Mutual Information Registration

Innoventix Consulting, Deon Joubert, 13 January 2015

The elephant survey system finds elephants within aerial images by first detecting possible targets in an Infrared (IR) image and then classifying the targets as being elephants or not by examining the corresponding visual image. As images from two different cameras are used, it is important that the matrix describing the alignment, or registration, between the IR and visual images be accurate, otherwise detected targets would not be classified using the correct region within the corresponding visual image.

The current system makes use of a fixed homography matrix, calculated using manual feature mapping on several IR and visual image pairs, to describe the alignment. The application of Mutual Information (MI) on a per image basis could improve the accuracy of the registration and thereby counteract the variations in alignment introduced by noise and vibration.

An investigation was conducted to determine whether MI registration would lead to a noticeable and viable improvement in classification results. Furthermore, the effect on accuracy of the size of the image area used during classification was also investigated.

Pre-processing

Direct application of the MI algorithm is difficult due to the difference in resolution of the two cameras as well as the non-linear relationship in the contrast values of objects in the images. The extent of the differences is illustrated in Figure 1.

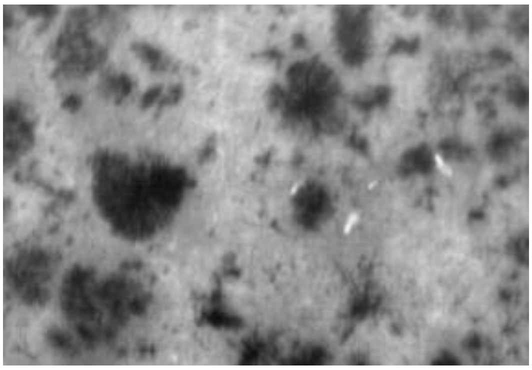

Figure 1. An IR (left) and visual (right) image pair of the same scene showing the variation brought on by resolution differences and non-linear contrast relationships. Note the lack of detail of the bushes in the IR image when compared to the visual image. Also note the difference in colour of the elephants.

To reduce the difficulty of the registration problem, the visual images were converted to a colour space closer in appearance to the IR image. The LAB colour space represents images in terms of luminosity and the A and B chroma values. The A chroma value indicates the colour value of a pixel on the red-green colour axis. As cool regions in an IR image of a nature reserve usually correspond to green vegetation and warmer areas correspond to exposed red-brown ground surfaces, the A chroma value is particularly useful for the current application.

The visual image was also scaled down to more closely match the resolution of the IR image. The resultant image is shown in Figure 2, where in comparison with Figure 1 it can be seen that the two image are far more similar in appearance and therefore less difficult to align using MI registration.

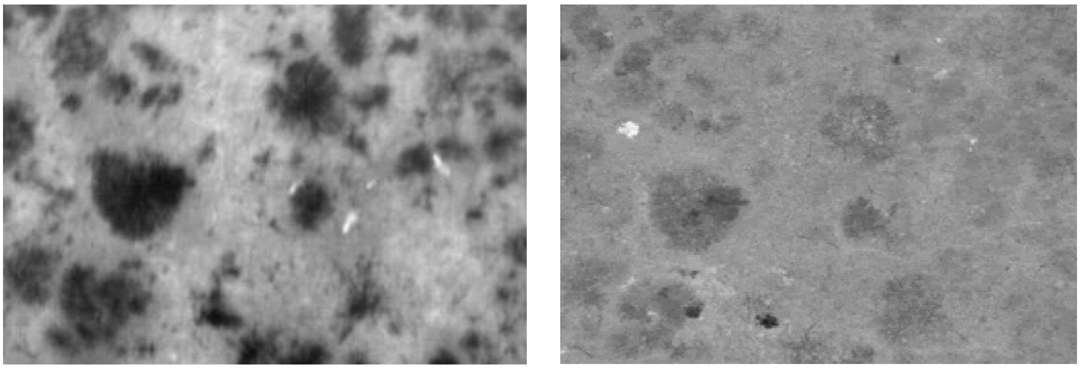

Figure 2. An IR (left) and LAB A chroma value (right) image pair of the same scene, showing the closer relationship between the two images.

Implementation

The MI registration algorithm was applied after the target detection stage of the system. The algorithm takes as input an initial transform from which to compute an estimate of the true transform. The initial transform was based on the homography matrix currently in use.

The visual image was first converted to the LAB colour space and the A Chroma value colour channel extracted after which the image was scaled down. The MI algorithm requires a reference position in the IR image from which to extract an image patch. The reference position should be selected for every image pair, according to the following criteria:

- The region should be close to the detected elephants, ensuring that the estimated transform is most accurate when applied to the positions of the elephants.
- The regions should look approximately the same in both image, thereby reducing the difficulty of the registration.
- The region should contain some object which is recognisable in both images and has a distinctive, oriented shape, further reducing the difficulty.

An example of such a point is shown in Figure 3. The current MI implementation uses the mean of the detected target positions as the reference point until a more sophisticated point selection algorithm satisfying all three criteria can be developed.

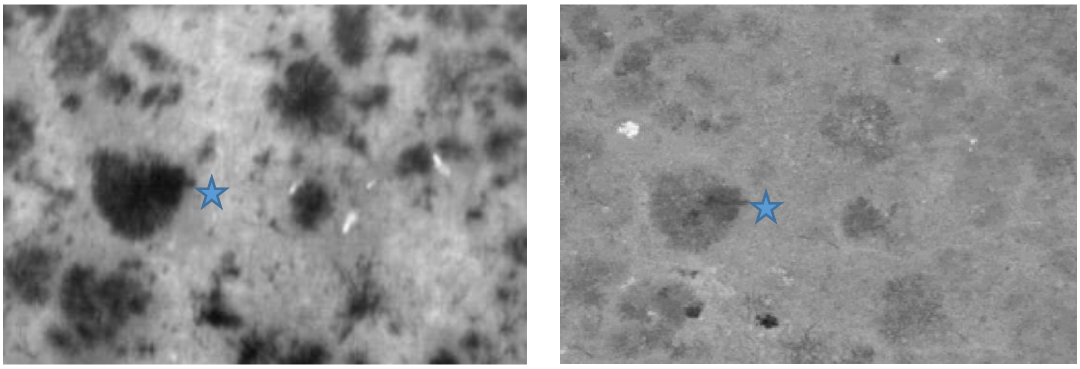

Figure 3. An IR (left) and LAB A (right) image pair with a selected reference point (blue star). The reference point was chosen as it is close to the elephants (as seen in the left image), the trees and bushes in the surrounding area appears similar in both images and have a distinctive half circle shape.

After the MI transform has been computed, it is used to project the position of the detected targets into the visual image. An image region surrounding each target is then extracted and used to classify whether the target is an elephant or not. The classification area is visualised as the blue square surrounding the yellow star target in Figure 4. Note that the standard RGB colour space is used during classification, and not the LAB colour space.

Evaluation

As the MI refinement is computationally expensive, it was decided to adapt the existing MATLAB code to use multiple processing threads. In future, parallel processing will allow faster evaluation of new modifications to the system.

The two-stage constant false alarm rate (CFAR) algorithm was used to detect targets. Two successive Viola-Jones cascade classifiers were used to identify the targets as being elephants or not, the first using histogram of gradient (HOG) features and the second using Haar-like features. Both classifiers were trained using upright-orientated data.

The classification subsystem, using either the homography-based or MI-refined registration, was evaluated using a small dataset of 121 IR and visual image pairs.

Results

Figures 4, 5 and 6 show examples of how the image registration of the dataset has been improved by indicating how the animals are more accurately segmented during the transformation of their detected positions.

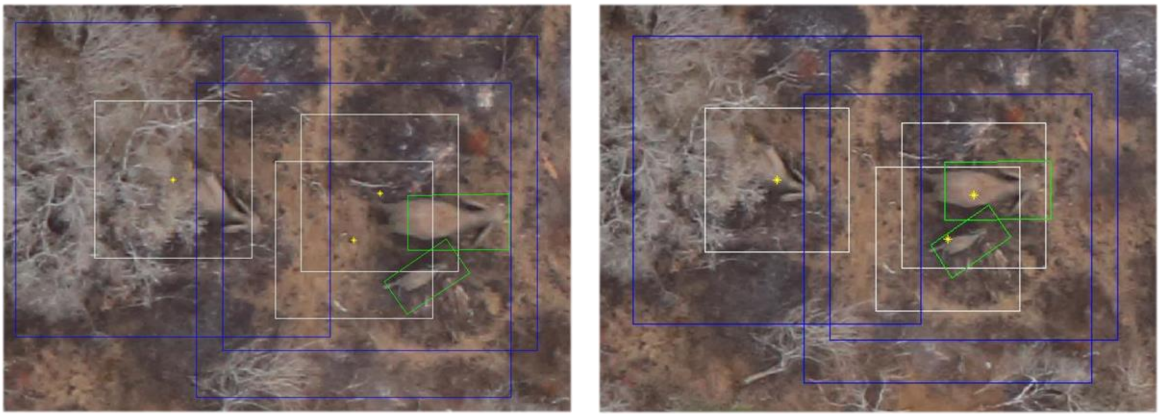

Figure 4. Detected position of the elephants (yellow dots) using homography matrix (left) and using MI registration (right).

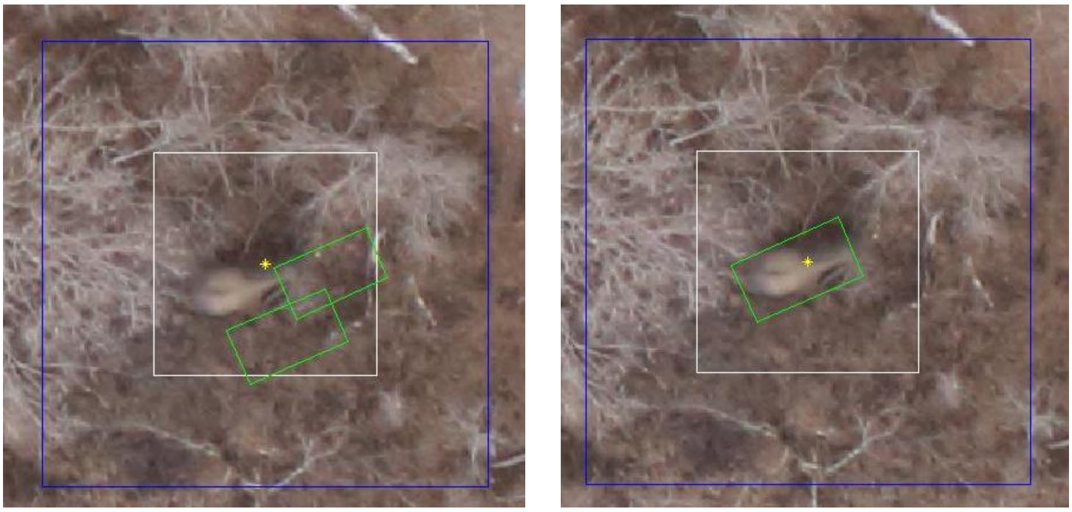

Figure 5. Detected position of the elephants (yellow dots) using homography matrix (left) and using MI registration (right). Note the lack of the double classification error in the MI Registered image.

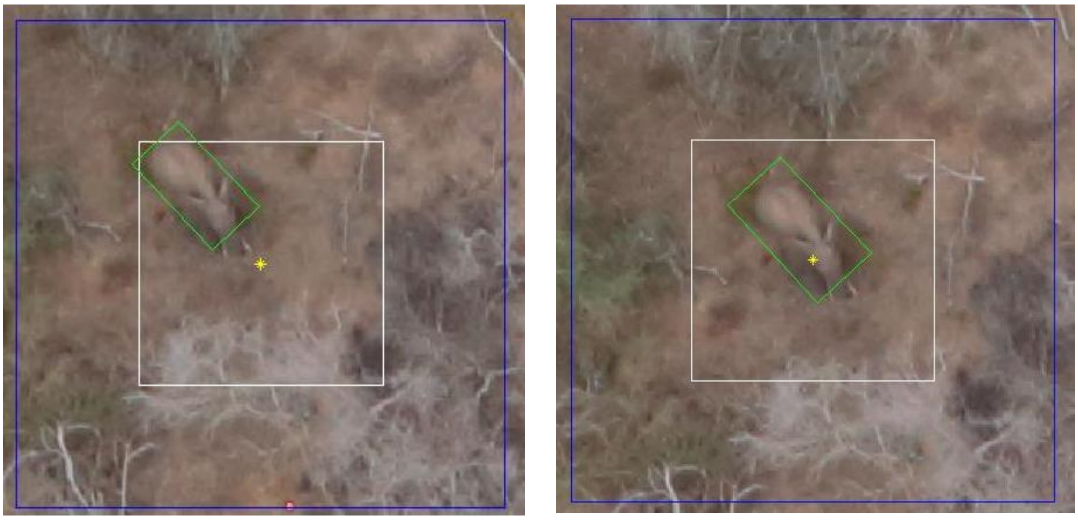

Figure 6. Detected position of the elephants (yellow dots) using homography matrix (left) and using MI registration (right).

Comparison between homography-based and MI-refined registration

From the examples it can be seen that there is a definite increase in the accuracy of registration. However, it is still required to investigate whether the improved registration leads to an increase in classification accuracy.

The classification results using homography-based and MI-refined registration are shown in Table 1. Here it can be seen that an increase of about 2% was achieved in the overall correctly classified targets using MI-refined registration with a slight decrease in the number of false classifications per hectare. Computation using the MI-refined registration was about 5 times longer.

	Homography-based	**MI-refined registration**	**Number of instances per category**
Visible Adults (%)	70.89	73.35	529
Visible Calves (%)	82.05	79.49	39
Partially Visible Elephants (%)	42.72	44.94	316
Total (%)	**61.31**	**63.46**	884
False classifications (per ha)	0.0693	0.0624	-
Computation Time (s)	604.78	2523.05	-

Table 1. The percentage correct classifications for a number of categories using homography-based and MI-refined registration, as well as the number for false classifications per hectare and the total computation time.

Classification region size

As can be seen in Figure 4, the size of the region used for classification is relatively large due to the fact that the homography-based registration was not as accurate. A smaller classification region should in theory lead to a more accurate and faster classification as less, potentially confusing, "non-elephant" data needs to be processed.

The size of the classification region was modified using an *expansion factor* of which the results can be seen in Table 2. An expansion factor of 1.0 was used to generate the original MI-refined result in Table 1. As can be seen, an expansion factor of 0.8 leads to an additional increase in accuracy of 1%. However, the increase comes at the cost of more false classifications. An expansions factor of 0.9 leads to greater accuracy without generating more false classifications than the homography-based registration seen in Table 1.

Expansion Factor	1.0	0.9	0.8	0.6
Visible Adults (%)	73.35	73.91	74.86	70.89
Visible Calves (%)	79.49	82.05	82.05	79.49
Partially Visible Elephants (%)	44.94	44.94	44.62	43.35
Total (%)	**63.46**	**63.91**	**64.37**	**61.43**
False classifications (per ha)	0.0624	0.06588	0.1040	0.1000
Computation Time (s)	2523.05	2011.16	2078.42	2262.33

Table 2. Classification results using MI-registration and varying sizes of the classification region.

Note that there is a great deal of variation in the computation time of the system, and therefore no conclusions will be made based on this data.

Conclusion

The results show that an MI registration algorithm can be used to improve the alignment of visual and IR images on a per image pair basis given that the appropriate pre-processing is carried out. However, the increase in the overall number of correct classifications is slight, and most likely not worth the much larger computational cost of using MI-refined registration. A reduction in processing time through optimisation could lead to a more viable MI-refined registration implementation