion involves the combining of data from a number of image sources in such a manner that the output image contains more information than that contained in any of the individual images. The image sources are often different in their properties or characteristics. An important use of image fusion is to trade temporal resolution for spatial resolution, by combining several low-resolution images to construct a high-resolution image [1-3].

The pixels of an image provide a set of samples of the world. The sampling density, or the spacing between the pixels, limits the achievable resolution. However if a series of images is captured, each with the samples in slightly different locations, the combined sample density is higher that that of any single image. For the ensemble to contain more information than any one of its constituent images, a single image must not contain all of the information. This implies that the sample frequency for each individual image must be below the Nyquist rate, and the images be subject to aliasing [3]. The process of constructing a higher resolution image untangles the aliased information, so that the output image contains more information than that available from any of the individual input images.

An important step within this reconstruction process is to register the low-resolution images with one another so that they may be combined. A range of different image registration methods can be used, and the performance of these was reviewed at last year's conference [4]. Four types of pixel based registration methods were compared: correlation methods; difference methods; phase methods; and predictive interpolation. High level, segmentation based methods were not discussed or compared. It can be shown that the correlation and difference methods are closely related [5], and they gave similar performance for sub-pixel registration accuracy [4].

The predictive interpolation method was a new method, that is computationally efficient, gives very good registration accuracy in the presence of aliasing, and degrades relatively gracefully in the presence of moderate noise. The purpose of this paper is to present a more detailed description of this new method, and to give an overview of some of its properties.

2. Using interpolation for registration

To give sub-pixel accuracy, some form of interpolation is required. Consider at this stage, determining the offset between a pair of images. Also, assume that the images have already been registered to the nearest pixel, so all that is required is to refine the accuracy to within a pixel. This interpolation may be approached in one of two ways.

The first is to interpolate one (or both) images to be registered and use standard pixel-level interpolation methods (correlation or difference methods) to evaluate the fit. There are at least two problems with this approach. The first is the computational overhead of performing the image interpolation at each offset tested. The second is that since interpolation is a form of low pass filtering, there may be a loss of some fine detail in the interpolated image. For this reason, both images may be interpolated by the same amount but in opposite directions so the filter characteristics do not affect the accuracy. A second problem with interpolation is that since the images contain aliased information, the interpolation filter further distorts this information. Subsequent resampling of the interpolated images will not recover the aliased information.

The second approach to interpolation is to use pixel level registration, and evaluate the fit on an integer grid. This fit surface is then interpolated to estimate the position of optimum fit. An appropriate surface model must be selected to accurately find the best fit [4].

2.1 Predictive interpolation for registration

The predictive interpolation method turns the problem around. Rather than interpolate and then evaluate the best fit, the interpolation method attempts to predict the pixel values of one image from the other, and from that to infer what the offset is.

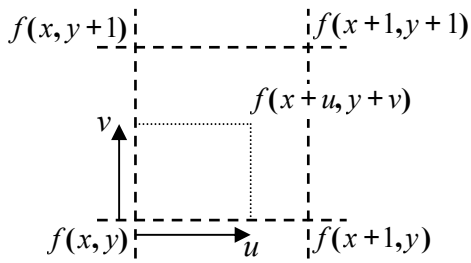

Figure 1: Interpolation on an integer grid.

Consider bilinear interpolation of an image f within an integer spaced grid as shown in figure 1. The value of the interpolated point, $f(x+u,y+v)$ is given by [6]

$$f(x+u,y+v) = (1-u)(1-v)f(x,y) + (1-u)vf(x,y+1) + u(1-v)f(x+1,y) + uvf(x+1,y+1) \quad (1)$$

Given that we have a pair of images, f and g, that have already been registered to the nearest pixel, we can use this relation to estimate the offset between images. The idea here is to try and predict the pixel values of one image from the other using a linear predictor:

$$g(x,y) = A_{00}f(x,y) + A_{01}f(x,y+1) + A_{10}f(x+1,y) + A_{11}f(x+1,y+1) \quad (2)$$

where the A_{ij} are the linear prediction coefficients. Further, from comparing equations (1) and (2), we want to apply the following contraint:

$$A_{00} + A_{01} + A_{10} + A_{11} = 1 \quad (3)$$

Substituting (3) into (2) gives:

$$\begin{aligned} g(x,y) - f(x,y) = {} & A_{01}\big(f(x,y+1) - f(x,y)\big) \\ & + A_{10}\big(f(x+1,y) - f(x,y)\big) \\ & + A_{11}\big(f(x+1,y+1) - f(x,y)\big) \end{aligned} \quad (4)$$

Least squares minimisation can then be used over the whole image to give the values of the coefficients that minimise the prediction error between the two images.

By equating equations (1) and (2) we have

$$A_{01} = v - uv \quad A_{10} = u - uv \quad A_{11} = uv \tag{5}$$

which can be solved to give the sub-pixel offset between the two images as

$$u = A_{10} + A_{11} \quad v = A_{01} + A_{11} \tag{6}$$

Therefore using bilinear interpolation filter as a prediction equation, we are able to obtain a direct estimate of the sub-pixel offset between two images.

2.2 Other considerations

In practise, simple application of least squares over equation (4) is not particularly effective because featureless regions within the images would give meaningless prediction coefficients, dominated by noise. This may be overcome by weighting the individual instances of equation (4) by the local variance. In this way regions which are relatively featureless are given little weight, while those regions that have more edge activity are given more weight.

One limitation of using predictive interpolation is that it requires the two images to be registered to the nearest pixel before it can be applied. There are two ways that this limitation may be overcome. The first is through scanning, where a range of different pixel level offsets is tried. The pixel level offset that has the minimum prediction error will also provide the sub-pixel offset. The second approach is to use a hierarchical method to determine the offset at coarser scales. With this approach, a series of increasingly lower resolution images can be created by averaging 2x2 blocks from the next higher resolution. Sub-pixel registration on the low-resolution images would indicate the nearest pixel-level registration for use at higher levels.

The procedure described in the previous section is computationally efficient when compared with the more conventional correlation and difference methods for sub-pixel registration.

3 Registration accuracy

To evaluate the accuracy of registration, it is necessary to work with a set of images where the offsets are known in advance. This was accomplished by taking a high-resolution image, and sub-sampling it to produce a series of low-resolution images with predetermined offsets. Predictive interpolation was then used to measure the offsets, and these results were compared with the known offsets to determine the registration error.

In the tests here, the original images were 512x512 pixels. These were filtered using a 3x3 linear filter and then sub-sampled by a factor of 2 by discarding every second pixel in the horizontal and vertical directions. Selecting the filter weights controlled the sub-pixel offset of each low-resolution image. This allowed a set of 256x256 images to be created that were offset by multiples of 5% of a pixel in both the horizontal and vertical directions.

The sample images that were used are shown in figure 2. The images were chosen for their varying level of detail, from a fully structured image (A), two general images (B and C) with varying levels of detail, a high resolution text image (D) and a low resolution text image (E) with inadequate resolution to be able to read the text. The final image is expected to present a considerable challenge because the fine detail within the image will result in considerable aliasing.

Figure 2: The sample low resolution images.

3.1 Registration error

To determine the registration error, each pairwise combination of images was taken, and the offset between the images was predicted using the central 100x100 pixels from the images. The RMS error between the predicted offsets and the true offsets was used as a measure of registration performance. The average registration errors are listed in table 1 for the sample images used.

Image	A	B	C	D	E
Error	3.48%	4.56%	4.28%	4.59%	8.65%

Table 1: RMS registration error for the different images as a fraction of a pixel.

Preliminary investigations [4] indicated that there might be a systematic bias in the errors. To check for the presence of such a bias, the horizontal and vertical components of the error were plotted as a function of the known offsets between the images. With horizontal and vertical offsets in steps of 5% of a pixel, this gave a clear indication of the form of the errors. Two components of the errors were determined: the systematic bias (shown in figure 3) and the random component (shown in figure 4).

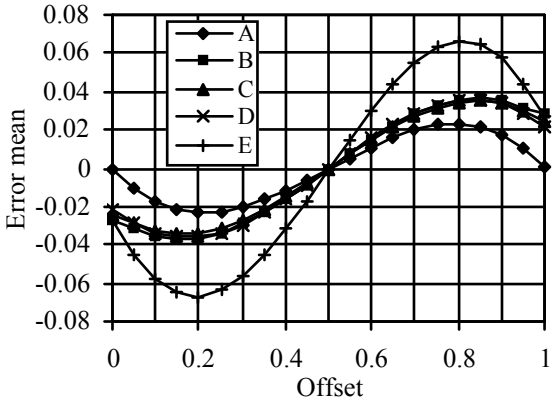

Figure 3: Non-random bias component. Error mean as a function of offset. Both scales are in pixels.

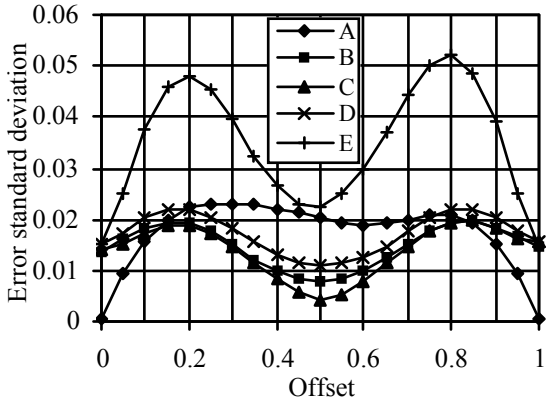

Figure 4: Random error component. Standard deviation as a function of offset. Both scales are in pixels.

3.2 Effect of window size

To investigate the effects of the window size on the two error components, the experiment was repeated, varying the size of the window used for registration from 10x10 pixels up to 160x160 pixels.

The systematic bias component did not vary significantly with window sizes greater than 20x20. This implies that the systematic bias is not introduced by the windowing, but by the method itself. Unfortunately, the bias appears to be image dependent (at least to some extent as illustrated in figure 3), so cannot be removed without knowing in advance what its level is.

For small window sizes, the random component of the error decreased with increasing window size, as would be expected. However, for window sizes above 60x60 to 100x100 (depending on the image), the reduction in error was negligible. For small window sizes, the error is distributed uniformly as a function of offset, with a slight peak at about 0.5 pixel. As the window size is increased, the double peaked nature of the error, as seen in figure 4, becomes pronounced. The error at 0.5 pixel offset reduces significantly while the peaks at 0.2 and 0.8 pixel offsets do not decrease significantly with windows sizes larger than 100x100 pixels. This implies that the limit of the technique is being reached and the variability remaining is inherent in the technique.

3.3 Enforcing consistency

The results presented so far in this paper are based on pairwise registrations. If more than two images are to be registered relative to one another, it is possible to reduce the errors between the pairs of images by enforcing consistency. If there are N images being registered then there are $\frac{1}{2}N(N-1)$ separate registration measurements. Enforcing consistency determines the $N-1$ independent offsets in such a way that the error with respect to the measurements is minimised. If M_{ij} is the measured offset between image i and image j, then the minimised offsets with respect to the first image are given by

$$\overline{M_{i1}} = \frac{1}{N}\left(\sum_{j=2}^{N} M_{j1} + \sum_{j=1}^{i-1} M_{ij} - \sum_{j=i+1}^{N} M_{ji}\right) \tag{7}$$

If the registration errors are independent and normally distributed, then the average error will be reduced by a factor of $\sqrt{2/N}$. Results presented at last year's conference [4] indicated that the average error was reduced by significantly more than this factor. The reason for this is the particular pattern of bias measured in section 3.1. In [4] the low resolution images were offset in the horizontal and vertical directions by exactly 1/3 and 2/3 pixels. In enforcing consistency, the

biases cancelled exactly, leaving only the random component of the error. As a consequence, the error improved significantly better than was expected from enforcing consistency.

A more realistic study to investigate the effects of enforcing consistency would be to create a set of images with random (but known) relative offsets, and measure the reduction in average error that results from enforcing consistency. The results shown in table 2 are the average improvements resulting from 10 runs on each image with a 100x100 window.

Image	A	B	C	D	E	Expected improvement
4 images	0.7190	0.7140	0.7160	0.7117	0.7109	0.7071
8 images	0.6992	0.5059	0.5069	0.5154	0.5439	0.5000
16 images	0.6352	0.3642	0.3605	0.4152	0.4302	0.3536

Table 2: The reduction in average error resulting from enforcing consistency.

In these experiments, the improvement in accuracy as a result of enforcing consistency was almost to the expected level. The notable exception to this was image A. It is surmised that because there are relatively few edges in this image that gathering more data does not improve the accuracy significantly.

4 Summary and Conclusions

Predictive interpolation is an efficient method for determining the relative offset between pairs of images. The accuracy of the method between pairs of images is typically 5% of a pixel or better. When more than two images are being registered relative to one another, the average accuracy may be further improved by enforcing consistency within the offsets of an image set.

There are two components to the error: a systematic bias component, and a random component. The bias component is independent of window size above 20x20, and the random component is independent of window size above 100x100 pixels. Therefore, unless there is noise present, the optimum window size for matching is approximately 100x100 pixels.

A flaw has been identified in the previous experiments [4] that resulted in a better than expected improvement from enforcing consistency. Enforcing consistency over multiple images with random offsets gives the expected improvement in accuracy.

References

[1] R.C. Hardie, T.R. Tuinstra, J. Bognar, K.J. Bernard, and E. Armstrong: High resolution image reconstruction from digital video with global and non-global scene motion. *Proc. IEEE International Conference On Image Processing*, Vol I (October 1997) 153-156.

[2] M.C. Hong, M.G. Kang, A.K. Katsaggelos: An iterative weighted regularized algorithm for improving the resolution of video sequences. *Proc. IEEE International Conference On Image Processing*, Vol II (October 1997) 474-477.

[3] D.G. Bailey: Super-resolution of bar codes. Proc SPIE Vol 3521, Machine Vision Systems for Inspection and Metrology VII (November 1998) 204-213

[4] D.G. Bailey and T.H. Lill: Image Registration Methods for Resolution Improvement, *Proceedings of Image and Vision Computing NZ* (August 1999) 91-96

[5] A.K. Jain: *Fundamentals of Image Processing*. Prentice Hall, Englewood Cliffs, New Jersey (1989) 400-407.

[6] G. Wolberg: *Digital Image Warping*. IEEE Computer Society, Los Alamitos, CA (1990) 58-59.