Defence Science Journal, Vol. 52, No. 3, July 2002, pp. 243-251

Algorithm for Fast Registration of Radar Images

Subrata Rakshit and Dipti Deodhare

Centre for Artificial Intelligence & Robotics, Bangalore – 560 001

ABSTRACT

Radar imagery provides an all-weather and 24 h coverage, making it ideal for critical defence applications. In some applications, multiple images acquired of an area need to be registered for further processing. Such situations arise for battlefield surveillance based on satellite imagery. The registration has to be done between an earlier (reference) image and a new (live) image. For automated surveillance, registration is a prerequisite for change detection. Speed is essential due to large volumes of data involved and the need for quick responses.

The registration transformation is quite simple, being mainly a global translation. (Scale and rotation corrections can be applied based on known camera parameters). The challenge lies in the fact that the radar images are not as feature-rich as optical images and the image content variation can be as high as 90 per cent. Even though the change on the ground may not be drastic, seasonal variations can significantly alter the radar signatures of ground, vegetation, and water bodies. This necessitates a novel approach different from the techniques developed for optical images. An algorithm has been developed that leads to fast registration of radar images, even in the presence of specular noise and significant scene content variation. The key features of this approach are adaptability to sensor/terrain types, ability to handle large content variations and false positive rejection. The present work shows that this algorithm allows for various cost-performance trade-offs, making it suitable for a wide variety of applications. The algorithm, in various cost-performance configurations, is tested on a set of ERS images. Results of such tests have been reported, indicating the performance of the algorithm for various cost-performance trade-offs.

Keywords: Image processing, radar imagery, battlefield surveillance, image registration, image matching, algorithms, radar images, guidance

1. INTRODUCTION

Radar satellite imagery is extensively used in defence because it provides an all-weather, and a 24 h coverage of target areas. However, from the image processing point of view, radar images pose a special challenge because these are often very noisy and do not contain the well-defined geometric features present in optical imagery. As a result, image processing algorithms that process radar imagery tend to be computation-intensive as they try to remove the noise before detecting corners, and lines and contours. On the other hand, there exists a class of applications which require fast registration of radar images. These applications arise in the context of guidance and battlefield surveillance. In the first application, it is necessary to register an image acquired by an onboard radar with a stored reference image. The objective of the registration is to obtain a position fix so as to enable a course correction of the vehicle. In order for this image processing module to be useful, the registration must be as fast as possible. Typically, the onboard navigation

Received 21 December 2001

system provides reliable information regarding height and orientation of the vehicle. Using this information, the two images can be corrected for rotation and magnification. A similar situation arises in satellite surveillance. Images obtained on successive visits to an area have little scale or rotation variations, but usually have significant translation offsets. Due to large volumes of data generated daily, it would be desirable if an automated process could cue a human to look at only those areas where there have been some changes. In order to do this, the images must first be registered with single-pixel accuracy. While this can be done manually, it is a very tedious and time-consuming task. The need for fast algorithms arises in this context due to sheer volume of data. It is to be noted that these applications do not require image interpretation–they only require image matching.

An excellent survey of various methods[1], which includes methods developed for satellite images, medical images (MRI, PET) and computer vision applications. The overall schema identified for any registration is: (i) choice of features, (ii) choice of metric, and (iii) a search algorithm for identifying the solution. The present work fits this schema. However, certain characteristics of the radar images rule out the use of the standard techniques used to implement the first two modules. These characteristics only will be highlighted here.

In most of the applications, the change in image content between two images being registered is implicitly assumed to be small. This is a good assumption for optical images, in general. View point changes, lighting and movements of some objects leave the overall structure of the image unchanged while changing the details of individual objects in the scene. Such a situation allows the use of global transforms (FT-based correlation, FT phase-based techniques) as well as multiresolution approaches (coarse-to-fine methods). However, in seasonally variant radar images, often, only a few small details remain invariant between images while appearance of the vast majority of the image changes. This rules out all global transforms as well as multiresolution techniques except for certain areas, special areas like big cities. Thus, the general solution for radar images must be able to ignore the majority of the image and focus on the few invariant structures present.

The above constraint indicates that a high level, feature-based approach may be suitable for the problem. As indicated, the desirable characteristics of features are: (i) invariance under expected sources

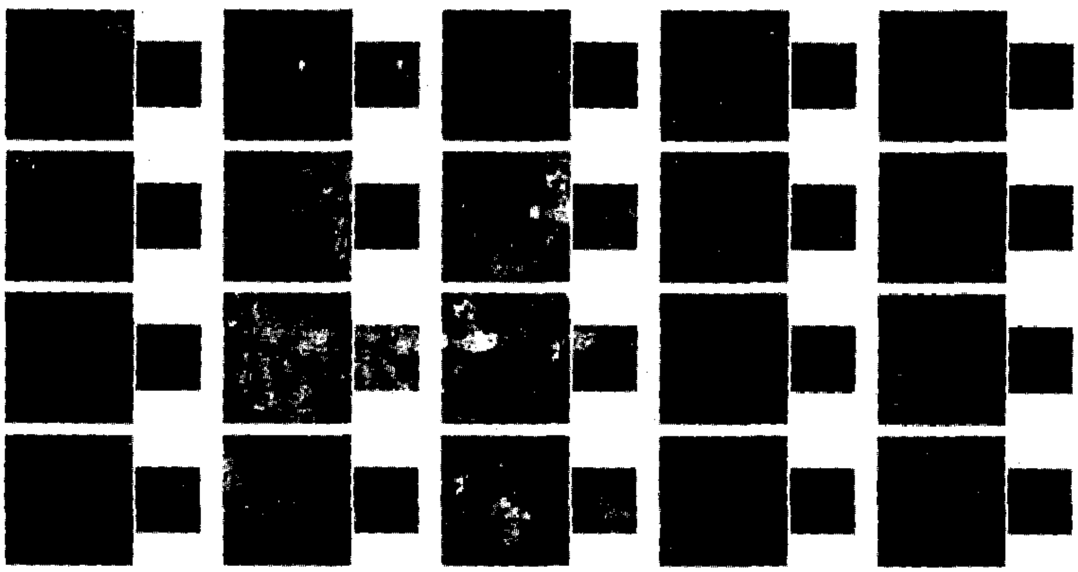


Figure 1. Sample ERS-1 radar images. The big images were used to compute the features using principal component analysis. Various terrain types are shown. From left: Rivers (col. 1), urban areas (cols. 2, 3), vegetation (col. 4) and hills (col. 5). The small images show the central areas of the big images, but in the dry season.

of distortions, (ii) relevance to the problem, and (iii) low computation costs. While the above criteria are relevant, the peculiarities of radar images taken across seasons rule out the usual features listed in the survey[1], such as points, edges, corners, lines, moments, centroids, and higher level syntactic description. Once again, these peculiarities are focussed rather than giving an exhaustive analysis of each possible feature reported in the literature.

Radar images are grainy due to specular noise inherent in any coherent imaging system. Thus, these can never be as smooth as optical images. The use of any point or line/curve-based feature requires expensive smoothing operations that can model the noise correctly. The radar images are characterised by bright spots: Corner reflectors (urban structures, bright linear features (mountain sides), dark regions (water bodies, such as rivers or lakes) and various types of textures (ground, vegetation). Thus, local, area-based texture features that capture bright spots as well, seem to be ideal. Higher level semantic structures are expensive and unreliable to compute. Geometric features based on large regions/objects are very much prone to variation due to seasonal effects. Finally, an invariance wrt bias and contrast in the images is desirable.

This study presents a registration technique that works under the constraints posed by seasonally variant radar images. A brief outline of the method is presented for completeness, though it has already been reported[3].

2. REGISTRATION ALGORITHM

The registration algorithm consists of three steps: (i) feature set construction, (ii) feature extraction, and (iii) feature matching. The first step is adaptive – it produces features tuned to a particular imaging radar. The second step extracts features from the stored reference images. For each reference image, feature vectors are computed for various sub-images of size equal to the live image. These sub-images are sampled from the reference image with a sampling rate, r, determined by the expected accuracy of the system. These two steps are computationally intensive but must be done only once. These are typically done as an offline set-up step. The feature vector matching step is the third and the last step. This step has to be performed online for each live image and is optimised for speed. A single-feature vector is computed for the live image as a whole. It is then compared with the pre-computed feature vectors of the reference image sub-images. This produces a match coordinate and a confidence measure which indicates whether the match coordinates should be accepted.

2.1 Feature Set Design

The decision to use area features was taken to deal with the noise present in radar imagery. In particular, one would like to do a feature-level characterisation without having to do any noise removal. This would imply that larger the feature size, the better. However, large features have an associated disadvantage in that they lead to a loss of resolution. If features are defined as 64×64 patterns, then the output of such a feature detector may not change if the input changes by only a few pixels. Thus, the dual considerations of noise immunity and resolution constrain the choice of feature size. Initially, a feature size of 32×32 had been chosen[3] and a feature set designed for that size. That analysis has been extended to feature sizes of 32×32, 24×24, 16×16, and 8×8 in the present work. Since the task at hand

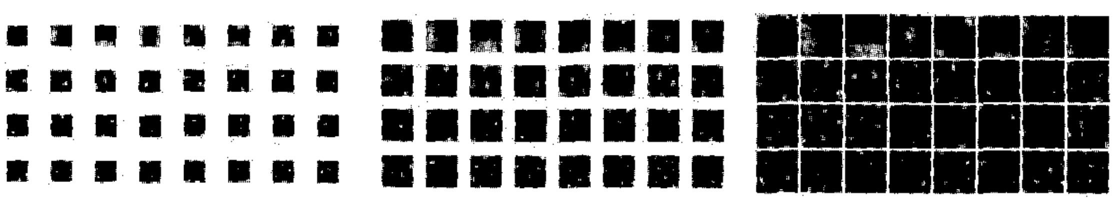


Figure 2. Features extracted from ERS images at 16 × 16, 24 × 24 and 32 × 32 resolutions. For each size, only first 32 principal components are shown.

requires discrimination rather than modelling/reconstruction, principal components (PCs) are more appropriate than eigenvalues as choice of features.

For each feature size (W), the feature set design proceeds as follows:

(a) Extraction of $W \times W$ sub-images from a data set of reference images

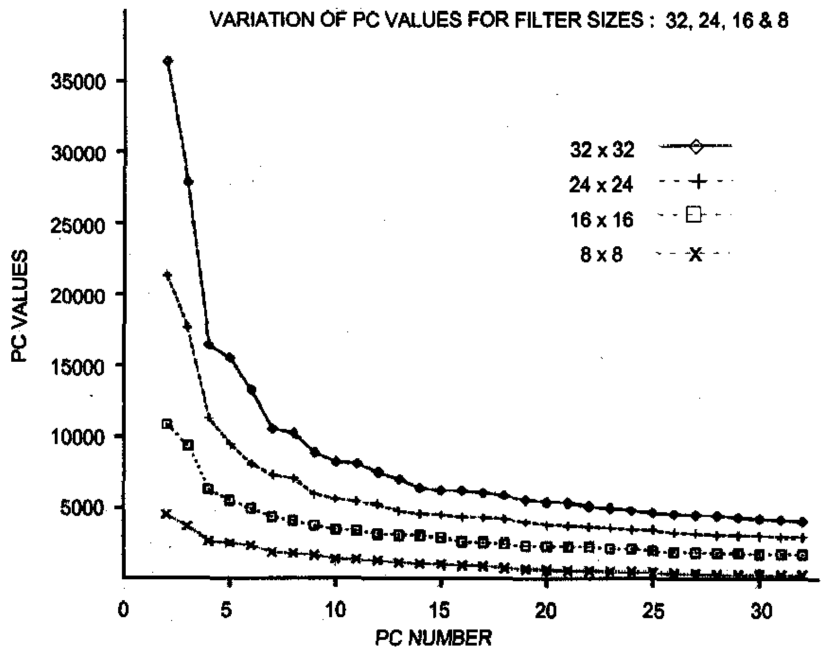
(b) Principal component analysis (PCA) of sub-images to obtain the PCs. These will consist of $W \times W$ PCs, each a $W \times W$ gray scale pattern.

(c) Examination of eigenvalues associated with each PC to determine how many PCs need to be used.

The PCA yields an eigenvalue associated with each PC. From a plot of these eigenvalues, one can determine how many features, M, are adequate to describe the images for a given W. Since larger features describe a greater area, it is expected that for larger W, the plot of eigenvalues associated with the PCs will fall off slower indicating the need for more features. Thus, the PCA produces a family of features of the desired size and an examination of the associated eigenvalues enables one to determine M, the number of those features that need to be used. This graph is shown for $W = 32$, 24, 16, 8 in Fig. 3. It can be seen that $W = 32 \Rightarrow M = 16$, $W = 24 \Rightarrow M = 12$, and $W = 16 \Rightarrow M = 8$. This completes the feature set design.

2.2 Computation of Feature Vectors

The large reference image is characterised as follows: (i) sub-images of sizes equal to the live images are extracted. These may be true reference and live images as in guidance applications. For surveillance or satellite time series alignment applications, these can be arbitrary, small size images (live) extracted from a single image for matching into a larger extracts (reference) from another image, (ii) for each sub-image, each $W \times W$ feature is used as a convolution kernel with a step size of W. Thus, if $W = 32$ and live image is 128×128, each feature produces $4 \times 4 = 16$ outputs. If M features are used, this will create a M*16 dimension feature vector for each sub-image, (iii) the dynamic range of the convolution outputs will be very large, though most values will be clustered around a mean.



(a)

Figure 3. Analysis of various feature sets: (a) shows the eigenvalues associated with various PCs for PC sizes of 32, 24 and 16 pixels.

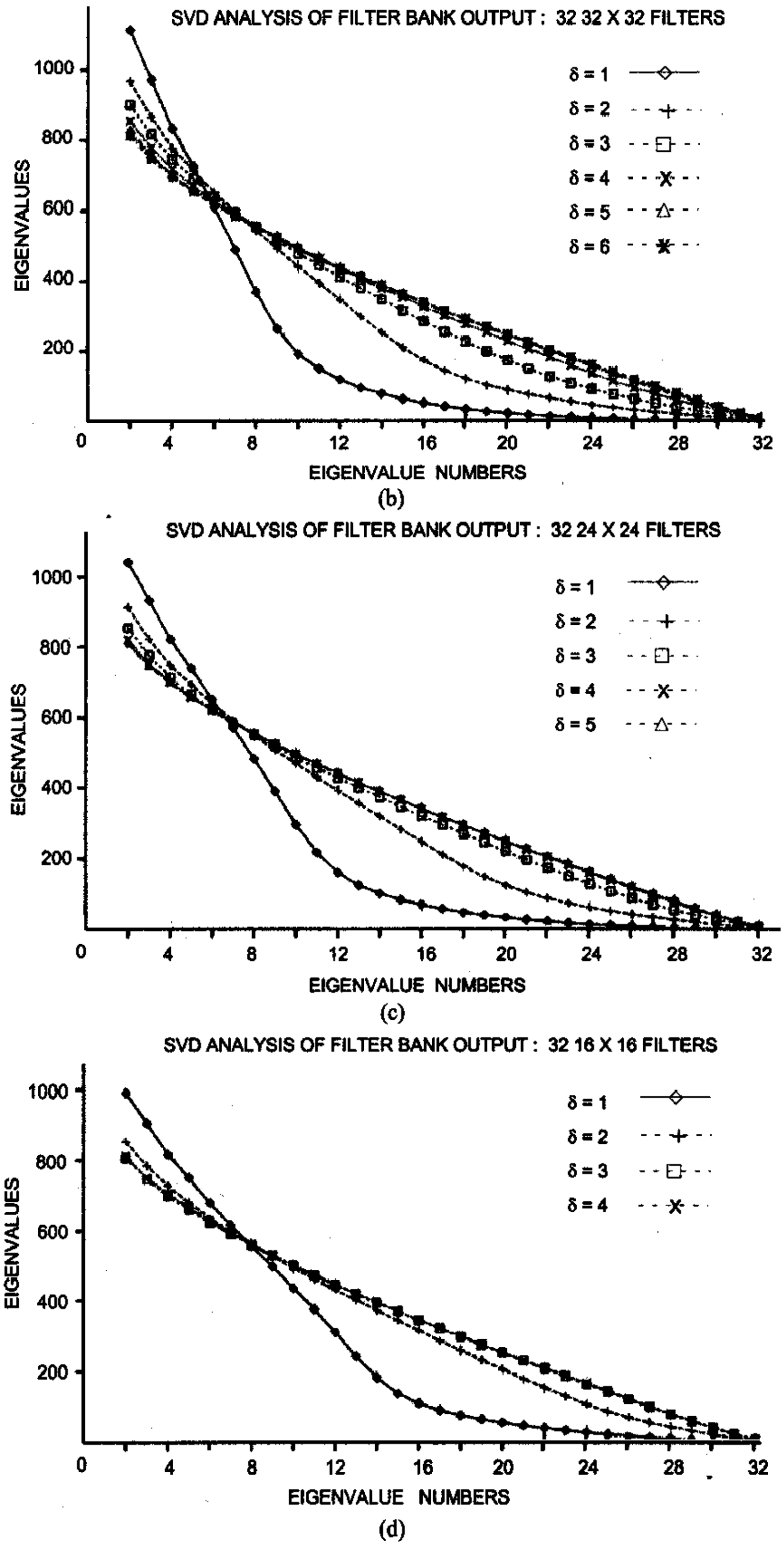


Figure 3. Analysis of various feature sets: The figures (b), (c) and (d) indicate the registration accuracy expected at PC size. δ is the sampling rate of the input to the filter banks. The lowest value of δ for which the graph saturates for each feature size indicates the minimum input separation required for change detection. This resolution is 4 pixels for 32×32, 3 pixels for 24×24 and 2 pixels for 16×16.

One can introduce nonlinearity at the convolution output that, in effect, does a (nonuniform) quantisation to give an 8-bit output that has a near-uniform distribution across its dynamic range. A sigmoid transfer function suffices in most cases to achieve this histogram equalisation[3].

The following characteristics of the above approach are worth noting. There is no preprocessing or noise removal step prior to feature extraction. The features are a low-level characterisation of the image, requiring no high-level understanding or modelling. The feature vector size is constant, independent of image content. The feature detector output values are always of equal significance, irrespective of their gray scale values.

2.3 Matching Process

The reference image is characterised by a large number of fixed-size feature vectors as described above. In an identical manner, a feature vector is computed for the live image. The metric used to compare the vectors is the normalised dot product, $(\vec{a}.\vec{b})/(|\vec{a}|.|\vec{b}|)$. This similarity measure is fast to compute, is not influenced by outliers, and also makes the system contrast-invariant[3]. If the feature kernels are chosen to have 0 bias (bipolar), the system as a whole becomes bias and contrast-invariant.

After the calculation of the normalised dot products, one has a measure of similarity between the live image and various parts of the reference image. The maximum in the distribution of dot product values gives the location of the best match. Reliability measures can be formulated based on the statistics of this distribution. The measure used in this study is based on the maximum, mean and standard deviation: If the maximum is more than five standard deviations from the mean, it is considered a significant peak and its location is declared the registration location. If the maximum is not sufficiently large, that match is rejected and a NULL output is given. Random, seasonally variant image pairs often have no identifiable invariant features that can be used for reliable registration, even by human beings. Thus, a good match reliability measure is very critical to eliminate false positives.

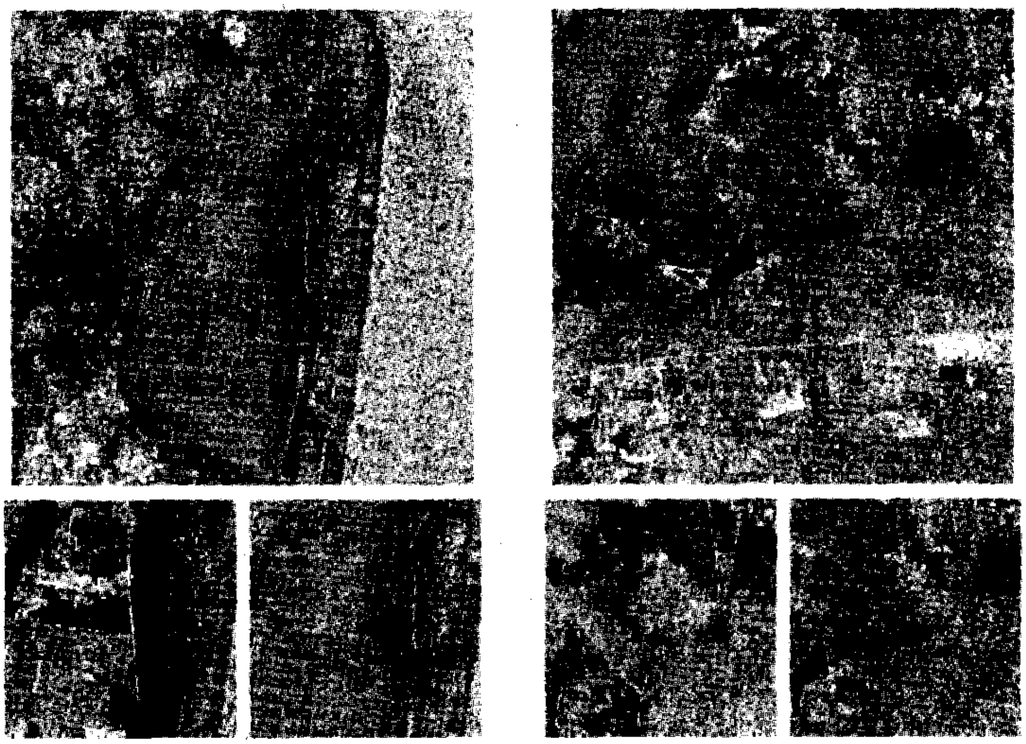


Figure 4. Two-sample image registration results show the big images are the reference images. Corresponding to each reference image, two smaller images are shown. For each pair, the left image is the live image and the right image is the computed matching region in the reference. Both matches are accurate to within 2 pixels.

2.4 Computation Complexity

The registration cost per image pair is dominated by the cost of generating feature vectors for reference image sub-images. The number of sub-images depends on r, the sampling step used in extracting the sub-images. This, in turn, depends on the expected accuracy of the system, determined by the feature size, W. The number of feature elements also depend on the feature size, W and the number of features, M. The choice of feature set then determines the parameters for the feature extraction and feature matching process. Thus, W effectively determines the over all cost of the algorithm.

3. COST-PERFORMANCE TRADE-OFFS

The feature size, W determines the expected resolution of the system. The expected resolution of the system, for a given W, is estimated as follows:

(a) A feature set is designed for a given feature size, W to get M principal componants of size $W \times W$.

(b) A filter bank with M of these PCs as filter kernels is constituted. The output of this filter bank for N $W \times W$ inputs taken from the reference images with a sampling of δ is recorded, where $N >> M$. This creates an $N \times M$ matrix for each δ.

(c) A singular value decomposition (SVD) is performed for each of these matrices and the singular values plotted.

The SVD analysis indicates the expected resolution (registration accuracy). Assuming that for a particular W, the resolution of the system is r pixels. Then for $\delta \geq r$, the rank of the output matrix should be M ($M < N$). However, if $\delta < r$, then the rank will be $< M$. By plotting the singular values for matrices corresponding to various δ's, one can determine the smallest δ for which the output matrix achieves full rank. These plots are shown in Fig. 3. These show that $W = 32 \Rightarrow r = 4$, $W = 24 \Rightarrow r = 3$ and $W = 16 \Rightarrow r = 2$.

The feature vectors are computed for sub-images taken from the reference images with a sampling of r pixels. Thus, the computation complexity scales as $(1/r)^2$. The computation cost also varies directly as M varies as it determines the size of the feature vector. It has been shown that both r and M are related to W, the feature vector size. The above analysis only gives an upper bound on M and a lower bound for r. One can pick a smaller M or larger r to reduce computation cost at the expense of performance. The above analyses show that, for a fixed W, increasing M or reducing r beyond a certain point will not provide better results even though they add to the cost. Alternately, given a performance criterion, one can determine the W that yields the most computationally efficient solution. This provides one with a rational basis for selecting feature size and number of features for an expected resolution.

4. RESULTS

Various configurations of the matching algorithm were tested on a set of 320 image pairs. These images are ERS-1 radar images and each image pair consists of two images of an area imaged in the dry season (April-June) and in the wet season (September-December). The 320 image pairs consisted of 16 sets of 20 image pairs. Each set was chosen along an arbitrary tract. As a result, there were only a few images in each sequence of 20 that contained enough invariant features to allow for reliable registration.

A total of 22 configurations of the algorithm have been tested. The configurations varied in the choice of feature size (W), the number of features used (M) and the sampling rate (r) of sub-images in the reference images. For each configuration, one can compute the computation cost for matching an image pair. Thus, this test produced a performance measure (number of confident matches out of 320) for various computation costs (millions of floating-point operations (MFLOP's)). The cost has been calculated for 256×256 reference and 128×128 live images. The last column shows the number of correct and confident matches. Most of the images lacked sufficient seasonally invariant features to enable confident matches. The results are shown in Table 1. Using this table, one may pick the cheapest configuration for achieving a given level

Table 1. Results of running a 320 image test set on various configurations of the registration algorithm. The parameters varied are the feature size (*W*), number of features (*M*) and sampling rate of sub-images from reference images (*r*)

Configuration	Computation costs (MFLOP's)	No of confident matches
$W = 16, M = 16, r = 2$	4.326	87
$W = 16, M = 12, r = 2$	3.245	86
$W = 16, M = 8, r = 2$	2.165	78
$W = 16, M = 4, r = 2$	1.082	61
$W = 16, M = 16, r = 4$	1.115	73
$W = 16, M = 12, r = 4$	0.836	72
$W = 16, M = 8, r = 4$	0.558	66
$W = 16, M = 4, r = 4$	0.279	53
$W = 24, M = 20, r = 2$	2.112	78
$W = 24, M = 16, r = 2$	1.690	74
$W = 24, M = 12, r = 2$	1.268	70
$W = 24, M = 20, r = 3$	0.925	72
$W = 24, M = 16, r = 3$	0.740	70
$W = 24, M = 12, r = 3$	0.555	66
$W = 32, M = 30, r = 2$	2.028	80
$W = 32, M = 24, r = 2$	1.622	75
$W = 32, M = 20, r = 2$	1.352	78
$W = 32, M = 16, r = 2$	1.082	70
$W = 32, M = 30, r = 4$	0.523	73
$W = 32, M = 24, r = 4$	0.418	72
$W = 32, M = 20, r = 4$	0.348	70
$W = 32, M = 16, r = 4$	0.278	65

of performance. This is shown as a graph of performance versus cost in Fig. 5. It may be noted that while various configurations of the algorithm produced confident and correct matches for approximately 20-25 per cent of the images, there were no incorrect but confident matches.

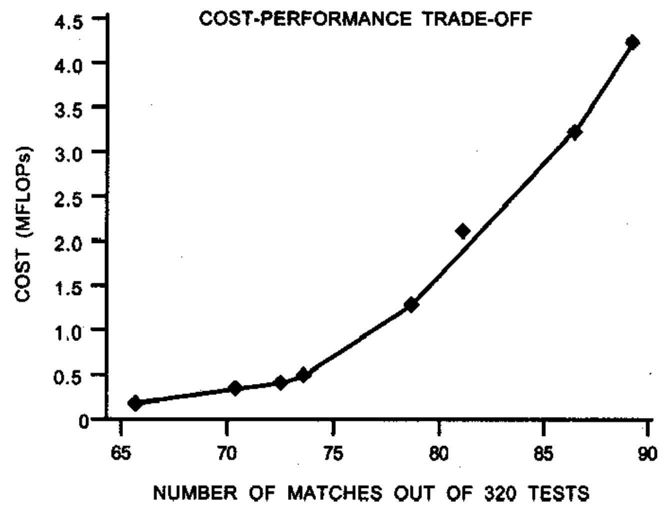


Figure 5. Proposed algorithm allows one to trade-off performance for cost. For the 320 image test set, only about 85 had urban and/or hill features. Hence, the cost escalates rapidly as one tries to get more than 80 of the images confidently registered.

5. CONCLUSIONS

An algorithm for fast registration of radar images has been presented. This algorithm is suitable for applications where scale and rotation corrections can be applied before hand or are not significant. The algorithm is robust to noise and seasonal variations. The inherent flexibility in the design of the algorithm allows for a cost-performance trade-off that enables one to tailor it for various applications. The false match rejection is very good, making this algorithm suitable for critical applications.

ACKNOWLEDGEMENTS

The authors acknowledge the support given by Director, Centre for Artificial Intelligence & Robotics, Bangalore, for this work.

REFERENCES

1. Brown, L.G. A survey of image registration technique. *ACM Comput Surv.*, 1992, **24**(4), 325-75.

2. Calway, A.D. *et al.* A multiresolution frequency domain algorithm for fast image registration. *In* Third International Conference on Visual Search, April 1992, pp 1-2.

3. Rakshit, S.; Deodhare, Dipti & Krishnan, K. Fast registration of satellite images in the presence of significant scene content variations. *In* SPIE Conference on Signal and Data Processing of Small Targets, 1999, Denver, USA. *Proceedings of SPIE*, **3809**, 42-50.

Contributors

Dr Subrata Rakshit received his BTech (Engg Phy) from the Indian Institute of Technology, Bombay. He received his MS and PhD from the Caltech, USA, in 1989 and 1994, respectively. He has been working at the Centre for Artificial Intelligence & Robotics (CAIR) since 1994. He currently heads the Computer Vision Group. His research interests are: Image processing, multiresolution, pattern recognition, neural nerworks, and information theory.

Ms Dipti Deodhare recieved her MSc(Computer Science) from Pune University and MS(Engg) in Computer Science and Automation from the Indian Institute of Science(IISc), Bangalore. She has been working at the CAIR, Bangalore since 1991. She currently heads the AI & ANN Group. Her research interests include: Neural networks, pattern recognition, and optimisation.