

Noise Characteristics of Higher Order Predictive Interpolation for Sub-pixel Registration

Andrew Gilman, Donald G Bailey, Stephen R Marsland

Institute of Information Sciences and Technology
Massey University

Palmerston North, New Zealand

agilman@iee.org, D.G.Bailey@massey.ac.nz, S.R.Marsland@massey.ac.nz

Abstract - Sub-pixel registration has application in many image processing tasks. Predictive interpolation, a novel registration technique, solves the problems of choosing a particular interpolation function and needing to search for the best offset. Predictive interpolation determines the optimum interpolation function for a given pair of images, and estimates the offset from the interpolation weights. The estimate of the offset between the images is biased, and this bias depends strongly on any noise present in the image. It is shown that the bias resulting from the noise is opposite from the bias from the image. This leads to the counter-intuitive result that the registration accuracy can improve significantly (by a factor of 10 for a second order filter) with the addition of moderate amounts of noise. A 5th order filter is accurate to better than 0.5% of a pixel over a wide range of noise levels. These results are verified by measuring the accuracy of registration on sample images.

Keywords - sub-pixel, super-resolution, registration, motion estimation, interpolation, linear prediction, imaging model, noise

A. Introduction

Image registration is an important step in many image processing applications. While pixel accurate registration is adequate for many applications, many techniques can benefit from registration to sub-pixel accuracy. These include: super-resolution [1]; motion compensation for video coding [2]; sensor fusion [3]; stereo imaging; image stitching [4]; motion detection using optical flow [5]; and image stabilisation.

Given two images (or subimages in some applications), f and g , which differ only by a translation (rotation and scaling are not considered in this paper), registration involves estimating the offset between the pixel locations of the two images. The general approach is to designate one of the images as the reference, and measure the offset of the other image relative to this reference.

Sub-pixel registration requires estimating the offset to an accuracy of a small fraction of a pixel. Often this takes place in two steps: first the images are registered to the nearest pixel, and then the offset within the pixel is estimated. In this paper, it is assumed that the images have been pre-registered to the nearest pixel (using any method described in [6] or [7]).

A novel linear approach is described for estimating the sub-pixel offset that avoids the non-linear minimisation associated with conventional methods. The accuracy with which the sub-

pixel offset may be estimated with this approach has been examined in previous work in the absence of noise [13]. The focus of this paper is to present a simplified analytic analysis of the effects of noise on the systematic bias associated with offset estimation and to verify the analytic results using experimental image data.

For simplicity, the analysis here will be restricted to 1-D (although it easily generalises to two and higher dimensions). Let $f(x)$ represent the pixel values for the reference image. The target image $g(x)$ is offset from f by u such that $f(x+u)$ is in registration with $g(x)$, i.e.

$$f(x+u) = f_u = g(x). \quad (1)$$

In practise, all imaging systems introduce noise from a range of sources: imaging (shot) noise, amplifier noise, quantisation noise, and so on. In the analysis here, we assume that the total combination of noise from all the sources can be considered as additive white Gaussian noise with zero mean and standard deviation σ . Let ξ and ζ to be noise vectors, such that:

$$\begin{aligned} \tilde{f}(x) &= f(x) + \xi \\ \tilde{g}(x) &= g(x) + \zeta \end{aligned} \quad (2)$$

The sub-pixel registration problem is then to estimate u given the noisy images \tilde{f} and \tilde{g} , where $0 < u < 1$. The performance outside this range is also of interest because of the possibility of error in the pixel-level registration.

Most analyses of image processing operations, including registration, assume that the images are band-limited. In particular, the Nyquist sampling criterion requires that the highest sinusoid frequency component within an image is less than twice the sampling frequency to prevent aliasing. In practice, many images of real world objects contain some degree of aliasing by virtue of the fact that objects have sharp edges or boundaries. Natural objects in particular have detail at a wide range of scales so will contain energy over a broad bandwidth. The only bandwidth limiting elements within image capture systems are the optical transfer function of the lens, and area sampling performed by the sensor. Therefore some degree of aliasing is inevitable. In fact, applications such as super-resolution require that the input images be aliased in

order to obtain more information from the image ensemble than is available within any single image [8]. Consequently, any sub-pixel image registration technique should be insensitive to, or at least tolerant of aliasing.

The rest of this paper is structured as follows: Section II defines predictive interpolation and derives the optimal interpolator. This is then used to estimate the sub-pixel offset between the target and reference images. The bias associated with the estimated offset, and the effects of noise on this bias are analysed in section III for step image models and different order predictive interpolation filters. Section IV compares the bias observed from real data with that from the theoretical analysis. The implications are discussed in Section V.

II. PREDICTIVE INTERPOLATION

While there are many different registration methods (see for example [1,6,7,9]), this paper focuses on a relatively new technique: predictive interpolation. This was first introduced in [1] and analysed in more detail in [10] and [11].

Conventional interpolation based methods for sub-pixel registration use an interpolator to create a continuous surface from the reference image samples. The continuous surface is then offset (and effectively resampled) and compared with the target image. The sub-pixel offset is determined by searching for the offset that gives the best fit with the target image:

$$u = \arg \min_u \left\| \tilde{g}_0 - \tilde{f}_u \right\|^2. \quad (3)$$

Interpolation can be considered as a linear filter. The resampled, offset reference image may be formed by using a sampled interpolation function, resulting in a discrete filter:

$$\begin{aligned} f_u &= \dots + h_{-1}\tilde{f}_{-1} + h_0\tilde{f}_0 + h_1\tilde{f}_1 + h_2\tilde{f}_2 + \dots \\ &= \sum_{i \in W} h_i \tilde{f}_i \end{aligned} \quad (4)$$

where W is the region of support for the interpolation kernel, and the filter weights h_i , depend on the desired offset u . Different interpolation kernels are derived based on making different assumptions about the image. As a result, they may have different regions of support, and will give different sets of weights derived from those assumptions.

Solving (3) is non-trivial because the difference between the target and reference image is a non-linear relationship of the offset, u , and doesn't necessarily have well-defined gradients. Consequently, the offset is usually found through iterative optimisation techniques.

Predictive interpolation avoids this search by turning the problem around. It does not choose a particular interpolation function, but instead uses the image data itself to determine the interpolation kernel. It uses the pixel values in the reference image to predict those in the target image, effectively determining the 'best' (that is, optimal in a least squares sense [12]) interpolation kernel ($h_i, i \in W$) that relates the target to the reference:

$$\tilde{g}_0 \approx \sum_{i \in W} h_i \tilde{f}_i. \quad (5)$$

Since (5) is linear, the optimal filter coefficients may be easily determined by least squares minimisation. As an interpolator, the prediction weights are subject to the constraint

$$\sum_{i \in W} h_i = 1. \quad (6)$$

If this was not the case, then the intensity within uniform regions would change. Eliminating h_0 from (5) gives:

$$\tilde{g}_0 - \tilde{f}_0 \approx \sum_{i \in W, i \neq 0} h_i (\tilde{f}_i - \tilde{f}_0). \quad (7)$$

Least squares minimisation can then be used over the whole image to give the values of the coefficients that minimise the prediction error from f to g . If we define

$$\begin{aligned} \tilde{g} &= \tilde{g}_0 - \tilde{f}_0 \\ \tilde{f}_i &= \tilde{f}_i - \tilde{f}_0 \end{aligned} \quad (8)$$

then (3) transforms to

$$h_i = \arg \min_{h_i, i \in W, i \neq 0} \sum_{\text{pixels}} \left\| \tilde{g} - \sum h_i \tilde{f}_i \right\|^2, \quad (9)$$

which has well defined gradients, and being quadratic in h_i , has a single global minimum that can be determined analytically. Taking the partial derivative with respect to each coefficient and solving for when these derivatives are equal to 0 gives a system of linear equations. The order of the interpolator is given by the number of independent coefficients fitted (which is 1 less than the size of W). For example, for a 3rd order predictive interpolator:

$$\begin{bmatrix} \sum \tilde{f}_{-1}^2 & \sum \tilde{f}_{-1}\tilde{f}_1 & \sum \tilde{f}_{-1}\tilde{f}_2 \\ \sum \tilde{f}_1\tilde{f}_{-1} & \sum \tilde{f}_1^2 & \sum \tilde{f}_1\tilde{f}_2 \\ \sum \tilde{f}_2\tilde{f}_{-1} & \sum \tilde{f}_2\tilde{f}_1 & \sum \tilde{f}_2^2 \end{bmatrix} \begin{bmatrix} h_{-1} \\ h_1 \\ h_2 \end{bmatrix} = \begin{bmatrix} \sum \tilde{f}_{-1}\tilde{g} \\ \sum \tilde{f}_1\tilde{g} \\ \sum \tilde{f}_2\tilde{g} \end{bmatrix} \quad (10)$$

or more generally

$$\tilde{\mathbf{F}}\mathbf{h} = \tilde{\mathbf{g}}. \quad (11)$$

This is then solved to give the optimal interpolation weights. The problem then is how to determine the offset from the interpolation coefficients. In previous work [11], we have matched the weights to those that would be obtained using conventional low order polynomial interpolators. This was valid for a first order filter (linear interpolation in 1-D) where the equations are the same. However, for higher orders the optimal interpolation function does not necessarily correspond with any standard interpolation method, so matching coefficients is harder to justify.

If we consider integer offsets, i.e. $g_0=f_i$, then in the absence of noise the optimum interpolation procedure will derive the corresponding filter coefficient, h_i , to be 1, and all of the remaining coefficients to be 0. In the general case, however, the target image is partway between pixels, so a mixture of reference pixels is used to predict the target. If we assume that the offset is a linear combination of the filter coefficients

$$\hat{u} = \sum_{i \in W} w_i h_i \quad (12)$$

then from the super-position principle, we get the weights as

$$w_i = i. \quad (13)$$

Note that this is exactly the same as the result obtained by matching the weights produced for a polynomial interpolator (for example a spline with finite support) and solving the resulting equations for the offset, u . It can be shown that (13) holds regardless of the particular interpolation kernel used.

The fact that h_0 makes no contribution to the offset, is why it is convenient to eliminate h_0 rather than one of the other weights in (7). Combining (11) and (12) we get

$$\hat{u} = \mathbf{w} \tilde{\mathbf{F}}^{-1} \tilde{\mathbf{g}}, \quad (14)$$

where \mathbf{w} is a row vector made of the sample locations within the region of support, eliminating the origin ($i=0$). Note that the $\tilde{\mathbf{F}}^{-1}$ term depend only on the reference image, so the matrix inverse only needs to be calculated once regardless of the number of target images that are registered. The rows of this inverse are then weighted (by \mathbf{w}) giving a single row vector that depends only on the reference image. This leads to an efficient implementation when there are a number of target images that need to be registered relative to one another (for example in super-resolution).

III. EFFECT OF NOISE ON BIAS

Any estimate of the offset between the two images will be subject to uncertainties resulting from the characteristics of the images. There will also be uncertainties resulting from underlying differences between the two images (for example noise or aliasing). It has been demonstrated [10] that when performing predictive interpolation with a sufficiently large image area the bias in the estimate dominates over the variance, making the bias of particular interest. The bias in the estimator can be defined as any systematic deviation of the measured offset from the actual offset:

$$\text{Bias} = E[\hat{u}] - u. \quad (15)$$

To obtain an analytic expression of the bias, it is necessary to have a mathematical model of the features within the images. Many images can be approximated by piecewise constant regions (with step edges in between). In forming the image, a step edge is blurred to a single pixel wide ramp by area sampling. This means that the edge pixels take on an intermediate value depending on the exact position of the edge relative to the sampling grid. Without loss of generality, in the following analysis we will consider an image with a single step edge of height 1.

Since the location of the edge is unknown relative to the sampling grid position, the expected value is calculated by assuming that all possible relationships are equally likely within the image. The sum in (9) is replaced by an integral over all possible edge positions within the region of support of the interpolator:

$$h_i = \arg \min_{h, j \in W, i \neq 0} \int_{\text{support region}} \left\| \tilde{\mathbf{g}} - \sum h_j \tilde{\mathbf{f}}_j \right\|^2. \quad (16)$$

This effectively considers, and averages over, all of the different configurations of the pixel values as a result of an edge being somewhere within the region of support. The advantage of explicitly enforcing the constraint of (6) is that there is no contribution from edges outside the region of support (because of the subtraction in (7)).

The elements of $\tilde{\mathbf{F}}$ (and similarly $\tilde{\mathbf{g}}$) are of the form:

$$\begin{aligned} \int \tilde{\mathbf{f}}_i \tilde{\mathbf{f}}_j &= \int (\tilde{f}_i - \tilde{f}_0)(\tilde{f}_j + \tilde{f}_0) \\ &= \int ((f_i + \xi_i) - (f_0 + \xi_0))((f_j + \xi_j) - (f_0 + \xi_0)) \end{aligned} \quad (17)$$

Expanding and taking expectations:

$$\begin{cases} \sum_{\text{pixels}} (f_i - f_0)(f_j - f_0) + \sigma^2 = \sum_{\text{pixels}} f_i f_j + \sigma^2 & \text{for } i \neq j \\ \sum_{\text{pixels}} (f_i - f_0)^2 + 2\sigma^2 = \sum_{\text{pixels}} f_i^2 + 2\sigma^2 & \text{for } i = j \end{cases}, \quad (18)$$

shows that the elements of $\tilde{\mathbf{F}}$ and $\tilde{\mathbf{g}}$ are sums of two components: one is data dependent and one is noise. The elements of $\tilde{\mathbf{g}}$ and the off-diagonal elements of $\tilde{\mathbf{F}}$ contain σ^2 , while the diagonal elements of $\tilde{\mathbf{F}}$ contain $2\sigma^2$. The consequence of this is that as the noise begins to dominate, the interpolation weights will tend to become equal regardless of the actual image offset. This will lead to an estimated offset of 0.5 pixels for odd order filters and 0 pixels for even order filters, regardless of the actual offset.

A. 1st Order

In previous work [11], we have analysed performance of the first order predictive interpolation on a step edge in presence of noise for interval $0 \leq u \leq 1$. The effect of noise was to add an additional term that depends on noise to the numerator and denominator.

$$E[\hat{u}] = \frac{\frac{3}{4}u + \frac{3}{4}u^2 - \frac{1}{2}u^3 + 3\sigma^2}{1 + 6\sigma^2} \quad (19)$$

For low or no noise the bias is dominated by the signal, but when σ is large, the expected value of the offset estimate tends towards 0.5.

It was shown that the bias in the absence of noise is towards the nearest grid location and is away from the centre of the interpolation region of support; whereas the bias introduced by the noise is in the opposite direction i.e. towards the centre. At intermediate levels of noise, the bias partially cancels out.

B. 2nd Order

Second order predictive interpolation has extended region of support of two pixels ($-1 < u \leq 1$). For the step edge model,

the expected value of the estimated offset for second order estimator is:

$$E[\hat{u}] = \frac{u - u^3 / 6}{\sigma^2 + 5/6}, \quad -1 \leq u \leq 1. \quad (20)$$

The bias from the deterministic signal can be found by substituting zero for σ in (20) and subtracting u :

$$\text{Bias} = E[\hat{u}] - u = \frac{u - u^3}{5}, \quad -1 \leq u \leq 1. \quad (21)$$

From (21) it can be seen that the bias from the deterministic signal is towards the edge of the region of support and away from the centre (i.e. u is overestimated when negative and underestimated when positive). This is similar to the result from the first order interpolator. Likewise, equation (20) indicates that bias due to noise is towards the centre of the region of support (as σ increases, the estimated offset tends toward zero). This creates a partial cancellation of bias for intermediate levels of noise. This can be clearly seen from Fig. 1, which shows the bias curve for a number of noise levels.

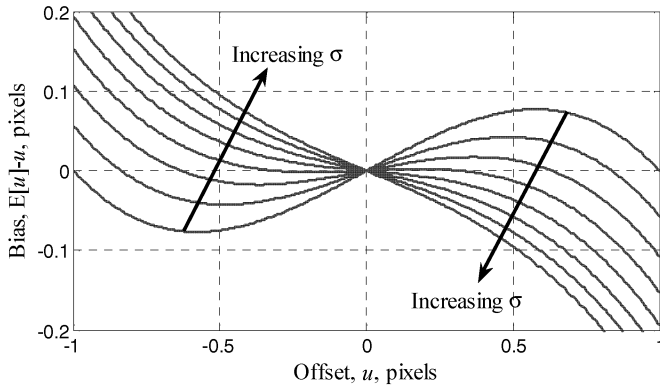


Fig. 1. Bias in estimating the offset using 2nd order predictive interpolation for a range of noise levels: $\sigma = 0:0.05:0.35$.

C. Higher Orders

We investigated predictive interpolation filters up to fifth order. Since the primary interest is sub-pixel registration, the bias is averaged over an offset interval of 1 pixel (to give indication of the performance on average) by calculating the RMS bias over that interval. Assuming the images are already pre-registered to within one pixel, for odd order filters offsets in the range of $0 \leq u \leq 1$ were considered. For even order filters, the symmetry about 0 suggests that it is more appropriate to consider offsets in the range $-0.5 \leq u \leq 0.5$. The results are shown in Fig. 2 for noise standard deviations up to the edge height.

As observed previously, the addition of modest levels of noise can significantly reduce the bias because the noise induced bias is opposite that from the image. This effect is even more marked with the higher order filters, with the bias being completely cancelled within the sub-pixel range for

$\sigma = \sqrt{1/6}$. Bias function for these higher orders is a fraction of two polynomials:

- The denominator is a function of σ^2 and is always positive,
- The numerator is a polynomial function of u and σ^2 ,
- For third and higher order filters $\sigma^2 = \sqrt{1/6}$ is a root of the numerator, making the bias zero at this level of noise for any offset, u .

Both the 3rd and 5th order interpolation filters give excellent (better than 1% of a pixel) registration performance over a wide range of noise levels. At higher noise levels, the higher the order of the filter, the better the performance.

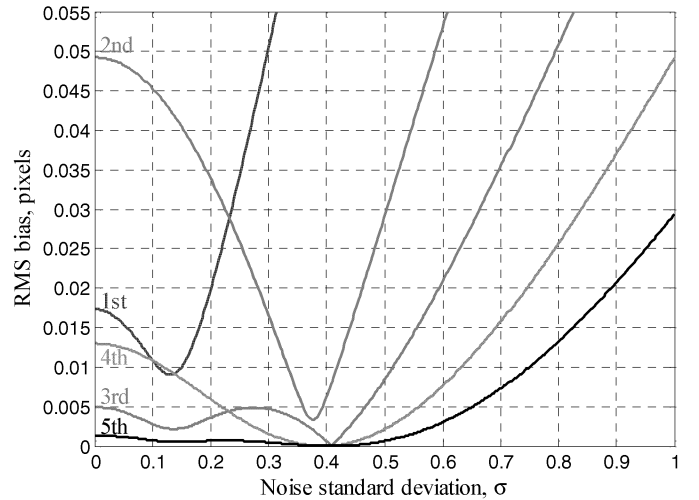


Fig. 2. The effect of noise on bias.

IV. VALIDATION WITH REAL IMAGES

To assess the registration accuracy, it is necessary to work with images where the offset is known in advance. Capturing a sequence of images with precisely known offsets is difficult, if not impossible. Therefore, a high-resolution image was used as the image source, and a simple imaging model was used to simulate the capture of the sample images. A 1700x1700 source image was filtered using a 1x20 horizontal box average filter to simulate 1D area integration. Shifting this blurred image by an integer number of pixels and sub-sampling horizontally by a factor of 20 (to 1700x85) produces a series of low-resolution images with known offsets in steps of 0.05 pixels. By taking different offset images as a reference, the above scheme provides 20 pairs of low-resolution images for each sub-pixel offset.

To calculate the bias for each sub-pixel offset, each of the 20 pairs of images had white Gaussian noise of required variance added to them and registered. This was repeated 10 times for each pair with different noise instances giving a total of 200 measurements of each offset. These measurements were averaged, and the known offset subtracted to give the bias. Finally the RMS average was obtained of the bias

measurements over the range of offsets of interest ($0 \leq u \leq 1$ for odd order filters and $-0.5 \leq u \leq 0.5$ for even orders).

Fig. 4 shows the measured bias obtained from the test image “Beach” (Fig. 3) as a function of added noise. The pattern of the bias is very similar to that obtained analytically from the step edge model (shown in Fig. 2). This implies that the piecewise constant model with area sampling provides a reasonable representation of the dominant image characteristics in terms of registration using predictive interpolation. The bias from the test image is slightly higher than that estimated from the model because the model does not exactly match the characteristics of the image.



Fig. 3. Sample low-resolution image, “Beach”, dominated by low frequencies, with some sharp edges resulting in a limited degree of aliasing. Here the image was downsampled by a factor of 20 in both directions for displaying purposes.

The minima in the bias characteristics with noise occur at different levels of standard deviation compared with the model. This is because the effect of noise will depend on both the number of significant edges within the image, and their height. It is the noise standard deviation relative to the edge height that is significant in determining the noise level that gives the minimum bias. This makes it difficult to know how much noise to add to an unknown pair of images to eliminate the registration bias.

V. DISCUSSION AND CONCLUSIONS

Different order filters have shown to have quite different noise characteristics. Comparing the performance of different order filters in the absence of noise, as done in [13], offers a helpful insight into the bias mechanism, but cannot give a definitive answer which filter is more favourable for a particular application. For example, the first order interpolator is significantly better than the second order filter for low noise conditions, yet it performs very poorly at moderate noise; and both of these are extremely poor in a high noise situation. A similar trade-off can be seen with the third and fourth order filters, where the third order interpolator performs better at lower noise and is worse at higher noise levels.

If the noise level is unknown, or possibly a whole range of noise levels can be encountered in the images to be registered – using a third or fifth order filter may be best, as these have low (less than 1% of a pixel) bias over a wide range of noise levels. At high noise levels, the higher the order of the filter

the better the performance. This can be explained by a wider region of support of the interpolation kernel, which offers additional averaging. This is similar to passing the signal through a low-pass filter with a lower cut-off frequency. As the natural images tend to have $1/f$ frequency content and the noise is white, the signal to noise ratio is increased.

This improved performance of higher order filters at higher noise levels comes at the expense of an increased computational cost. The most time consuming step is performing the summations to form $\tilde{\mathbf{F}}$ and $\tilde{\mathbf{g}}$. The number of elements calculated grows with the square of the filter order. This is offset by the fact that only a single pass is required through each of the reference and target images. A further advantage of using higher order filters is the initial pre-registration of the images to the nearest pixel is also relaxed.

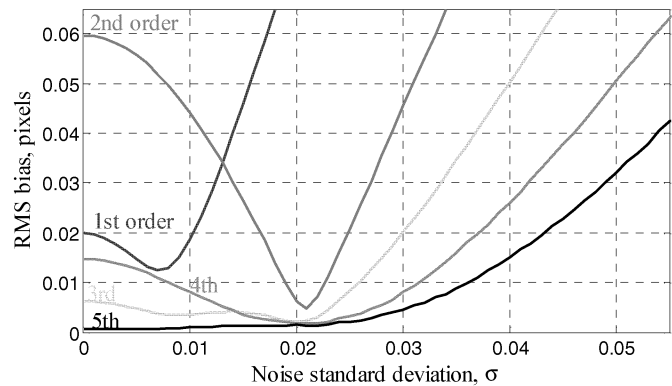


Fig. 4. Bias characteristics measured from the “Beach” test image for different order interpolation filters (first to fifth order).

The results have been validated with a typical scene, giving results that closely match those obtained theoretically from a simple piecewise constant with area sampling image model. This implies that such a model closely approximates the important image characteristics from the point of view of registration. The method needs to be tested with a wider range of images, particularly to investigate the effects of significant aliasing on registration accuracy. Preliminary experiments indicate that when significant aliasing is present, the bias increases and the bias pattern changes subtly.

The analysis here needs to be extended from 1-D to 2-D registration. While in principle the predictive interpolation filters are straight forward to implement, in general the optimal filters are not separable, making the analysis significantly more complex. For example the 2-D equivalent of a 1-D 3rd order filter has 15 degrees of freedom (independent filter coefficients).

For lower order filters, it would be useful to have a method of estimating the level of noise that needs to be added to minimise the bias. The bias for the second order filter in particular improves by a factor of 10 with the right level of noise added.

To conclude, predictive interpolation offers a way to register translated images in the presence of noise with very

little bias (less than 0.5% of a pixel for 5th order for a wide range of noise levels). The method also offers good flexibility in the trade-off between performance and computational cost via a selection of different orders, which may be more appropriate for a particular application.

REFERENCES

- [1] D.G. Bailey and T.H. Lill, "Image Registration Methods for Resolution Improvement", in *Image and Vision Computing NZ*, Christchurch, New Zealand, 91-96 (30-31 August, 1999).
- [2] G. Dane and T.Q. Nguyen, "The effect of global motion parameter accuracies on the efficiency of video coding", in *IEEE International Conference on Image Processing*, **5**: 3359-3362 (2004).
- [3] R. Sharma and M. Pavel, "Multisensor image registration", in *Society for Information Display*, **XXVIII**: 951-954 (1997).
- [4] C.Y. Chen and R. Klette, "Image Stitching - Comparisons and New Techniques", in *8th International Conference on Computer Analysis of Images and Patterns, Lecture Notes in Computer Science*, **1689**: 615-622 (1999).
- [5] J.L. Barron, D.J. Fleet, S.S. Beauchemin, and T.A. Burkitt, "Performance of optical flow techniques", in *1992 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '92)*, 236-242 (1992).
- [6] L.G. Brown, "A Survey of Image Registration Techniques", *ACM Computing Surveys*, **24**:(4) 325-376 (1992).
- [7] B. Zitova and J. Flusser, "Image Registration Methods: a Survey", *Image and Vision Computing*, **21**: 977-1000 (2003).
- [8] D.G. Bailey, "Super-Resolution of Bar-Codes", *Journal of Electronic Imaging*, **10**:(1) 213-221 (2001).
- [9] Q. Tian and M.N. Huhns, "Algorithms for sub-pixel registration", *Computer Vision, Graphics and Image Processing*, **35**:(2) 220-233 (1986).
- [10] D.G. Bailey, "Predictive Interpolation for Registration", in *Image and Vision Computing Conference NZ*, Hamilton, New Zealand, 240-245 (27-29 November, 2000).
- [11] D.G. Bailey, A. Gilman, and R. Browne, "Bias Characteristics of Bilinear Interpolation Based Registration", in *IEEE Region 10 Conference (IEEE Tencon'05)*, Melbourne, Australia, (21-24 November, 2005).
- [12] A. Gilman and D.G. Bailey, "Near optimal non-uniform interpolation for image super-resolution from multiple images", in *Image and Vision Computing New Zealand (IVCNZ'06)*, Great Barrier Island, NZ, 31-36 (27-29 November, 2006).
- [13] D.G. Bailey and A. Gilman, "Bias of Higher Order Predictive Interpolation for Sub-pixel Registration", appearing in *Sixth International Conference on Information, Communications and Signal Processing*, Singapore, (10-13 December 2007).