A Framework for Parameter Optimization in Mutual Information (MI)-based Registration Algorithms

Girish Gopalakrishnan, S. V. Bharath Kumar†, Rakesh Mullick, Ajay Narayanan and Srikanth Suryanarayanan

Imaging Technologies Lab, GE Global Research, Bangalore, India.

ABSTRACT

In this paper, we present a framework that one could use to set optimized parameter values, while performing image registration using mutual information as a metric to be maximized. Our experiment details these steps for the registration of X-ray Computer Tomography (CT) images with Positron Emission Tomography (PET) images. Selection of different parameters that influence the mutual information between two images is crucial for both accuracy and speed of registration. These implementation issues need to be handled in an orderly fashion by designing experiments in their operating ranges. The conclusions from this study seem vital towards obtaining allowable parameter range for a fusion software.

Keywords: Image Registration, Parameter Optimization, Mutual Information, Multi-resolution, Computer Tomography, Positron Emission Tomography, Multi-modality Fusion, Design of Experiments.

1. INTRODUCTION

Image fusion (registration), a classical problem in medical image understanding for patient care, involves mapping/associating data points on different images for diagnosis. Registration involves the search for a suitable transform so as to maximize/minimize a chosen metric given certain input parameters. A key metric for registration is Mutual Information (MI): an information theoretic measure, suitable especially for multi-modality registration. The final convergence of MI is dependent upon different parameters that are involved in the process of registration. For a good quality of registration, we need to use the optimal combination of these parameters to maximize MI. Though, there have been studies describing the influence of such parameters on MI, specific to MR and SPECT brain images,[1] they are not comprehensive as they assume the parameters influencing MI to be independent. Zhu *et al.*[1] emphasize the selection of modules such as optimizer, interpolation and registration strategy by using one implementation at a time on a 6-parameter rigid-body registration.

To overcome the same, in this paper, we present a framework that could be used to set optimized parameter values, while performing image registration using MI as a metric to be maximized. The study is motivated by the desire to compute an optimum set of parameters for a given modality by conducting Design of Experiments (DOE). The proposed framework involves DOE, which studies the interaction among these parameters while maximizing MI and compute the optimum setting for these parameters. This framework can be used in the parameter optimization for any set of multi-modality data-sets with appropriately chosen parameters. Though we demonstrate the flow on a 12-parameter affine registration, it is possible to optimize these parameters for any registration methodology (deformable etc.) that uses MI as its metric. In our study, we considered the problem of CT-PET fusion and the experiments are conducted on 5 subjects wherein we have computed an optimum parameter setting for CT-PET fusion. We have used a slight extension of MI as our metric to measure the performance change between different settings so that it can be cross-verified using phantoms[2] and radiologists.

Some of the parameters of interest are number of levels of multi-resolution decomposition, number of samples chosen to compute probability density function (pdf) and hence entropy, histogram bin size, interpolation methods, optimization algorithms and preprocessing. Different implementation strategies were evaluated by changing

†The author is presently a Ph. D. student in the Department of Electrical and Computer Engineering at the University of California, San Diego, USA.

Send correspondence to Girish Gopalakrishnan, Plot #122, EPIP, John F Welch Technology Center Pvt. Ltd., Hoodi Village, Whitefield Road, Bangalore-560066, India. Phone: +91 (80) 2503 3160. E-mail: girish.gopalakrishnan@ge.com.

each of these parameters systematically. Since the information content varies across the data-sets, it is not possible to predict the final value of the metric across cases. This also rules out the option of determining one common regression equation relating the metric to influencing parameters for all possible cases. The outcome of this experiment is a narrowed parameter space that guarantees improvement in the registration quality. The final settings within this space can be chosen based upon the time that is rationed for the algorithm.

The paper is organized as follows. In Section 2, we present the concepts of Shannon entropy and MI before introducing the metric that is used to drive our design of experiments. By choosing an ITK[3] based PET-CT registration algorithm that uses MI as a driver, we provide the steps in selection of variance drivers in Section 3. In Section 4, we explain design optimization and selection of optimal values for these drivers. We finally conclude with some future directions of this effort in Section 5.

2. SHANNON ENTROPY AND MUTUAL INFORMATION

Claude Shannon's 1948 seminal paper "A Mathematical Theory of Communication" marked the beginning of information theory, wherein the fundamental laws of data compression and transmission are laid out. Information theory deals with the processing of information and is used in multitude of areas. In the mid 1990s, information theoretic methods made its entry into the field of medical image registration.[4,5] In this section, we review the basic definitions of information theory[6] and their applications to medical image registration.

The *Shannon entropy* $H(X)$ of a discrete random variable X with a set of possible values, $S_X = \{s_1, s_2, \cdots, s_N\}$, having probabilities $\{p_1, p_2, \cdots, p_N\}$, with $P(X = s_i) = p_i$, $p_i \geq 0$ and $\sum_{x \in S_X} P(X = x) = 1$ is defined as

$$H(X) = -\sum_{x \in S_X} P(X = x) \log P(X = x) = -\sum_{i=1}^{N} p_i \log p_i \tag{1}$$

with the convention for $P(X = x) = 0$ that $0 \times \log \frac{1}{0} \equiv 0$, since $\lim_{\theta \to 0^+} \theta \log \frac{1}{\theta} = 0$. The entropy measures the average information content or 'uncertainty' of X. It is to be noted that $H(X) \geq 0$ with equality if, and only if, $p_i = 1$ for any one i.

If we consider another discrete random variable Y with a marginal probability distribution q, corresponding to values in the set $T_Y = \{t_1, t_2, \cdots, t_M\}$, the *joint entropy* of a pair of discrete random variables X and Y is defined as

$$H(X,Y) = -\sum_{x \in S_X, y \in T_Y} P(X = x, Y = y) \log P(X = x, Y = y) = -\sum_{i=1}^{N} \sum_{j=1}^{M} p_{ij} \log p_{ij} \tag{2}$$

where $p_{ij} = P(X = s_i, Y = t_j)$ is the joint probability of X and Y. The *conditional entropy* of X and Y is defined as

$$H(X|Y) = -\sum_{x \in S_X, y \in T_Y} P(X = x, Y = y) \log P(X = x|Y = y) = -\sum_{i=1}^{N} \sum_{j=1}^{M} p_{ij} \log p_{i|j} \tag{3}$$

where $p_{i|j} = P(X = s_i|Y = t_j)$ is the conditional probability.

The *mutual information* between two discrete random variables X and Y is defined as

$$I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) \tag{4}$$

which can be written as

$$\begin{aligned} I(X;Y) &= \sum_{x \in S_X, y \in T_Y} P(X = x, Y = y) \log \frac{P(X = x, Y = y)}{P(X = x)P(Y = y)} = \sum_{i=1}^{N} \sum_{j=1}^{M} p_{ij} \log \frac{p_{ij}}{p_i q_j} \\ &= H(X) + H(Y) - H(X,Y) \end{aligned} \tag{5}$$

MI represents the amount of information that one random variable, here Y contains about the second random variable, here X and vice-versa. $I(X;Y)$ is the measure of shared information or dependence between X and Y. It is to be noted that $I(X;Y) \geq 0$ with equality if, and only if, X and Y are independent.

From the perspective of medical image registration, the random variables X and Y represent two images A and B respectively for which the marginal $[p_i, q_j]$ and joint distributions $[p_{ij}]$ of gray value pairs (i, j) of corresponding image grey values are considered. MI measures the dependence of the images and the maximal dependence is assumed to occur when the images are aligned. Registration is, therefore, achieved by finding the geometrical transformation that yields the highest MI value.

In this paper wherein our aim is to optimize the parameters involved in the registration process, we choose the percentage change in MI after registration w.r.t. the value before registration instead of using MI itself as the metric in driving the design of experiments which leads to parameter optimization. In other words, if $I_b(X;Y)$ and $I_a(X;Y)$ represent the MI between the discrete random variables X and Y before and after registration respectively, then metric, $M(X;Y)$ used for driving the design of experiments is

$$M(X;Y) = \frac{I_a(X;Y) - I_b(X;Y)}{I_b(X;Y)} \times 100 \tag{6}$$

The motivation behind using the proposed metric $M(X;Y)$ is that we need to choose the optimum values from a set of different combinations of parameters over which $M(X;Y)$ is positive. There might be certain combination of parameters, which result in more mis-registration thus making $I_a(X;Y) < I_b(X;Y)$ and hence have to be discarded.

3. THE VARIANCE DRIVERS

There are several parameters that are required to be set before registration while maximizing MI. Among these, some can be decided with the existing domain knowledge and others require some experimentation. The selection of an appropriate optimizer, interpolator, and transform form the initial step before optimizing registration parameters such as the number of levels of multi-resolution decomposition, the percentage of the volume that needs to be considered as spatial samples for the computation of pdf and entropy, the number of histogram bins into which these samples are distributed, and finally some application specific settings that can improve the metric.

Choice of the transform represents the mapping of points from one image space to the other and can be application specific. For example, neurological images are usually well registered with rigid or at most an affine transformation. On the other hand, registration of images from the thoracic region will require a deformable registration to take into account motion-artifacts such as breathing. With increase in the parameters for registration, there is an added penalty in the form of time and implementation.

In our experiments, we considered the problem of CT-PET fusion. The experiment was conducted on 5 subjects, who have undergone CT and PET (with both transmission and emission) examinations of thorax on LightSpeed QXi/Plus (GE Healthcare, Milwaukee, USA) multi-detector CT scanner (Matrix Size 512×512, DFOV 48–63 cm, Slice Thickness 2.5–5.0 mm) and PET (transmission and emission) data-sets (Matrix Size 128×128, DFOV 55–66 cm, Voxel Thickness 4.25-5.14 mm). For each of the subjects, CT and PET data-sets need to be registered for which we used a 12-parameter affine transformation along with a tri-linear interpolator and a regular step descent gradient optimizer while maximizing MI. These drivers comprise of both discrete and continuous factors of which the significant ones are chosen from the cause and effect diagram as shown in Figure 1.

The set of parameters that are chosen to find their optimum setting for the maximization of MI are 1) number of levels of multi-resolution (l) 2) percentage number of spatial samples (p_s) in the computation of pdf and hence the entropy 3) transmission weighting factor (α) to generate a composite PET image which is the convex combination of transmission and emission images and 4) CT grid size (g), which is essentially the size to which CT data has to be down-sampled prior to registration to improve the speed and memory utilization without compromising on the quality of registration. Though, the number of histogram bins (n_b) is also an important parameter, we chose it to be 50, since the metric value was relatively insensitive to changes in n_b

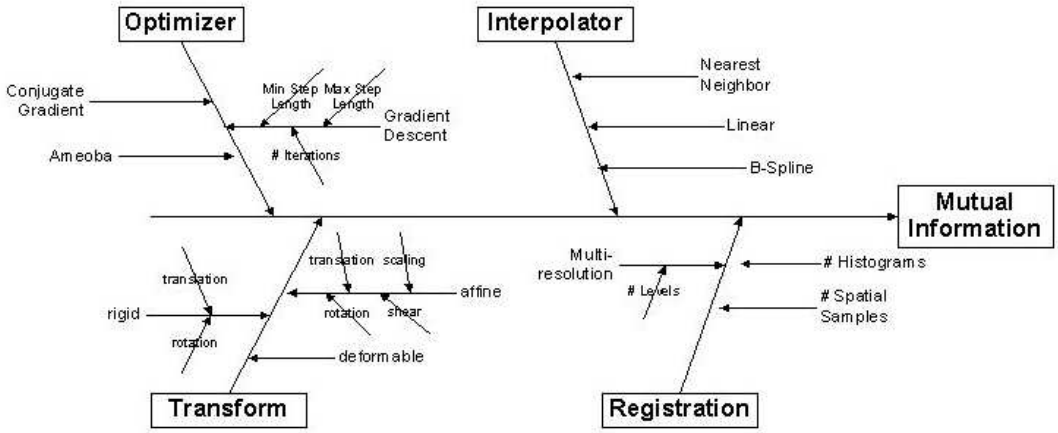


Figure 1. Cause and Effect Diagram.

beyond a particular setting. Depending on the application, some of the above parameters may not be significant and different set of parameters may have to be considered.

3.1. Transmission Weighting Factor,[7] α

Though the given PET images comprise of both transmission and emission images, it is customary to use only the transmission image to drive registration with the CT image. Such an approach appears intuitive because the transmission image is considered a poor man's CT and provides low resolution anatomical information unlike the emission image. In addition, the emission image is considered to be hardware registered to the transmission image. But the fact that motion can occur between the transmission and emission scans is overlooked. Also, respiratory motion is a key source of mis-registration in the thorax. To reduce the dependance on the highly variable PET image characteristics, a weighted combination of the two was first implemented by Slomka *et al.*[7]

It has been experimentally observed that by considering the moving image as a convex combination of transmission (TR) and emission (EM) images, higher value of MI can be achieved than that is achievable by using PET-TR alone. This is especially so in cases where there are higher number of pathologically significant hot-spots in PET-EM which then complement the PET-TR in adding more "information content" to the moving image. This factor, α, $0 \leq \alpha \leq 1$ is a continuous weighting factor that combines the transmission and emission images. Thus the moving image is defined as

$$\text{Moving Image} = \alpha \times \text{PET-TR} + (1 - \alpha) \times \text{PET-EM} \tag{7}$$

The basic hypothesis at this stage is that there is a point of optimum weighting factor for emission at which the increase in the average information content of the moving image due to PET-EM is just balanced by the loss of information content due to reduced contribution from PET-TR. After this point any gain from PET-EM is overshadowed by the increased loss from reduction in contribution from PET-TR to the overall information content.

3.2. CT Grid Size, g

Since CT (Matrix Size 512×512, DFOV 48–63 cm, Slice Thickness 2.5–5.0 mm) and PET (transmission and emission) data-sets (Matrix Size 128×128, DFOV 55–66 cm, Voxel Thickness 4.25–5.14 mm) have a mismatch in resolution, this led to the observation that the resolution quality of the registration can only be as good as the lower resolution of the two images. Down-sampling the CT image to the PET image resolution prior to the registration improves both speed (time to converge) and memory utilization without compromising on the quality of the registration.

α	l	g	p_s
0.25	1	128	0.5
0.50	2	64	1.0
0.75	3	32	1.5
1.0	4	–	2.0

Table 1. Experimental design space for (α, l, g, p_s).

3.3. Number of Levels of Muti-resolution Decomposition, l

In a multi-resolution approach, the input images are registered starting from a coarse scale and moving to finer scales. Each finer level receives its initialization parameters from the previous coarse level. Such an approach is fast due to the initialization process, more accurate by eliminating local optima at coarser levels and in turn robust[3]. In our experiment, we have varied l from 1 through 4. This means that the coarsest scale ($l = 4$) for an $X \times Y \times Z$ image would be of size $\frac{X \times Y \times Z}{8 \times 8 \times 8}$.

3.4. Percentage of Spatial Samples, p_s

p_s implies the number of random samples that are selected as a percentage of the total voxels in the volume. These samples are used to calculate the entropies and their selection is critical for both speed and statistical significance.

4. DESIGN OPTIMIZATION & RESULTS

Each of the aforementioned drivers, when varied through a range that is chosen out of domain knowledge, produce different $M(X;Y)$. Table 1 shows the different values that are being considered for each of the drivers thus creating a design space of $C_n = 192$ combinations per data-set of which we need to choose the optimum one based on the value of $M(X;Y)$. The attempt here is to come up with a regression equation of the form

$$I_a(X;Y) = f(\alpha, l, g, p_s) \tag{8}$$

for a given pair of modalities. Algorithm 1 provides the framework (in pseudo code) to determine the optimal parameter setting for the registration of data-sets from the given modalities. For each of the parameter combinations (total of C_n), $I_b(X;Y)$ and $I_a(X;Y)$ are computed for the given data-sets. These parameter combinations are then down-sampled based on the positive percentage improvement in MI before and after registration (basically using $M(X;Y) > 0$). As per Algorithm 1, we generate Q and P which contain the down-sampled combinations and their respective $M(X;Y)$ values. Since P has to be normal to analyze the interaction among parameters,

Algorithm 1 Optimal parameter setting

for $i = 1$ to C_n **do**
  compute $I_b(X;Y)$, $I_a(X;Y)$ and $M(X;Y)$
end for
$P = \left\{M(X;Y) : M(X;Y)|_{(\alpha,l,g,p_s)} > 0\right\}$
$Q = \left\{(\alpha, l, g, p_s) : M(X;Y)|_{(\alpha,l,g,p_s)} > 0\right\}$
$x = \underline{normality}(P)$ /* $x \in \{0, 1\}$ */
while $x = 0$ **do**
  $P = P - \{\text{outliers of } P\}$
  $x = \underline{normality}(P)$
end while
$Q = \{(\alpha, l, g, p_s) : M(X;Y)|_{(\alpha,l,g,p_s)} = y, \forall y \in P\}$
$R = I_a(X;Y)(Q)$
conduct *DOE (interaction analysis)* of (α, l, g, p_s) and *regression analysis* using Q and R
choose $(\alpha, l, g, p_s)_{opt}$

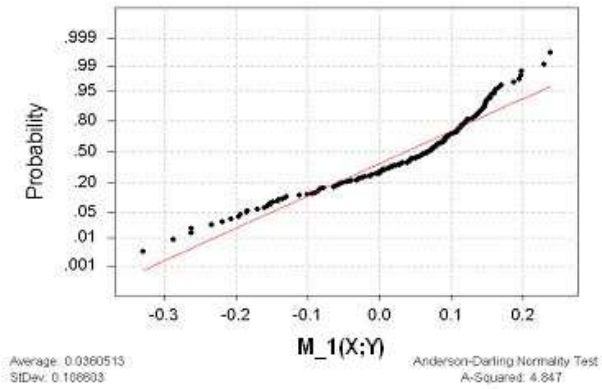


Figure 2. Normal probability plot of $M(X;Y)$ for 192 combinations of (α, l, g, p_s).

α	l	g	p_s
0.25	1		
0.50	2	128	1.0
0.75	3		1.5
1.0	4		2.0

Table 2. Experimental design space for (α, l, g, p_s) after eliminating $p_s < 1.0$ and $g < 128$.

we test the normality of P where one may use *Anderson-Darling* normality test. This is indicated in the pseudo code as $normality(P)$. We prune the set P by removing any outliers and repeat the normality test until we get a normal data. Finally, we choose only those combinations of (α, l, g, p_s) and corresponding $I_a(X;Y)$ values to conduct *DOE* and *regression analysis* for which the normality of $M(X;Y)$ is established. Using these combinations, we study the interaction among parameters and choose the optimum parameter setting, $(\alpha, l, g, p_s)_{opt}$.

In our experiments, to start with, $I_a(X;Y)$ computed for complete C_n combinations is not normal and we defined $M(X;Y)$ as a measure to down-sample the data so as to remove the outliers and establish the normality of data. Figure 2 shows the normality plot (based on *Anderson-Darling* test) of $M(X;Y)$ computed for 192 combinations of (α, l, g, p_s) for one data-set (say $\mathcal{D}$) of CT-PET (out of 5 data-sets) and it can be seen that the distribution of $M(X;Y)$ is not normal. If we carefully observe Figure 2, it is clear that the data can be modelled as a mixture of Gaussians, one for the set of $M(X;Y)$ for which it is negative and the other for which $M(X;Y)$ is positive. Since $M(X;Y) < 0$ is not acceptable as $I_a(X;Y) < I_b(X;Y)$, we choose only those combinations of (α, l, g, p_s) for which $M(X;Y) > 0$ so that the outlier combinations can be culled out. We observed that negative values for $M(X;Y)$ occur when the number of samples selected from the volume are too few for statistical significance and also when the CT image was down-sampled below the PET image resolution. A worst case example would be down-sampling a $512 \times 512 \times n_z$ CT image to $128 \times 128 \times n'_z$ (n_z and n'_z are the number of slices along z-direction for CT and PET images respectively) and then further multi-scaling it by 4 levels would bring it down to $16 \times 16 \times n_z$. From this volume, if we were to select 1% of its voxels as samples, we end up with $2.56 \times n_z$ voxels which is clearly not sufficient to capture the statistics of the data and the statistics computed out of this data would be in error. Hence, we revisit the dynamic ranges and set them as in Table 2.

Figure 3 shows that $M(X;Y)$ is normal for the new parameter settings. Figure 3(a) and Figure 3(b) show the normality plots of $M(X;Y)$ and $I_a(X;Y)$ for the data-set $\mathcal{D}$ in the down-sampled design space and it is obvious to see that the data is normal. The experiments were carried out on all 5 CT-PET cases but we show the normality plots, main-effects and interaction plots for only $\mathcal{D}$ as a proof of concept. Though the down-sampling of combinations based on $M(X;Y)$ gave normal data in our experiments, it might also not remove all the outliers in other cases for which we need different measures to remove the outliers and make the data ultimately normal before proceeding to interaction and regression analysis.

Having found the valid combinations of (α, l, g, p_s) along with their corresponding $I_a(X;Y)$ values, we conduct

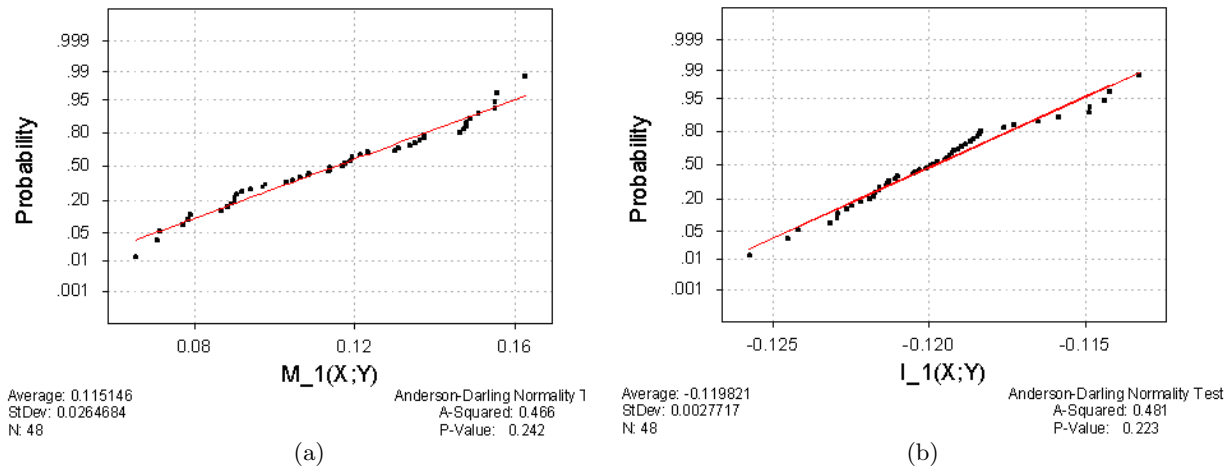


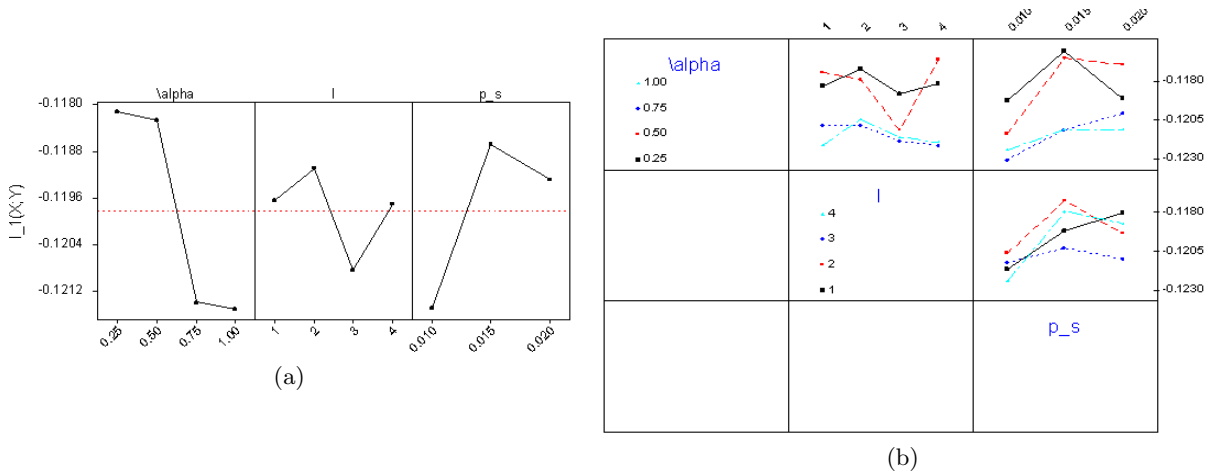
Figure 3. Normal probability plot of (a) $M(X;Y)$ and (b) $I_a(X;Y)$ for down-sampled 48 combinations of (α, l, g, p_s), shown in Table 2.

main-effects and interaction studies of the parameters and find the optimum parameter combination that provides improved registration performance across the data-sets. Figure 4(a) and Figure 4(b) show the main-effect and interaction plots for the parameters from which we can find the optimum parameter setting that maximizes $I_a(X;Y)$. We would like to mention that in our experiments, instead of maximizing MI we used negative of MI as the cost function to be minimized. Hence the values of $I_a(X;Y)$ shown in Figure 3 and Figure 4 are negative. Since the main-effects plot does "one factor at a time" analysis, it is clear from Figure 4(a) that $(\alpha, l, g, p_s) = (1, 3, 128, 1)$ is the optimum setting. If the first-order interaction between the parameters are considered, it can be seen from Figure 4(b) that $(\alpha, l, g, p_s) = (0.75, 4, 128, 1)$ is the optimal setting. Since the interaction plots consider the interaction amongst parameters, we prefer the operating point obtained from interaction plots than main-effects plot and consider that as the optimum operating point and so $(\alpha, l, g, p_s)_{opt} = (0.75, 4, 128, 1)$. The fact that the main-effects and interaction analysis give different optimal settings indicate that parameters are not independent as considered in Zhu *et al.*[1] but are dependent on one another and so interaction study is the right thing to do instead of "one factor at a time" analysis.

As our ultimate goal is generate a regression equation relating the variance drivers and the MI value for a given modality-pair, we conducted this analysis for all the five CT-PET data-sets. But we were not successful in generating a regression equation for CT-PET modality because of significant variations that are present across the data-sets. To find a regression for $I_a(X;Y)$, it is necessary to perform a similarity test such as the 2-sample t-test across data-sets. The 2-sample t-test performs a hypothesis test and computes a confidence interval of the difference between two population means when the population standard deviations are unknown. Comparing $I_a(X;Y)$ for different cases (for a 95% confidence interval) produced results that implied that the data is statistically dissimilar across cases. Using the operating point that is determined, Table 3 shows $I_b(X;Y)$, $I_a(X;Y)$ and $M(X;Y)$ for all the 5 data-sets and it can be seen that the optimum parameter combination shows significant improvement for 3 cases. This is impressive because we have used the optimum parameter setting of one of the data-sets for remaining of them and still get a good improvement in registration performance. This experiment can be repeated by using the optimum parameter setting of each of the data-sets for the remaining cases and study the improvement in registration performance. This framework can be generalized and improved to other modalities by choosing appropriate and adaptive variance drivers.

5. CONCLUSIONS AND FUTURE DIRECTIONS

We have presented a framework that could be used to determine optimum parameter configuration while performing image registration by maximizing mutual information. The framework was demonstrated on the registration of CT and PET data-sets and could be used in any multi-modality setting. Combination of transmission and emission images with $\alpha = 0.75$, CT-grid size (g) matched to the PET resolution with 4-levels of multi-resolution decomposition and percentage of samples chosen for entropy calculations being at least 1% provides significant



(a) (b)

Figure 4. (a) Main-effects plot (data-means for $I_a(X;Y)$) assumes that the parameters are independent and so the optimum operating point would be $(\alpha, l, g, p_s) = (1, 3, 128, 1)$. (b) Interaction plot (data-means for $I_a(X;Y)$) analyzes the dependency amongst parameters and provides the operating point to be $(\alpha, l, g, p_s) = (0.75, 4, 128, 1)$.

Case	$I_b(X;Y)$	$I_a(X;Y)$	$M(X;Y)$
1	-0.110	-0.126	14.62
2	-0.078	-0.086	10.20
3	-0.166	-0.215	29.30
4	-0.074	-0.077	4.42
5	-0.226	-0.232	2.62

Table 3. $M(X;Y)$ shows improvement for all 5 cases with the optimum parameter combination of $(\alpha, l, g, p_s)_{opt} = (0.75, 4, 128, 1)$. In our experiments, we minimized the negative of MI and so the values of $I_b(X;Y)$ and $I_a(X;Y)$ are negative.

improvement in $M(X;Y)$ for CT-PET registration. The final parameters chosen were verified on the test cases for positive $M(X;Y)$. Though, we were interested in obtaining a regression equation relating the variance drivers and MI for CT-PET modality, wide variation in the data-sets prevented us from achieving the same. But, an experiment such as this aids in narrowing down the parameter space to guarantee consistent and positive $M(X;Y)$. A reliability test such as this is essential before the algorithms are deployed for diagnosis. Generalization of this framework to other modalities by considering specific and adaptive variance drivers would be interesting.

REFERENCES

1. Y.-M. Zhu and S. M. Cochoff, "Influence of implementaion parameters on registration of MR and SPECT brain images by maximization of mutual information," *Journal of Nuclear Medicine* **43**(2), pp. 160–166, 2002.
2. G. Gopalakrishnan, T. Poston, N. Nagaraj, R. Mullick, and J. Knoplioch, "Implicit function-based phantoms for evaluation of registration algorithms," in *Proceedings of SPIE, Medical Imaging*, pp. 149–155, (San Diego, USA), 2005.
3. L. Ibáñez, W. Schroeder, L. Ng, and J. Cates, *The ITK Software Guide*, 2003.
4. F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Transactions on Medical Imaging* **16**, pp. 187–198, 1997.
5. P. Viola and I. W. M. Wells, "Alignment by maximization of mutual information," *International Journal of Computer Vision* **24**(2), pp. 137–154, 1997.
6. T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons Inc., 1991.
7. P. J. Slomka, D. Dey, C. Przetak, U. E. Aladl, and R. P. Baum, "Automated 3-dimensional registration of stand-alone ^{18}F-FDG whole-body PET with CT," *Journal of Nuclear Medicine* **44**(7), pp. 1156–1167, 2003.