

GRADIENT CROSS CORRELATION FOR SUB-PIXEL MATCHING

N. A. Campbell and X. Wu

CSIRO Mathematical and Information Sciences, 65 Brockway Road, Floreat, WA, 6014, Australia
Phone +61 8 9333 6162, Fax +61 8 9333 6121, Xiaoliang.Wu@csiro.au

WgS-PS, WG VII/6

KEY WORDS: Correlation, Matching, Registration, Fusion, Multisensor

ABSTRACT:

Sub-pixel matching is one of the key components for image registration and image fusion. Ideally, image matching should allow for offsets in the target image, and for scaling and rotation. Offsets allow for sub-pixel shifts in the two images, while scaling is necessary when matching images from different sensors or images taken from different distances using the same camera. Rotation allows for the matching between rectified and un-rectified images or images taken from different viewpoints.

This paper presents a novel matching method called gradient cross correlation, which has been derived from the well-known normalised cross correlation coefficient formulation. In experimental evaluations, the method has been applied to image matching from various satellites (Landsat MSS and TM, QuickBird and other sensors). For comparison, an alternative method for estimating sub-pixel shifts and scaling and orientation parameters was applied – the least squares matching method. The mathematical details of the gradient cross correlation method, the experimental results, and some aspects of how to implement the approach in practice will be described and discussed in this paper.

Various models of the gradient cross correlation have been derived from the relationships between the affine transformation parameters. A hierarchy of relationships between the affine transformation parameters can be specified in practice as follows: Model-I: different scale, different rotation; IIA: different scale, common rotation; IIB: common scale, different rotation; III: common scale, common rotation; and IV: fixed scale, fixed rotation.

These models lead to a more natural interpretation of the resulting parameters, especially when matching images which have considerable scaling and rotation differences. The particular formulation of the affine transformation adopted leads to useful insights into the image matching. The experiments showed that Model-IV is the worst model for matching all kind of points. It is essential to choose an appropriate geometric transformation depending on different image characteristics and types of points.

The gradient cross correlation method and the least squares matching method with an offset and gain are equivalent from a theoretical point of view. Both methods can achieve sub-pixel matching accuracy and, when appropriate models are chosen, they give very similar results. However, from an implementation point of view, the gradient cross correlation method is superior to the least squares matching method because radiometric correction and geometric correction can be achieved using only scaling and rotation parameters. Furthermore, incorporating a line search strategy with either the gradient cross correlation method or the least squares matching method shows that improved cross correlation coefficients may be achieved within a few extra iterations.

Experiments were conducted to compare methods applied to a range of images. The matching correlation results from the gradient cross correlation are nearly identical (both the matching results and the number of iterations) to that of the least squares matching. However, the gradient cross correlation method combines radiometric correction and geometric correction into a single step, which makes its parameter estimation and practical computation implementation simple. Both the gradient cross correlation method and the least squares matching method require good initial approximations or a small pull-in range in order to find the minimisation points (1 to 2 pixels in average from our experience).

For the matching of raw to rectified TM images, the scaling is about 0.83 (25m/30m) and is the same for line and pixel, while the angle of rotation is common for line and pixel, at around 10°. For the matching of raw MSS to rectified TM images, the angle of rotation is common for line and pixel, again at around 10°, while the scaling is different for line and pixel, agreeing closely with the expected values of 0.44 (25m/57m) and 0.32 (25m/79m), respectively. Reasonably good results were also obtained when points were matched from QuickBird to SPOT and to TM, given the huge pixel resolution change (more than 40 times between QuickBird and TM). For matching of a stereo pair of high-resolution images, the flexibility of varying the scaling and/or orientation gives a better matching correlation. It could be valuable to use bootstrap procedures to establish the typical range of variation for the matching correlation for Model-I against which to judge the adequacy of the simpler models.

Limited experience of experimental DEM generation using the gradient cross correlation with line search suggests that incorporating a quadratic line search with Model-I often improves the convergence and leads to a higher matching correlation, but requires some additional computing time. Given that editing a DEM requires considerable operator intervention, it may be desirable to ensure the best possible matching, at the expense of increased computing time.

1. INTRODUCTION

Matching pixels in two images is a fundamental operation in image rectification and DEM generation.

The standard approach for area matching for two images to the nearest pixel maximises the cross-correlation coefficient when the second image is shifted systematically relative to the first over a regular grid (Ackermann, 1984).

Ideally, the matching should allow for offsets in the target image, and scaling and rotation. Offsets allow for sub-pixel shifts in the two images, while scaling is necessary when matching images from different sensors (e.g. Landsat TM, Landsat MSS) and rotation allows the matching between rectified and un-rectified images.

The need to carry out the correlation matching to sub-pixel accuracy lead to a number of authors considering so-called least squares matching, including Forstner, 1982; Ackermann, 1984; Gruen, 1985; Rosenholm, 1987; Norvelle, 1992 and Zhaltov and Sibiryakov, 1997.

The essence of least squares matching is to determine offset, scaling and rotation parameters to produce interpolated grey-level values for the second image which match as closely as possible the grey-level values for the first image. This is achieved by choosing the parameters to minimise the sum of squared differences between the grey-level values for the first image and the interpolated grey-level values for the second image. The parameters are estimated by iterative least squares after linearising by a standard Taylor expansion (Gruen, 1985). An affine transformation is usually adopted to determine the predicted line and pixel coordinates for the second image (Gruen, 1985; Rosenholm, 1987). Rosenholm has also suggested including parameters to compensate for differences in the grey-level contrast between the two images.

This paper gives details of an implementation of sub-pixel matching using the normalised cross-correlation coefficient formation as the objective function, and allowing for offsets, scaling and rotation. The adoption of cross-correlation as the objective function automatically allows for a possible linear radiometric transformation between the two images. The implementation uses first and second derivatives to estimate these parameters efficiently.

Section 2 presents the details of the proposed gradient cross correlation method, including the gradient vector and the matrix of second derivatives. Section 2 also outlines the calculation of the interpolated grey-level values for the second image and how to estimate parameters. Section 3 shows the equivalence of least squares matching and gradient cross correlation. Sections 4 and 5 discuss the implementation and present some results. Finally Section 6 gives some conclusions and discussions and future work.

For the sake of convenience, the following abbreviations are used to represent the different matching methods: GCC for

gradient cross correlation and LSM for least squares cross matching.

2. GRADIENT CROSS CORRELATION (GCC)

The formulation of the cross correlation coefficient is as:

$$R = \frac{\sum (g_1 - \bar{g}_1)(g_2 - \bar{g}_2)}{\sqrt{\sum (g_1 - \bar{g}_1)^2 \sum (g_2 - \bar{g}_2)^2}} = \frac{C_{12}}{\sqrt{C_{11}C_{22}}}$$

where g_1, g_2 denote the left and right image intensity values, $\bar{g}_1, \bar{g}_2$ denote the left and right image average intensity values within the left and right patches, C_{11}, C_{22}, C_{12} denote the left and right image variances and covariance, respectively.

An affine transformation to calculate the line and pixel in the second image as a function of six parameters can be written as:

$$\begin{cases} x = x_0 + a + Sx \cdot \cos(Rx) \cdot x + Sx \cdot \sin(Rx) \cdot y \\ y = y_0 + b - Sy \cdot \sin(Ry) \cdot x + Sy \cdot \cos(Ry) \cdot y \end{cases} \quad (1)$$

where x_0, y_0 denote the pixel and line coordinates for the best whole-pixel match on the second image; a, b denote the pixel and line offset or shift; Sx, Sy denote the pixel and line scaling; Rx, Ry denote the pixel and line rotation angles.

The full model in (1) involves six parameters, which are usually re-parameterised as:

$$\begin{cases} a_1 = Sx \cdot \cos(Rx), a_2 = Sx \cdot \sin(Rx) \\ b_1 = Sy \cdot \sin(Ry), b_2 = Sy \cdot \cos(Ry) \end{cases}$$

The formulation in (1) is adopted here, as it leads to a more natural interpretation of the resulting parameters, especially when matching un-rectified and rectified satellite images.

In the approach adopted here, bilinear interpolation is used to calculate the grey values of the second image at the predicted line and pixel coordinates:

$$\begin{cases} dx = x_0 + a + Sx \cdot \cos(Rx) \cdot x + Sx \cdot \sin(Rx) \cdot y - \text{int}(x) \\ dy = y_0 + b - Sy \cdot \sin(Ry) \cdot x + Sy \cdot \cos(Ry) \cdot y - \text{int}(y) \end{cases}$$

$$g = (1 - dx)(1 - dy)g_{i,j} + dx(1 - dy)g_{i,j+1} + (1 - dx)dyg_{i+1,j} + dxdy g_{i+1,j+1}$$

The first-order derivatives of the grey-level value g with respect to image coordinates (x, y) and the gradients are given as follows:

$$\begin{cases} \frac{\partial g}{\partial x} = g_{i,j+1} - g_{i,j} + dy \cdot g_d \\ \frac{\partial g}{\partial y} = g_{i+1,j} - g_{i,j} + dx \cdot g_d \\ g_d = g_{i,j} - g_{i,j+1} - g_{i+1,j} + g_{i+1,j+1} \end{cases}$$

The first-order derivatives of the grey-level value g with respect to a, Sx, Rx, b, Sy, Ry are given as follows:

$$\begin{bmatrix} \frac{\partial g}{\partial a} \\ \frac{\partial g}{\partial Sx} \\ \frac{\partial g}{\partial Rx} \\ \frac{\partial g}{\partial b} \\ \frac{\partial g}{\partial Sy} \\ \frac{\partial g}{\partial Ry} \end{bmatrix} = \begin{bmatrix} \frac{\partial g}{\partial x} = g_{i,j+1} - g_{i,j} + dy \cdot g_d \\ \frac{\partial g}{\partial x} \frac{\partial x}{\partial Sx} = \frac{\partial g}{\partial x} (\cos(Rx)x + \sin(Rx)y) \\ \frac{\partial g}{\partial x} \frac{\partial x}{\partial Rx} = \frac{\partial g}{\partial x} Sx(-\sin(Rx)x + \cos(Rx)y) \\ \frac{\partial g}{\partial y} = g_{i+1,j} - g_{i,j} + dx \cdot g_d \\ \frac{\partial g}{\partial y} \frac{\partial y}{\partial Sy} = \frac{\partial g}{\partial y} (-\sin(Ry)x + \cos(Ry)y) \\ \frac{\partial g}{\partial y} \frac{\partial y}{\partial Ry} = \frac{\partial g}{\partial y} Sy(-\cos(Ry)x - \sin(Ry)y) \end{bmatrix}$$

The second-order derivatives of x, y with respect to a, Sx, Rx, b, Sy, Ry are given as follows:

$$\begin{bmatrix} \frac{\partial^2 x}{\partial a \partial a} & \frac{\partial^2 x}{\partial a \partial Sx} & \frac{\partial^2 x}{\partial a \partial Rx} & \frac{\partial^2 x}{\partial a \partial b} & \frac{\partial^2 x}{\partial a \partial Sy} & \frac{\partial^2 x}{\partial a \partial Ry} \\ \frac{\partial^2 x}{\partial Sx \partial a} & \frac{\partial^2 x}{\partial Sx \partial Sx} & \frac{\partial^2 x}{\partial Sx \partial Rx} & \frac{\partial^2 x}{\partial Sx \partial b} & \frac{\partial^2 x}{\partial Sx \partial Sy} & \frac{\partial^2 x}{\partial Sx \partial Ry} \\ \frac{\partial^2 x}{\partial Rx \partial a} & \frac{\partial^2 x}{\partial Rx \partial Sx} & \frac{\partial^2 x}{\partial Rx \partial Rx} & \frac{\partial^2 x}{\partial Rx \partial b} & \frac{\partial^2 x}{\partial Rx \partial Sy} & \frac{\partial^2 x}{\partial Rx \partial Ry} \\ \frac{\partial^2 x}{\partial b \partial a} & \frac{\partial^2 x}{\partial b \partial Sx} & \frac{\partial^2 x}{\partial b \partial Rx} & \frac{\partial^2 x}{\partial b \partial b} & \frac{\partial^2 x}{\partial b \partial Sy} & \frac{\partial^2 x}{\partial b \partial Ry} \\ \frac{\partial^2 x}{\partial Sy \partial a} & \frac{\partial^2 x}{\partial Sy \partial Sx} & \frac{\partial^2 x}{\partial Sy \partial Rx} & \frac{\partial^2 x}{\partial Sy \partial b} & \frac{\partial^2 x}{\partial Sy \partial Sy} & \frac{\partial^2 x}{\partial Sy \partial Ry} \\ \frac{\partial^2 x}{\partial Ry \partial a} & \frac{\partial^2 x}{\partial Ry \partial Sx} & \frac{\partial^2 x}{\partial Ry \partial Rx} & \frac{\partial^2 x}{\partial Ry \partial b} & \frac{\partial^2 x}{\partial Ry \partial Sy} & \frac{\partial^2 x}{\partial Ry \partial Ry} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -\sin(Rx)x + \cos(Rx)y & -\cos(Ry)x - \sin(Ry)y & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -\sin(Rx)x + \cos(Rx)y & -\cos(Ry)x - \sin(Ry)y & 0 & 0 & 0 & 0 \\ Sx(-\cos(Rx)x - \sin(Rx)y) & Sy(\sin(Ry)x - \cos(Ry)y) & 0 & 0 & 0 & 0 \end{bmatrix}$$

The matrix of second-order (partial) derivatives of the grey-level value g with respect to a, Sx, Rx, b, Sy, Ry can be explicitly expressed as follows:

$$\begin{bmatrix} \frac{\partial^2 g}{\partial a \partial a} & \frac{\partial^2 g}{\partial a \partial Sx} & \frac{\partial^2 g}{\partial a \partial Rx} & \frac{\partial^2 g}{\partial a \partial b} & \frac{\partial^2 g}{\partial a \partial Sy} & \frac{\partial^2 g}{\partial a \partial Ry} \\ \frac{\partial^2 g}{\partial Sx \partial a} & \frac{\partial^2 g}{\partial Sx \partial Sx} & \frac{\partial^2 g}{\partial Sx \partial Rx} & \frac{\partial^2 g}{\partial Sx \partial b} & \frac{\partial^2 g}{\partial Sx \partial Sy} & \frac{\partial^2 g}{\partial Sx \partial Ry} \\ \frac{\partial^2 g}{\partial Rx \partial a} & \frac{\partial^2 g}{\partial Rx \partial Sx} & \frac{\partial^2 g}{\partial Rx \partial Rx} & \frac{\partial^2 g}{\partial Rx \partial b} & \frac{\partial^2 g}{\partial Rx \partial Sy} & \frac{\partial^2 g}{\partial Rx \partial Ry} \\ \frac{\partial^2 g}{\partial b \partial a} & \frac{\partial^2 g}{\partial b \partial Sx} & \frac{\partial^2 g}{\partial b \partial Rx} & \frac{\partial^2 g}{\partial b \partial b} & \frac{\partial^2 g}{\partial b \partial Sy} & \frac{\partial^2 g}{\partial b \partial Ry} \\ \frac{\partial^2 g}{\partial Sy \partial a} & \frac{\partial^2 g}{\partial Sy \partial Sx} & \frac{\partial^2 g}{\partial Sy \partial Rx} & \frac{\partial^2 g}{\partial Sy \partial b} & \frac{\partial^2 g}{\partial Sy \partial Sy} & \frac{\partial^2 g}{\partial Sy \partial Ry} \\ \frac{\partial^2 g}{\partial Ry \partial a} & \frac{\partial^2 g}{\partial Ry \partial Sx} & \frac{\partial^2 g}{\partial Ry \partial Rx} & \frac{\partial^2 g}{\partial Ry \partial b} & \frac{\partial^2 g}{\partial Ry \partial Sy} & \frac{\partial^2 g}{\partial Ry \partial Ry} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & g_d & g_d \frac{\partial y}{\partial Sy} & g_d \frac{\partial y}{\partial Ry} \\ 0 & 0 & \frac{\partial g}{\partial x} \frac{\partial x}{\partial Sx} & g_d \frac{\partial x}{\partial Sx} & g_d \frac{\partial x}{\partial Sx} \frac{\partial y}{\partial Sy} & g_d \frac{\partial x}{\partial Sx} \frac{\partial y}{\partial Ry} \\ 0 & \frac{\partial g}{\partial x} \frac{\partial x}{\partial Rx} & \frac{\partial g}{\partial x} \frac{\partial x}{\partial Rx} & g_d \frac{\partial x}{\partial Rx} & g_d \frac{\partial x}{\partial Rx} \frac{\partial y}{\partial Sy} & g_d \frac{\partial x}{\partial Rx} \frac{\partial y}{\partial Ry} \\ g_d & g_d \frac{\partial x}{\partial Sx} & g_d \frac{\partial x}{\partial Rx} & 0 & 0 & 0 \\ g_d \frac{\partial y}{\partial Sy} & g_d \frac{\partial y}{\partial Sx} \frac{\partial x}{\partial Sx} & g_d \frac{\partial y}{\partial Sx} \frac{\partial x}{\partial Rx} & 0 & 0 & \frac{\partial g}{\partial y} \frac{\partial y}{\partial Sy} \frac{\partial x}{\partial Sx} \\ g_d \frac{\partial y}{\partial Ry} & g_d \frac{\partial y}{\partial Ry} \frac{\partial x}{\partial Sx} & g_d \frac{\partial y}{\partial Ry} \frac{\partial x}{\partial Rx} & 0 & \frac{\partial g}{\partial y} \frac{\partial y}{\partial Ry} \frac{\partial x}{\partial Sx} & \frac{\partial g}{\partial y} \frac{\partial y}{\partial Ry} \frac{\partial x}{\partial Rx} \end{bmatrix}$$

Assuming α_i represents one of parameters (a, Sx, Rx, b, Sy, Ry) which need to be solved, the first-order and second-order derivatives of the cross correlation coefficient R with respect to each parameter α_i are given as follows:

$$\frac{\partial R}{\partial \alpha_i} = \frac{N_{2i}}{M}$$

where

$$\begin{cases} M = C_{11} C_{22} \sqrt{C_{11} C_{22}} \\ N_{2i} = C_{11} C_{22} \sum (g_1 \frac{\partial g_1}{\partial \alpha_i}) - C_{12} C_{11} \sum (g_2 \frac{\partial g_2}{\partial \alpha_i}) \end{cases}$$

The second-order derivatives of cross correlation coefficient R with respect to each parameter α_i are given as follows:

$$\frac{\partial^2 R}{\partial \alpha_i \partial \alpha_j} = \frac{M \frac{\partial N_{2i}}{\partial \alpha_j} - N_{2i} \frac{\partial M}{\partial \alpha_j}}{M^2}$$

where

$$\begin{cases} \frac{\partial M}{\partial \alpha_j} = 3C_{11} \sqrt{C_{11} C_{22}} \sum (g_2 \frac{\partial g_2}{\partial \alpha_j}) \\ \frac{\partial N_{2i}}{\partial \alpha_j} = C_{11} C_{22} \sum (g_1 \frac{\partial^2 g_1}{\partial \alpha_i \partial \alpha_j}) + 2C_{11} \sum (g_2 \frac{\partial g_2}{\partial \alpha_i} \sum (g_1 \frac{\partial g_1}{\partial \alpha_j}) - C_{12} C_{11} \sum (g_2 \frac{\partial^2 g_2}{\partial \alpha_i \partial \alpha_j}) - C_{12} C_{11} \sum (\frac{\partial g_2}{\partial \alpha_i} \frac{\partial g_2}{\partial \alpha_j}) - C_{11} \sum (g_1 \frac{\partial g_1}{\partial \alpha_i}) \sum (g_2 \frac{\partial g_2}{\partial \alpha_j}) \end{cases}$$

A full Newton-Raphson implementation (Chambers, 1977; Adbys and Dempster, 1974) using first and second derivatives was implemented:

$$[d\alpha_i] = -H^{-1} \cdot G = -\left[\frac{\partial^2 R}{\partial \alpha_i \partial \alpha_j} \right]^{-1} \cdot \left[\frac{\partial R}{\partial \alpha_i} \right]$$

where H is the Hessian matrix and G is the Jacobian gradient.

3. THE EQUIVALENCE OF LSM AND GCC

Least squares matching assumes that the left and right image grey-level values should be identical between two small patches surrounding the left and right points:

$$g_1(x, y) = g_2(x, y)$$

A radiometric correction and a geometric correction for the right images are applied:

$$\begin{aligned} g_1(x, y) + n_1(x, y) &= c_0 + c_1 g_2(x_2, y_2) + n_2(x_2, y_2) \quad (2) \\ \begin{cases} x_2 = x_{20} + a + Sx \cdot \cos(Rx) \cdot x + Sx \cdot \sin(Rx) \cdot y \\ y_2 = y_{20} + b - Sy \cdot \sin(Ry) \cdot x + Sy \cdot \cos(Ry) \cdot y \end{cases} \end{aligned}$$

where n_1, n_2 are the left and right image random noises, c_0, c_1 are the radiometric correction coefficients and x_{20}, y_{20} are the starting image locations for the right point.

The least squares observation equation after linearisation (2) is:

$$\begin{aligned} v &= dc_0 + g_2 dc_1 + \frac{\partial g_1}{\partial a} da + \frac{\partial g_1}{\partial Sx} dSx + \frac{\partial g_1}{\partial Rx} dRx + \frac{\partial g_1}{\partial b} db + \frac{\partial g_1}{\partial Sy} dSy + \frac{\partial g_1}{\partial Ry} dRy - dg \\ dg &= g_1(x, y) - g_2(x, y) \end{aligned} \quad (3)$$

Of course, the radiometric correction can be treated as either in a separate prior step or estimated with other affine parameters simultaneously.

The matrix version of least squares matching (2) is:

$$L = AX - V$$

where X is the unknown vector, L is the observation vector and A is the design matrix. The least squares normal equation and its solution are:

$$\begin{aligned} A^T AX &= A^T L \\ X &= [A^T A]^{-1} L = N^{-1} A^T L \end{aligned}$$

In order to show the equivalence of least squares matching and gradient cross correlation, firstly, that the correlation coefficient R is invariant with respect to a linear radiometric correction.

Assume after applying a linear radiometric correction, that the right image value is:

$$g'_2 = c_0 + c_1 g_2$$

The formulation of the new cross correlation coefficient R' is:

$$R' = \frac{\sum (g_1 - \bar{g}_1)(g_2' - \bar{g}_2')}{\sqrt{\sum (g_1 - \bar{g}_1)^2 \sum (g_2' - \bar{g}_2')^2}} = \frac{C'_{12}}{\sqrt{C'_{11} C'_{22}}} \quad (4)$$

Replacing g_2' with $c_0 + c_1 g_2$ in (4) and after a series of steps can be reduced to:

$$R' = \frac{c_1 \sum (g_1 - \bar{g}_1)(g_2 - \bar{g}_2)}{\sqrt{\sum (g_1 - \bar{g}_1)^2 c_1^2 \sum (g_2 - \bar{g}_2)^2}} = \frac{C_{12}}{\sqrt{C_{11} C_{22}}} = R$$

Secondly, the following shows that the least squares matching and gradient cross correlation use the same criterion to estimate the unknowns.

Least squares techniques minimise the sum of squares of observation errors or image intensity differences (3):

$$\sum vv = \min$$

Assuming g_1 and g_2 are normalised, then the linear radiometric correction coefficients can be obtained:

$$c_0 = 0, c_1 = \frac{\sum g_1 g_2}{\sum g_2^2} - 1 \quad (5)$$

$\sum vv$ can then be expanded using (3) and (5):

$$\sum vv = \sum (g_2 \frac{\sum g_1 g_2}{\sum g_2^2} - g_1)^2 = \sum g_1^2 - \frac{(\sum g_1 g_2)^2}{\sum g_2^2} = \sum g_1^2 (1 - R^2)$$

The relationship between $\sum vv$ and R can be described using the following equation:

$$\frac{\sum vv}{\sum g_1^2} = 1 - R^2, R = \sqrt{1 - \frac{\sum vv}{\sum g_1^2}} \quad (6)$$

(6) means that finding the minimisation of the sum of squares of intensity differences between the left and right image patches is equivalent to maximising the cross correlation coefficient between the two patches.

4. IMPLEMENTATION

A hierarchy of relationships between the affine transformation parameters can be specified in practice:

- Model-I: different scale, different rotation (6 unknowns: two offsets, two scales, two rotations)
- Model-IIA: different scale, common rotation (5 unknowns: two offsets, two scales, one rotation)
- Model-IIB: common scale, different rotation (5 unknowns: two offsets, one scale, two rotations)
- Model-III: common scale, common rotation (4 unknowns: two offsets, one scale, one rotation)
- Model-IV: fixed scale, fixed rotation (2 unknowns: two offsets)

In order to investigate the behavior of different models for various images, the above models were also implemented within two matching methods (GCC and LSM). Further, a quadratic line search strategy (Adby and Dempster, 1974) is applied to

both GCC and LSM matching. The further line search suggests improved cross correlation coefficients may be achieved with a few extra iterations. The duration of computation time is recorded for comparison purposes. The experiments were conducted on a DELL Pentium III personal computer with CPU clock speed of 1.70GHz and memory of 512MB.

5. EXPERIMENT RESULTS

The performance of the algorithm is examined for three pairs of images. The first pair (Figure 1) relates to the registration of a Landsat TM image from February 1992 (the middle image in Figure 1) to a rectified Landsat TM image from March 1995 (the left image in Figure 1). The original TM image pixel size is 30m, and the rotation of the original image is about 9° from true north. The rectified Landsat TM image is in AMG (Australian Map Grid) zone 50, and the pixel size is 25m.

The second pair relates to the registration of a Landsat MSS image from January 1987 (the right image in Figure 1) to the 1995 Landsat TM image (the left image in Figure 1). The original MSS image pixel size is 57m×79m, and the rotation is again about 9°.

Three control points were chosen around the large patch of bush in the rectified TM images: Point 1 is at the top right of the patch, Point 2 is at the bottom right of the patch and Point 3 is at the top left of the patch. Their corresponding points in the raw TM and MSS images were roughly located as the initial start points for registration. The correlation window size used is 41 pixels by 41 pixels.

The third pair (Figure 2) relates to the matching of two QuickBird high-resolution satellite images, which were flown on June 19, 2003; the rotation between two raw images is about 13°. Figure 2 shows a small isolated forest patch and the surrounding shadows. One point was selected on the treetops among the forest patch at the middle of the image and another point was selected at the shadow edges on the bare ground. The correlation window size used is 21 pixels by 21 pixels.

Table 1 summarises the results for all models for the Landsat TM registration. For this example, Model-III (common pixel scaling and common rotation angle) should be appropriate. This is confirmed in Table 1, where the matching correlation coefficient for each control point for Model-III is similar to that for Model-I, Model-IIA and Model-IIB. Model-IV gives the worst matches. The results also indicate that GCC and LSM give similar results.

Table 2 summarises the results for the Landsat MSS registration. For this example, Model-IIA (different pixel scaling, common rotation) should be appropriate. This is confirmed in Table 2, where the matching correlation coefficient for each control point for Model-IIA is similar to that for Model-I. Model-IV gives the worst matches. Again, GCC and LSM give similar results for the appropriate model. The estimated line and pixel scaling are roughly consistent with the expected values: the line scaling should be about $25/57 = 0.44$, while the observed values in Table 2 are 0.40, 0.44 and 0.45; and the pixel scaling should be about $25/79 = 0.32$, while the observed values in Table 2 are 0.27, 0.30 and 0.30.

Table 3 summarises the results for the QuickBird image matching. For both points (Points 1 and 2 in Table 3), the best

matches are always given by Model-I for all three matching methods, and Model-IIB gives the second best matches. Model-IV is the worst model for matching treetop point 1 and ground point 2. The correlation coefficients also show that the treetop point is more difficult to be matched in comparison with the ground point. The reason may be due to the mixed texture and complicated geometry at the treetop.

More QuickBird matching experiments (not presented in this paper) also confirm that, due to the view angle changes of QuickBird sensors and the changing angle between an object and its shadow (in particular between tree shadow and tree), high-resolution satellite imagery such as QuickBird requires a well-defined geometric model for their image registration and matching; in this case, Model-I seems the appropriate choice.

Incorporating a quadratic line search with GCC or LSM matching often improves the convergence and leads to a higher matching correlation. From both GCC and LSM line search results, it shows a very slight improvement of matching (cross correlation coefficient) within a few extra iterations.

The function maximisation procedures require a tolerance which indicates when successive function values are sufficiently similar. Tables 1–3 also list the number of iterations (maximum is 50). This very limited comparison suggests that a tolerance of 0.002 gives similar results to those obtained from a more stringent convergence tolerance, in about one third of the number of iterations.

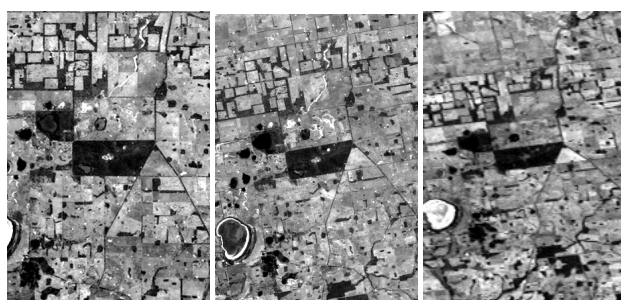


Figure 1: Segments of Landsat scenes (path 111, row 84). Left image: segment of the rectified TM scene for Band 3, March 1995 (map grid: AMG, pixel size: 25m). Middle image: segment of the raw TM scene for Band 3, February 1992 (pixel size: 30m×30m). Right image: segment of the raw MSS scene for Band 2, January 1987 (pixel size: 57m×79m).

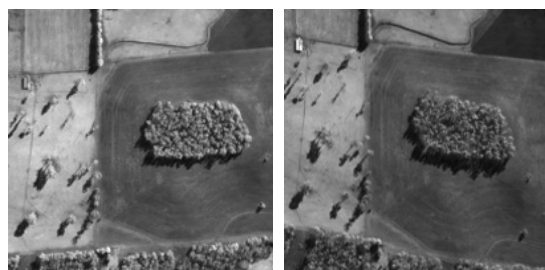


Figure 2: Left and right images are two segments from a raw QuickBird stereo pair for multiple-spectral band 4, June 2003 (pixel size: approximately 3m×3m).

Point	Model	GCC		LSM	
		Corr. Score	Iter.	Corr. Score	Iter.
1	Model-I	0.87377	50	0.87377	11
	Model-IIA	0.87245	11	0.87245	11
	Model-IIB	0.86975	10	0.86975	11

2	Model-III	0.86961	12	0.86961	50
	Model-IV	0.86844	10	0.86844	11
	Model-I	0.89668	14	0.89669	13
	Model-IIA	0.89618	12	0.89618	12
	Model-IIB	0.89665	16	0.89665	14
	Model-III	0.89609	13	0.89609	15
3	Model-IV	0.89367	10	0.89367	11
	Model-I	0.97144	11	0.97144	11
	Model-IIA	0.97139	11	0.97139	11
	Model-IIB	0.97034	9	0.97034	9
	Model-III	0.97041	12	0.97030	8
	Model-IV	0.96727	10	0.96727	10
	Model-IV	0.74198	3	0.74198	2

Table 1: GCC and LSM sub-pixel matching a raw Landsat TM image from February 1992 (centre image in Figure 1) to a rectified and resampled TM image March 1995 (left image in Figure 1) for three ground control points (average computing time is 0.03 second per point).

Point	Model	GCC		LSM	
		Corr. Score	Iter.	Corr. Score	Iter.
1	Model-I	0.88635	20	0.88635	18
	Model-IIA	0.88634	20	0.88634	18
	Model-IIB	0.84573	31	0.84575	31
	Model-III	0.84558	28	0.84558	19
	Model-IV	0.75940	22	0.75940	22
2	Model-I	0.84540	44	0.84540	23
	Model-IIA	0.84542	21	0.84542	20
	Model-IIB	0.82646	21	0.82647	23
	Model-III	0.82466	40	0.82467	29
	Model-IV	0.75606	32	0.75604	31
3	Model-I	0.93416	15	0.93416	15
	Model-IIA	0.93360	15	0.93360	15
	Model-IIB	0.88439	16	0.88439	16
	Model-III	0.88151	17	0.88149	16
	Model-IV	0.85979	19	0.85979	19

Table 2: GCC and LSM sub-pixel matching an original Landsat MSS image from January 1987 (right image in Figure 1) to a rectified and resampled TM image March 1995 (left image in Figure 1) for three ground control points (average computing time is 0.04 second per point).

Point	Model	GCC		LSM	
		Corr. Score	Iter.	Corr. Score	Iter.
1	Model-I	0.75367	9	0.75367	13
	Model-IIA	0.63904	50	0.63933	50
	Model-IIB	0.67154	7	0.67205	50
	Model-III	0.58464	50	0.58461	13
	Model-IV	0.48116	5	0.48116	4
2	Model-I	0.93372	22	0.92689	14
	Model-IIA	0.92933	50	0.92693	50
	Model-IIB	0.92534	50	0.92358	50
	Model-III	0.92357	50	0.92320	50
	Model-IV	0.89629	4	0.89629	5

Table 3: GCC and LSM sub-pixel matching of two QuickBird bush images for treetop point 1 and ground point 2 (average computing time is 0.03 second per point).

6. CONCLUSION AND DISCUSSION

The correlation results from the gradient cross correlation are nearly identical (both the matching results and iterations) to those of the least square matching. However, the gradient cross correlation method combines radiometric correction and geometric correction into a single step, which makes its parameter estimation and practical computation implementation

simple. Both the gradient cross correlation method and the least squares matching method require good approximation or small pull-in range in order to find the minimisation points (1 to 2 pixels in average from our experience).

The particular formulation of the affine transformation in Equation 2-2 leads to useful insights into the image matching. Model-IV (shift only, not allowing scaling and rotation) is the worst model for matching all kind of point, which means that it is essential to choose an appropriate geometric transformation for certain kind of sub-pixel matching.

For the matching of TM images, the scaling is about 0.83 (25m/30m) and is the same for line and pixel, while the angle of rotation is common for line and pixel, at around 10°.

For the matching of TM and MSS images, the angle of rotation is common for line and pixel, again at around 10°, while the scalings are different for line and pixel, agreeing closely with the expected values of 0.44 (25m/57m) and 0.32 (25m/79m), respectively.

For matching of a stereo pair of high-resolution images, the flexibility of varying the scaling and/or orientation gives a better matching correlation. It could be valuable to use bootstrap procedures (Efron and Gong, 1983; Efron and Tibshirani, 1993) to establish the typical range of variation for the matching correlation for Model-I (i.e. confidence limits) against which to judge the adequacy of the simpler models.

Limited experience of experimental DEM generation using the gradient cross correlation with line search suggests that incorporating a quadratic line search with Model-I often improves the convergence and leads to a higher matching correlation, but requires some additional computing time. Given that editing a DEM requires more operator intervention, it may be desirable to ensure the best possible match, at the expense of increased computing time.

7. REFERENCES

- Ackermann, F., 1984. Digital Image Correlation: Performance and Potential Application in Photogrammetry. *Photogrammetric Record*, 11, pp.429-439.
- Adby, P., R. and Dempster, M. A. H., 1974. *Introduction to Optimisation Methods*. Chapman and Hall, London.
- Chambers, J., M., 1977. *Computational Methods for Data Analysis*. Wiley, New York.
- Efron, B. and Gong, G., 1983. A Leisurely Look at the Bootstrap, the Jackknife, and Cross-validation. *American Statistician*, 37, 36-48.
- Efron, B. and Tibshirani, R., J., 1993. *An Introduction to the Bootstrap*. Chapman and Hall, New York.
- Forstner, W., 1982. On the Geometric Precision of Digital Correlation. *International Archives of Photogrammetry and Remote Sensing*, Symposium Helsinki Commission III, 24-Part 3, 176-189.
- Gruen, A., W., 1985. Adaptive Least Squares Correlation: a Powerful Image Matching Technique. *South African Journal of Photogrammetry, Remote Sensing and Cartography*, 14, pp.175-187.
- Norville, F., R., 1992. Stereo Correlation: Window Shaping and DEM Corrections. *Photogrammetric Engineering and Remote Sensing*, 58, No 1, pp.111-115.
- Rosenholm, D., 1987. Least Squares Matching Method: Some Experimental Results. *Photogrammetric Record*, 12, pp.493-512.
- Zhaltov, S., Y. and Sibiryakov, A., V., 1997. Adaptive Subpixel Cross-correlation in a Point Correspondence Problem. In A. Gruen and H. Kahmen (eds.), *Optical 3-D Measurement Techniques IV*, Wichmann Verlag, Heidelberg, pp.86-95.