SYNTHETIC APERTURE SONAR SNIPPET IMAGE REGISTRATION

S Dugelay Centre for Maritime Research and Experimentation (CMRE), La Spezia, Italy.
dugelay@cmre.nato.int

1 INTRODUCTION

In mine-hunting operations, the use of autonomous systems removes men and ships from the minefield. In order for these systems to perform their missions effectively, the systems themselves and the collected data should be exploited efficiently in order to maximize performance. Automatic target recognition (ATR) is a key enabling component of achieving this goal. Whilst immense progress has been made to design optimal ATR algorithms, they still require further improvement to perform at the level of an operator in all environments. One way to increase ATR performance is to fuse multiple pieces of information on a given target, either at data level (early integration), at feature level (intermediate integration) or at classification level (late integration). This paper details an algorithm for the registration and incoherent fusion of multiple views of synthetic aperture sonar snippets of targets at data level. The algorithm, based on minimizing variation in the fused image, works on the assumption that the strongest highlights belong to the target and viewing angles are less than or equal to 90°. The first assumption can be partially mitigated by only taking small snippets around the target and minimising the probability of stronger highlights within the image. The second assumption sits well within the CMRE MUSCLE Autonomous Underwater Vehicle (AUV) *modus operandi* (shown in Figure 1) where a first view of a target is acquired in a track-adaptive lawn mower pattern and a second view is acquired at 90° or less in a fully autonomous reacquisition mode.

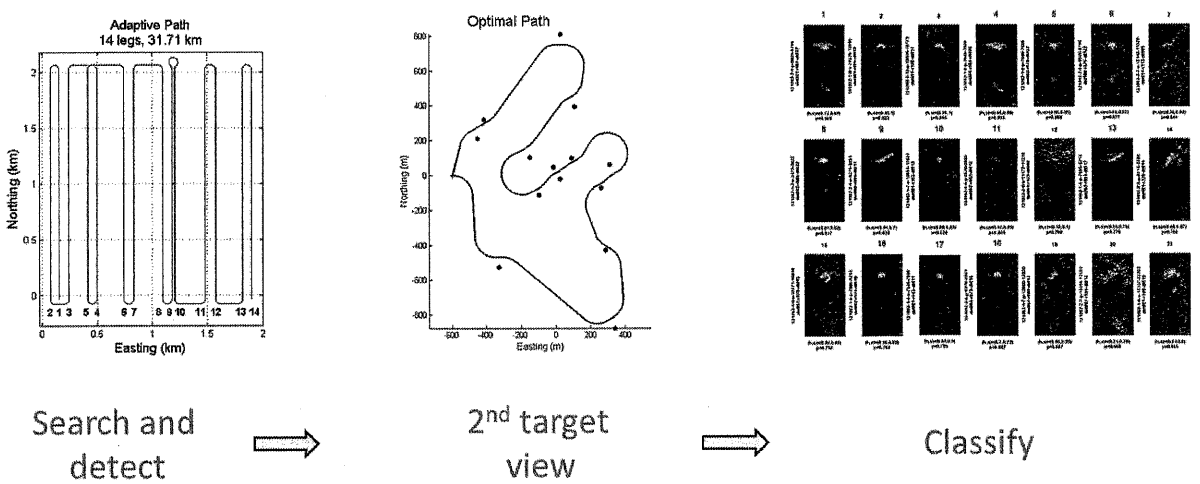

Figure 1: *Modus operandi of the CMRE MUSCLE system for mine search and classification.*

Results obtained on data collected at sea by a synthetic aperture sonar (SAS) system mounted on an autonomous underwater vehicle (AUV) demonstrate the clear potential of the approach. This paper also includes initial results of image quality and content assessment to possibly explain registration performance.

This paper is organized as follows. Section 2 describes the algorithm for registration and fusion as well as some measures of image quality. Section 3 presents some of the results obtained from the incoherent fusion of snippets alongside the measures of quality. Finally conclusions and ideas for future work are given in Section 4.

2 REGISTRATION, FUSION AND QUALITY

2.1 Registration and Fusion

The research undertaken at CMRE on image registration focuses on the alignment of image snippets of associated contacts automatically detected by the CMRE on-board automatic target recognition algorithms [1]. An example of two such snippets is given in Figure 2. These two images collected on the same object at 90° difference in aspect illustrate the typical registration problem to be solved.

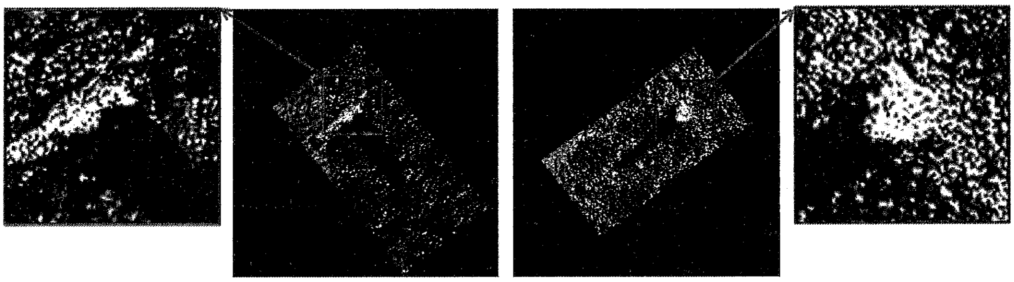

Figure 2: *Examples of image snippets collected on the same object at 90° difference in aspect angles to be registered to build an improved highlight picture of the object.*

Image registration is a broadly researched topic for a wide variety of applications (see [2] for a survey of techniques). However, due to the particular nature of sonar images, and the specific angular diversity of images to fuse, most previously developed approaches cannot be directly applied to sonar snippets of detected objects. The most common approaches can be divided into three categories: feature-based techniques, area-based methods and model-based approaches.

Feature-based techniques typically rely on automatically detected salient anchor points within two images to align them. Whilst this is a highly popular and effective technique in optical applications, it is virtually impossible to use this approach for sonar applications as salient features in one image at a given azimuthal aspect may look different or simply not appear in an image of the same scene at a different viewing angle. This is indeed the case in the example of Figure 2 where a landmark is clearly visible at one aspect angle (in the small image on the left, just above the target highlight), but is not seen at a perpendicular viewing angle (small image on the right).

Area-based methods usually use a correlation-based estimation between smaller areas of the two images to perform the registration. Again, because of the nature of sonar images the use of correlation is impractical when aligning images of an object viewed at largely different aspect angles. The most obvious obstacle is that the goal of the registration is to fuse object highlights whilst discarding object shadows. Additionally, as shown in Figure 3 using the examples from Figure 2, the registration result shows a superposition of the highlights using a correlation-based approach instead of a gradual build-up of the object contour as required.

A model-based approach will firstly compare the image to a set of a priori known objects, match the best option and register the images using models of the objects. This is an extremely robust and viable method if all objects are known in advance. The goal of the registration in this study is to build a better picture of the object before assigning a label. In the case of the examples shown in Figure 2 one aspect (on the left) would correlate highly with the model of a cylinder, whilst the other aspect (on the right) would correlate well with a truncated cone. Hence, model-based approaches are ill-suited for fusing sonar snippets when the goal is to increase object information prior to labelling.

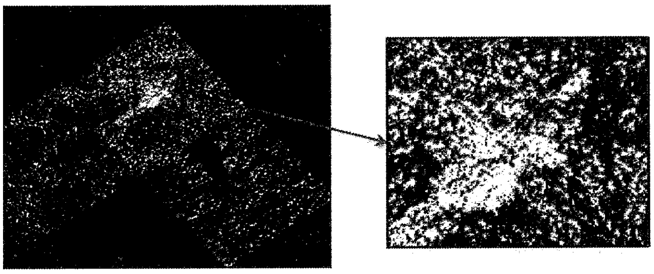

Figure 3: *Correlation-based snippet registration.*

Williams [3] first presented an alternative approach for snippet image registration which combined the use of sonar imagery and bathymetric maps. The approach is based on minimising the variation within the fused image in order to estimate translation between snippets. In this study, solely sonar images have been employed in the registration process using an adaptation of the method described in [3]. In the early stages of this research, two assumptions are made:

1. The difference in viewing angles between two detection snippets does not exceed 90 degrees.

2. The rotation to be applied is estimated from the sonar heading and snippets are rotated accordingly to a common reference frame. Whilst this is not optimal, it is considered as a first step in a chain of registration estimates and the current focus is on estimating translation.

Locations of high value information in the SAS snippets are the detection highlights which tend to present strong image gradients. The registration approach therefore exploits this characteristic by employing the Laplacian transform of the image. The Laplacian of an image is a useful operator to detect rapid changes and can give a measure of the total variation contained within an image. The Laplacian of an image $I(x,y)$ is given by

$$L(I) = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2}$$

And the total variation in an image is given by:

$$TV(I) = \sum_x \sum_y |L(I)|$$

The estimation of the translation between two images is effected in the following manner. For a given row and column translation $\tau(r,c)$, the fused image FI is calculated by taking the maximum pixel value at each pixel location, i.e.:

$$FI_\tau(x,y) = \max\big(I_1(x,y), I_2(x+r, y+c)\big), \forall(x,y)$$

The total variation of the fused image is then calculated for every possible translation τ:

$$TV(FI_\tau) = \sum_x \sum_y |L(FI_\tau)|$$

The best estimation of translation is assumed to be achieved when the total variation of the fused image is minimum. The problem is then to find the best translation $\hat{\tau}$ that minimises the total variation of the fused image:

$$\hat{\tau} = \mathrm{argmin}(TV(FI_\tau))$$

2.2 Image Quality

2.2.1 DPCA coherence rate

The Displaced Phase Centre Antenna (DPCA) algorithm produces a cross-correlation between two consecutive sonar pings which is used to estimate vehicle motion (and subsequently compensate for motion in the synthetic aperture sonar focusing), but is also used to estimate maximum imaging range achieved during survey as a quality criterion [4]. The cross-correlation for a window $t = [t_i, t_j]$ is given by:

$$\varsigma_{k,k+1}(\tau) = \sum_{t=t_i}^{t_j} p_k^*(t) p_{k+1}(t+\tau)$$

Where τ is a time shift. The normalised peak correlation ρ is then:

$$\rho = \max_{\tau} \left| \frac{\varsigma_{k,k+1}(\tau)}{\sqrt{\varsigma_{k,k}(0)\varsigma_{k+1,k+1}(0)}} \right|$$

An average DPCA coherence rate is calculated for each snippet as an indicator of overall image quality.

2.2.2 Sharpness

The sharpness of an image is calculated using a simple and basic approach based on the estimation of gradients. It is not meant to be used as an absolute value but rather as a measure of comparison to assess whether relative sharpness has an effect on the performance of image registration. The gradient of a function of two variables $F(x,y)$ is defined as [5]:

$$\nabla F = \frac{\partial F}{\partial x}\vec{i} + \frac{\partial F}{\partial y}\vec{j}$$

The sharpness is then calculated as:

$$S = \frac{1}{M}\sum_x \sum_y \sqrt{\left(\frac{\partial F}{\partial x}\right)^2 + \left(\frac{\partial F}{\partial y}\right)^2}$$

where M is the number of pixels in the image. It is to be noted that this measure of sharpness is closely correlated with the Laplacian of the image (double derivative) which is employed in the registration process. Sharpness in the results sections is quoted in dB (i.e. $20.\mathrm{Log}_{10}(S)$).

2.2.3 Contrast

Contrast is typically defined as the difference in luminance or colour which makes an object distinguishable by eye. In this study the contrast between key parts of the image are calculated as the ratio of the signal mean (target highlight) over the standard deviation of the noise (here background amplitude). This is also defined as the imaging signal-to-noise ratio for target highlights. The *Rose* criterion [6] states that to be able to distinguish image features at 100% certainty, a SNR of at least 5 (or ~7dB) is required. Two contrasts are defined. The signal-to noise (SNR) ratio (contrast between the highlight and the seabed) is given by:

$$SNR = \frac{\frac{1}{H}\sum_{i=1}^{H} x_i^H}{\sqrt{\frac{1}{B}\sum_{i=1}^{B}\left(x_i^B - \frac{1}{B}\sum_{i=1}^{B} x_i^B\right)^2}}$$

where in this case the superscript H relates to highlight only pixels and the superscript B relates to background pixels. The contrast between shadow and seabed (SHNR) is expressed by a similar definition:

$$SHNR = \frac{\frac{1}{S}\sum_{i=1}^{S} x_i^S}{\sqrt{\frac{1}{B}\sum_{i=1}^{B}\left(x_i^B - \frac{1}{B}\sum_{i=1}^{B} x_i^B\right)^2}}$$

3 RESULTS

3.1 Truncated cone

Figure 4 shows two examples of fusion performed with snippets of a truncated cone. In the first example at the top, two views taken at angles differing by approximately 40° are registered and fused correctly despite a reduced image quality (see measures in Table 1). Coherence is below the standard 2/3 accepted as an indicator of good quality and SNRs are below the 7dB Rose criterion making the highlight difficult to distinguish. Tactics may dictate that given the inferior image quality, the next target view should be acquired at angles less than 45°. When the image quality is improved such as in example 2, viewing angles can differ by up to 90° and the fusion process will still provide excellent results.

3.2 Cylinder

Figure 5 illustrates two examples of fusion performed with snippets of a cylinder. In both examples, the two views are of similar quality and in both cases the views have long edge of the cylinder in common. The registration process correctly aligns the two views; further in house results have shown that this may not always be the case if the only the short edge is common to both views. Additional constraints are now being investigated to improve registration.

3.3 Rock

Finally, Figure 6 exemplifies the registration and fusion of snippets of a rock. Image qualities are again comparable despite a reduced sharpness for the second example. Registration results are good and it is anticipated that in these cases the additional information will help operators and eventually automatic target recognition to differentiate between man-made and natural objects with increased certainty.

4 CONCLUSIONS AND FUTURE WORK

This paper has presented a method for SAS snippet registration to build up an object highlight for improved operator and eventual ATR classification performance. The registration approach is based on minimising the variation in the fused image. The fusion is achieved by taking the maximum pixel amplitude between images in order to emphasise the highlight construction. Additionally, some image quality measures are calculated to assess influencing factors on registration process. Initial results have shown good performance of both registration and fusion when image quality is high or when viewing angles differ less than 45°. Further in-house results have demonstrated a need for additional constraints when angles differ by more than 45° and image quality is reduced. Further work at CMRE will now focus on a large scale performance assessment of registration and fusion with the current method but also will investigate the introduction of additional constraints for difficult environments and for viewing angles that differ by more than 90°.

5 REFERENCES

1. Williams, D.P., Groen, J., Fox, W.L.J., *A Fast Detection Algorithm for Autonomous Mine Countermeasures*, NATO Unclassified, NATO Underwater Research Centre, NURC-FR-2011-006, 2011, 44 pages.

2. Zitova, B., Flusser, J., Image registration methods: a survey, *Image and Vision Computing*, Vol. 21, pp 977-1000, 2003.
3. Williams, D.P., SAS and Bathymetric Data Fusion for Improved Target Classification, Proceedings of the International Conference on Underwater remote Sensing, October 2012.
4. Couillard, M., Fox, W.L.J., Groen, J., and Davies, G.L., *Synthetic Aperture Sonar Performance Estimation: Measurements and Modelisation*, NATO Unclassified, Centre for Maritime Research and Experimentation, Report CMRE-FR-2012-004, 2012, 40 pages.
5. Kumar, J., Chen, F., Doermann, D., *Sharpness Estimation for Document and Scene Images*, Conference proceedings 21st International Conference on Pattern recognition, 2012, pp 3292-3295
6. Pawley, J., *Handbook of Biological Confocal Microscopy*, Chapter 2, Third Edition, ISBN: 987-0-387-25921-5, Springer, New York.

Table 1: Image measures for each view of a truncated cone.

Example	View	Coherence	H/B SNR	B/S SNR	Sharpness
1	1	0.65	5.9	-5.5	60.6
	2	0.66	5.6	-5.0	61.2
2	1	0.89	10.	-6.	64.5
	2	0.86	9.4	-5.5	66.8

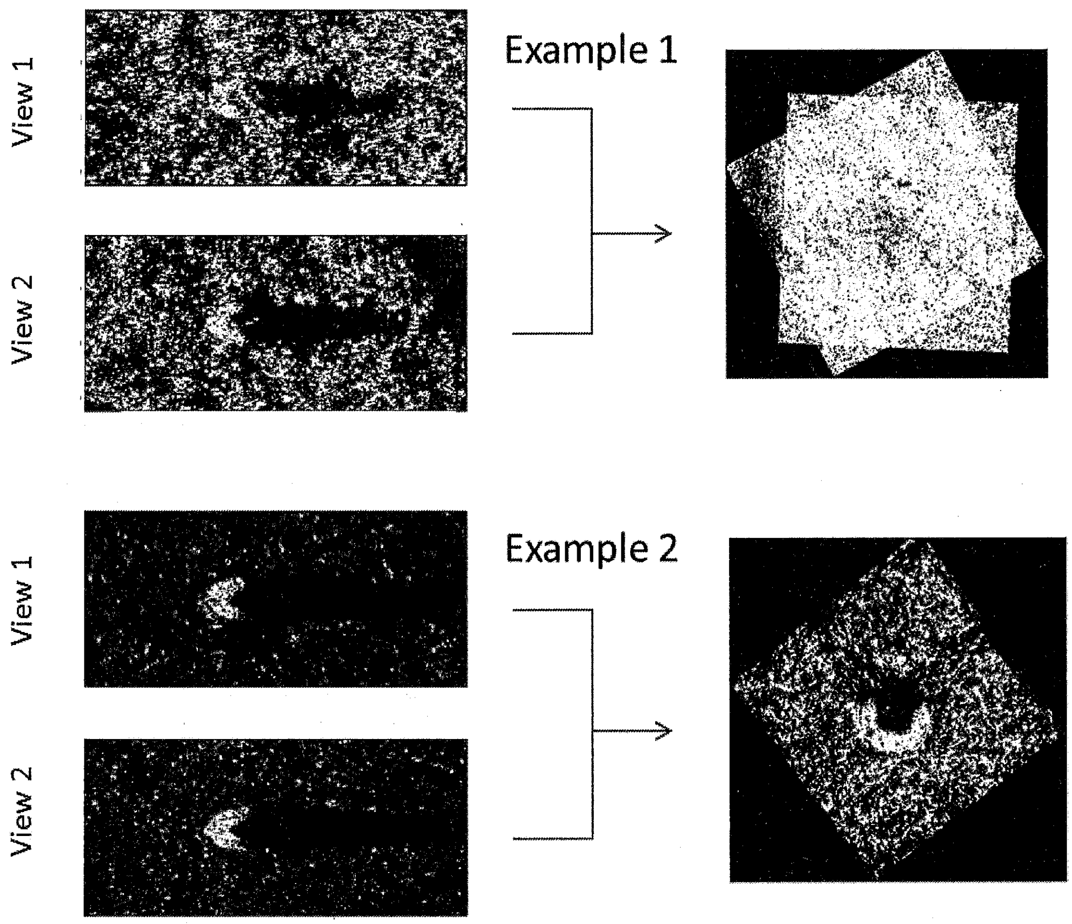

Figure 4: *Two examples of registration and fusion of two views on a truncated cone. The first two views in the top set are angled at ~40°. The second set at the bottom, viewing angles differ by 90°.*

Table 2: Image measures for each view of a cylinder.

Example	View	Coherence	H/B SNR	B/S SNR	Sharpness
1	1	0.84	10.	-6	64.
	2	0.79	7.5	-5.	67.
2	1	0.83	11.	-7.	63.
	2	0.7.	12.	-4.	63.

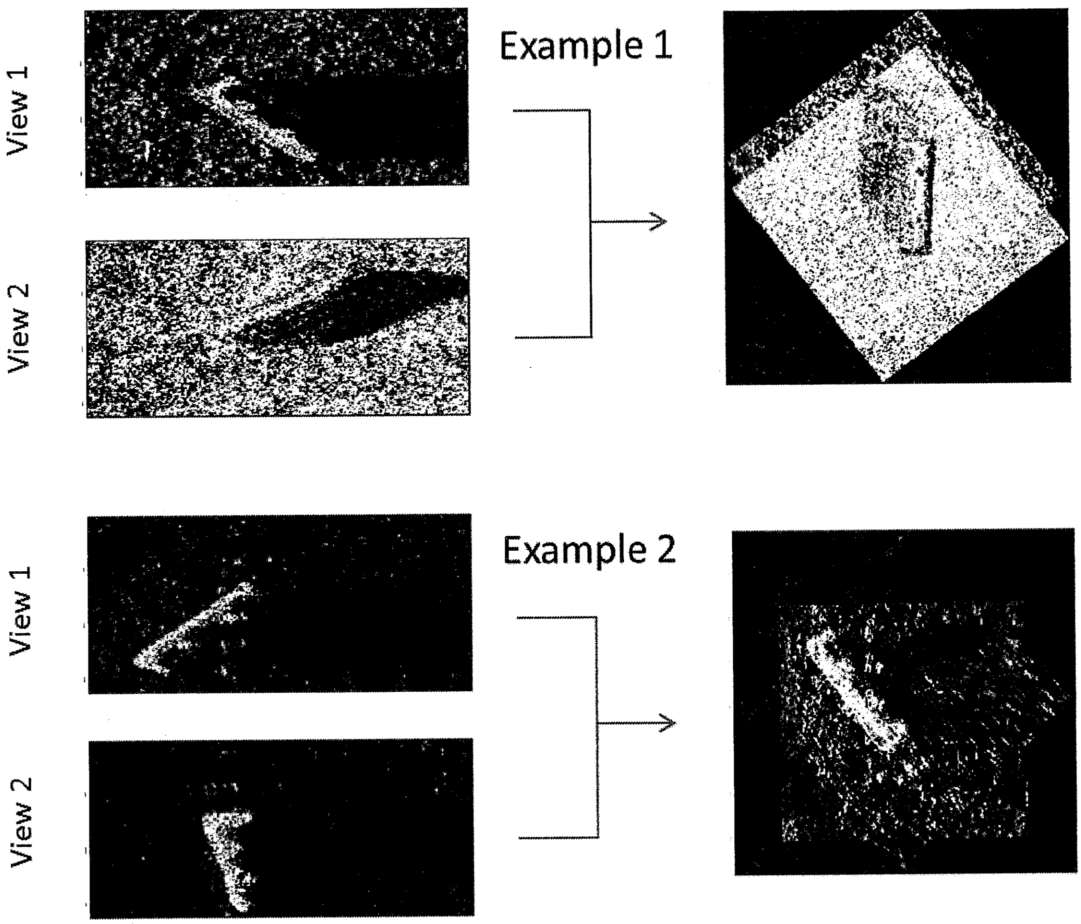

Figure 5: *Two examples of registration and fusion of two views on a cylinder. The first two views in the top set are angled at 90 degrees. The second set at the bottom, viewing angles differ by 45 degrees.*

Table 3: Image measures for each view of a rock.

Example	View	Coherence	H/B SNR	B/S SNR	Sharpness
1	1	0.77	12.	-4.	62.
	2	0.86	10.	-7.	67.
2	1	0.85	11.	-5.3	54.6
	2	0.81	9.8	-5.8	58.

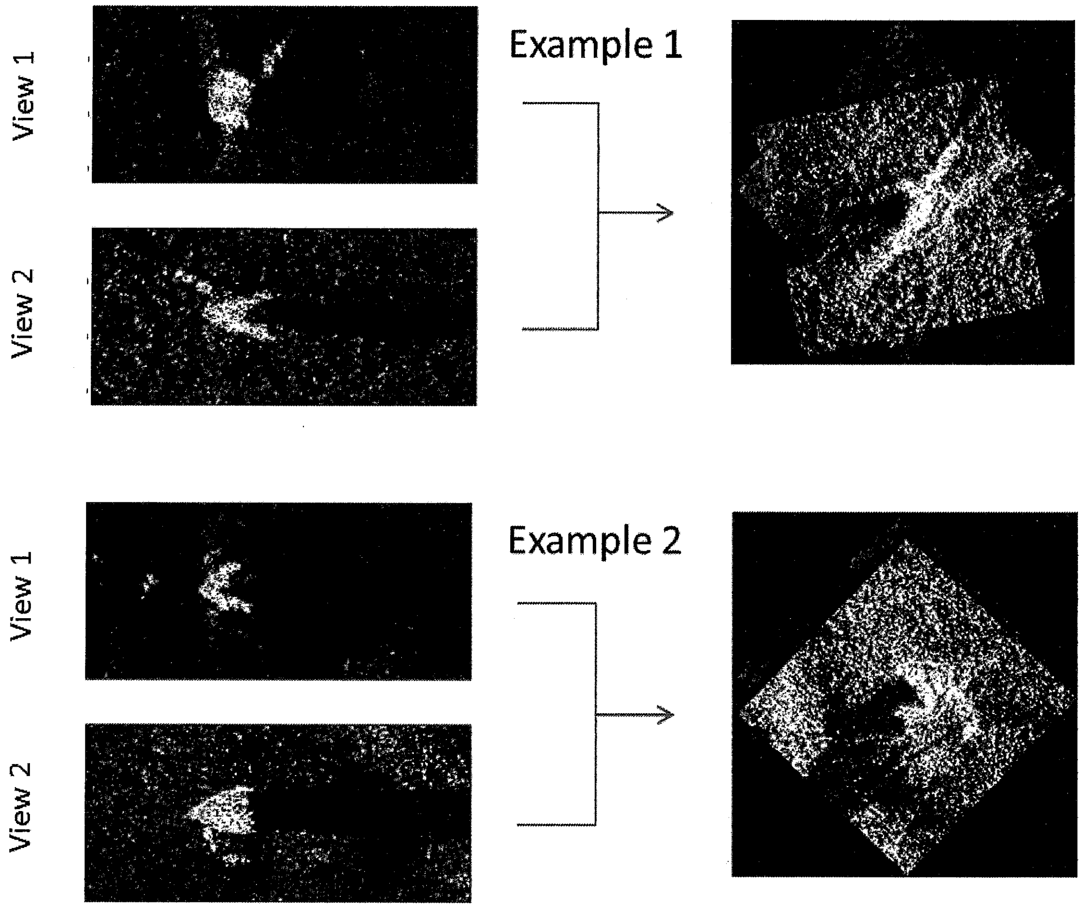

Figure 6: *Two examples of registration and fusion of two views on a rock. The first two views in the top set are angled at ~40°. The second set at the bottom, viewing angles differ by 90°.*