This article can be cited as S. M. Moorthi, R. K. Gambhir, R. Ramakrishnan and R. Sivakumar, Performance Study of Optimization Methods for Intensity Based Automatic Satellite Image Registration, International Journal of Imaging and Robotics, vol. 8, no. 2, pp. 101-110, 2012.

Performance Study of Optimization Methods for Intensity Based Automatic Satellite Image Registration

Subbiah Manthira Moorthi[1], Rajdeep Kaur Gambhir[1], Rajagopalan Ramakrishnan[1] and Ramamoorthy Sivakumar[2]

[1]Data Products Software Group, Space Applications Centre, ISRO, Ahmedabad, India
Email: {smmoorthi, rajdeep, rama}@sac.isro.gov.in

[2]Department of Civil Engineering, SRM University, Kattankulathur, TN, India
Email: Sivakumar.r@ktr.srmuniv.ac.in

ABSTRACT

Satellite Image registration plays an important role in remote sensing data processing and applied in wide variety of tasks such as image fusion, image overlay and change detection using different images of the same region. It is one of the challenging image processing tasks due to imaging by different sensors, at view angles or and at different times. Satellite images also poses unique challenges for registration with issues like cloud pixels, noise in the images, systematic errors, multispectral images, terrain induced distortions etc. It requires building an elaborate computational framework to handle specific problems of satellite image registration as the automatic image registration is very important requirement for voluminous data sets. Conventional approaches in satellite image registration involve a feature collection strategy manually or automatic, measuring similarity measures to find the best feature matches and further use spatial coordinates of the best matches to estimate a transform. Recent advances in medical image registration topic have suggested employing intensity based non rigid image registration framework that engages a sampling strategy, a similarity metric, a transform and an optimization procedure in an iterative manner. This procedure finds optimal transform parameters by maximizing the chosen similarity measure criteria and thus minimizing a cost function, providing a robust image registration framework. Satellite image registration can be treated as an optimization problem with the goal of finding the spatial mapping that will bring the two images into alignment. So a suitable choice of optimizer plays a key role in registration process. Ways of employing and comparing the performances of different optimization methods such as Evolution Strategy, Conjugate Gradient, Gradient Descent, Simultaneous Perturbation, Robbins-Monro, Adaptive Stochastic Gradient Descent, and Quasi Newton is reported here for the intensity based satellite image registration. Elastix, a public domain tool developed for doing intensity based medical image registration has been used in this study to perform non-rigid satellite image registration.

Keywords: Satellite image processing, optimization methods, registration

Mathematics Subject Classification: 68U10

Computing Classification System: I.4.3

1. INTRODUCTION

Image registration is a fundamental image processing task to match and align physically two images which could have been imaged by different sensors, view angles or and at different times. It plays an important role in remote sensing and applied in wide variety of tasks such as image fusion, image overlay and change detection using different images of the same region. When there are multiple images of same object world and they are not geometrically confirming, registration is required and unless corrected or modelled for relative geometric errors, further use of the images are subjective. Satellite images pose

This article can be cited as S. M. Moorthi, R. K. Gambhir, R. Ramakrishnan and R. Sivakumar, Performance Study of Optimization Methods for Intensity Based Automatic Satellite Image Registration, International Journal of Imaging and Robotics, vol. 8, no. 2, pp. 101-110, 2012.

unique challenges for registration with issues like cloud pixels, noise in the images, systematic errors, multispectral channels and terrain induced distortions (Leprince et al. 2007).

Automatic image registration framework should accomplish the mandatory steps like collecting samples, features or land marks, establish transformation, and warp the input image to target image geometry by resampling the input image completely in unassisted manner (Leprince et al., 2007, Yao et al., 2001). Feature based registration methods has few inherent difficulties such as detection and ensuring uniform distribution of feature points and dependency on detected feature points to be used in the estimating the transform model parameters whereas in the intensity based method the whole image can participate in the process depending on the necessity. Though, similar steps are used in feature based and intensity based methods, of finding feature points or sample points, measuring similarity to decide best match, and establishing transforms, the problem definitions are different. Feature based methods simply find feature points to find best matches to be used in establishing the correspondence in parts (Zhen et al., 2010). But, the intensity based method finds optimal transform parameters by maximizing the chosen similarity measure criteria and thus minimizing a cost function, providing a robust integrated image registration framework (Klein et al., 2010).

Feature based image registration methods are elaborately discussed in (Brown, 1992, Zitova et al., 2003, Zhen et al., 2010) and the intensity based image registration framework is described briefly in the next section. Description of all optimization methods are covered in *Section* 3 and evaluation of performance in *Section 4*. Adaptive stochastic gradient descent optimization method is separately discussed due to its better performance in the chosen satellite image registration tasks. Stefan Klein's work in medical image registration and *Elastix* development based on *ITK* library has inspired us to use *Elastix* for satellite image registration.

2. INTENSITY BASED IMAGE REGISTRATION

In intensity based image registration, image to be registered is called the moving image M(x), is deformed to fit the other image, the fixed image F(x). In other words, registration is the problem of finding a coordinate transformation T(x) that makes M (T(x)) spatially aligned with F(x). The quality of alignment is defined by a cost function C(T; F, M).The optimal coordinate transformation is estimated by minimizing the cost function with respect to T, usually by means of an iterative optimization method embedded in a hierarchical (multiresolution) scheme depicted in Fig.1. Intensity based image registration is discussed in detail in Klein et al., 2010 and Moorthi et al., 2011.

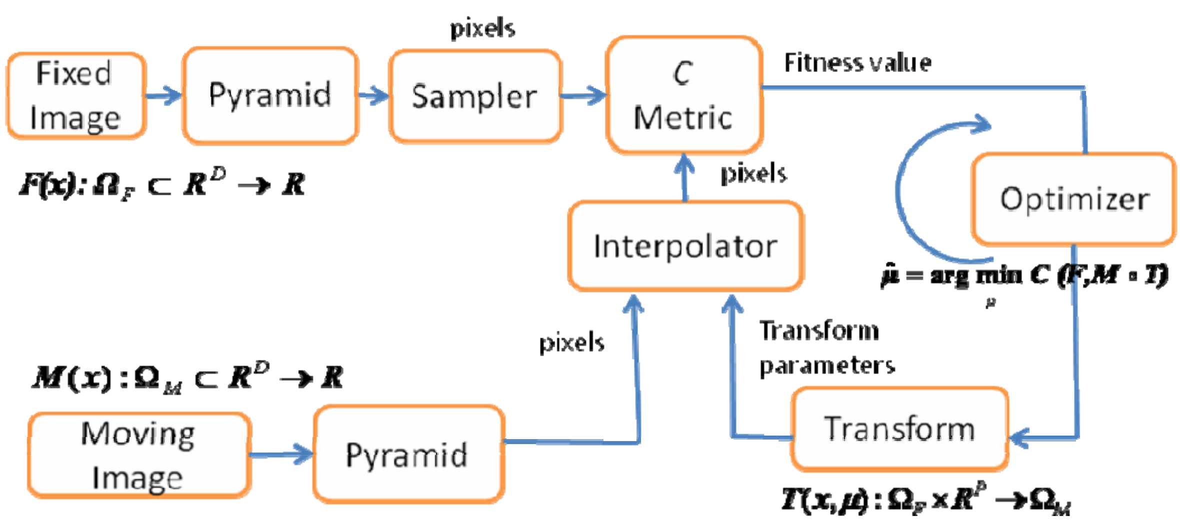


Figure 1. Intensity based image registration framework

This article can be cited as S. M. Moorthi, R. K. Gambhir, R. Ramakrishnan and R. Sivakumar, Performance Study of Optimization Methods for Intensity Based Automatic Satellite Image Registration, International Journal of Imaging and Robotics, vol. 8, no. 2, pp. 101-110, 2012.

Application of an image registration method requires many choices to be made, such as the optimization method, the multiresolution strategy, the method of image interpolation to evaluate M(T(x)), the coordinate transformation model, and the definition of the cost function. Several possibilities for the optimization method are discussed as the main theme of this paper. For the cost function C many options have been proposed in the literature. Commonly used intensity-based cost functions are the mean squared difference, normalized correlation, mutual information (Klein et al., 2010, Th´evenaz, et al., 1996, Zitova et al., 2003).

Intensity-based image registration is usually treated as a nonlinear optimization problem. Define the fixed image F(x) : $\Omega_F \subset R^D \rightarrow R$, the moving image M(x) : $\Omega_M \subset R^D \rightarrow R$, and a parameterized coordinate transformation T(x,μ) : $\Omega_F \times R^P \rightarrow \Omega_M$, where $\mu \in R^P$ represents the vector of transformation parameters. The following minimization problem is considered:

$$\hat{\mu} = \arg\min_{\mu} C(F, M \circ T) \qquad (1)$$

Where, C is the cost function that measures the similarity of the fixed image and the deformed moving image. The solution $\hat{\mu}$ is the parameter vector that minimizes that cost function. Henceforth, we use the short notation

$$C(\mu) \equiv C(F, M \circ T) \qquad (2)$$

To determine the optimal set of parameters, an iterative optimization strategy is employed as follows.

$$\mu_{k+1} = \mu_k + a_k d_k \text{, k = 0,1,2,....(3)}$$

Where d_k is the search direction at iteration k, and a_k a scalar gain factor controlling the step size along the search direction. The search directions and gain factors are chosen such that the sequence μ_k converges to a local minimum of cost function C. Many optimization methods can be found in literature, differing in the way a_k and d_k are computed. But no extensive comparison is done for non-rigid satellite image registration tasks. In this exercise, several optimization methods are compared with respect to time consumption, accuracy, precision and robustness.

3. OPTIMIZATION METHODS

The optimization techniques can be overall represented by (3). Different strategies can be followed to estimate a_k and d_k .

1. Estimation of gain factor. Assessment of gain factor a_k can be done in following different ways:
   - Just set as a constant
   - Decaying function of k
   - Exact line search, where in each iteration, Cost function C is minimized
   - Inexact line search, which finds gain factor with sufficient reduction of C
2. Estimation of search direction. Search direction d_k is basically derivative of cost function $\partial C / \partial \mu$, known as g.

This article can be cited as S. M. Moorthi, R. K. Gambhir, R. Ramakrishnan and R. Sivakumar, Performance Study of Optimization Methods for Intensity Based Automatic Satellite Image Registration, International Journal of Imaging and Robotics, vol. 8, no. 2, pp. 101-110, 2012.

Brief overview of some optimizers is given in table 1. For detailed description refer to (Klein et al., 2007, Nocedal et al., 1999).

Table 1: Optimization methods for image registration

Optimizer	Gain Factor (a_k)	Search Direction (d_k)	Model	Notes
Gradient Descent (GD)	Determined in two ways namely GDD and GDL	Negative gradient of cost function	$\mu_{k+1} = \mu_k - a_k g(\mu_k)$ GDD : $a_k = a/(k+A)^{\alpha}$ GDL : More-Thuente Algorithm	Convergence properties of this method is guaranteed on the ground that there exists theoretical bounds on distance to solution at iteration *k.*
Quasi Newton (QN)	More-Thuente Algorithm (inexact line search)	Negative inverse of Hessian matrix $L_k \approx [H(\mu_K)]^{-1}$	$\mu_{k+1} = \mu_k - a_k L_k g(\mu_k)$	This method basically uses 2nd order information which gives better convergence than GD. To ensure super linear convergence $a_k = 1$ should be tried first. (Dennis et al., 1977)
Nonlinear Conjugate Gradient (NCG)	More-Thuente Algorithm(inexact line search)	Linear combination of $g(\mu_k)$ and previous search direction d_{k-1}	$\mu_{k+1} = \mu_k + a_k d_k$ $d_k = -g(\mu_k) + \beta_k d_{k-1}$	In contrast to QN where unit gain has to be tried first, here no such bound is there, like $a_k = a_{k-1}$ can be used as initial guess (Dai,Y.H., 2001)
Stochastic Gradient Descent (SGD)	$a_k = \dfrac{a}{(k+A)^{\alpha}}$	$g(\mu_k)$ is replaced by $\tilde{g}_k$	$\mu_{k+1} = \mu_k - a_k \tilde{g}_k$	It is applied when computation of Exact derivative is costly, an approximation can be used in this situation but it may lead to decrease in speed of

This article can be cited as S. M. Moorthi, R. K. Gambhir, R. Ramakrishnan and R. Sivakumar, Performance Study of Optimization Methods for Intensity Based Automatic Satellite Image Registration, International Journal of Imaging and Robotics, vol. 8, no. 2, pp. 101-110, 2012.

				convergence.
Simultaneous Perturbation (SP)	$a_k = \dfrac{a}{(k+A)^{\alpha}}$	Derivative estimation of this method (Radac et al., 2011), is based on approximate evaluation of cost function.	$\tilde{C}_k^+ = C(\mu_k + c_k\Delta_k) + \varepsilon_k^+$ $\tilde{C}_k^- = C(\mu_k - c_k\Delta_k) + \varepsilon_k^-$ $[\tilde{g}_k]_i = \dfrac{\tilde{C}_k^+ - \tilde{C}_k^-}{2c_k[\Delta_k]_i}$	In N-dimensional estimation problem, this method requires 2 evaluations per iteration, independent of N. (Spall, J, C., 1992)
Robbins Monro (RM)	$a_k = \dfrac{a}{(k+A)^{\alpha}}$	As compared to SP, RM assumes that approximation of derivative of cost function is available	$\tilde{g}_k = g(\mu_k) + \varepsilon_k$	The approximation $\tilde{g}_k$ does not necessarily vanish to $\hat{\mu}$,so convergence must be forced by $a_k \to 0$ as $k \to \infty$. (Robbins, et al., 1951)
Evolution Strategy (ES)	$a_k = \dfrac{(k+A)^{\alpha}}{(k+A+1)^{\alpha}}$	A weighted average of P selected trial direction is computed.	$d_k = \sum_{p=1}^{P} w_p d_k^{(p;\lambda)}$	After each iteration a_k and C_k are updated automatically based on d_{k-1} and $d_k^{(p;\lambda)}$.
Adaptive Stochastic Gradient Descent (ASGD)	the γ function is not evaluated at the iteration number k, but at the 'time' t_k adapted depending on the inner product of the gradient $\tilde{g}_k$ and the previous gradient $\tilde{g}_{k-1}$	Same as SGD	$t_{k+1} = [t_k + f(-\tilde{g}_k^T \tilde{g}_{k-1})]^+$ $\mu_{k+1} = \mu_k - \gamma(t_k)\tilde{g}_k$	It implements an adaptive step size mechanism. (Klein et al., 2009)

Elastix (Klein et al., 2010) is a public domain tool which has a collection of optimizers and metrics implemented to experiment image registration tasks of this kind. This software was originally meant to perform medical image registration, but recently used to register remote sensing images in non-rigid

This article can be cited as S. M. Moorthi, R. K. Gambhir, R. Ramakrishnan and R. Sivakumar, Performance Study of Optimization Methods for Intensity Based Automatic Satellite Image Registration, International Journal of Imaging and Robotics, vol. 8, no. 2, pp. 101-110, 2012.

image registration category (Moorthi et al., 2011). All experiments in this paper have been achieved using *Elastix* implementations only.

4. EVALUATION OF PERFORMANCE OF OPTIMIZATION METHODS

This evaluation is based on maximization of mutual information in combination with parameterized deformation field. Precision and convergence properties of these methods are studied and observations are recorded in proper manner with sufficient number of experiments on different image datasets. The consistency of results is one of the deciding factors in choosing a suitable optimizer for satellite image registration. Satellite image registration is a computation and memory intensive task, so the choice of optimizer should be such that it can minimize these resources while giving desired high (order of 0.25 pixels) accuracy. To compare the methods in terms of registration accuracy, visual as well as overlap and swipe measures are used. However, the main focus of these experiments is to test the convergence criteria.

The optimization methods are tested on image registration tasks that involve Indian Remote Sensing Satellite (IRS) sensors such as LISS-3, AWIFS, and LISS-4 images. Experiments were performed with different image sizes, number of iterations, B-Spline control point spacing and number of histogram bins with multi resolution (pyramidal approach) strategy. Elastix allows configurability to choose parameters and models with an ASCII file interface containing text in keyword and parameter value syntax.

Typical sensor data used for the study is obtained from Resourcesat-1 AWIFS of spatial resolution 50 m, acquired on 11th January, 2011 (moving image) and a reference image (fixed image) acquired during the year 2008 in the month of November covering the same geographical area, which was already corrected for geometric infidelities with a size of 4065 X 4184 pixels. Number of iterations for experiment was set to 500 for all methods.

Earlier mentioned optimization techniques can be divided into two groups, namely deterministic and stochastic. In stochastic optimization algorithms such as SP, RM or ES, new subset of pixels chosen by a sampling strategy are used in every iteration for better performance. So, these techniques can be compared by changing number of sample pixels required for optimization.

In the analysis and comparison presented here, description legends like SP-2048 indicates that 2048 samples are used in the iteration with SP optimization. Same conventions are used for other methods lying in this category.

In deterministic optimization algorithms (such as GDD, GDL, QN, NCG), single subset of sample pixels are used on regular grid with downsampling factors. For e.g. QN-2 indicates a downsampling factor of 2 and this convention is followed in this category where always 2048 samples were used for experimentation.

The two deciding factors for comparing different optimization methods are the required number of iterations and the computation time per iteration. Computation time is dominated by time taken for calculating cost function and its derivative. In this case the metric chosen was mutual information, with number of histogram bins 32.

For experimental purposes no multi resolution scheme is adopted so that the differences can be observed at every iteration at a particular resolution level. *Elastix* offers to combine transformations one after another to achieve registration. We employed affine first and BSpline transformations subsequently for robustness in the actual registration tasks. However, for comparison purposes, among all transformations only affine was employed for the test of convergence. The convergence of registration is assumed when change in affine coefficients becomes insignificant in the subsequent iterations. This is somewhat different from the comparison shown in Stefan Klein's work (Klein et al. 2007) considered as our

This article can be cited as S. M. Moorthi, R. K. Gambhir, R. Ramakrishnan and R. Sivakumar, Performance Study of Optimization Methods for Intensity Based Automatic Satellite Image Registration, International Journal of Imaging and Robotics, vol. 8, no. 2, pp. 101-110, 2012.

reference, where the image deformations similar to residuals were used to establish the convergence. However we prove here with alternate but equivalent measure, the residual shift in sample direction which can be picked from affine transformation estimated in every iteration or experiment itself instead of computing deformations separately. It saves time and effort with no compromise on rigorous analysis, we intended to show. Pixel displacement versus iterations plots for all particular optimization techniques are shown with different configurations such as number of samples or downsampling factors in Fig-2 to Fig.8.

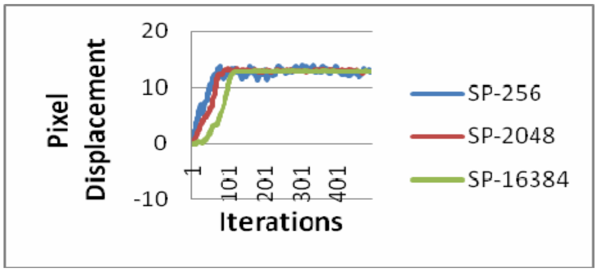


Figure 2. Convergence of SP

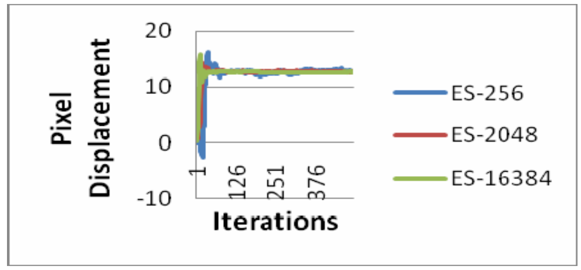


Figure 3. Convergence of ES

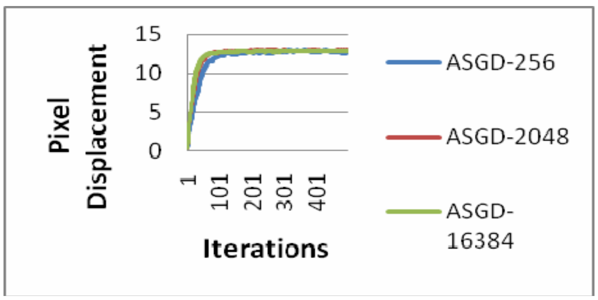


Figure 4. Convergence of ASGD

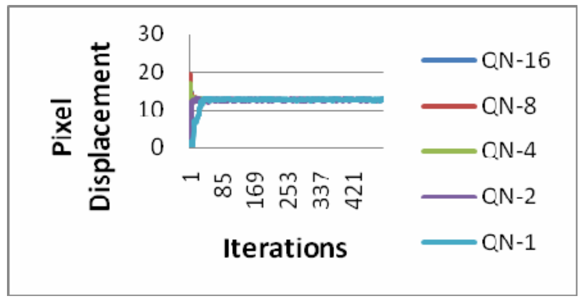


Figure 5. Convergence of QN

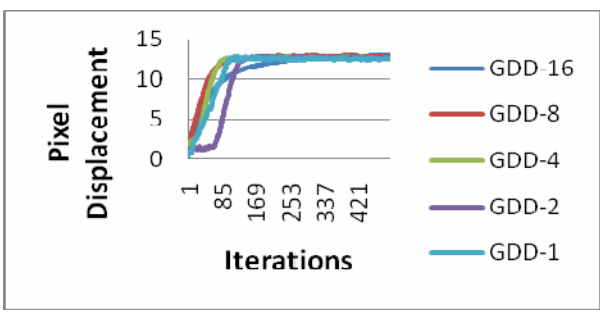


Figure 6. Convergence of GDD

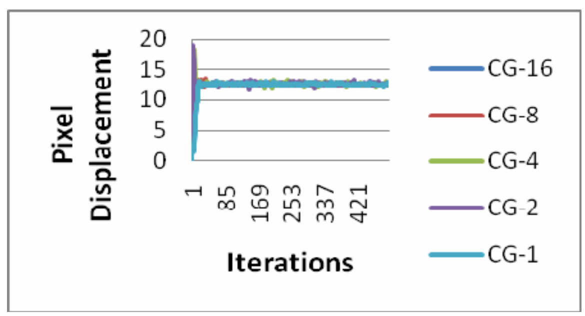


Figure 7. Convergence of CG

This article can be cited as S. M. Moorthi, R. K. Gambhir, R. Ramakrishnan and R. Sivakumar, Performance Study of Optimization Methods for Intensity Based Automatic Satellite Image Registration, International Journal of Imaging and Robotics, vol. 8, no. 2, pp. 101-110, 2012.

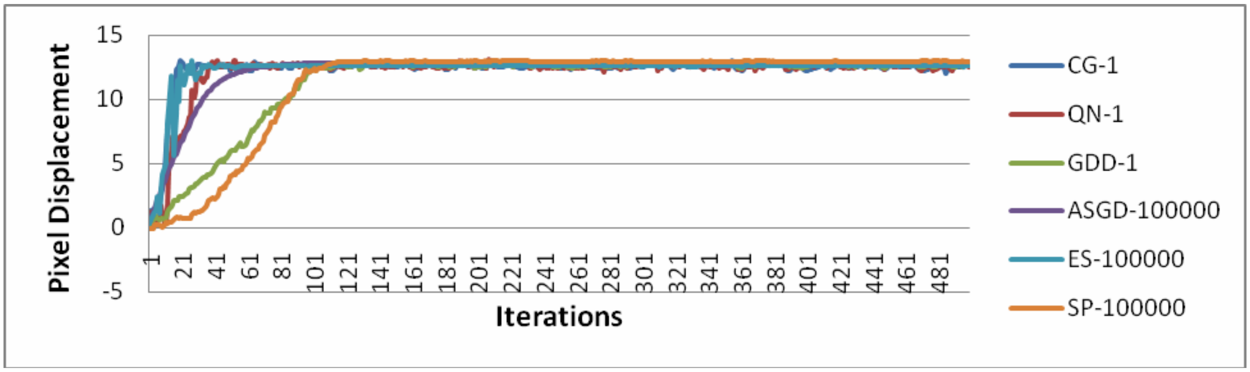


Figure 8. Convergence of All Methods

As can be seen from Fig.2 to Fig. 4, for stochastic optimization techniques that 2048 samples are sufficient for convergence, but to demand better convergence 16384 samples can be used at the cost of computing time. ASGD (improved version of RM), shows less fluctuations and its convergence towards solution is fast as compared to SP and ES.

In deterministic methods depicted in Fig.5, Fig.6 and Fig.7, QN and CG converges faster as compared to GDD and both of them stops as soon as Wolfe condition (reference) is not satisfied. The difficulty with QN and CG is we need to introduce a regularization term application wise to avoid unrealistic deformations. This makes these methods more vulnerable to changes as compared to robust ASGD method, in which no such setting is required. Fig-7 compares all optimization techniques for convergence with different configurations for each technique and all of them converge beyond 150 iterations. It can be seen that some methods fluctuate heavily before settling unlike ASGD method where a smooth transition occurs.

We compared six optimization methods based on maximization of mutual information. From the experimental results it can be seen that Stochastic Gradient Descent method namely RM gives better results as compared to others in terms of robust convergence towards solution. With this method computation time can be extremely lowered by usage of random sampling per iteration. Minimum number of samples required is found to be around 2048. QN and CG gives better precision than RM but application wise setting a regularization term is major drawback. SP and GDD method's convergence is quite low as compared to other methods. QN and CG achieve slight higher precision at the cost of large computational time with an overhead of a regularization term to be set in every registration task. Time taken for every individual optimization method is listed in table 2.

Table 2: Time taken by optimization methods for 500 iterations

Method	Time (sec)
CG-1	11.796
QN-1	11.093
GDD-1	11.078
ASGD-2048	7.328
SP-2048	5.781
ES-2048	10.329

It can be easily observed from table 2 that mostly deterministic methods take more time as compared to stochastic ones in a given sample range.

5. PERFORMANCE OF ASGD

ASGD is a variation of SGD optimization method with adaptive step size overcoming the need of predetermined step size (Klein et al., 2009). After studying ASGD for convergence, and performance,

This article can be cited as S. M. Moorthi, R. K. Gambhir, R. Ramakrishnan and R. Sivakumar, Performance Study of Optimization Methods for Intensity Based Automatic Satellite Image Registration, International Journal of Imaging and Robotics, vol. 8, no. 2, pp. 101-110, 2012.

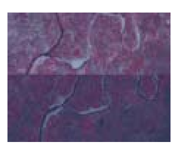
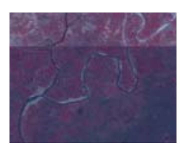
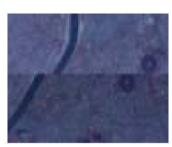
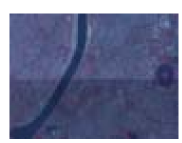
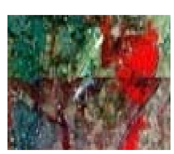
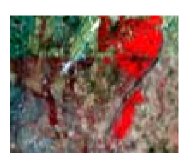
many satellite image registration tasks were run with this method, and evaluated for the registration accuracies which are presented in table 3.

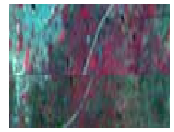
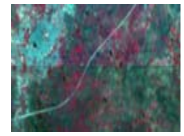
Our experiments indicate that ASGD gives better performance for satellite image registration tasks chosen here. ASGD computation time can be decreased by using few set of image samples to compute the derivative of cost function w.r.t transform parameters. Despite our apprehensions, even images of highly undulating terrains could be registered with satisfactory results (usually these images need terrain relief correction for better image registration) without using the georeferencing information (Moorthi et al., 2008). This approach has been validated by experimenting with more than 1000 image pairs of different acquisitions. From results obtained, it can be safely assumed that this method is robust enough in defined parameter range as reported elaborately about ASGD in Moorthi et al., 2011.

Performance of ASGD was impressive in line with the reported results of Klein et al., 2007 (Moorthi et al., 2011) though this experiment was done with satellite images rather than medical images where imaging modalities and geometries are different.

In table 3, registration performance is graphically represented using image swipes of registered moving and fixed images from LISS-3 and AWIFS multi temporal images. Table 3 shows how the images are relatively placed in geometry before registration and after at various image magnification levels. The horizontal line in the images shows the swipe cutline between moving and fixed images. We also estimated the registration accuracy by evaluating global translation parameters between fixed and registered images going through the registration process once again with specific parameter choices. Registration error is found to be less than 0.1 pixels in both the cases which are sufficient for the many remote sensing applications. Residual errors computed for LISS-3 and AWIFS images are presented in table 3.

Table 3: Image registration performance and accuracies with ASGD optimizer

Sensor	Fixed Image +Path/Row/Date of imaging	Moving Image +Path/Row/Date of imaging	Image Swipe Before Registration	Image Swipe After Registration	Residual Error Moving - Fixed	
					Pix	*Scan*
LISS-3	093/056/02-05-11	093/056/04-01-06			0.012	0.005
LISS-3	093/056/02-05-11	093/056/31-01-10			0.012	0.004
AWIFS	092/052/07-01-09	090/054/18-12-10			0.009	-0.013

AWIFS	098/052/26-11-08	096/051/30-11-10			-0.010	0.034

6. CONCLUSION

Ways of employing and comparing the performances of different optimization methods such as Evolution Strategy, Conjugate Gradient, Gradient Descent, Simultaneous Perturbation, Robbins-Monro, Adaptive Stochastic Gradient Descent, and Quasi Newton is reported here for the intensity based satellite image registration. This comparison is based on maximization of mutual information in combination with parameterized deformation field. Precision and convergence properties of these methods are studied and observations are recorded with sufficient number of experiments on different image datasets.

We compared six optimization methods based on maximization of mutual information. Our experiments indicate that ASGD gives better performance for satellite image registration tasks chosen here. ASGD computation time can be decreased by using few set of image samples to compute the derivative of cost function w.r.t transform parameters. Despite our apprehensions, even images of highly undulating terrains could be registered with satisfactory results (usually these images need terrain relief correction for better image registration) without using the georeferencing information which is used by default in satellite image registration tasks. Further experiments and results obtained with *Elastix* tool in satellite image registration tasks will be published later.

REFERENCES

Brown, L.G., 1992, A survey of image registration techniques, ACM Computing Surveys, **24(4)**, 325–376.

Dai, Y.H., 2001, An efficient hybrid conjugate gradient method for unconstrained optimization, Annals of Operations Research, **103(1–4)**, 33–47.

Dennis, Jr., J.E., Moré, J.J., 1977, Quasi-Newton methods, motivation and theory, SIAM Review, **19(1)**, 46–89.

Fonseca, L.M.G., Manjunath, B.S., 1996, Registration techniques for multisensor remotely sensed imagery, Photogrammetric Engineering & Remote Sensing, **62(9)**, 1049-1056.

Klein, S., Staring, M., Pluim, J.P.W., 2007, Evaluation of optimization methods for non rigid medical image registration using mutual information and B-splines, IEEE Transactions on Image Processing, **16(12)**, 2879–2890.

Klein, S., Pluim, J.P.W., Staring, M., Viergever, M.A., 2009, Adaptive stochastic gradient descent optimisation for image registration, International Journal of Computer Vision, **81(3)**, 227–239.

Klein, S., Staring, M., Murphy, K., Viergever, M.A., Pluim, J.P.W., 2010, Elastix: A toolbox for intensity-based Medical image registration, IEEE Transactions on Medical Imaging, **29(1)**, 196-205.

Leprince, S., Sylvain, B., Francois, A., Avouac, J.P., 2007, Automatic and precise orthorectification, co registration, and subpixel correlation of satellite Images, application to ground deformation measurements, IEEE Transactions on Geoscience and Remote Sensing, **45(6)**, 1529-1558.

Moorthi, S.M., Kayal, R., Ramakrishnan, R., Srivastava, P.K., 2008, RESOURCESAT-1 LISS-4 MX bands on ground co-registration by in-flight calibration and attitude refinement, International Journal of Applied Earth Observation and Geoinformation, **10(2)**, 140–146.

Moorthi, S.M., Gambhir, R.K., Misra, I., Ramakrishnan, R., 2011, Adaptive stochastic gradient descent optimization in multi temporal satellite image registration, Proceedings of 2011 IEEE Conference on Recent Advances in Intelligent Computational Systems (RAICS), Trivandrum, India, 373–377.

Nocedal, J., Wright, S.J., 1999, Numerical Optimization. New York: Springer-Verlag.

Robbins, H., Monro, S., 1951, A stochastic approximation method, Annals of Mathematical Statistics, 22**(3)**, 400–407.

Radac, M. B., Precup, R. E., Petriu, E.M., Preitl, S., 2011, Application of IFT and SPSA to servo system control, IEEE Transactions on Neural Networks, **22(12)**, 2363-2375.

Spall, J, C., 1992, Multi variate stochastic approximation using a simultaneous perturbation gradient approximation, IEEE Transactions on Automatic Control, **37(3)**, 332-341.

Thévenaz, P., Unser, M., 1996, A pyramid approach to sub-pixel image fusion based on mutual information, Proceedings of IEEE International Conference on Image Processing, Lausanne, Switzerland, 265–268.

Yao, J., Tien C.C., 2001, The practice of automatic satellite image registration, Proceedings of 22nd Asian Conference on Remote Sensing (ACRS 2001), **1**, 221–226.

Zhen, X., Yun, Z., 2010, A critical review of image registration methods, International Journal of Image and Data Fusion, **1(2)**, 137–158.

Zitova, B., Flusser, J., 2003, Image registration methods: a survey, Image and Vision Computing, **21(11)**, 977–1000.