# DEM REGISTRATION BASED ON MUTUAL INFORMATION

M. Ravanbakhsh, C. S. Fraser


Cooperative Research Center for Spatial Information, Department of Infrastructure Engineering
University of Melbourne, VIC 3010, Australia
[m.ravanbakhsh, c.fraser]@unimelb.edu.au


**KEY WORDS:** DEM image, registration, mutual information, gradient map

**ABSTRACT:**

The registration process forms an important step in the integration of Digital Elevation Models (DEMs) of differing accuracy and resolution. DEM registration invariably involves what is effectively an 'image matching' process utilising similarity measures. Of a number of prospective image matching approaches, Mutual Information (MI) offers potential for application to DEM registration. MI has proven to be a robust and accurate method of image registration, especially in the field of medical imaging. However, a lack of variation within the terrain models potentially limits the utility of MI for the matching of multi-resolution DEMs. This paper investigates the potential of MI for DEM registration, and proposes an improved Gradient-based MI (GMI) in which the matching is applied to raster DEMs where the elevation value at each grid point is replaced by the surface gradient magnitude at that point. Experimental results demonstrate that MI-based methods can outperform the often employed Cross Correlation Function (CCF). Furthermore, the GMI approach has been found to produce a more accurate and robust registration than MI.

## 1. INTRODUCTION

Within the integration process to form a multi-resolution Digital Elevation Model (DEM) from two overlapping raster DEMs of differing resolution and accuracy, the registration process forms an important step. Registration effectively constitutes a surface-to-surface matching in which horizontal and vertical offsets between the two DEMs are determined. Van Niel et al. (2008) have reported on the impact of misregistration on DEM image differences and shown that image differences between DEMs are sensitive to even small amounts of misregistration. They concluded that misregistration is a source of strong correlation between elevation difference and aspect.

Of the range of image similarity-based registration methods, the Cross Correlation Function (CCF) has often been employed for the registration of raster DEMs (Li & Bethel, 2008; Costantini et al., 2006; Roth et al., 1999). Mutual Information (MI) offers another similarity measure that has been adopted for a variety of applications (Viola & Wells III 1996), including the registration of medical imagery. MI is integrated within the recently developed semi-global matching method for DEM generation from aerial and space imagery (Hirschmueller, 2008, Hirschmueller et al., 2005 and Krauß et al., 2008). Also, the MI measure has been used as a local matching cost aggregated into a global energy function in the determination of object surface models from airborne video sequences (Gerke, 2008). The energy function was iteratively minimised in an optimisation process that led to optimal alignment between images.

Although MI has been adopted for image matching, it has to date not been applied to the registration of DEMs. This paper investigates the potential of MI and a variation of it, namely Gradient-based MI (GMI), for horizontal registration of multi-resolution DEMs. The CCF method is also evaluated, primarily for comparative assessment of the performance of MI and GMI. The three image similarity-based matching methods are described in Section 2 and the results of registering four different DEM datasets to a LiDAR DEM are presented in Section 3.

## 2. METHOD

When applied to surface models, an image similarity method utilises a metric to quantify the degree of correspondence between two Regions of Interest (ROIs) extracted from a reference and target DEM within the coordinate transformation to register the DEMs. All the DEMS considered in the reported investigation were georeferenced in the same geodetic datum and were assumed to be rotationally aligned. Thus, the transformation model adopted here includes translations only and is identical for all similarity methods. The optimum match between ROIs occurs when similarity metrics reach their maximum values. In the following sections, similarity metrics and their characteristics are briefly described.

### 2.1 Cross correlation function

CCF is used as a metric to measure the degree of similarity between two images. It is mathematically defined as

$$CCF(R, T) = \frac{\sum (R(r) - \overline{R}(r))(T(t) - \overline{T}(t))}{\sqrt{\sum (R(r) - \overline{R}(r))^2 \sum (T(t) - \overline{T}(t))^2}} \qquad (1)$$

where $\overline{R}(r)$ and $\overline{T}(t)$ denote the average intensities of the reference and target windows. CCF provides an effective measure of alignment between two images when subtle intensity changes between the images occur or the same area is captured with the same sensor at different times (Zitova & Flusser, 2003).

### 2.2 Mutual information

The MI measure of two images is a combination of the entropy values of the images, both separately and jointly. One interpretation of entropy is as a measure of dispersion of a probability distribution. A distribution with only a few large probabilities has a low entropy value and maximum entropy is reached for a uniform distribution. One of the most commonly used entropy measures was introduced by Shannon (1948). Given $n$ events occurring with probabilities $p_1 ... p_n$, the Shannon entropy is defined as

$$H = \sum_{i=1}^{n} p_i \log 1/p_i \qquad (2)$$

For an image, the entropy is calculated from the intensity histogram in which the entries comprise the probabilities. The joint entropy of images quantifies the combined information content in the two images. The joint entropy H (X, Y) can be calculated using the joint histogram of the two images. If the images are totally unrelated, then the joint entropy will be the sum of the entropies of the individual images. The more similar the images, the lower is the joint entropy compared with the sum of the individual entropies.

As the images become misaligned, dispersion of their joint histogram increases. Therefore, image-to-image registration can be accomplished by minimizing the joint entropy, though MI proves to be a better criterion when marginal entropies H (X) and H (Y) are also taken into account. MI is related to joint and marginal entropies of two images of X and Y by the equation

$$\begin{aligned} I\,(X,Y) &= H\,(X) - H\,(X/Y) \\ &= H\,(X) + H\,(Y) - H\,(X,Y) \end{aligned} \qquad (3)$$

Optimal registration can be carried out by maximising the MI of the two images. The MI metric is robust against changes caused through use of different sensors, allowing registration even when the images have differing appearance, as long the intensities of overlapping pixels are statistically correlated when the images are registered (Olson, 2001).

## 2.3 Gradient-based mutual information

MI has been shown to be a robust and accurate method of image registration, especially in the field of medical imaging (Viola & Wells III, 1996). However, when applied to DEM registration it can sometimes fail, partly due to the lack of variation within the surface model, which is assumed here to be in the form of an intensity image. Corresponding images representing the distribution of surface gradient magnitudes, effectively slope maps, can supply supplementary 'intensity' information for inclusion in the MI process. The gradient images, formed by means of differential gradient operators in two orthogonal axes within the image, clearly indicate areas with high topographic variation or steepness (Figure 1). It has been shown that the use of gradient magnitude images used in the calculation of MI can improve the quality of registration (Penny et al. 1999 & Maintz et al. 1996).
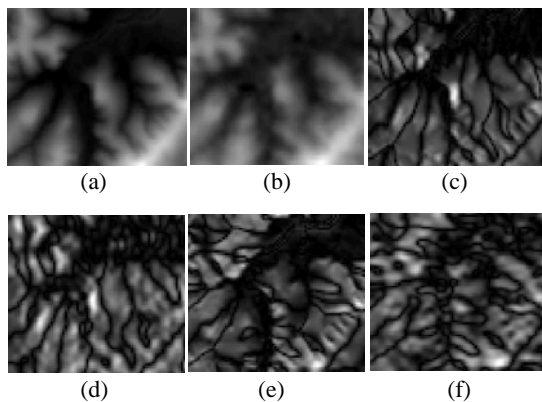


(a)       (b)       (c)



(d)       (e)       (f)

Figure 1. (a) and (b) show reference and target DEMs at 30m post spacing at the size of $70\times80$ pixels; (c) and (e) show gradient magnitude images of the reference DEM in x and y directions; and (d) and (f) display gradient magnitude images of the target DEM in x and y

directions. Gradient magnitude images are normalised within [0, 1].

Given a 2D image $F$ with intensity $f(x,y)$, its spatial gradient field $G_F(x,y)$ can be computed by

$$G_F(x,y) \;=\; \frac{\partial f(x,y)}{\partial x}\vec{\imath} \;+\; \frac{\partial f(x,y)}{\partial y}\vec{\jmath} \qquad (4)$$

where $\vec{\imath}$ and $\vec{\jmath}$ are the unit vectors along the x and y axis, respectively. Since it is assumed that DEMs are rotationally aligned, the orientation of gradient vectors is ignored. Instead, gradient magnitude images of the reference and target DEMs in the x and y directions, denoted by $f_x^R$, $f_x^T$, $f_y^R$ and $f_y^T$, are used to calculate MI measures. To avoid computational complexity, the reference and target DEMs should have identical post spacing. Thus, the reference DEM, which is often of higher horizontal resolution, has to first be resampled to the target DEM grid size.

GMI is the resultant of MI measures of gradient magnitude images, separately in x and y directions. In mathematical terms, GMI is defined as

$$GMI = I\,(f_x^R, f_x^T) + I\,(f_y^R, f_y^T) \qquad (5)$$

Figure 3 shows the behaviour of the magnitude of the similarity functions described above on a DEM image when the image is translated separately in the x, y and diagonal directions. The sharper peak from a lower similarity base value indicates that the GMI function is more stable and will thus lead to a more accurate registration.
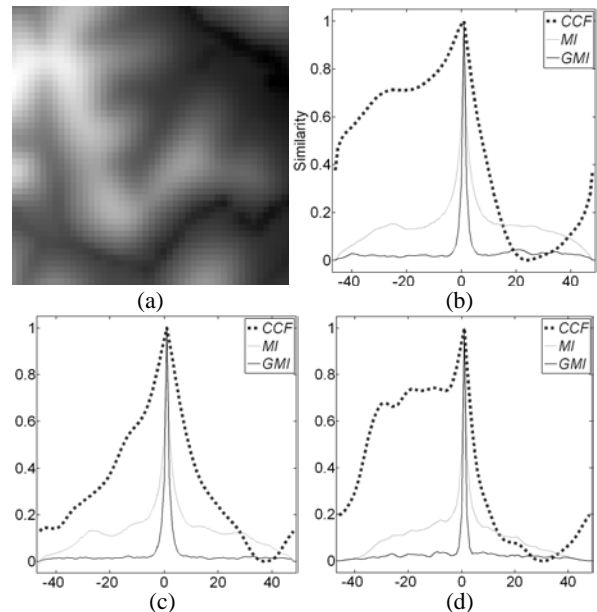


(a)       (b)



(c)       (d)

Figure 3. (a) DEM image; (b) ,(c) and (d) show the trace of CCF, MI and GMI functions while translations are applied along the x, y and diagonal directions. Horizontal axes show translations in pixels.

The CCF has a wider range about the global maximum compared to the MI and GMI measures, indicating that in the case of adopting an optimisation strategy to minimise an energy function (Hirschmueller, 2008 and 2005), the function can be initialised at a large distance from the global peak. This characteristic gives flexibility to the initialisation. However, it does not affect the quality of the registration because the search to find the global maximum covers the entire area.

Figure 4 shows, for a 24×24 pixel template within the LiDAR DEM, an example of where the CCF and MI methods determined erroneous matching positions within an SRTM DEM, but GMI found the correct matching position. The main reason for the success of GMI in this case is that it makes use of gradient magnitude information, which is dominant and ignores other information that contributes little to describing the shape of the terrain.



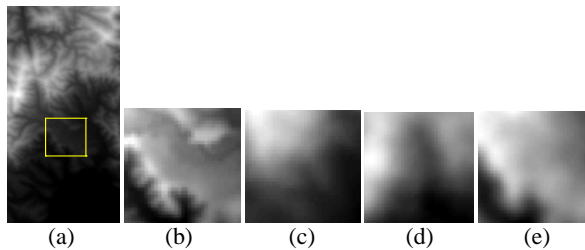(a)      (b)      (c)      (d)      (e)

Figure 4. (a) LiDAR DEM and marked template; (b) LiDAR template; (c) and (d) Erroneous patches found incorrectly by the CCF and MI; and (e) Correct DEM window found via GMI.

## 3. RESULTS AND EVALUATION

The outline of the test areas and the corresponding reference LiDAR DEM to which four different DEM datasets are to be registered are shown in Figure 5. The area at the top (Area 1) is hilly and has a coverage of dense coastal forest, and that below (Area 2) comprises low-lying urban terrain. Both areas have a size of 18 km². The specifications for the target DEMs, along with their quality measures as quantified by representative RMS elevation discrepancy values relative to the reference LiDAR DEM, are shown in Table 1.
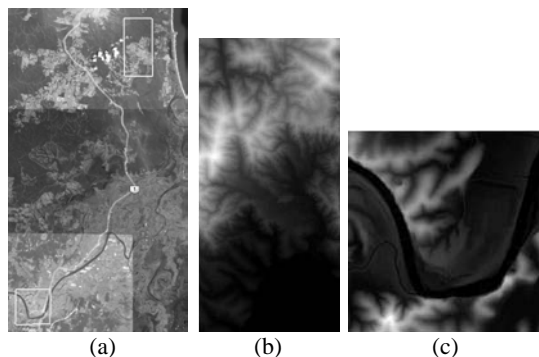


(a)       (b)       (c)

Figure 5. (a) Outline of the two test areas; with corresponding LiDAR DEMs displayed in (b) and (c).

Table 1. Target DEMs, with resolution and representative vertical accuracy (RMS) compared to LiDAR reference DEM.

| Dataset | Technology | Grid size (m) | RMS (m) Area1 | RMS (m) Area2 |
|---------|-----------|---------------|-------|-------|
| SRTM | Space-borne IfSAR | 30 | 6.1 | 3.4 |
| SPOT5 | Space photogrammetry | 30 | 6.6 | 5.8 |
| InSAR | Air-borne IfSAR | 5 | 2.1 | 1.8 |
| ADS40 | Aerial photogrammetry | 8 | 2.2 | 1.9 |

In order to assess the accuracy and robustness of DEM registrations conducted, the four target DEMs were first translated by 30m with respect to the reference DEM. Then, within the experimental registrations, two template sizes of 0.4 x 0.4km and 1 x 1km were selected. In each template size category, the entire reference DEM was sampled at the template size interval, resulting in 83 and 10 samples for Area1, and 90 and 16 for Area2. These sampled templates move over the target DEMs at one-pixel increments and CCF, MI and GMI measures between the sampled template from the reference DEM and its target counterpart are computed. The measures are accumulated in a so-called global similarity matrix until the entire test area has been searched.

The maximum value of the global similarity matrix generated for each similarity metric corresponds to the best achievable alignment between the templates for that metric. If the ground coordinate discrepancies of the template centres fall within the tolerance range of [-1.5, 1.5] pixels at their respective grid size, the translations in x and y are maintained. The translations estimated for each matching method at each template size are then averaged and used to transform the target DEMs. The transformed DEMs were subsequently compared to the LiDAR DEM for quality assessment (Table 2). The success rate, SR%, of each method was computed to quantify robustness, with SR% being defined as the percentage of matched templates to the total number of reference templates.

The results summarised in Table 2 demonstrate that GMI achieved a higher overall success rate than both MI and CCF. The robustness of all matching approaches generally increased in proportion to the template size, since larger templates contain more information, whether it is intensity or gradient values. The difference in success rates is more significant in the following cases:

a) When significant terrain variations or gradient information exist allowing for a higher success rate for GMI. This is seen in Area 1 where SR% values of five to ten times those computed for the CCF and MI are obtained by GMI for the SRTM DEM with the larger template size (1 km²). In contrast, in Area 2 the measures are roughly the same because the terrain is relatively flat.

b) For DEMs with low horizontal resolution, which do not (and cannot) exhibit fine-scale variations in surface topography. Thus, the information content required for a correct matching via MI or CCF does not exist. In Area1, with the SPOT5 dataset and a 1×1 km² template size, the success rate of GMI is 3 - 4 times that obtained with MI CCF, but for the higher resolution ADS40 and InSAR DEMs, GMI shows only a 15% higher average success rate than MI.

The results listed in Table 2 reveal that, for the smaller template size, MI was unable to achieve a matching of the reference and target templates in the low resolution DEMs in Area1, whereas CCF found correct matches, but only for a few samples. This highlights the fact that GMI outperforms both CCF and MI, particularly in the registration of low-resolution DEMs. However, the performance of MI improves in proportion to horizontal resolution, such that MI produces similar results to GMI in Area 2 for the ADS40 DEM, for both template sizes.

Computed estimates for the translation parameters were consistent with the true datum shift in x and y directions for all three methods. The imposed 30m horizontal translation is equivalent to one pixel for the SRTM and SPOT5 DEMs, six pixels for the InSAR DEM and four pixels for the ADS40 DEM.

Table 2. Accuracy evaluation results against the reference DEM after the registration. SR% indicates the success rate, $x_{mean}$ and $y_{mean}$ are the computed translations in pixels, and RMS is the root mean square height discrepancy value after registration

| Site | DEM | Method | Template size (km$^2$) | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | 0.4*0.4 | | | | 1*1 | | | |
| | | | SR% | $x_{mean}$ | $y_{mean}$ | RMS (m) | SR% | $x_{mean}$ | $y_{mean}$ | RMS (m) |
| Area 1 | SRTM | CCF | 5 | 1.5 | 1.0 | 6.1 | 20 | 0.0 | 0.0 | 6.1 |
| | | MI | - | - | - | - | 10 | 2.0 | 0.0 | 6.1 |
| | | GMI | 12 | 1.2 | 0.6 | 6.0 | 100 | 0.8 | 0.4 | 6.0 |
| | SPOT5 | CCF | 2 | 1.0 | 2.0 | 6.3 | 30 | 1.0 | 0.7 | 5.6 |
| | | MI | - | - | - | - | 20 | 1.5 | 1.0 | 5.8 |
| | | GMI | 14 | 1.4 | 1.7 | 6.1 | 90 | 1.1 | 1.0 | 5.6 |
| | InSAR | CCF | 1 | 7.0 | 6.0 | 1.6 | - | - | - | - |
| | | MI | 39 | 6.9 | 7.2 | 1.5 | 30 | 7.0 | 7.3 | 1.5 |
| | | GMI | 39 | 7.0 | 7.1 | 1.5 | 40 | 6.8 | 7.5 | 1.6 |
| | ADS40 | CCF | 2 | 3.0 | 3.0 | 2.0 | 10 | 3.0 | 5.0 | 2.0 |
| | | MI | 7 | 4.0 | 4.0 | 1.7 | 80 | 3.5 | 4.0 | 1.8 |
| | | GMI | 18 | 3.6 | 3.6 | 1.8 | 100 | 3.4 | 4.0 | 1.9 |
| Area 2 | SRTM | CCF | 6 | 0.8 | 1.0 | 2.7 | 56 | 0.9 | 0.8 | 2.7 |
| | | MI | 17 | 0.7 | 1.1 | 2.7 | 75 | 0.3 | 1.1 | 2.8 |
| | | GMI | 27 | 0.7 | 1.1 | 2.7 | 69 | 0.5 | 1.2 | 2.7 |
| | SPOT5 | CCF | 13 | 0.8 | 1.0 | 5.3 | 69 | 0.9 | 1.5 | 5.3 |
| | | MI | 29 | 0.8 | 1.4 | 5.3 | 81 | 0.5 | 1.0 | 5.3 |
| | | GMI | 39 | 0.8 | 1.5 | 5.3 | 75 | 0.4 | 1.3 | 5.3 |
| | InSAR | CCF | 8 | 6.6 | 7.3 | 1.0 | 25 | 6.3 | 6.3 | 1.0 |
| | | MI | 36 | 6.5 | 7.1 | 1.0 | 50 | 6.2 | 7.0 | 1.0 |
| | | GMI | 11 | 6.6 | 7.0 | 1.0 | 42 | 6.8 | 6.8 | 1.0 |
| | ADS40 | CCF | 9 | 4.4 | 3.9 | 1.0 | 33 | 4.3 | 4.0 | 1.0 |
| | | MI | 12 | 3.9 | 3.5 | 1.1 | 92 | 4.4 | 3.8 | 1.0 |
| | | GMI | 10 | 4.5 | 4.0 | 1.0 | 92 | 4.4 | 3.8 | 1.0 |

Improvements in the quality of registration indicated by the RMS elevation discrepancy values relative to the reference DEM for Area 2, as shown in Table 2, over the equivalent values in Table 1, were 0.7, 0.5, 0.8 and 0.9 m for the SRTM, SPOT5, InSAR and ADS40 datasets, respectively. The difference in RMS values for each target DEM obtained through the use of different matching criteria, namely CCF, MI and GMI, was negligible. For Area1, however, except for high resolution ADS40 and InSAR DEMs where MI and GMI both outperformed CCF, the GMI method produced a more accurate registration, this being 0.1 m and 0.2 m superior to that from MI for SRTM and SPOT5 DEMs, respectively.

## 4. CONCLUSION AND OUTLOOK

An investigation into the potential of mutual information-based horizontal registration of multi-resolution DEMs has been reported, the motivation for the research being in part due to the proven success of the MI approach for registration of multi-sensor medical imagery. Both the MI and GMI methods have been shown to produce robust and accurate results in the registration of four raster DEM datasets, with post spacings ranging from 5 to 30m, to a 1m LiDAR DEM. The performance of the MI-based methods was also compared with the CCF approach and MI and GMI were shown to yield more robust registrations. Furthermore, the registration success rate was significantly improved under the GMI method, particularly for the lower resolution DEMs.

In the reported MI-based registration approach, the entire test area was scanned in the template localisation process. However, a more efficient search strategy could restrict the search area to only templates with rich information content. For example, a surface roughness criterion could be used to assess whether sufficient terrain information is available to support robust MI-based matching.

## REFERENCES

Costantini, M., Malvarosa, F., Minati, F. and Zappitelli, E., 2006. A Data Fusion Algorithm for DEM Mosaicking: Building a Global DEM with SRTM-X and ERS Data. Geoscience and Remote Sensing Symposium (IGARSS). pp. 3861-3864.

Gerke, M., 2008. Dense Image Matching in Airborne Video Sequences. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Beijing, China, Vol. XXXVII, Part B3b, pp. 639-644.

Hirschmueller, H., 2008. Stereo processing by Semi-Global Matching and Mutual Information. IEEE Transactions on Pattern Analysis and Machine Intelligence 30(2), pp. 328-341.

Hirschmueller, H., Scholten, F. and Hirzinger, G., 2005. Stereo vision based reconstruction of huge urban areas from an airborne push-broom camera (HRSC). In: Lecture Notes in Computer Science: Pattern Recognition, Proceedings of the 27th DAGM Symposium, Vol. 3663, Springer, Berlin, pp. 58-66.

Krauß, T., Lehner, M., Reinartz, P., 2008. Generation of Coarse Models of Urban Areas from High Resolution Satellite Images. The International Archives of the Photogrammetry, Remote

Sensing and Spatial Information Sciences, Beijing, China, Vol. XXXVII, Part B1, pp. 1091-1098.

Li, Z., Bethel, J., 2008. DEM registration, alignment and evaluation for SAR interferometry. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B1, pp. 11-115.

Maintz, J. B. A., van den Elsen, P. A. and Viergever M. A.,1996. Comparison of edge-based and ridge-based registration of CT and MR brain images. Medical Image Analysis, Vol. 1, No. 2, pp. 151–161.

Olson, Clark F., 2001. Image Registration by Aligning Entropies. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'01), Vol. 2, pp.331-336.

Penney, G. P., Weese, J., Little, J. A., Desmedt, P., Hill, D. L. G. and Hawkes, D. J.,1999. A comparison of similarity measures for use in 2D-3D medical image registration. IEEE Transactions on Medical Imaging, Vol. 17, No. 4, pp. 586–595.

Roth, A., Kntipfle, W., Rabus, B., Gebhard, S. and Scales, D., 1999. GEMOS –A system for the geocoding and mosaicking of interferometric digital elevation models. Geoscience and Remote Sensing Symposium (IGARSS), pp. 1124-1127.

Shannon, C.E., 1948. A mathematical theory of communication. Bell System Technical Journal, Vol. 27, pp. 379–423 and 623-656.

Van Niel, T.G., Mcvicar, T.R., Li, L., Gallant, J.C. and Yang, Q., 2008. The impact of misregistration on SRTM and DEM image differences. Remote Sensing of Environment journal, 112(5), pp. 2430-2442.

Viola, P. and Wells III, W. M., 1996. Alignment by maximization of mutual information. *International Journal of Computer Vision*, Vol. 24, No. 2, pp. 137-154.

Zitova, B., Flusser, J., 2003. Image Registration Methods: A Survey. Image and Vision Computing (21), No. 11, pp. 977-1000.