SUB-PIXEL PRECISION IMAGE MATCHING FOR DISPLACEMENT MEASUREMENT OF MASS MOVEMENTS USING NORMALISED CROSS-CORRELATION

Misganu Debella-Gilo¹ and Andreas Kääb²

¹²Institute of Geosciences, University of Oslo, P. O. Box 1047, Oslo, Norway
¹m.d.gilo@geo.uio.no (corresponding author), ²kaeb@geo.uio.no,

KEY WORDS: Normalised cross-correlation, Sub-pixel, Image matching, Displacement measurement, Rockglacier, Glacier, Rock slide

ABSTRACT:

This study evaluates the performance of two fundamentally different approaches to achieve sub-pixel precision of normalised cross-correlation when measuring surface displacements on mass movements from repeat optical images. In the first approach, image intensities are interpolated to a desired sub-pixel resolution using a bi-cubic interpolation scheme prior to the actual displacement matching. In the second approach, the image pairs are correlated at the original image resolution and the peaks of the correlation coefficient surface are then located at the desired sub-pixel resolution using three techniques, namely bi-cubic interpolation, parabola fitting and Gaussian fitting. Both principal approaches are applied to three typical mass movement types: rockglacier creep, glacier flow and rock sliding. Their performance is evaluated in terms of matching accuracy and in reference to the images of the resolution they are expected to substitute. Our results show that intensity interpolation using bi-cubic interpolation (first approach) performs best followed by bi-cubic interpolation of the correlation surface (second approach). Both Gaussian and parabolic peak locating perform weaker. By increasing the spatial resolution of the matched images by intensity interpolation using factors of 2 to 16, 40% to 80% reduction in mean error could be achieved in reference to the same resolution original image.

1. INTRODUCTION

Present climatic change shifts geomorphodynamic equilibriums and intensifies related mass movement processes such as landslides and permafrost creep (Haeberli and Beniston 1998; Rebetez et al. 1997). Extension and intensification of human activities in areas affected by such mass movements increase the probability of connected adverse impacts like natural hazards or building stability problems. The consequently growing needs for monitoring mass movements are complemented by growing remote sensing opportunities for doing so. The increasing number of available stacks of multi-temporal space-borne, air-borne and terrestrial images, and the improvements in remote sensing and image processing in general significantly enhance the potential for applying matching techniques to detect and quantify earth surface mass movements from repeat remotely sensed data. All the above needs and developments call for continued efforts to improve terrain displacement matching methods based on repeat images.

Image matching is a group of techniques of finding corresponding features or image patches in two or more images taken of the same scene from different viewing positions, at different times and/or using different sensors (Zitová and Flusser 2003). Image matching is among others used for a large variety of applications such as image (co-)registration, stereo parallax matching for generation of digital elevation models, particle image velocimetry (PIV), or displacement measurements.

The group of area-based matching techniques is the most widely used method due to its relative simplicity (Zitová and Flusser 2003). A number of similarity criteria can be used for the matching process. Cross-correlation, in particular its normalised form which accounts for intensity variations in image sequences, is the most widely used due to its reliability and simplicity (Lewis 1995). The normalised cross-correlation (NCC) algorithm has been used to investigate earth surface mass movements such as glacier flow, rockglacier creep and land sliding in many empirical studies (Haug et al. 2010; Kääb and Vollmer 2000; Scambos et al. 1992).

Although NCC has been documented to be simple and reliable, a number of drawbacks have been reported as well: Firstly, its precision is, in principle, limited to one pixel, and thus varying with the pixel size of the image data used. Secondly, NCC is sensitive to noise in the images. Such noise may result in wrong correlation maxima leading to mismatches. This problem is often partly addressed through image transformation (e.g. Fourier) in case of images with low signal-to-noise ratio (Lewis 1995). Thirdly, NCC is sensitive to significant scale, rotation or shearing differences between the images to be correlated (Zhao et al. 2006). Due to this limitation, NCC is recommended in cases where the movement is mainly due to translation with limited rotation, scaling or shearing. A way to partly overcome this limitation of NCC is to orthorectify the images used before the matching (Kääb and Vollmer 2000). Fourthly, for the measurement to be reliable the displacement has to be greater than the mean error of the image (co-)registration. Improving NCC precision improves displacement accuracy twofold: firstly, it reduces image registration error; secondly, it improves the matching accuracy directly.

To achieve sub-pixel precision in NCC, three approaches can be used. One is to improve the imaging system towards a higher spatial resolution. This approach is complicated by a number of
financial and technological limitations. The second option is to resample the image intensity to a higher spatial resolution through interpolation. The third option is to interpolate the cross-correlation surface after the matching process to a higher spatial resolution in order to locate the correlation peak with sub-pixel precision. Since the intensity interpolation approach and the correlation interpolation approach are both generic, and independent of image resolution, they are subject to this study. A number of investigations have explored these two approaches applied to medicine (Althof et al., 1997; mechanics (Westerweel 1993; Willert and Gharib 1991; Zhou and Goodson 2001), and stereo matching and motion tracking (Karyabli et al. 2008; Yamaguchi et al. 2003). However, there is no study available that compares the relative performance of the two approaches when measuring the displacement of earth surface mass movements from repeat images.

Many earth surface mass movements such as landslides, glacier flow, and rockglacier creep are characterized by displacement rates of the order of magnitude of cm$^{-1}$ or ma$^{-1}$ which is often less than the spatial resolution of the space-borne or air-borne imagery typically available for their measurement. Sub-pixel accuracy of image matching algorithms, here NCC, has therefore a large potential to improve the signal-to-noise ratio of the measurements. Using NCC as an example, this study compares the performance of two fundamentally different approaches to reaching sub-pixel precision in mass movement detection and measurement from repeat remotely sensed images.

In the method section of this contribution we describe the dataset used, ways of reaching sub-pixel precision, quantification of matching accuracy, and our experimental set-up. In further sections the matching results and their accuracy are presented and discussed. Short conclusions terminate our contribution.

2. METHODS

2.1 Image data and pyramid

For this study, three different types of mass movements were selected: landslide, glacier flow, and rockglacier creep. The selection of these mass movement types was made based on their frequency in high mountain areas. Three temporal pairs of images each covering one of these types of earth surface mass movements were used (Table 1). These images were accurately orthorectified prior to displacement matching. Additionally, one image pair was created from just one original image and the movement applied was only translation the pair (offset between the images) was computed by matching stable grounds. The computation revealed that a maximum of 1 pixel offset exists in each dimension. The positions of the templates with valid matches in the original resolution were then used in the matching of the coarser resolution images.

Matching and displacement measurement were in a next step performed on all resolution levels of the image pyramid pairs for all those locations saved from the reference matching. The absolute sizes and positions of the reference templates and the search windows were kept constant metrically throughout the image pyramid by adjusting the number of pixels according to the resolution. In other words, the ground area covered by the templates remained the same, the respective image resolution changed. This was done in order to avoid variations in signal content as a result of inclusion or exclusion of ground features. The area covered by the images range from 0.25km$^2$ to just over 3km$^2$. The size of the template was kept at around 26m and 65m for the originally 0.2m and 0.5m resolution images, respectively. The size of the search window was kept at around 102m and 265m for the originally 0.2m and 0.5m resolution images, respectively, so that it certainly included the expected maximum surface displacement.

<table>
<thead>
<tr>
<th>Type</th>
<th>Location</th>
<th>Pixel size</th>
<th>Older</th>
<th>Recent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rockglacier</td>
<td>Muragl (Swiss)</td>
<td>0.2m</td>
<td>1981</td>
<td>1994</td>
</tr>
<tr>
<td>Glacier</td>
<td>Ghiacciaio del Belvedere (Italian)</td>
<td>0.5m</td>
<td>Sep. 2006</td>
<td>Oct. 2006</td>
</tr>
<tr>
<td>Rock slide</td>
<td>Aletsch (Swiss)</td>
<td>0.2m</td>
<td>1976</td>
<td>2006</td>
</tr>
<tr>
<td>Control (Glacier)</td>
<td>Manually translated motion</td>
<td>0.2m</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Brief description of the image data used
2.3 The sub-pixel precision approaches

2.3.1 Intensity interpolation
The coarse resolution images within the above-computed resolution pyramids were back-interpolated to different finer resolutions using the MATLAB-based ‘imresize’ module. Again the bi-cubic interpolation was used for the same reason. After such back-interpolation, the NCC algorithm was applied using the same templates and search windows as used in the original reference image pairs. The interpolation is done on the fly for each reference template and search window, and not for the entire image before the matching process. This was done due to memory restriction by MATLAB.

2.3.2 Similarity interpolation
Bi-cubic interpolation. To find the sub-pixel position, one can interpolate the cross-correlation surface to higher resolution using two dimensional bi-cubic interpolation algorithms. The algorithm uses a two-dimensional cubic convolution of the correlation coefficients to the resampled grid. The peak is then relocated.

Curve fitting. As an alternative to peak interpolation, one can also create a continuous function that optimally fits the correlation coefficient data and compute the precise location of the peak from the maximum of the function. The challenge is that no single function can usually perfectly describe the cross-correlation surface. However, the fact that the correlation surface around its peak often approaches a bell shape can be exploited. Therefore, two dimensional polynomial functions can approximate the surface. A number of interpolation models have been tested in empirical and theoretical researches, particularly in particle image velocimetry (PIV), though with varying successes (Nobach and Honkanen 2005; Westerweel 1993; Willert and Gharib 1991). Some of the well performing ones will be tested here for mass movement analysis. These are parabola fitting and Gaussian fitting, as these have shown successes especially in PIV.

In parabola fitting, the shape of the correlation surface is assumed to fit two separable orthogonal parabolic curves. The location of the ‘actual’ peak is computed by independently fitting one dimensional quadratic function and computing the location of the peak (Nobach and Honkanen 2005; Westerweel 1993).

In Gaussian fitting, the bell shape of the correlation surface is assumed to fit a 2D Gaussian function (Nobach and Honkanen 2005; Westerweel 1993; Willert and Gharib 1991). It is assumed that the two dimensions are separable and orthogonal. Thus, the sub-pixel peak location is calculated separately for the two directions by fitting a second-order polynomial to the logarithm of the maximum sample and the direct neighbours.

2.4 Evaluation of different levels of sub-pixel detail
Section 2.3 evaluates which sub-pixel approach performs best in improving the precision and accuracy of NCC-based image matching. It is also important to know how far sub-pixel interpolation of coarse resolution image intensities or the correlation surface is able to substitute pixel-level matching of images of the corresponding but original resolution. In other words, what is the sub-pixel detail at which the interpolation to achieve sub-pixel precision can no longer sufficiently substitute image of that resolution. The approach used here to resolve this issue is to compute the sub-pixel precision matching at different levels of the image pyramid and evaluate its performance in reference to the pixel-level matching of images with the same but original resolution. This issue becomes clearer with an example. Suppose we want to know the performance of sub-pixel precision matching at the level of half a pixel. This can be achieved by taking an image of, for instance, 8m resolution, compute the sub-pixel precision matching to 4m and compare the latter sub-pixel performance to the performance of pixel-level matching of an image with 4m original resolution. Or else, take a 4m resolution image, compute its sub-pixel resolution matching to 2m and compare the performance of the latter in relation to a 2m resolution original image, and so forth including the entire pre-processed image pyramid and all resolution steps included in it.

2.5 Performance evaluation
As an indicator of accuracy, we used the shift in matching position instead of the often-used difference in displacement magnitude. The matching positions obtained during the correlation of the original images were considered as references. All the matching positions at the different coarser or back-interpolated resolutions were compared to these reference positions. The magnitude of this offset (deviation) is here used as measure for the accuracy of the image and algorithm used. The deviation between the matching position of the interpolated image and that of the same resolution original image is used to assess the relative performance of the sup-pixel approaches.

3. RESULTS

3.1 Displacements of the different mass movement types
Table 2 summarises displacement statistics for the three mass movements investigated. The results are produced from the analyses of the original ortho-images after filtering all the mismatches. One can well see that the glacier moves very fast as compared to the rockglacier and the even slower moving rockslide. Figure 1 and Figure 2 present the displacement vectors of the three mass movements. Image matching showed that all the areas in the scene show non-zero displacements due to the presence of systematic image (co-)registration error. However, after filtering of the vectors based on the estimated overall image (co-)registration error of one pixel, thresholding of the correlation coefficients and excluding upslope movements, only the remaining vectors presented in the figures are considered to be valid and useful as reference.

3.2 Accuracy of the sub-pixel algorithms
Figure 3 and Figure 4 depict the mean deviation of the matching positions against the sub-pixel precisions of each of the sub-pixel approaches for the control set and the three mass movements respectively. The magnitudes of Figure 4 are created by averaging the values obtained for the three mass movement types as the trend is very similar for all the three. Both figures show that interpolation of the image intensity before matching results in the best matching accuracy. If one looks at the interpolation of the correlation surface, the bi-cubic approach follows the intensity interpolation. The curve fitting using parabola and Gaussian models perform only better than bi-cubic interpolations to one half of the original pixel size.
For the real mass movements, there is very little accuracy gain by interpolating to lower than 0.1 pixels. As the result of the control shows, when the movement is only translation, the magnitude of the deviation is very low. Besides, it seems that, for the control set, interpolation to more detail level improves the accuracy further.

<table>
<thead>
<tr>
<th>Mass movement</th>
<th>Mean displacement (m)</th>
<th>Maximum displacement (m)</th>
<th>Standard deviation of displacement (m)</th>
<th>Maximum velocity (m/α)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aletsch Rockslide</td>
<td>1.5</td>
<td>4.2</td>
<td>0.45</td>
<td>0.14</td>
</tr>
<tr>
<td>Muragl Rockglacier</td>
<td>2.4</td>
<td>5.8</td>
<td>1.20</td>
<td>0.45</td>
</tr>
<tr>
<td>Ghiacciaio del Belvedere Glacier</td>
<td>12.22</td>
<td>18.83</td>
<td>5.0</td>
<td>226</td>
</tr>
<tr>
<td>Control</td>
<td>7.50</td>
<td>7.50</td>
<td>0</td>
<td>7.50</td>
</tr>
</tbody>
</table>

Table 2. Summary statistics for the displacement magnitudes and average velocity of the mass movements and the translation-only control image as estimated from the matching of the high-resolution original ortho-images

For example, a 16m resolution image interpolated to 2m using bi-cubic interpolation before matching performs exactly as a 2m resolution image pair as long as there is no other source of difference between the image pairs than rigid translation. But when the level of detail goes beyond 1/8th, there appears deviation between the two. The magnitude of these numbers depends, of course, on the translation magnitude applied in the control set. However, the test shows the better performance of bi-cubic intensity interpolation over the other sub-pixel algorithms tested.

For all the real mass movement types (Figure 6), as the difference in pixel size between the coarse resolution and the reference resolution increases, the deviation of the sub-pixel matching position from the matching position of the same (but original) image resolution increases regardless of the algorithm. This means, not surprisingly, that the sub-pixel algorithm resembles images of comparable resolution less and less as the sub-pixel detail increases. At every resolution, the mean deviation is the lowest when intensity interpolation is used before matching followed by the bi-cubic interpolation of the correlation surface. The parabola- and Gaussian-based peak localisations perform poorer and alike. This confirms the above results.

Remarkably, at a certain level of sub-pixel detail (about 1/16th), the deviation between the sub-pixel algorithm and same resolution original image gets so high that interpolating beyond that level has no meaningful advantage although the control set gives less deviation even at greater level of detail.

Figure 3  Accuracy of the different sub-pixel precision approaches for the control set expressed as the mean deviation of the matching positions from that of the reference high-resolution original ortho-images

Figure 4  Accuracy of the different sub-pixel precision approaches expressed as the mean deviation of the matching positions from that of the reference high-resolution original ortho-images (averaged for the three mass movement types)

3.3 Relative performance of the sub-pixel precisions

Figure 5 shows the mean deviation between the matching position of the interpolated image pairs and that of the same resolution (but original) reference image pairs plotted against sub-pixel precision for the control set. When the difference between the images is only the here-applied translation, sub-pixel interpolation of the image intensities up to 1/8th of the original pixel size prior to matching can perfectly substitute images of comparable original resolution. This perfect substitution can be achieved by using bi-cubic interpolation of the correlation surface only up to 1/4th of the original pixel size.
Figure 5  Relative performance of the different sub-pixel approaches for the control set expressed as the mean deviation of the matching positions from that of the same resolution original image.

Figure 6  Relative performance of the different sub-pixel approaches expressed as the mean deviation of the matching positions from that of the same resolution original image (averaged from the three mass movement types investigated).

4. DISCUSSION

The results show that intensity interpolation outperforms all the other algorithms of similarity interpolation. There can be two explanations to this. Firstly, in correlation interpolation the positions of the correlation values on which the interpolation is based, and which are computed based on coarse resolution images, influence the position of the recomputed correlation peak. Secondly, the number of pixels in an entity is higher when intensity interpolation is applied leading to the suppression of noise. Fewer numbers of pixels in an entity makes the entity more susceptible to chance-based, i.e. erroneous matching results. This explains the increased difference between intensity interpolation and similarity interpolation at very detailed levels of sub-pixel precision.

The bi-cubic interpolation scheme that was used for the intensity interpolation is known to replicate the reference data better than most interpolation schemes (Keys 1981), and it is known to approximate the sinc interpolation that is ideal in image interpolation (Dodgson 1992). This has led to the fact that the images re- interpolated from coarser resolutions were found to have high correlation with the aerial images of corresponding original resolution. For example, when the down- sampled rock glacier image of resolutions 0.4m, 0.8m, 1.6m, 3.2m and 6.4m were re- interpolated to a resolution of 0.2m (1/2 to 1/32 of a pixel respectively) their global correlation coefficients with the reference image of 0.2m resolution were 0.98, 0.96, 0.93, 0.90 and 0.86 respectively. Although the images deteriorate due to resampling noise, they still remain well-correlated with the reference image due to the good performance of the interpolation algorithm. Correlation is, in fact, one of the quality measures of image interpolation (Lehmann et al. 1999).

The same interpolation algorithm, bi-cubic, performed best in the similarity interpolation approach although not as good as in the intensity interpolation. The better performance in comparison to the Gaussian and parabola fitting is partially ascribed to the reasons explained above. In addition to that, parabola fitting is reported in many occasions to have a systematic bias known as “pixel locking”, which forces the estimated sub-pixel locations to approach integer values (Nobach and Honkanen 2005; Prasad et al. 1992). The presence of a systematic bias is testified by the fact that both parabola and Gaussian fitting could not fully substitute the same resolution original images in the case of the control set unlike the other two algorithms (Figure 5). Although reports from PIV state that Gaussian peak finding does not have that kind of bias and performs better (Westerweel 1993; Willert and Gharib 1991), it performed no better than parabola fitting in the present study. We believe the underlying reason is the fact that the cross-correlation surfaces of the mass movements cannot be perfectly modelled by either parabolic or Gaussian functions. The image resolutions used in the present study are not so high to be compared to that of particle images used in mechanics which is high enough to be approximated by, for example, Gaussian. Besides, noise that is present in the images due to temporal surface changes and other sources contribute to the deviation of the correlation shape from both Gaussian and parabolic.

Finally, two important points regarding the size of the matching entities: Firstly, in this study the absolute metric size of the matching entities was kept constant across image resolutions. This means that the number of pixels in each entity varies with the pixel resolution, leading to a variable signal-to-noise ratio. This was done for the sake of comparison. In reality, the size of matching entities will vary with the resolution of the image pair to keep a good signal-to-noise ratio. Secondly, the size of the matching entities was kept the same for the entire scene. In reality matching entities vary in size.

5. CONCLUSIONS

This study has clarified a number of questions around the relation between accuracy and pixel or sub-pixel resolution when matching terrain displacements such as glacier flow, land sliding or permafrost creep from repeat optical images by using pixel-precision correlation measures, here namely the normalized cross correlation (NCC). That way the study contributes, on the one hand, to better exploiting the unexploited archives of repeat remotely sensed images that exist over actual or potential earth surface mass movements, and on the other hand, to better meeting the increasing needs to quantify and monitor mass movements, in particular when they are accompanied by adverse effects.

This study has in particular evaluated the performance of two different approaches to sub-pixel precision in NCC for displacement measurement based on repeat images. When sub-pixel accuracy is aimed for, interpolating image intensities to a
higher resolution using bi-cubic interpolation prior to the actual image correlation performs better than both interpolation of the correlation surface using the same algorithm and peak localisation using curve fitting. Correlation peak localisation using Gaussian and polynomial algorithms are inferior in such applications.

Therefore, we conclude that more precise and accurate displacement measurements are obtained by interpolating the available images to a higher resolution using bi-cubic interpolation prior to matching. In such approaches, one can gain over 40% reduction in mean error by interpolating the images up to 1/16th of a pixel. Interpolating to a more detailed sub-pixel resolution than 1/16th of a pixel does not add much. Or in other words, when matching low-resolution images using normalized cross-correlation with intensity-interpolation based sub-pixel precision, 40% or better accuracy increment can be achieved compared to pixel-precision matching of images in reference to the same original resolution as the interpolated one. When real low-resolution images are used together with varying sizes of the matching entities, as opposed to the approach used in this study, even better precision and accuracy might be obtained as the noise due to resampling will not be present, and template and search window sizes will be adjusted with the pixel size.

It should also be noted that although the relative performances of the algorithms is expected to be valid at least for other spatial domain matching approaches and for other applications, the magnitudes given here are strictly only valid for the similarity measure and test sites used in this paper. Further research is needed for their validity outside the conditions described in this study.

REFERENCES


Haug, T., Kääb, A., & Skvarca, P. (2010). Monitoring ice shelf velocities from repeat MODIS and Landsat data – a method study on the Larsen C ice shelf, antarctic Peninsula, and 10 other ice shelves around Antarctica. The Cryosphere Discussions (in review), 4, 35-75


ACKNOWLEDGEMENTS

The research was conducted at the Geosciences department of the University of Oslo and financially supported by the Norwegian Research Council (CORRIA project). The authors are very grateful to both institutions.