# Efficient Global Optimization for Image Registration

Ying Chen, Richard R. Brooks, *Member*, *IEEE Computer Society*,
S. Sitharama Iyengar, *Fellow*, *IEEE*, Nageswara S.V. Rao, *Senior Member*, *IEEE*, and
Jacob Barhen, *Member*, *IEEE*

**Abstract**—The image registration problem of finding a mapping that matches data from multiple cameras is computationally intensive. Current solutions to this problem [3], [4], [5] tolerate Gaussian noise, but are unable to perform the underlying global optimization computation in real time. This paper expands these approaches to other noise models and proposes the Terminal Repeller Unconstrained Subenergy Tunneling (TRUST) method, originally introduced by Cetin et al. [7], as an appropriate global optimization method for image registration. TRUST avoids local minima entrapment, without resorting to exhaustive search by using subenergy-tunneling and terminal repellers. The TRUST method applied to the registration problem shows good convergence results to the global minimum. Experimental results show TRUST to be more computationally efficient than either tabu search or genetic algorithms.

**Index Terms**—Image registration, sensor fusion, global optimization, terminal repellers, subenergy-tunneling, genetic algorithms, tabu search.

## 1 Introduction

Automated control systems react to the environment through the data provided by the sensors. Hence, the transformation of imprecise sensor data into a reliable and timely model of the environment is an important and challenging problem. In particular, systems with multiple sensors (as in Fig. 1) must fuse data from individual sensors to interpret it correctly. The past 10 years have seen extensive research in methods for fusing various types of sensor data [5]. Since real life sensor data contains noise, has finite accuracy and limited dependability, implementing multisensor systems is challenging, and often is significantly more difficult than single sensor systems.

In autonomous systems with more than one sensor input, such as a stereo vision, the ability to correlate and calibrate readings from both sensors is important for a number of tasks. *Image registration* refers to finding a function that maps one image, the *observed* image, onto another image, the *reference* image. This step is sometimes referred to as image correlation. To avoid confusion, we prefer the term registration, since the former has a more general meaning in the area of signal processing [10], [13]. Surveys of image registration can be found in [5], [6], [9]. Our previous work [3], [4], [5] on image registration formulated it in terms of a global optimization problem where images are matched using pixel information. This approach is justified since high-level features may not always be present and image noise is best modeled at the pixel level. We have previously explored the use of genetic algorithms [3], tabu search [5], and simulated annealing [4] as global optimization paradigms for image registration. Unfortunately, these approaches are all computationally intensive, making them unsuitable for real-time applications. Registering sensor images in real time requires a new technique.

In this paper, we propose an efficient global optimization technique for image registration based on the Terminal Repeller Unconstrained Subenergy Tunneling (TRUST) technique proposed by Cetin et al. [7] (and generalized to lower semicontinuous functions by Baren et al. [1], [2]). We formulate the registration problem as a global optimization problem to be solved by TRUST, which avoids both local minima entrapment and exhaustive search. In particular, TRUST uses subenergy-tunneling and terminal repellers techniques to converge to the global minimum. We present a detailed analysis of TRUST's implementation for registering images along with examples. Our experimental results show the search time required by TRUST to be much lower than for other methods, such as tabu search and genetic algorithms. The main advantage of this approach over our earlier efforts [3], [4], [5] is its increased computational efficiency. We also further address the basic issues of limited accuracy and image corruption with noise by discussing the use of this approach under multiple noise models.

This paper is organized as follows: Section 2 presents an overview of image registration, a statement of the problem, and functions which can be used in the optimization problem. The TRUST technique and details of its operation are presented in Section 3. An approach to image registration based on TRUST is presented in Section 4. Experimental results of the proposed algorithm are given in Section 5.


- *Y. Chen is with Motorola Incorporated, Shaumburg, IL.*
- *R.R. Brooks is with the Information Science and Technology Applied Research Laboratory, The Pennsylvania State University, P.O. Box 30, State College, PA. 16804-0030. E-mail: rrb@acm.org.*
- *S.S. Iyengar is with the Department of Computer Science, Louisiana State University, 298 Coates Hall, Baton Rouge, LA 70803. E-mail: iyengar@bit.csc.lsu.edu.*
- *N.S.V. Rao and J. Barhen are with the Center for Engineering Science Advanced Research, Oak Ridge National Laboratory, P.O. Box 2008, Building 6010, MS-6355, Oak Ridge, TN 37831-6355.*


*Manuscript received 4 Apr. 1998; revised 22 Jan. 1999; accepted 18 Feb. 1999; posted to Digital Library 12 June 2001.*
*For information on obtaining reprints of this article, please send e-mail to: tkde@computer.org, and reference IEEECS Log Number 106449.*

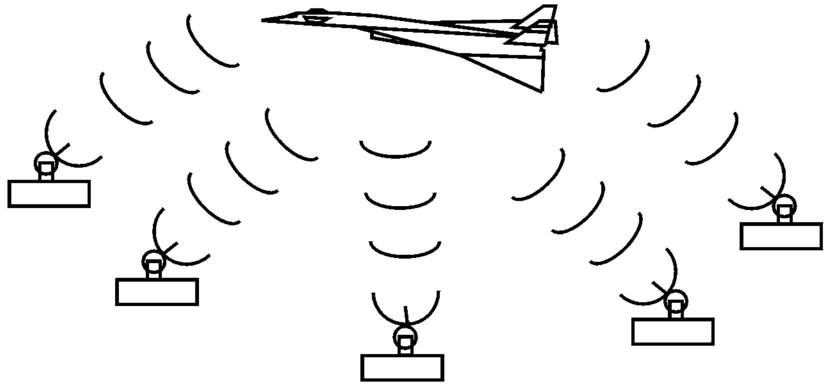


Fig. 1. Multisensor system (adapted from [5]).

## 2 Image Registration

Multiple sensor systems receive readings from different positions and orientations as in Fig. 1. To fuse multiple sensor readings, they must first be registered into a common coordinate system. Thus, image registration has been the focus of a number of studies over the past decades [4], [5], [6], [9]. Here, we consider only two-dimensional images without occlusion and projection effects, although the method proposed could easily be extended to account for projection. The class of transformations considered is gruences (translation and rotation). An addition of scaling would make the class of transformations include all affine transformations. The survey in [5] discusses registration methods that have been proposed for use with other transformations. Occlusion and projection would require a higher-level knowledge of the scene structure and interpretation of the sensor data before registration. It is doubtful that this type of interpretation is feasible in real time.

Given two overlapped sensor readings as shown in Fig. 2, we wish to find the position and orientation of the second sensor relative to the first sensor. $F$ is a function that maps a reading of sensor 2 to a reading of sensor 1. Sensor 1, represented by a function $S_1(x_1, x_2, \ldots, x_n)$, returns scalar values within a finite range of $x_1, x_2, \ldots, x_n$. Similarly, sensor 2 corresponds to the function $S_2(x'_1, x'_2, \ldots, x'_n)$. If $F$ is the mapping function, and sensors 1 and 2 are free of noise, we have [5]:

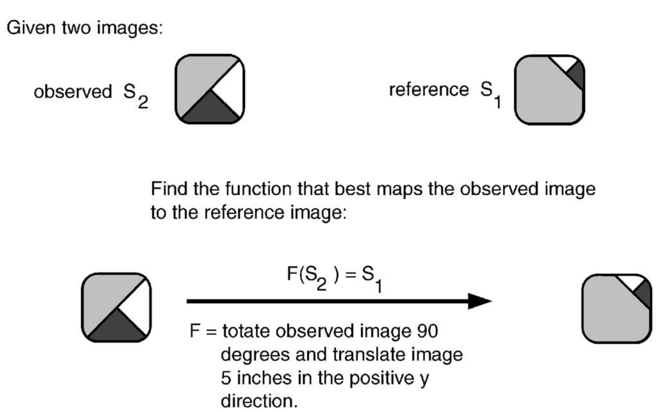



Fig. 2. Registration is finding the mapping function F(S2) (adapted from [3]).

$$F(S_2) = S_1. \tag{2.1}$$

In practice, this equation is almost never exactly satisfied since sensor readings are corrupted by noise and measurement error. In Section 2.2, we utilize this equation to derive a *fitness function* (objective function) that provides a metric for global optimization.

### 2.1 Problem Statement

The problem addressed in this paper was originally posed in [3]: Given two noisy overlapping sensor readings, compute the optimal gruence (i.e., translation and rotation) which maps one to the other. The sensors return two-dimensional gray level data from the same environment. Both sensors have identical geometric characteristics. The sensors cover circular regions such that the two readings overlap. Since the size, position, and orientation of the overlaps are unknown, traditional image processing techniques are unsuited to solving the problem. For example, the method of using moments of a region is useless in this context [10], [13].

Readings from sensors 1 and 2 are both corrupted with noise. In [3], [5], we use a Gaussian noise model to derive the fitness function since it is applicable to a large number of real-world problems. It is also the limiting case when a large number of independent sources of error exist. Our goal is to find the optimal parameters $(x_T, y_T, \theta)$ which define the relative position and orientation of the two sensor readings. The search space is a three-dimensional vector space defined by these parameters, where a point is denoted by the vector $w = [x_T, y_T, \theta]^T$. Fig. 3 shows the transformation between the two images given (2.2):

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & X_T \\ \sin\theta & \cos\theta & Y_T \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}. \tag{2.2}$$

### 2.2 Fitness Function

When noise in the sensor data has a Gaussian distribution, we have derived the fitness function in (2.3):

$$\frac{\sum (read_1(x,y) - read_2(x',y'))^2}{K(W)^2} = \frac{\sum ((gray_1(x,y) - gray_2(x',y')) + (noise_1(x,y) - noise_2(x',y'))^2}{K(W)^2}, \tag{2.3}$$

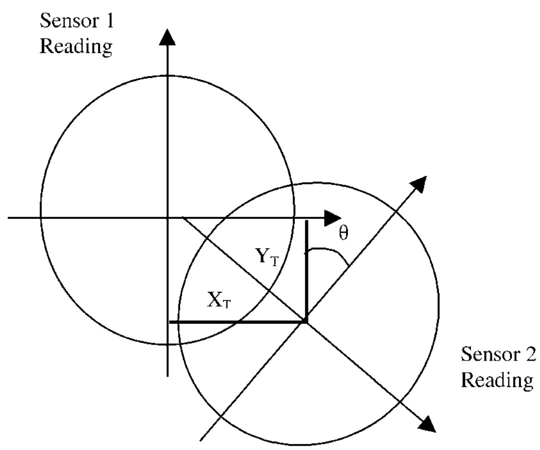



Fig. 3. Geometric relationship of two sensor readings.

| | |
|---|---|
| Where: $w$ | is a point in the search space |
| $K(w)$ | is the number of pixels in the overlap for $w$ |
| $(x',y')$ | is the point corresponding to $(x,y)$ for |
| $read_1(x,y)(read_2(x',y'))$ | is the pixel value returned by sensor 1 (2) at point $(x,y)((x',y'))$ |
| $gray_1(x,y)(gray_2(x',y'))$ | is the noiseless value for sensor 1 (2) at $(x,y)((x',y'))$ |
| $noise_1(x,y)(noise_2(x',y'))$ | is the noise in the sensor 1 (2) reading at $(x,y)(x',y')$). |

This function has been shown to reflect the problem adequately when the noise at each pixel follows a Gaussian distribution of uniform variance. In our experiments, this is not strictly true. The gray scale is limited to 256 discrete values. Because of this, when the gray-scale value is 0 (255) the noise is limited to positive (negative) values. For large intersections, however, this factor is not significant.

Equation (2.3) is derived by separating the sensor reading into information and additive noise components. This means the fitness function is made up of two components: 1) lack of fit and 2) stochastic noise. The lack of fit component has a unique minimum when the two images have the same gray-scale values in the overlap (i.e., when they are correctly registered). The noise component follows a Chi-squared distribution, whose expected value is proportional to the number of pixels in the region where the two sensor readings intersect. Dividing the difference squared by the cardinality of the overlap, we make the expected value of the noise factor constant. Dividing by the cardinality squared favors large intersections. For a more detailed explanation of this derivation, see [3], [4], [5].

Other noise models can be accounted for by simply modifying the fitness function. Another common noise model is the salt-and-pepper noise. Either malfunctioning pixels in electronic cameras or dust in optical systems commonly causes this type of noise. In this model, the correct gray-scale value in a picture is replaced by a value of 0 (255) with an unknown probability $p(q)$. An appropriate fitness function for this type of noise is (2.4):

$$\sum_{\substack{read_1(x,y)\neq 0\\ read_1(x,y)\neq 255\\ read_2(x',y')\neq 0\\ read_2(x',y')\neq 255}} \frac{(read_1(x,y)-read_2(x',y'))^2}{K(W)}. \tag{2.4}$$

A similar function can be derived for uniform noise by using the expected value $E[(U_1-U_2)^2]$ of the squared difference of two uniform variables $U_1$ and $U_2$. An appropriate fitness function is then given by (2.5):

$$\sum \frac{(read_1(x,y)-read_2(x',y'))^2}{E\left[(U_1-U_2)^2\right]K(w)}. \tag{2.5}$$

For the rest of this paper, we use (2.3). Depending upon the noise present in the sensor data, another function may be more appropriate.

In optimization literature, the function to be optimized is often called the *objective function* [1], [2], [7]. In this problem, we seek the mapping that minimizes the value given by the *fitness function* (2.3). The term fitness function is used in genetic algorithms literature [3], [4], [5], [12] in a manner similar to the use of objective function in optimization literature. Mappings are defined by (2.2), which is an affine transform without scaling. Given two sensor readings and a mapping, the value of the fitness function (2.3) is well-defined. Therefore, we seek the values of $x_T, y_T$, and $\theta$ that provide the globally minimal value for (2.3). For this reason, the terms objective function and fitness function will be used interchangeably in the rest of this paper.

## 3 Trust Method

The global minimization problem can be stated as follows: Given a function $f$ over some (possibly vector valued) domain $D$, compute $\bar{x}_{GM\in} D$ such that

$$f(\bar{x}_{GM}) \leq f(\bar{x}), \forall\, \bar{x} \in D.$$

Usually, but not necessarily, $f$ is assumed to be continuous and differentiable.

One strategy for global minimum determination is shown in Fig. 4. Given an initial starting position in the search space, we find the local minimum closest to this value. This can be done using gradient descent or a probabilistic search [8], such as genetic algorithms and simulated annealing. Once a local minimum is found, it is reasonable to assume that it is not globally optimal. Therefore, we attempt to escape the local basin of attraction.

Global optimization research concentrates on finding methods for escaping the basin of attraction of a local minimum. Often, the halting condition is difficult to determine. This means there is a trade-off between accepting a local minimum as the solution and performing an exhaustive search of the state space. In certain cases, global optimization can become prohibitively expensive when the search space has a large number of dimensions. Thus, there is a natural trade-off between solution accuracy and the time spent searching. Additionally, if an analytic form of the cost function is unknown, as is typically the case, then this problem is more pronounced.

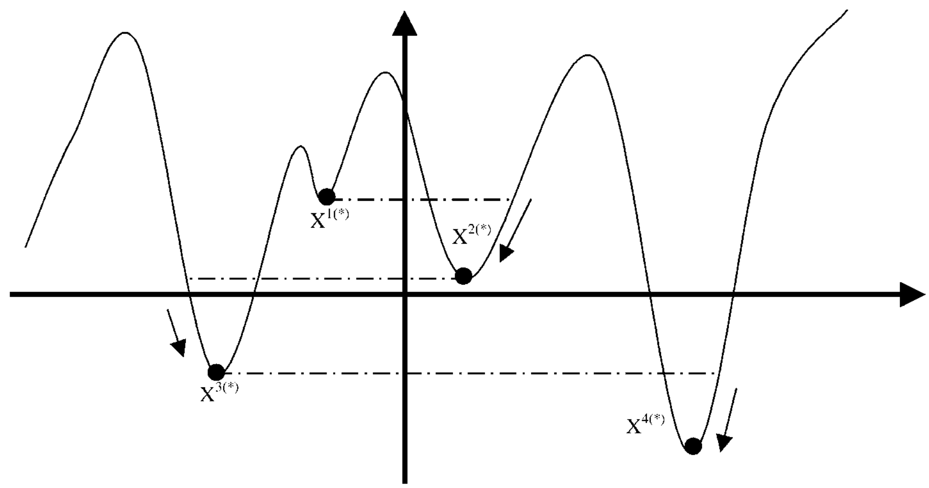



Fig. 4. Conventional tunneling method [14] for determining global minimum.

In general, the following characteristics are desirable in a global optimization method:

- Avoid entrapment in local minimum basins.
- Avoid performing an exhaustive search on the state space.
- Minimize the number of object function evaluations.
- Have a clearly defined stopping criteria.

In practice, a good global optimization method judiciously balances these conflicting goals.

A number of approaches have been proposed for solving global optimization problems (see references in [2]). For registration, we have explored the use of genetic algorithms, simulated annealing, tabu search, and artificial neural networks [3], [4]. Here, we show improved performance using the TRUST method.

### 3.1 Convergence Properties

The TRUST method [2], [7] is a deterministic global optimization method, which avoids local minima entrapment and exhaustive search. This method defines a dynamic system that uses two concepts: subenergy tunneling and non-Lipschitz terminal repellers to avoid being stuck in local minima. TRUST makes the following assumptions about the cost function $f$ and its domain $D$ [1]:

1. $f : D \rightarrow R$ is a lower semicontinuous function with a finite number of discontinuities.
2. $D \subseteq R^n$ is compact and connected.
3. Every local minima of $f$ in $D$ is twice differentiable except at a finite number of points. Furthermore, for any local minima $\boldsymbol{x}_{LM}$, we have

$$\left.\frac{\partial f}{\partial x_i}\right|_{x_{LM}} = 0, \forall\, i = 1, 2, \ldots, d \tag{3.1}$$

and

$$\boldsymbol{y}^T \frac{\partial^2 f(x_{LM})}{\partial x^2} \boldsymbol{y} \geq 0, \forall\ y\ \in D,$$

where $\frac{\partial^2 f(x_{LM})}{\partial x^2}$ is the Jacobian matrix given by

$$\left[ \left.\frac{\partial^2 f}{\partial x_i \partial x_j}\right|_{x_{LM}} \right].$$

In contrast with the conventional tunneling approach [14], tunneling is performed in TRUST by transforming the function $f$ into a new function $E(\boldsymbol{x}, \boldsymbol{x}^*)$ with similar extrema properties such that the current local minimum of $f$ at some value $\boldsymbol{x}^*$ is a global maximum of $E(\boldsymbol{x}, \boldsymbol{x}^*)$. A value of $f$ strictly less than $f(\boldsymbol{x}^*)$ is then found by applying gradient descent. The general algorithm is as follows:

1. Use a gradient descent method to find a local minimum at $\boldsymbol{x}^*$.
2. Transform $f$ into the following *virtual objective function*

$$E(\boldsymbol{x}, \boldsymbol{x}^*) = E_{sub}(\boldsymbol{x}, \boldsymbol{x}^*) + E_{rep}(\boldsymbol{x}, \boldsymbol{x}^*) \tag{3.2}$$

$$E_{sub}(x, x^*) = \log\left(\frac{1}{1 + e^{-(f(x) - f(x^*)) + a}}\right), \tag{3.3}$$

where

$$E_{rep}(x, x^*) = -\frac{3}{4}\rho(x - x^*)^{\frac{4}{3}}\, u(f(x) - f(x^*)) \tag{3.4}$$

and $E_{sub}$ is the *subenergy tunneling term* and is used to isolate the function range of $f$ less than the functional value $f(x^*)$. $E_{rep}$ is the *terminal repeller term* and is used to guide the search in the next step. In this term, $u(y)$ is the Heaviside function. Note that $E(x, x^*)$ is a well-defined function with a global maximum at $x^*$.

3. Apply gradient descent to $E(x, x^*)$. This yields the dynamical system

$$\begin{aligned} \dot{x} = -\frac{\partial E}{\partial x} = -\frac{\partial f}{\partial x}\frac{1}{1 + e^{(f(x) - f(x^*)) + a}} + \\ \rho(x - x^*)^{\frac{1}{3}}\, u(f(x) - f(x^*)). \end{aligned} \tag{3.5}$$

An equilibrium state of (3.5) is a local minimum of $E(x, x^*)$, which in turn, is a local or global minimum of the original function $f$.

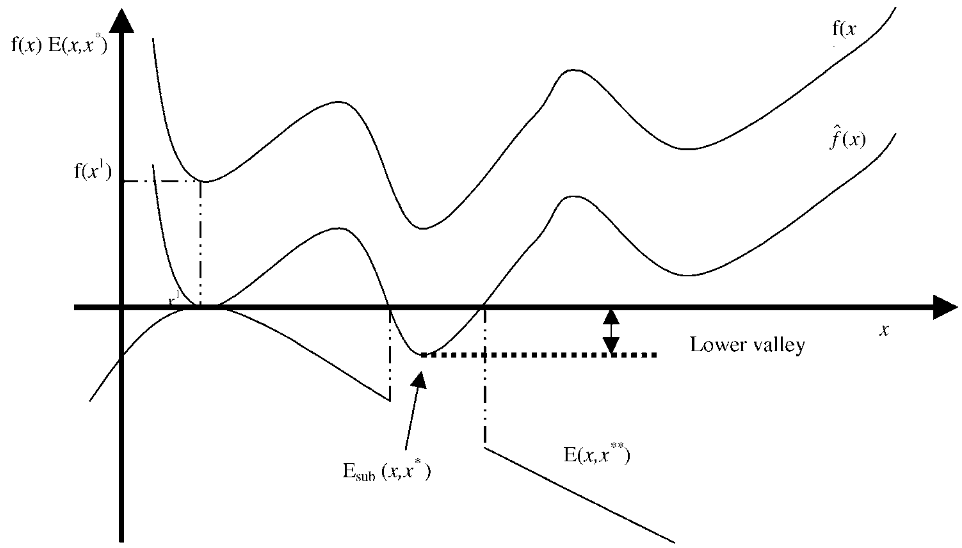



Fig. 5. TRUST in one-dimensional case.

4. Until the search boundaries are reached, repeat Step 2 with the new local minima found in Step 3 as the initial extrema.

Complete theoretical development of $E_{sub}$ and $E_{rep}$, including a discussion of terminal repeller dynamics, can be found in Barhen and Protopopescu [1].

### 3.2 Convergence and Computational Properties

In the one-dimensional case, TRUST guarantees convergence to the globally optimal value. The reason for this is: The transformation of $f$ into $E(x, x^*)$ and subsequent gradient descent search generates monotonically decreasing minimal values of $f$ (see Figs. 5 and 6). This behavior is a result of the compactness of the domain $D$ (for details, see [1]. When the last local minimum is found and the method attempts to search $E(x, x^*)$, the difference $f(x) - f(x^*)$ becomes equivalent to the $x$-axis. Thus, $E(x, x^*) = E_{rep}(x, x^*)$, and the subsequent search on this curve will proceed until the endpoints of the domain $D$ are reached.

In the more difficult multidimensional case, there is no theoretical guarantee TRUST will indeed converge to the globally optimal value. To address this problem, Barhen and Protopopescu [1] present two strategies. The first strategy involves augmenting the repeller term $E_{rep}$ with a weight based on the gradient behavior of $f$ (basically, this is the same concept as momentum in conjugate gradient descent). The effect is to guide the gradient descent search in Step 3 to the closest highest ridge value on the surface of $E(x, x^*)$. The second strategy is to reduce the multidimensional problem into a one-dimensional problem for which

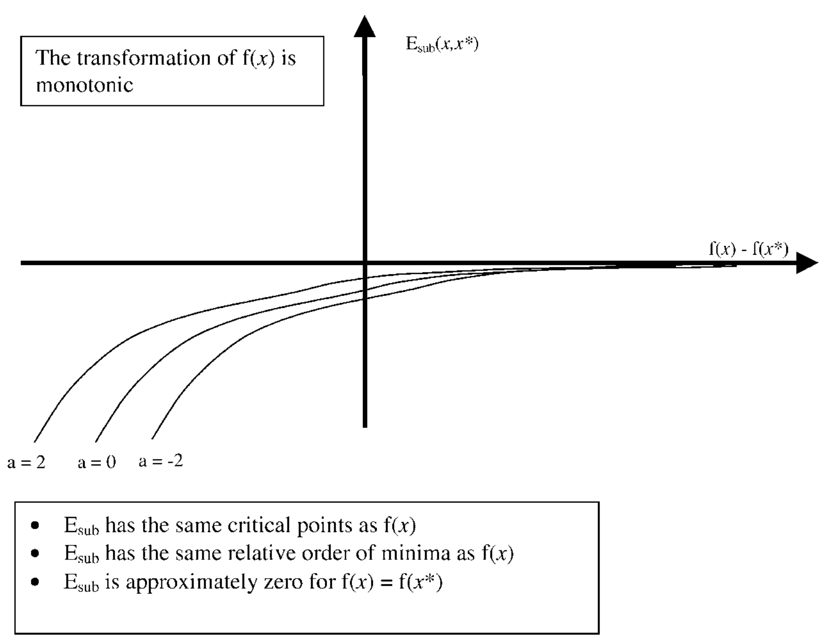



Fig. 6. TRUST: Subenergy tunneling function.

TRUST is guaranteed to converge by using hyperspiral embedding from differential geometry.

In summary, the characteristics of TRUST that are desirable for image registration include the following:

1. TRUST is computationally efficient in the number of function evaluations made during the convergence process. Usually, the most computationally demanding aspect of global optimization is evaluating the cost function [12]. In Barhen et al. [2], TRUST was compared to several other global optimization methods and was found to need far fewer function evaluations to converge.
2. Popular methods of eluding local minima basins other than conjugate gradient descent include simulated annealing and genetic algorithms [11], [12]. Both of these are probabilistic since their entrapment avoidance mechanism requires randomness. TRUST, on the other hand, is deterministic and has a well-defined stopping criteria.

## 4 Registration Algorithm

For the registration problem, we seek the optimal parameters $(\mathrm{X_T}, \mathrm{Y_T}, \theta)$ that define the correct mapping of sensor 2 to sensor 1. The following is an outline of TRUST applied to image registration:

Algorithm: TRUST-image-Registration
Inputs: Two sets of noise corrupted sensor readings $\mathrm{s_1, s_2}$
Output: Parameters x, y, and $\theta$ that map of sensor 2 to sensor 1.
Procedures:
Step 1: Read in sensor readings s1, s2.
Step 2: Set initial search position.
The initial position as the best position $(\mathrm{X_T}, \mathrm{Y_T}, \theta)$
Step 3: At each position $(\mathrm{x},\ \mathrm{y}, \theta)$
a: calculate new object function value
b: calculate the derivatives on the position $(\mathrm{x},\ \mathrm{y}, \theta)$
c: if the derivatives $(\partial \mathrm{f}/\partial \mathrm{x}, \partial \mathrm{f}/\partial \mathrm{y}, \partial/\mathrm{f}\partial\theta)$ exceed the threshold (such as 0.01)
c.1: Gradient descent phase, new position $(x + \partial \mathrm{f}/\partial \mathrm{x}, \mathrm{y} + \partial \mathrm{f}/\partial \mathrm{y}, \theta + \partial \mathrm{f}/\partial\theta)$
c.2: if $\mathrm{f}(\mathrm{x}^*, \mathrm{y}^*, \theta^*) < \mathrm{f}(\mathrm{X_T}, \mathrm{Y_T}, \theta_\mathrm{T})$, $\mathrm{f}(\mathrm{X_T}, \mathrm{Y_T}, \theta_\mathrm{T}) = \mathrm{f}(\mathrm{X}^*, \mathrm{Y}^*, \theta^*)$
d: if new position exceed the upper bound go to Step 4 else go to Step 3.
Step 4: Print out the results and terminate the program.

Assumptions and approximations described in the algorithm are as follows:

1. The variables X and Y are discrete offsets; $\theta$ is continuous.
2. Three-dimensional problem is reduced to one-dimensional case. Namely, for each objective function calculation, the changes to $\mathrm{X_T}, \mathrm{Y_T}$, and $\theta$ are calculated, respectively. The results in each dimension simulate three multidimensional cases.
3. The stopping criterion is the border of the search space. $\mathrm{X} \in (\mathrm{Xmin}, \mathrm{Xmax}), \mathrm{Y} \in (\mathrm{Ymin}, \mathrm{Ymax})$, no limit for $\theta$. The starting point is at the lower corner of the search space $(\mathrm{Xmin}, \mathrm{Ymin}, 0)$. The perturb parameter $(\varepsilon)$ for TRUST is (1,1, 0.01), which is one pixel on $\mathrm{X}, \mathrm{Y}$ direction and 0.01 degree on $\theta$.
4. To calculate the partial derivatives with respect to X and Y, five points are taken at one pixel intervals. Three points are taken to approximate the derivative with respect to $\theta$, at 0.01 degree intervals. The large number of samples increases the computational effort, but it also makes the results less sensitive to noise.
5. The strength parameter $\rho$ (kappa in program) is computed from

$$\begin{aligned}
&\rho = (1.0/(2^{(1/6)} * (1 + \mathrm{e^a})) \\
&\qquad \mathrm{sqrt}(\mathrm{x_dot_x}^2 + \mathrm{x_dot_y}^2 + \mathrm{x_dot_theta}^2) \\
&\mathrm{x_dot_x}: \text{ derivative at x direction} \\
&\mathrm{x_dot_y}: \text{ derivative at y direction} \\
&\mathrm{x_dot_theta}: \text{ derivative at 0 direction.}
\end{aligned}$$

Since we have more than one dimension, the magnitude of E and gradient vectors are taken with the F—defined above, the sum of the squares of the elements is 2, so to get the cube root of the magnitude, we take the sixth root. We also use $\rho = 10, 5, 1, 0.1$ in the program.
6. Since x and y are discrete variables, $\partial \mathrm{f}/\partial \mathrm{x}, \partial f/\partial \mathrm{y}$ are very small, sometimes less than 1. In which case, we change x and y by the sign of $\partial \mathrm{f}/\partial \mathrm{x}, \partial \mathrm{f}/\partial \mathrm{y}$.

## 5 Experimental Results

We have conducted 76 experiments using TRUST image registration using the synthesized image given by the equation (Fig. 7):

$$\begin{aligned}
100.0 + &\left({}^1\!/_{100}\right) \\
&\Big(-40x + 45y - 0.003xy + 0.02\,x^2 - 0.01\,y^2 - 20y\sin \\
&\left({}^x\!/_{18}\right) + 35y\cos(y/29) - 35\sin\left({}^x\!/_4 - {}^y\!/_{12}\right) \\
&+ 12x\cos\left({}^{xy}\!/_{100}\right)\Big).
\end{aligned}$$

This image has several periodic and nonperiodic components. In these tests, TRUST showed convergence to the global minimum within the specified parameters. The results show the influence of both the gradient and repeller parts of the algorithm. We have run similar tests on a nonsynthesized image of vegetation, with similar results. We present results using the synthesized image here, in order to allow comparison with our previously published results [3], [5].

### 5.1 One-Dimensional Case

In the one-dimensional case, TRUST is guaranteed to find the global minimum of $f$. We chose to use the rotation parameter $\theta$ for the one-dimensional test. $\theta$ is a continuous variable. Two sensor readings are cut from the terrain image (Figs. 8a and 8b). Sensor one is centered at (256,256), sensor two is centered

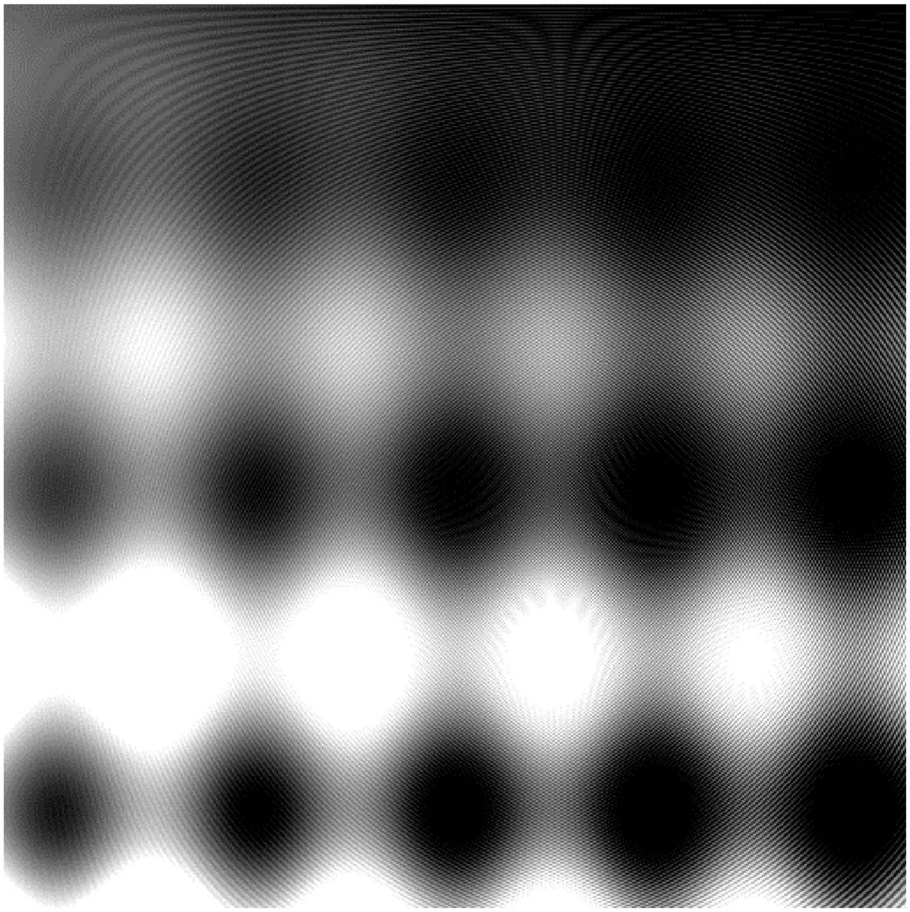
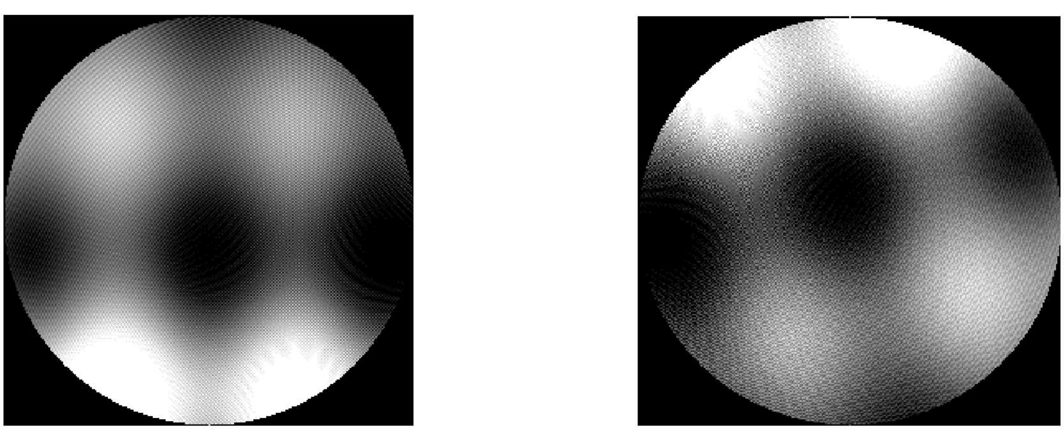


Fig. 7. Terrain model.

at (347,347) with rotation 2.74889 radians (157.5 degrees). In this report, we discuss the use of TRUST to register the data and compare these results with existing data for tabu search [5] and an elitist genetic algorithm [3].

The parameter values in Table 1 are the best answers found by TRUST. From the results (Table 1 and Fig. 9), we can see that TRUST finds local minima and climbs out of their basin of attraction quickly to find the global minimum. The result has been compared to tabu search and elitist genetic algorithm using the same starting position, 0 variance noise and sensor readings (Fig. 14). After 300 runs, tabu search is not even close to the global minimum. TRUST (Table 4) and the elitist genetic (Table 5) algorithm have almost the same quality except TRUST has reached its termination criteria. The fact that TRUST has definite termination criteria, increases our confidence that the algorithm has found the global minimum.



(a) (b)

Fig. 8. (a) Sensor 1 Reading (256,256,0). (b) Sensor 2 Reading (256,256,2.87979).

TABLE 1
TRUST Search Results

| | | |
|---|---|---|
| Sensor 1 position: | x = 0, y = 0, 0 = 0.0 | (x = 256, y = 256 in terrain model) |
| Sensor 2 relative position: | x = 0, y = 0, 0=2.87979 | (164.999817) |
| Starting point: | x = 0, y = 0, 0 = 0 | |
| α(Pole strength): | 2 | |
| ρ (Repeller strength): | 1 | |
| noise level* : | 0 | |

| *Iteration*** | *Current θ* | *Current Error* | *Best θ* | *Best error* |
|---|---|---|---|---|
| 1 | 0.060603 | 0.173321 | 0.000000 | 0.172828 |
| 15 | 21.579254 | 0.227272 | -0.76610 | 0.170642 (local min) |
| 30 | 50.870132 | 0.232877 | -0.76610 | 0.170642 (local min) |
| 45 | 58.825268 | 0.221911 | -0.76610 | 0.170642 (local min) |
| 60 | 66.101639 | 0.213194 | -0.76610 | 0.170642 (local min) |
| 75 | 88.053581 | 0.212771 | -0.76610 | 0.170642 (local min) |
| 90 | 121.599899 | 0.200787 | -0.76610 | 0.170642 (local min) |
| 105 | 152.440674 | 0.048031 | 152.440674 | 0.048031 |
| 120 | 164.931702 | 0.019233 | 164.169876 | 0.018294 |
| 135 | 164.940079 | 0.019328 | 164.617737 | 0.016797 |
| 150 | 164.642975 | 0.016843 | 164.617737 | 0.016797 |
| 165 | 163.842865 | 0.018778 | 164.625298 | 0.016780 (global min) |
| 180 | 179.566132 | 0.056133 | 164.625298 | 0.016780 (global min) |
| 190 | 489.318420 | 0.170217 | 164.625298 | 0.016780 (global min) |
| 192 | Out of bounds | | | |

*noise level is measured by variance
**Number of iterations of the algorithm

Fig. 9 shows the objective (fitness) function, which is being minimized, versus the angle of rotation $\theta$. The objective function is given by (2.3), the mean square error of the pixels in the intersection of the two images divided by the number of pixels in the intersection. Since the terrain model has significant periodic components, there are several local minima in the diagram. The valley containing the global minimum is very deep.

The search space for the one-dimensional case is from 0 to 360 degrees, which covers the entire range of $\theta$. The starting point $(x^*)$ was chosen as $\theta = 0$, the lower bound of $\theta$. The search is started by placing a repeller at $x^*$ and adding a disturbance $\varepsilon$. This gives a dynamic system with initial conditions $x^* + \varepsilon$. Table 1 gives the results. $\theta$ is given in degrees.

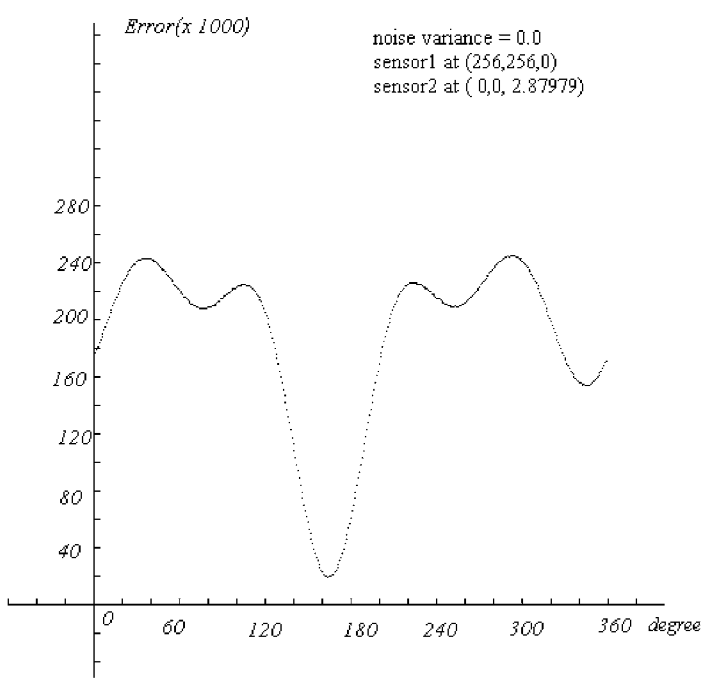



Fig. 9. Objective function.

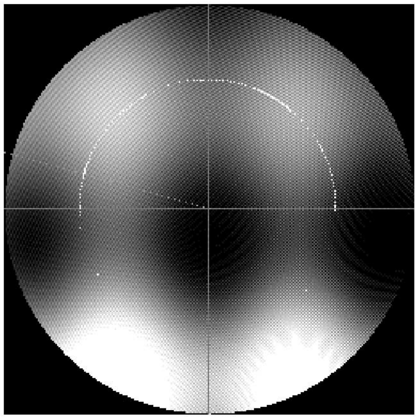


Fig. 10. Path taken by TRUST search sensor at 164.999817.

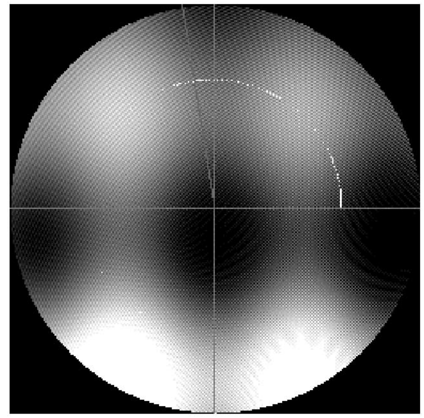


Fig. 11. Path taken by TRUST search sensor 2 at 100 degree.

Since initially $f(x^* + \varepsilon) > f(x^*)$ in the neighborhood of $x^*$, the repelling term of TRUST is active. The algorithm climbs out of a valley. The gradient of the objective function is uphill (positive), but the subenergy surface is relatively flat. With the help of a sufficiently large repeller strength $\rho$, the repeller at $x^* \mathrm{forces}$ the system uphill. This, in turn, forces the system out of the basin of attraction of the local minimum. In Table 1, the system is in the basin of attraction of the local minimum from $\theta = 21.579254$ to $121.599899$. It then enters the basin of attraction of the local minimum located at approximately 65 degrees. The search does not fall into the basin of attraction around 65 degrees, since this basin is more shallow than the one located at -0.76610. The search space remains flattened. The system remains in the repelling phase until it reaches a point where $f(x) < f(x^*)$. At which point, it has reached a basin of attraction deeper than any already encountered.

In Table 1, this occurs at approximately 100 degrees. The gradient descent phase becomes active. The algorithm then tries to minimize $f(x)$. The system finds a new minimum, and this becomes $x^*$. The function value at that point is the new $f(x^*)$. The value of the new $f(x^*)$ is less than the previous one. In this example, this occurs from $\theta = 100$ to the global minimum of 164.625298.

After reaching a minimum at 164.625298, the system recommences a cycle of tunneling. This continues until a lower basin of attraction is reached or the upper bound is exceeded. In this example, the upper bound at $2\pi$ is exceeded and the algorithm stops.

The stopping criteria of TRUST occur when the bounds of the search space are reached. In this case, $2\pi$ is the upper bound. When the system passes this point, the entire search space has been searched. The last local minimum is therefore the global minimum.

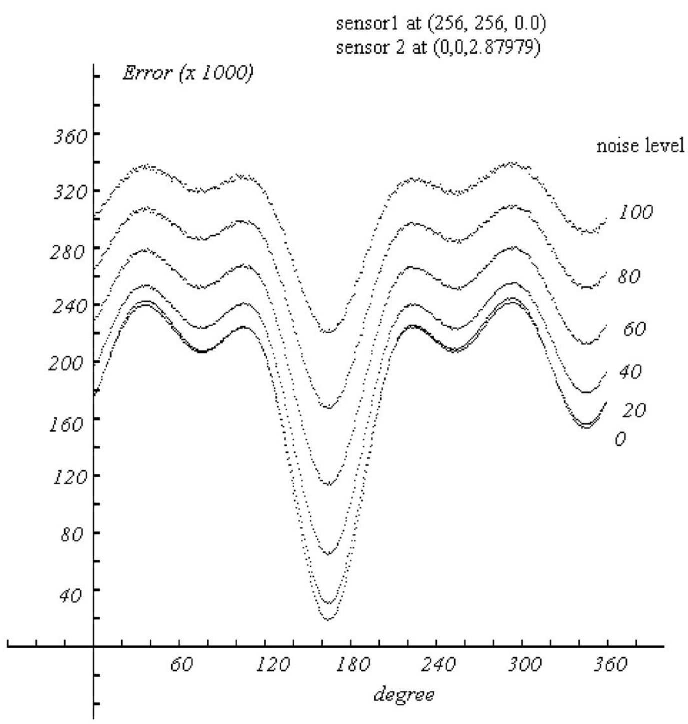



Fig. 12. Objective function under different noise levels.

Fig. 10 shows the path taken by TRUST in this example. The white dotted line denotes the rotation angle of sensor 2. The gray straight line crosses the image where the best match was found. Fig. 11 shows another path taken, when the sensor 2 was rotated by 100 degrees. These images illustrate the suitability of TRUST for image registration problems. The system moves from starting point $\theta = 0$. It moves faster in the tunneling phase, as shown in the image by there being fewer white dots. During the gradient descent phase, the system moves more slowly and more points are tested. After the global minimum is found, the subenergy surface is flattened from that point until the upper bound is reached. The system re-enters a tunneling phase and quickly reaches the upper bound. In Figs. 10 and 11, few points are tested after the global minimum is located.

Using the same parameters as in Table 1, several tests have been run with varying levels of noise. Fig. 12 shows the objective function under various noise levels. As noise increases, so do the objective function values. The increase is faster in regions with lower values. As a result, the objective function surface flattens as noise increases. The minima become less significant and the search becomes more difficult. But as shown in Table 2, TRUST handles noise well in the one-dimensional case. The search results deviate little even in the presence of sizable amounts of noise.

TABLE 2
TRUST One-Dimensional Search Results under Different Noise Levels

| *Noise level** | *$\theta$ value* | *Function value* |
|---|---|---|
| 0 | 164.625198 | 0.016780 |
| 20 | 164.226196 | 0.030003 |
| 40 | 164.576996 | 0.063540 |
| 60 | 164.847275 | 0.112951 |
| 80 | 184.856637 | 0.166876 |

**noise level is defined by variance*

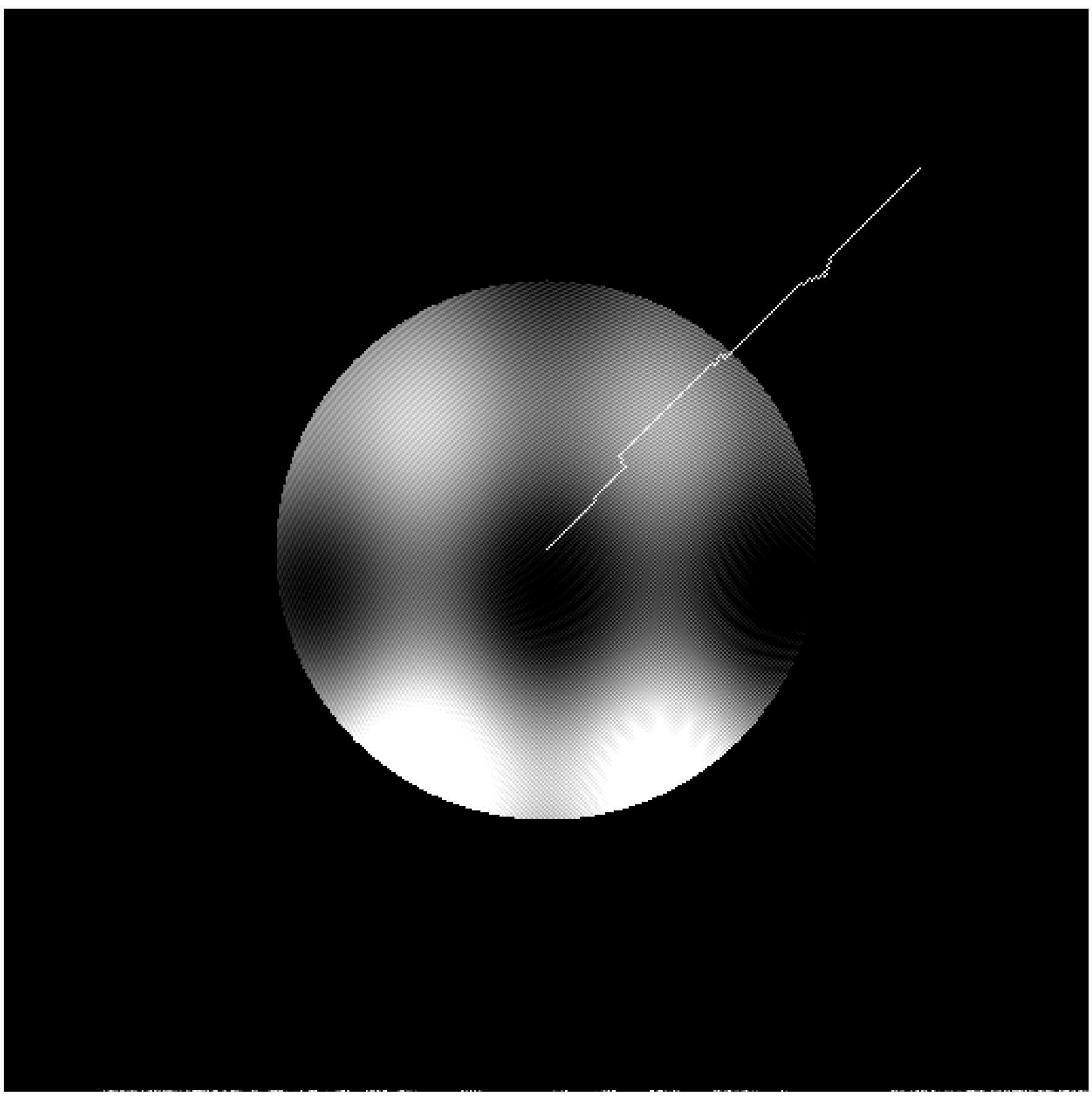


Fig. 13. Path taken by TRUST in the three-dimensional case with noise variance 0.0.

TABLE 3
Multidimensional TRUST Optimization

| | |
|---|---|
| Sensor 1 position: | x = 0, y = 0, 0 = 0.0 (x = 256, y = 256 in terrain model) |
| Sensor 2 relative position: | x = 91, y = 91, 0=2.748891 |
| Starting point: | x = 0, y = 0, 0 = 0 |
| α(Pole strength): | 2 |
| ρ (Repeller strength): | 1 |
| noise*: | 0 |

| *Iteration*** | *X value* | *Y value* | *θ value* | *Function value* |
|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0.338814 |
| 20 | 0 | 0 | 0 | 0.338814 |
| 40 | 16 | 16 | 0.734952 | 0.237387 (local min) |
| 60 | 16 | 16 | 0.734952 | 0.237387 (local min) |
| 80 | 16 | 16 | 0.734952 | 0.237387 (local min) |
| 100 | 16 | 16 | 0.734952 | 0.237387 (local min) |
| 120 | 99 | 95 | 2.704663 | 0.066328 (local min) |
| 140 | 92 | 88 | 2.746054 | 0.039967 (global min) |
| 160 | 92 | 88 | 2.746054 | 0.039967 (global min) |
| 180 | 92 | 88 | 2.746054 | 0.039967 (global min) |
| 200 | 92 | 88 | 2.746054 | 0.039967 (global min) |
| 220 | 92 | 88 | 2.746054 | 0.039967 (global min) |
| 240 | 92 | 88 | 2.746054 | 0.039967 (global min) |
| 256 | Out of bounds | | | |

** noise level is defined by variance*
***Number of iterations of the algorithm.*

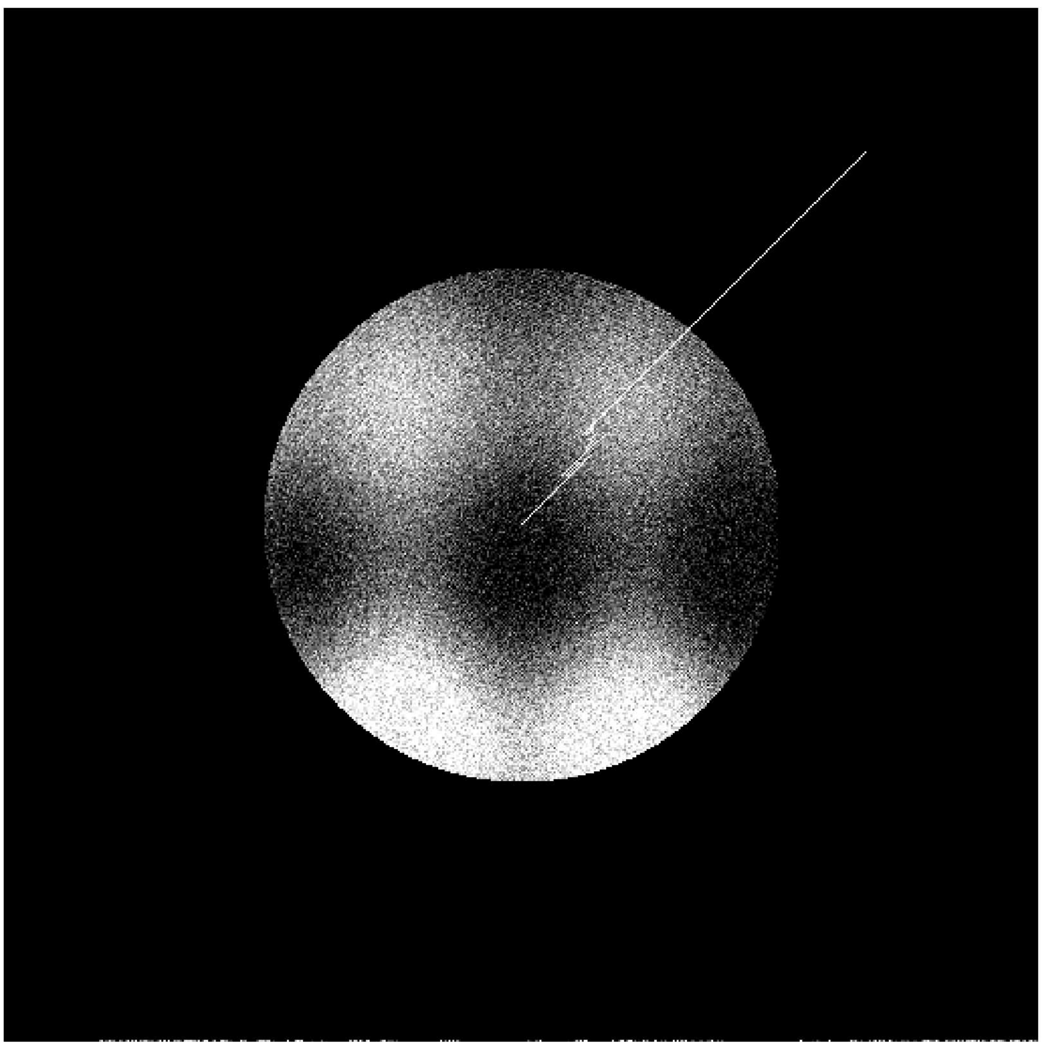


Fig. 14. Path taken by TRUST in the three-dimensional case with noise variance 30.0.

TABLE 4
TRUST Search Results under Different Noise Levels

| *Noise level** | *X value* | *Y value* | *θ value* | *Function value* |
|---|---|---|---|---|
| 0 | 92 | 88 | 2.746054 | 0.039967 |
| 10 | 94 | 88 | 2.710349 | 0.052626 |
| 20 | 124 | 126 | -0.378215 | 0.181336 |
| 30 | 93 | 87 | 2.703995 | 0.102064 |
| 50 | 93 | 89 | 2.723450 | 0.195911 |
| 70 | 92 | 92 | 2.651406 | 0.325892 |
| 90 | 0 | 0 | 0 | 0.369969 |

** noise level is defined by variance*

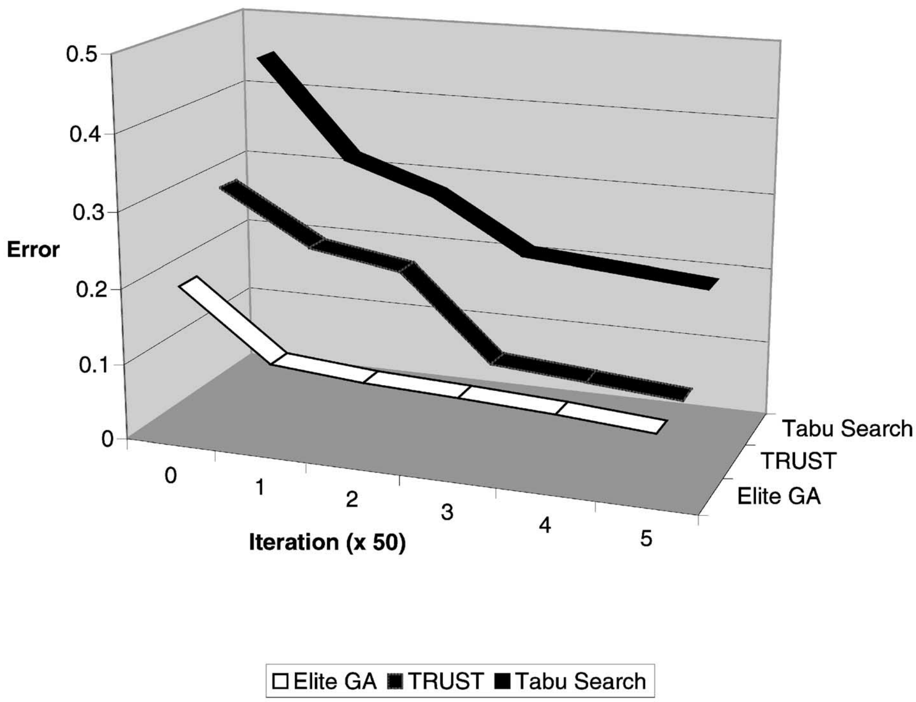



Fig. 15. Average answer value for Elite GA, TRUST, and Tabu search in three-dimensions.

### 5.2 Multidimensional Case

The search space in the three-dimensional case is: $-255 < x < 255, -255 < y < 255, 0 < \theta < 2\pi$. Fig. 13 shows the path taken by TRUST under the conditions given in Table 3. The line from the center (0,0) goes to bound (255,255) and the search stops. Table 3 and Fig. 15 illustrate the ability of TRUST to find local minima and quickly climb out of their basins of attraction. After 300 iterations, the results of tabu search are not even close to the global minimum. The results from TRUST and Elite Genetic Algorithms are almost identical, except TRUST has concrete stopping criteria that it has satisfied. It has finished examining the search space. We can be more certain that a global minimum has been located.

We have tested TRUST using noise with seven different variances under different conditions. We compare the results with the elitist genetic algorithm. The results in Table 4 use the same parameters as Table 3, except for the amount of noise. These results are compared with those from an elitist genetic algorithm in Table 5.

From Tables 4 and 5, both the elitist genetic algorithm and TRUST can handle noise with a variance of up to 30. The algorithms do not always find the global minima. But the TRUST results show the optimal value can be found even in the presence of large amounts of noise. When the noise reaches levels such as 70, or 90 as shown in Tables 3 and 4, it obscures the images and it becomes impossible to find the correct answer.

## 6 Conclusions

This paper discussed use of the TRUST method (proposed by Barhen et al. [2], [7]) as a deterministic optimization algorithm

TABLE 5
Elite Genetic Search Results under Different Noise Levels [3], [5]

| *Noise level** | *X value* | *Y value* | *θ value* |
|---|---|---|---|
| 0 | 89 | 91 | 2.74744 |
| 10 | 92 | 92 | 0 |
| 20 | 91 | 91 | 2.74744 |
| 30 | 89 | 89 | 2.74744 |
| 50 | 86 | -18 | 2.79768 |
| 70 | -48 | 6 | 6.02138 |
| 90 | 0 | 5 | 1.23297 |

** noise level is defined by variance*

to solve an image-processing problem. TRUST shows very good results in optimizing one-dimensional optimization problems. It approaches optimization by phrasing the problem as a differential equation. It uses terminal repellers to create a subenergy tunneling equation. The subenergy tunneling equation creates a dynamic system, which finds the globally optimal answer. We use TRUST to solve a multi-dimensional problem, with very encouraging results. TRUST finds globally optimal answers even in the presence of large amounts of noise. This is especially noteworthy since the effect of noise is usually magnified by differentiation. It is possible that the objective function's evaluation of error using a large number of pixels performs an implicit integration, which counteracts this effect.

Though convergence to a global minimum in the multi-dimensional case has not been proven mathematically, our results indicate that it performs at least as well as elite genetic algorithms. The computation required by TRUST is significantly less than required by genetic algorithms, and it has a well-defined halting criteria. Our previous work [3], [5] has compared elite genetic algorithms with classic genetic algorithms, tabu search, and simulated annealing, and found the results returned by elite genetic algorithms superior. TRUST was found to work well in the presence of large amounts of noise.

## Acknowledgments

Part of this research was funded by ONR grant number N00015-94-I-0343 to Professor Iyengar. Research for S.S. Iyengar was partially funded by Oak Ridge National Laboratory. Research for N.S.V. Rao and J. Barhen was sponsored by the Engineering Research Program of the Office of Basic Energy Sciences, US Department of Energy, under contract no. DE-AC05-00OR22725 with UT Battelle, LLC. The authors would like to thank the anonymous reviewers whose comments improved the presentation of this paper.

## References

[1] J. Barhen and V. Protopopescu, "Generalized TRUST Algorithms for Global Optimization," *State of the Art in Global Optimization,* C.A. Floudas and P.M. Pardalos, eds., pp. 163-180, Kluwer Academic, 1996.

[2] J. Barhen, V. Protopopescu, and D. Reister, "TRUST: A Deterministic Algorithm for Global Optimization," *Science,* vol. 276, May 1997.

[3] R.R. Brooks, S.S. Iyengar, and J. Chen, "Automatic Correlation and Calibration of Noisy Sensor Readings Using Elite Genetic Algorithms," *Artificial Intelligence,* vol. 84, pp. 339-354, 1996.

[4] R.R. Brooks, "Robust Sensor Fusion Algorithms: Calibration and Cost Minimization," PhD dissertation, Louisiana State Univ., Aug. 1996.

[5] R.R. Brooks and S.S. lyengar, *Multi-Sensor Fusion: Fundamentals and Applications with Software,* p. 8, Upper Saddle River, N.J.: Prentice Hall, 1998.

[6] L.G. Brown, "A Survey of Image Registration Techniques," *ACM Computing Surveys,* vol. 24, no. 4, pp. 325-376, Dec. 1992.

[7] B.C. Cetin, J. Barhen, and J.W. Burdick, "Teminal Repeller Unconstrained Subenergy Tunneling (TRUST) for Fast Global Optimization," *J. Optimization Theory and Applications,* vol. 77, no. 1, pp. 97-126, 1993.

[8] J.S. De Bonet, C.L. Isbell Jr., and P. Viola, "@IC: Finding Optima by Estimating Probability Densities," *Advances in Neural Information Processing Systems.* Cambridge, Mass: MIT Press, 1997.

[9] L. Grewe and R.R. Brooks, "Efficient Registration in the Compressed Domain," *Proc. SPIE Wavelet Applications, VI,* H. Szu, ed., vol. 3723, 1999.

[10] J.C. Russ, *The Image Processing Handbook.* Boca Raton, Fla.: CRC Press, 1995.

[11] Y. Shang and B.W. Wah, "Global Optimization for Neural Network Training," *Computer,* pp. 45-54, Mar. 1996.

[12] M. Srinivas and L.M. Patnaik, "Genetic Algorithms: A Survey," *Computer,* vol. 27, no. 6, pp. 17-26, June 1994.

[13] A.C. Weeks, *Fundamentals of Electronic Image Processing.* Bellingham, Wash.: SPIE Press, 1996.

[14] A.V. Levy and A. Montalvo, "The Tunnelling Algorithm for the Global Minimization of Functions," *SIAM J. Scientific and Statistical Computing,* vol. 6, pp. 15-29, 1985.

**Ying Chen** received the BS degree from the University of Science and Technology of China in 1992. She received the MS degree in computer science from Louisiana State University, Baton Rouge in 1998. Currently, she is working a software engineer for Motorola Incorporated in Shaumburg, Illinois.

**Richard R. Brooks** received the BA degree in mathematical sciences from the Johns Hopkins University and performed studies in computer science and operations research at the Conservatiore National des Arts et Metiers in Paris, France. He received the PhD degree in computer science from Louisiana State University, Baton Rouge in 1996. Curently, he is a full time researcher in the Information Science and Technology division of the Applied Research Laboratory at The Pennsylvania State University. Formerly, he was an assistant professor witht the Institute of Communications Science and Technology at California State University, Monterey Bay. His work experience includes being manager of systems and applications programming for Radio Free Europe/Radio Liberty in Munich, Germany. As a consultant, he has consulted on the implementation of a stock trading network for the French Stock Exchange Authority and the expansion of the World's Bank internal computer netwrok to Africa and the former Soviet Union. His areas of research expertise include intelligent indexing, reasoning with uncertainty, data fusion, robotics, image processing, multimedia imformation retrieval, and distributed systems. He is a member of the IEEE Computer Society.

**S. Sitharama Iyengar** received the MS degree from the Indian Institute of Science in 1970 and the PhD degree from Mississippi State in 1974. Since receiving his PhD degree, he has been involved with research in high-performance algorithms and data structures. Currently, he is the chairman of the Computer Science Department, a professor in the Computer Science Department, and has directed more than 28 PhD dissertations at Louisiana State University, Baton Rouge. He has served as a principal investigator on research projects supported by the US Office of Naval Research, The Defense Advanced Research Projects Agency, the Department of Defenses's multiuniversity Research initiative, the National Aeronautics and Space Administration (NASA), the US National Science Foundation, California Institute of Technology's Jet Propulsion Laboratory (JPL), the Department of Navy-NORDA, The US Department of Energy, LEQFS-Board of Regents, and the US Army Office. His publications include several books (Prentice Hall, CRC Press, IEEE CS Press, etc.) and more than 250 research papers in the areas of high-performance parallel and distributed algorithms and data structures for image processing and pattern recognition, autonomous navigation, and distributed sensor networks. He has also been a visiting professor at the Jet Propulsion Laboratory, Oak Ridge National Laboratory (ORNL), and the Indian Institute of Science. He was the series editor for *Neuro-Computing of Complex Systems* and an editor for the *Journal of Computer Science and Information.* He has served as a guest editor for the *IEEE Transactions on Knowledge and Data Engineering*, *IEEE Transactions and SMC*, *IEEE Transactions on Software Engineering*, *Journal of Theoretical Computer Science*, *Journal of Computer and Electrical Engineering*, and the *Journal of the Franklin Institute*. He is a fellow of the IEEE and was a distinguished visitor of the IEEE Computer Society (1995-1998). Since 1985, he has served as an ACM national lecturer and is a member of the New York Academy of Sciences. He has been the program chairman for many national/international conferences. Also, he has served on the prestgious NIH-NLM Review Committee in the area of medical informatics. In 1997, he was the winner of the prestigious IEEE Technical Achievement Award for outstanding contributions to data structures and algorithms in image processing and sensor fusion problems. He was also awarded the LSU Distinguished Faculty Award for Excellence in Research, and the LSU Tiger Athletic Foundation Teaching Award in 1996. He has been a consultant to several industrial and government organizations (JPL, NASA, etc.)

**Nageswara S.V. Rao** received the BTech degree in electronics and communications engineering from Regional Engineering College, Warangal, India in 1982, the ME degree from the School of Automation, the Indian Institute of Science, Bangalore in 1984, and the PhD degree in computer science from Louisiana State University, Baton Rouge in 1988. In 1993, he joined the Center for Engineering Systems Advanced Research (CESAR), Oak Ridge National Laboratory (ORNL) where he is currently a senior research staff member. From 1988 to 1993, he was an assistant professor in the Department of Computer Science at Old Dominion University, Norfolk, Virginia. His research interests include multiple sensor systems, computer network protocols, robot navigation in unknown terrains, neural networks, fault diagnosis, and adaptive algorithms. He has published more than 150 research articles in various journals, conference proceedings, and books. He was a co-chair for the Workshop on Foundations of Information/Decision Fusion, Washington, DC in 1996. He was a member of the editorial board for the *Journal of the Franklin Institute*. He is a senior member of the IEEE.

**Jacob Barhen** received the DSc degree from the Technion-Israel Institute of Technology, Haifa in 1978. He is the director of the Center for Engineering Science Advanced Research (CESAR) at the Oak Ridge National Laboratory (ORNL). He is also the head of the Intelligent and Emerging Computational Systems section in the Computer Science and Mathematics division. From 1987 to 1994, he was the head of the Nonlinear Science Information Processing Group at the Jet Propulsion Laboratory (JPL), California Institute of Technology. From 1978 to 1987, he was with ORNL, where he headed the Machine Intelligence Group. He has been the principal investigator and manager of numerous basic and applied research projects funded by the US Departments of Defense and Energy, NASA, and other US government agencies. His research interests include global optimization, neural networks, intelligent systems, computational methods for seismic imaging, and quantum computing and communication. He has authored more than 150 scientific papers and serves on the editorial boards of *Neural Networks*, *Neural Processing Letters*, *Mathematical and Computer Modeling*, and *Concurrency*. He holds eight US patents. He has received the 1998 R and D 100 Award for his invention of the TRUST global optimization method (*Science*, vol. 276, 1094-1097, 1997). He was also the recipient of 12 NASA Awards for Technical Innovation and was honored in 1993 and 1995 with two NASA Major Monetary Awards for his contribution to the National Space Program. He is a member of the IEEE, SPIE, the Planetary Society, and the International Neural Networks Society.