Fast and accurate image registration using Tsallis entropy and simultaneous perturbation stochastic approximation

S. Martin, G. Morison, W. Nailon and T. Durrani

The Tsallis measure of mutual information is combined with the simultaneous perturbation stochastic approximation algorithm to register images. It is shown that Tsallis entropy can improve registration accuracy and speed of convergence, compared with Shannon entropy, in the calculation of mutual information. Simulation results show that the new algorithm achieves up to seven times faster convergence and four times more precise registration than using a classic form of entropy.

Introduction: Image registration involves finding the optimum transformation T, which will best align a floating image $F(i, j)$ to a reference image $R(i, j)$, where i and j are their co-ordinates, such that $F(T(i, j))$ fits $R(i, j)$. By maximising the information associated with the pixel intensity distribution of the images, a similarity measure between two images can be computed. The definition of mutual information is based on the relative entropy, or Kullback-Leibler distance [1], which is described as:

$$I(X; Y) = \sum_{x,y} p(x, y) \log \frac{p(x, y)}{p(x)p(y)} \tag{1}$$

where X and Y are two discrete random variables representing the pixel intensity, $p(x)$ and $p(y)$ are their marginal probability density functions, respectively, and $p(x, y)$ is their joint probability density. Information theoretic similarity measures based on mutual information, combined with powerful optimisation methods, have been used in the field of image registration [1, 2]. The majority of these methods use only Shannon entropy in the calculation of the mutual information. Here, a fully automated algorithm for image registration based on Tsallis entropy [3] has been developed that uses Spall's simultaneous perturbation stochastic approximation (SPSA) algorithm [4] to provide accurate image registration and rapid convergence. The performance of the new approach is compared with the classic similarity measure of mutual information [1].

Registration algorithm: There are four key stages in an image registration algorithm. They are: a similarity measure between the images; an optimisation method; a transformation scheme; and an interpolation technique for the transformed image to fit the grid of the reference. The major contribution of this work is the development of an image registration algorithm utilising Tsallis mutual information and SPSA optimisation technique.

To test the algorithm, images were transformed by a rigid-body transformation scheme given a vector T equal to $[t_x, t_y, \theta]$, which corresponds to translations over x- and y- axes and a rotation through the centre of the image with an angle θ. Spline interpolation was used to fit the floating image $F(T(i, j))$ to the grid of the reference image $R(i, j)$ after each transformation.

Measures of divergence and mutual information: The divergence is a measure of the distance between two discrete random variables distributions P and Q. In the case of Kullback-Leibler divergence, it is given by:

$$D(P\|Q) = \sum_i p_i \log \frac{p_i}{q_i} \tag{2}$$

with p_i and q_i the probability distributions associated with the distributions P and Q. The mutual information is a particular case of the divergence measure defined between the joint probability density and the product of the factorised marginal probabilities. Using the previously defined pixel intensity random variable X and Y, the mutual information is the distance the joint probability density $p(x, y)$ and $p(x) \cdot p(y)$. In other terms, the mutual information can be represented as $I(X; Y) = D(p(x, y)\|p(x)p(y))$ which gives (1).

Tsallis has presented a form of nonextensive entropy to describe a vast class of physical phenomena [2]. The divergence measure proposed by Tsallis is defined by:

$$D_{Ts}(P\|Q) = \frac{1}{1-\alpha}\left(1 - \sum_i \frac{p_i^\alpha}{q_i^{\alpha-1}}\right) \tag{3}$$

with $\alpha \in \Re - \{1\}$. Replacing p_i and q_i, respectively, by the joint probability $p(x, y)$ and by the product of the marginal densities $p(x) \cdot p(y)$, the mutual information based on Tsallis definition of entropy is obtained, which is described as:

$$I^\alpha(X; Y) = H^\alpha(X) + H^\alpha(Y) - (1-\alpha) \times H^\alpha(X)H^\alpha(Y) - H^\alpha(X, Y) \tag{4}$$

with $H^\alpha = (1-\alpha)^{-1}(\sum_x p(x)^\alpha - 1)$ being Tsallis entropy of order. It is important to note that when a tends to 1, using L'Hopital's rule, Tsallis definition tends towards Shannon entropy.

SPSA algorithm: The simultaneous perturbation stochastic approximation (SPSA) algorithm was developed by Spall [4] and since its development has been used in a number of different fields, e.g. neural network training, traffic management and recently in the blind source separation problem [5]. The goal of the algorithm is to minimise a scalar valued cost function $J(t_x, t_y, \theta)$, assuming only that noisy measurements of the cost function are available, and that the cost function is differentiable. Differing from standard Robbins Monro stochastic approximation algorithms that require either explicit calculation or measurement of the gradient of the cost function, the SPSA algorithm utilises only two measurements of the cost function and a random simultaneous perturbation to estimate the gradient of the cost function. This is calculated as follows:

$$\nabla(\theta) = \frac{J(t_x, t_y, \theta + c\xi_i) - J(t_x, t_y, \theta)}{c\xi_i} \tag{5}$$

where c is a small perturbation value, ξ is the element in the ith row of a sign vector with elements taking a value $+1$ or -1 generated from a Bernoulli distribution. The vector elements have the following property:

$$E\{\xi_{ij}\} = 0 \tag{6}$$

Taking expectations of the Taylor series expansion of (5) the following expression is obtained:

$$E\{\nabla(\theta)\} = \frac{dJ(t_x, t_y, \theta)}{d\theta} \tag{7}$$

Thus to obtain more accurate estimates of the required gradient a number of estimates of the gradient are made and the expectation of the gradient estimates is used in a standard stochastic approximation gradient update equation. The sample expectation is taken over N values of the gradient estimates:

$$\theta(k+1) = \theta(k) - a_k\left\{N^{-1}\sum_{n=1}^{N-1}(\nabla\theta)_n\right\} \tag{8}$$

a_k represents the step size parameter for the system. To improve system convergence the step size is exponentially decayed with increasing data points that follow. The above analysis is also performed for the parameters t_x and t_y.

Implementation results: Three different sets of images have been used to test the new registration algorithm, including T1 and T2 magnetic resonance images (MRI) and computed tomography images (CT) with all 16-bit grey levels to provide reference images. The floating images were the same images corrupted with uniformly distributed random multiplicative noise with $\mu = 0$ and $s = 0.04$. Prior to starting the optimisation process, each floating image was transformed with an initial vector $T = [t_y, t_y, \theta] = [10, -5, 15]$. Complete registration is obtained when $T = [t_y, t_y, \theta] = [0, 0, 0]$.

Table 1: Registration results

Type of entropy		Results			Convergence iterations
Type	a	t_x	t_y	θ	
Classic		0.0081	0.0118	0.0316	300
Tsallis	0.9	0.0069	0.0086	0.0118	100
Tsallis	0.8	0.003	0.008	0.0028	42
Tsallis	0.7	2.121	0.1239	0.6712	14
Tsallis	0.6	3.1021	1.0196	1.6459	None
Tsallis	0.5	4.2386	10.8246	−99.6618	None

Table 1 provides a summary of the simulation results obtained for the experiment after 500 optimisation iterations for the T2 MRI. It shows that the Tsallis method can achieve a more accurate registration than Shannon based mutual information. Fig. 1 represents the behaviour of the θ transformation parameter during the 500 iterations of the optimisation process, where the three transformation parameters converge to the solution approximately at the same time. From Fig. 1, the fastest data set to converge was observed when using Tsallis mutual information, with $a = 0.7$, in approximately 14 iterations when tested with T2 MRI but the optimisation algorithm oscillates around the solution. With the exception of Tsallis with $a < 0.7$, the registration algorithm converges to a global solution. Therefore if a is too low, the algorithm may diverge and if $a = 1$, then the approach becomes that of Shannon mutual information. In practice, a is chosen by trading convergence speed and registration precision. For the T1 MRI and CT images, the results obtained show that the optimum value of a may slightly vary around 0.7 as with the T2 MRI data set.

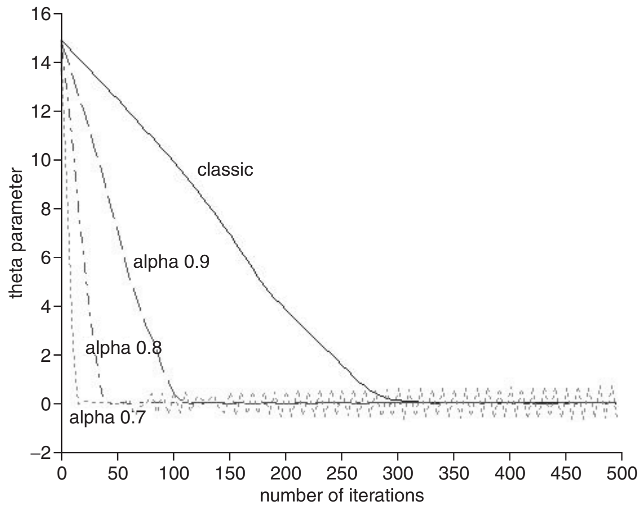

Fig. 1 *Convergence of θ transformation parameter against number of iterations of optimisation algorithm for Tsallis mutual information with different values of a parameter when initial vector T = [0, 0, 0], compared with classic Shannon information measure for T2 MRI*

Conclusions: Tsallis entropy based image mutual information, combined with a stochastic optimisation algorithm, has been proposed leading to a fast and accurate image registration algorithm. It is shown that the convergence speed of the algorithm can be improved by a factor up to seven, and the accuracy of the registration could be increased by an order of four compared to classic entropy with the a parameter correctly set. An adaptive SPSA algorithm is currently under investigation to increase the convergence of the optimisation algorithm in combination with alternative improved entropy measures to increase the overall convergence of the system. This concept is being extended to register multimodal images together with techniques that select the appropriate a parameter for their fast and accurate registration.

Acknowledgments: The authors wish to thank Epson Scotland Design Centre.

3 March 2004
Electronics Letters online no: 20040375
doi: 10.1049/el:20040375

S. Martin, G. Morison, W. Nailon and T. Durrani (*Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow G1 1XW, United Kingdom*)

E-mail: stefan@spd.eee.strath.ac.uk

References

1 Antoine Maintz, J.B., and Viergever, M.A.: 'Mutual-information-based registration of medical images: a survey', *IEEE Trans. Med. Imaging*, 2003, **22**, (8), pp. 986–1004
2 Cole-Rhodes, A., and Adenle, A.: 'Automatic image registration by stochastic optimization of mutual information'. IEEE Int. Conf. on Acoustics, Speech and Signal Processing, Hong Kong, 2003
3 Borland, L., Plastino, A.R., and Tsallis, C.: 'Information gain within nonextensive thermostatistics', *J. Math. Phys.*, 1998, **39**, (12), pp. 6490–6501
4 Spall, C.: 'Multivariate stochastic approximation using a simultaneous perturbation stochastic gradient approximation', *IEEE Trans. Autom. Control*, 1992, **37**, (3), pp. 332–341
5 Morison, G., and Durrani, T.: 'SPSA for noisy non-stationary blind source separation'. IEEE Int. Conf. on Acoustics, Speech and Signal Processing , Hong Kong, 2003, Vol. 5, pp. V_285–V_288