2D/3D Registration with the CMA-ES Method

Ren Hui Gong[a], Purang Abolmaesumi*[a,b]
[a]School of Computing, [b]Department of Electrical and Computer Engineering,
Queen's University, Kingston, Ontario, Canada
*purang@cs.queensu.ca

ABSTRACT

In this paper, we propose a new method for 2D/3D registration and report its experimental results. The method employs the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) algorithm to search for an optimal transformation that aligns the 2D and 3D data. The similarity calculation is based on Digitally Reconstructed Radiographs (DRRs), which are dynamically generated from the 3D data using a hardware-accelerated technique - Adaptive Slice Geometry Texture Mapping (ASGTM). Three bone phantoms of different sizes and shapes were used to test our method: a long femur, a large pelvis, and a small scaphoid. A collection of experiments were performed to register CT to fluoroscope and DRRs of these phantoms using the proposed method and two prior work, i.e. our previously proposed Unscented Kalman Filter (UKF) based method and a commonly used simplex-based method. The experimental results showed that: 1) with slightly more computation overhead, the proposed method was significantly more robust to local minima than the simplex-based method; 2) while as robust as the UKF-based method in terms of capture range, the new method was not sensitive to the initial values of its exposed control parameters, and has also no special requirement about the cost function; 3) the proposed method was fast and consistently achieved the best accuracies in all compared methods.

Keywords: 2D/3D registration, CMA-ES, simplex, unscented Kalman filter, digitally reconstructed radiograph, adaptive slice geometry, texture mapping

1. INTRODUCTION

2D/3D registration is a fundamental task in computer assisted surgery (CAS). In such surgeries, in order to use pre-operative CT to guide the surgical procedure during the intervention, the CT must first be mapped to the physical patient in the operating room, and this can be done through registering the 3D CT to a set of intra-operative 2D fluoroscope images. Another important application is in computer assisted radiotherapy, where registration of CT to a few portal images is used to focus the harmful treatment beams on the lesion area thus minimizing the damage to the surrounding healthy tissues.

The goal of 2D/3D registration is to find a spatial transformation that transforms one data set (usually a 3D data set) from its local coordinate space to the other's (usually a 2D data set consisting of a set of 2D images) coordinate space so that the two data sets are aligned in terms of some similarity metric. A 2D/3D registration method generally involves determining three components: a transformation that spatially correlates the two data sets, a similarity metric that evaluates how good the two data sets are aligned under a particular transformation, and an optimization technique that iteratively searches for an optimal solution of the transformation.

A variety of methods have been proposed for 2D/3D registration [1], [5], [8]. Most of the methods have focused on defining an accurate and efficient similarity metric, and have relied on simple search algorithms such as simplex, gradient-descent, etc., to find the final solution. Early methods [8] have used geometry features, e.g., edges and surfaces, to define the similarity metric for obtaining acceptable computation speed. The main drawback of those methods is the need for an accurate feature extraction, where the errors in segmentation propagate through the registration process. Due to the fast increase in computation power in recent years, most current methods [1], [5] compute the similarity directly from image intensities to achieve better robustness and accuracy. This group of methods dynamically generates the simulated 2D data, called Digitally Reconstructed Radiographs (DRRs), from the 3D data and computes the similarity from a set of 2D image pairs. A variety of functions, including normalized correlation (NC), variance-weighted correlation (VWC), gradient correlation (GC), gradient difference (GD), pattern intensity (PI), mutual information (MI), etc., have been used to define the similarity between a 2D image and its corresponding DRR. To achieve interactive computation performance, hardware-accelerated techniques [3] are usually employed to speed-up the DRR-generation

Medical Imaging 2008: Visualization, Image-guided Procedures, and Modeling, edited by
Michael I. Miga, Kevin Robert Cleary, Proc. of SPIE Vol. 6918, 69181M, (2008)
1605-7422/08/$18 · doi: 10.1117/12.770331

process. Finally, some recent work reconstructs a 3D data from the 2D data to take advantage of the variety of existing 3D registration techniques. However, one limitation of this type of methods is that they need a large number of 2D images or the statistical information about the studied object for accurate 3D reconstruction.

While simple optimization techniques are known for their ease of use, they are sensitive to local minima. When used in 2D/3D registration, they work well only if a good initial guess of the solution can be found. The main reason is that, due to different dimensionalities and modalities involved in this registration problem, the 2D/3D similarity metrics are usually highly nonlinear and have rugged search landscape. In real applications, finding such an initial alignment usually involves using a user interface, which is a time-consuming task, or using known geometry objects in the field. To develop a more robust approach, in our previous work an Unscented Kalman Filter (UKF) based method was proposed and the UKF was used as the optimization strategy [1], [2]. The method demonstrated significant improvement in capture range compared to a commonly used simplex-based method. It also provided a possibility to estimate the registration errors using a closed-form solution [7] after the registration is finalized. However, the method is only suitable for the situations that the knowledge about the system noises is easily obtained and the similarity metric has a known target value.

In this work, we propose a fast and more general method that uses the Covariance Matrix Adaption Evolution Strategy (CMA-ES) technique [4] as the optimization strategy to achieve high robustness and better usability. In Section 2, we provide the details of the algorithm. In Section 3, we validate the proposed method, and compare it with two prior work, i.e. our previous UKF-based method and a simplex-based method. Finally, the conclusions will be given in Section 4.

2. METHOD

2.1 Algorithm overview

The proposed method determines the three components of the aforementioned general scheme of 2D/3D registration. Fig. 1. shows the overall diagram of the method as well as the interactions between its components. The inputs are the two data being registered and an initial guess for the registration transformation. Then the optimizer, i.e. CMA-ES, iteratively refines the transformation according to the similarity between the 2D data and the dynamically generated DRRs. Without losing generality, in the subsequent discussions we assume orthopaedic surgery as the common application of CT to fluoroscope registration in CAS. In this case, the 3D data is the pre-operative CT, quantized in the coordinate space of the CT machine, and the 2D data is a series of intra-operative fluoroscopic images, captured from different orientations in the fluoroscope coordinate space. We briefly describe the transformation and similarity metric in the paragraphs that follow. The details about searching for an optimal transformation with CMA-ES will be given in Section 2.2.

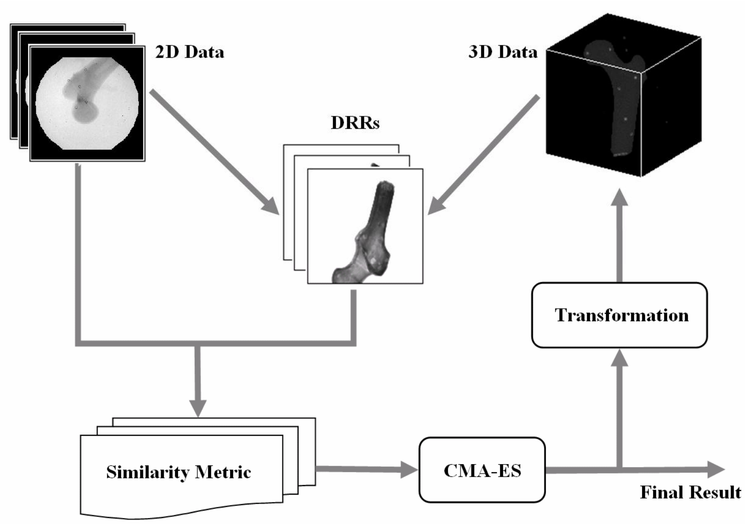


Fig. 1. CMA-ES based 2D/3D registration method.

The transformation takes the 3D data from its local coordinate space to the 2D data's coordinate space. Depending on the application, the transformation can be of any type including rigid, similarity, affine, non-rigid, or a combination of them. In the context of CT to fluoroscope registration, this is usually a 3D rigid transformation consisting of rotational and translational components. The translation has 3 degrees-of-freedom (DOFs), while the rotation can have various representations with different DOFs, e.g., Euler angles with 3 DOFs, quaternion with 4 DOFs, angle-axis with 4 DOFs, versor with 3 DOFs, and so on. Our method does not tend to a particular representation. For compactness and intuitiveness, the form of Euler angles was selected in this work.

We have adopted the intensity-based approach in the proposed method to take advantage of its robustness and high accuracy. For each fluoroscope, one DRR is generated from the CT using the current transformation and the fluoroscope's image settings, and then a similarity is computed between each fluoroscope and DRR pair. The final similarity between the 2D and 3D data is formulated as a linear combination of the similarities between each pair. All current similarity metrics, e.g. NC, VWC, GC, GD, MI, PI, etc., can be used with our method, and the selection usually depends on the image quality or content. Because DRR-generation is the dominant operation during the registration process, a hardware-accelerated technique, named Adaptive Slice Geometry Texture Mapping (ASGTM) [3], is used to speed-up the task. ASGTM is an improvement of the commonly used 3D texture mapping technique. It excludes the non-interesting voxels of the 3D data from rendering and performs hardware-supported intensity transfer to further accelerate the DRR generation process.

2.2 Optimization with CMA-ES

The main contribution of this work is to use the CMA-ES optimization technique in 2D/3D registration for improved robustness and better usability. CMA-ES [4] is a sampling-based search algorithm known for robust and efficient operation in a rugged search landscape. The method requires no calculation of derivatives; instead the learning is done through taking random samples around the current solution according to a multivariate normal distribution. In each iteration of the optimization process, the solution is refined by sampling, selection and recombination, and the search distribution is adaptively deformed according to both new information from the selected samples and the information from previous steps. Fig. 2 shows the key steps of the CMA-ES algorithm.

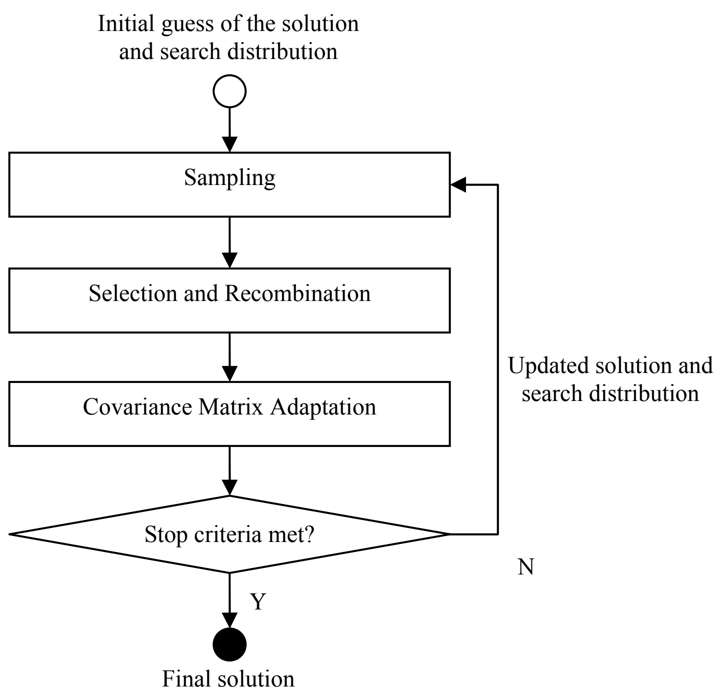


Fig. 2. The CMA-ES algorithm.

In the above figure, the initial guess of the solution and search distribution are provided by the user, which are the initial position of the 3D data and its uncertainty. The search distribution is represented using a scalar σ, indicating the distribution size, and a covariance matrix C, indicating the distribution shape. Initially, only σ is specified and the search distribution has a spherical shape with an isotropic standard deviation in all directions. The population size of each sampling (λ) is determined by the dimension of solution parameters (n). The selection and recombination are based on function values of individual samples, and are controlled by the parameters μ and $\{w_i\}_{i=1,...,\mu}$. The update of search distribution, or covariance matrix adaptation, is based on three sources: the search distribution of the previous iteration, the accumulated evolution path of the solution from the first iteration till the current iteration, and the distribution of the selected samples at the current iteration. Each source is assigned a weight and the assignment is controlled by two parameters c_c and c_σ. Except for the initial guess, all control parameters can be automatically determined and have been appropriately suggested [4]. The stopping criteria are user-defined. The commonly used ones are the maximum number of iterations, the tolerance of function update (w.r.t. similarity value), and the tolerance of parameter update (w.r.t. the solution parameters). In summary, our CMA-ES based 2D/3D registration method works as follows:

Inputs: 2D data, 3D data, initial transformation T_0, initial search distribution size σ_0

Output: final transformation T

Algorithm:

1. Initialize the solution T to be T_0, the search distribution N to be $N(T, \sigma_0 I)$, and the evolution path p to be null; compute the selection size μ, the recombination weights $\{w_i\}_{i=1,...,\mu}$, and parameters c_c and c_σ.
2. Until stopping criteria are met, do the following:
    a. Generate a population of samples $\{T_i\}_{i=1,...,\lambda}$, according to distribution N;
    b. For each sample T_i, transform the 3D data, generate DRRs, and compute the similarity metric;
    c. Select μ best samples according to similarity values;
    d. Update T by recombining the μ selected samples with the weights $\{w_i\}_{i=1,...,\mu}$;
    e. Update N by linearly combining the following three components with the parameters c_c and c_σ: the previous N, the covariance of the μ selected samples, and the covariance of p;
    f. Update p (see [4] for more details).

3. EXPERIMENTAL RESULTS

We used three bone phantoms of different sizes and shapes, including a long femur, a large pelvis, and a small scaphoid, to evaluate the proposed method. Four pairs of 2D and 3D data were acquired or synthesized from the phantoms. The 3D data were CTs, captured using a GE LightSpeed Plus machine. The 2D data were simulated fluoroscopic images and real fluoroscopic data. The simulated fluoroscopes were generated from CTs along coordinate axes using the ASGTM technique. The real fluoroscopes were acquired using an OEC-9800 fluoroscopy device. Table 1 lists the specifications of the data used in this study.

Three types of experiments were conducted. First, registration of CT to simulated fluoroscopes was performed for each phantom using the proposed method and two prior work, i.e. our previous UKF-based method and a commonly used simplex-based method. Next, registration of CT to real fluoroscopes was performed for the pelvis phantom. Finally, additional experiments for studying the impact of the initial search distribution of the CMS-ES algorithm were conducted. All experiments were done on a Dell OptiPlex GX270 computer equipped with 2GB RAM and an ATI Radeon X800 (256MB video RAM) graphics card.

Mean target registration error (mTRE), capture range, accuracy and computation time were used to evaluate each method. mTRE was used to measure the initial and final misalignments, and was calculated using the CT surface points of the corresponding bone. Capture range measures the robustness of a method under a collection of experiments. It was chosen as the range of initial mTRE that 95% of registrations would success. A registration was defined *successful* if the final mTRE ≤2mm for experiments using simulated fluoroscopes, and ≤4mm for real fluoroscopes. The accuracy was measured using the mean and standard deviation of the final mTREs of the successful registrations. The computation time was measured as the mean and standard deviation of the time required achieving successful registration.

3.1 Registration of CT to simulated fluoroscope

For each phantom three simulated fluoroscope images were generated from CT. The CT data was placed at the origin, and the simulated fluoroscopes were generated along coordinate axes with focal length of 920mm and origin being at the half focal length. Thus, in this group of experiments the gold standards were known and have the values of all zeros. 100 experiments with random initial CT positions were conducted for each phantom and each method. The initial CT positions were obtained by applying small perturbations to the gold standard. Table 2 lists the magnitude of perturbations for each phantom. NC was used as the similarity metric. Table 3 shows the initial and final mTREs. Table 4 shows the capture ranges, accuracies, and computation time.

Table 1. Data specifications.

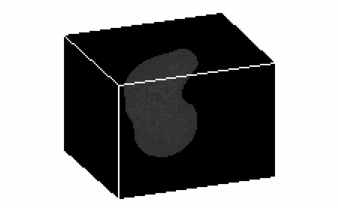
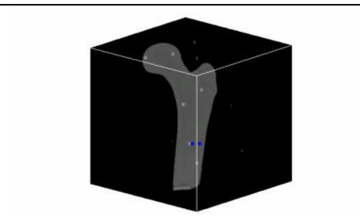
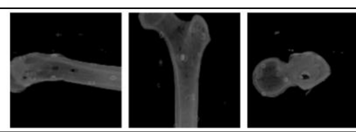
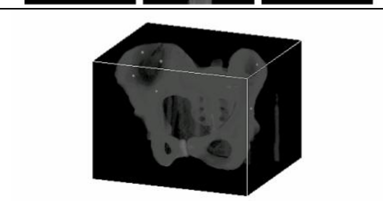
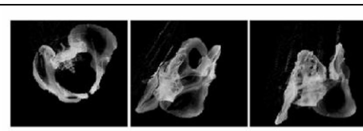
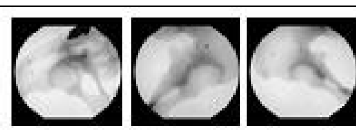
Phantom		**Image**	**Resolution (pixels)**	**Pixel Size (mm^3)**
Scaphoid	**CT**		256 x 256 x 64	0.375 x 0.375 x 0.525
	Simulated Fluoroscope		256 x 256	0.836 x 0.836
Femur	**CT**		256 x 256 x 256	1.176 x 1.176 x 0.766
	Simulated Fluoroscope		256 x 256	0.836 x 0.836
Pelvis	**CT**		256 x 256 x 128	0.625 x 0.625 x 1.445
	Simulated Fluoroscope		256 x 256	0.836 x 0.836
	Real Fluoroscope		256 x 256	0.836 x 0.836

Table 2. Perturbations used to generate random initial CT positions for CT to simulated fluoroscope registrations. The perturbations were made around the six components (3 rotational, 3 translational) of the gold standard.

	Rotational Components	Translational Components
Femur	±20°	±20mm
Pelvis	±50°	±20mm
Scaphoid	±30°	±10mm

Table 3. Experimental results of CT to simulated fluoroscope registrations: initial mTREs vs final mTREs (unit: mm).

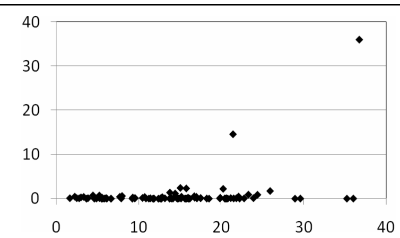
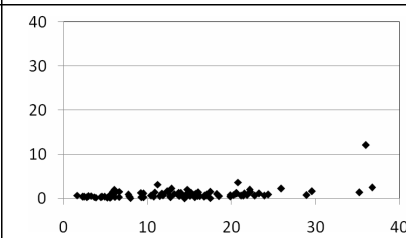
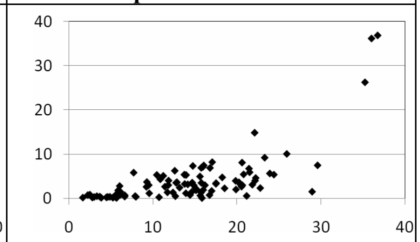
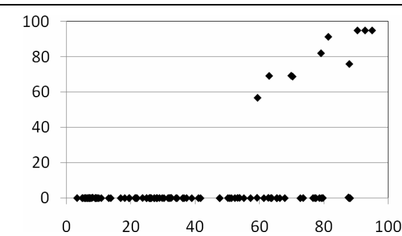
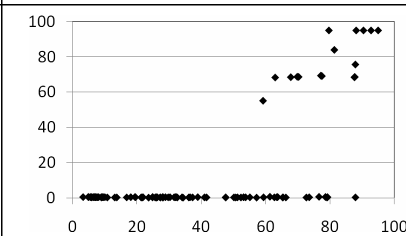
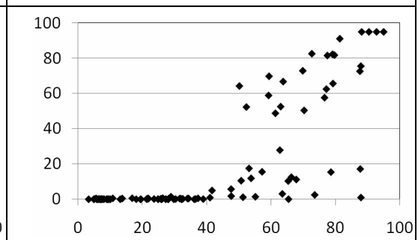
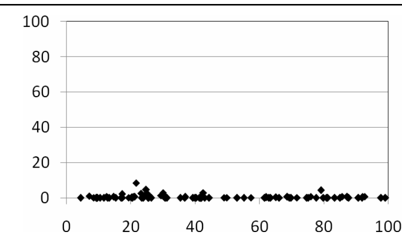
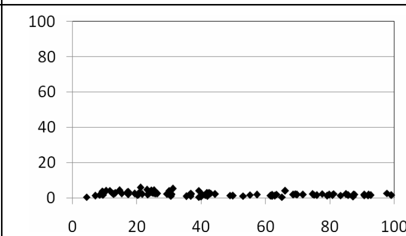
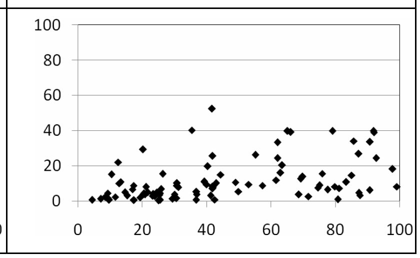
	CMA-ES-based Method	**UKF-based Method**	**Simplex-based Method**
Scaphoid			
Pelvis			
Femur			

Table 4. Experimental results of CT to simulated fluoroscope registrations: capture range (mm), accuracy (mm), and computation time (s).

Phantom		**CMA-ES**	**UKF**	**Simplex**
Scaphoid	**Capture Range**	>40	>40	9
	Accuracy	0.26±0.48	0.87±0.54	1.24±0.99
	Computation Time	147±53	187±16	95±27
Pelvis	**Capture Range**	80	72	50
	Accuracy	0.07±0.07	0.40±0.10	0.30±0.47
	Computation Time	94±26	154±38	86±22
Femur	**Capture Range**	>100	>100	10
	Accuracy	0.42±0.63	1.87±0.63	1.31±0.82
	Computation Time	99±27	124±14	65±14

From Tables 3 and 4, we have the following observations: 1) the proposed method achieved similar or better capture ranges as the UKF-based method for all testing phantoms, and both methods were significantly more robust than the simplex-based method in terms of capture range; 2) the CMA-ES based method consistently achieved the best accuracy; 3) the CMA-ES based method took slightly longer time than the simplex-based method to converge, but on average the difference was within one minute for a single registration.

3.2 Registration of CT to real fluoroscope

To examine the methods' performance under simulated surgical environment, in this test the real fluoroscopes of the pelvis phantom were used. Three fluoroscope images were acquired from arbitrary viewing directions that were apart about 45 degrees each other, and the pose information were reported by the tracking camera. Four embedded fiducials, visible in CT and tracked during fluoroscope acquisition, were used to obtain the gold standard. Similar to the CT to simulated fluoroscope registration experiments, 100 experiments with random initial CT positions around the gold standard were performed for each of the three methods. The perturbations were ±15° for rotational components and ±20mm for translational components. VWC was used as the similarity metric. The results are shown in Tables 5 and 6.

Table 5. Experimental results of CT to real fluoroscope registrations for pelvis: initial mTREs vs final mTREs (unit: mm).

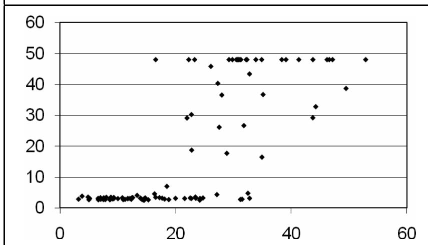
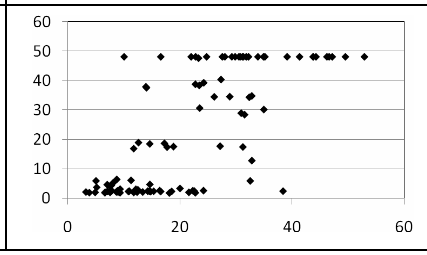
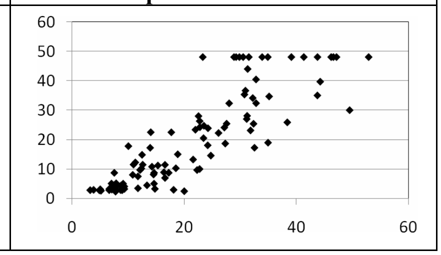
CMA-ES-based Method	UKF-based Method	Simplex-based Method
		

Table 6. Experimental results of CT to real fluoroscope registrations for pelvis: capture range (mm), accuracy (mm), and computation time (s).

	CMA-ES-based Method	UKF-based Method	Simplex-based Method
Capture Range	22	10	9
Accuracy	3.19±0.44	2.56±0.74	3.24±0.65
Computation Time	114±40	156±40	90±8

Obviously, the CMA-ES based method achieved the largest capture range. However, the UKF-based method achieved the best accuracy but did not show much improvement in capture range. This can be explained by one important property of the UKF-based method, i.e. it requires a good understanding about the error sources in the system to work robustly and efficiently. In CT to fluoroscope registration, a variety of sources, e.g., CT acquisition, fluoroscope acquisition, DRR generation, outliers in fluoroscope, etc., have caused the generated DRRs not exactly matching the corresponding fluoroscope images. Finding the statistics about the combined errors is usually not an easy task. In this study, they were determined by trial and error, and have not been carefully chosen. In our previous work [1], it was demonstrated that the UKF-based method was able to achieve significant improvements in capture range and computation time if such knowledge can be accurately obtained.

3.3 The impact of initial search distribution

While most control parameters of the CMA-ES algorithm have been suggested before [4], the user has to provide an initial search distribution in the form of σ. The parameter indicates the uncertainty about the user-supplied initial transformation, and the value of 1.0 has been recommended. Here, we analyze the sensitivity of the algorithm to this parameter. The CT and simulated fluoroscope of the pelvis phantom and three different values of σ, i.e. 0.5, 1.0, and 2.0, were used in this testing. Similar to the previous cases, 100 experiments were performed for each value of σ. The results are shown in Table 7 and Table 8.

The tables demonstrate that the recommended value of 1.0 gives the best balance between capture range, accuracy and computation time; however, the differences caused by the parameter were not significant. This observation was anticipated because the CMA-ES algorithm is able to adaptively change its search distribution according to the local search landscape.

4. CONCLUSIONS

In this paper, we presented a new 2D/3D registration method that takes advantage of the CMA-ES searching algorithm to achieve improved robustness, fast computation speed and better usability. From the experimental results, we have the following conclusions:

1. The proposed method is able to achieve highly accurate results and is significantly more robust w.r.t. local minima than the simplex-based method. It is a fast method as most registrations can be finished in 1-2 minutes.
2. The UKF-based method is able to achieve the same capture range as the CMA-ES based method if a good understanding about the system errors can be obtained. However, the CMA-ES based method is a more general solution because most of its control parameters can be automatically determined and it is not sensitive to the only exposed parameter.

In the future work, we would like to use the CMA-ES and UKF in combination to provide a complete and practical solution for 2D/3D registration: the CMA-ES can be used first to find an accurate solution, and then the UKF can be used to estimate the registration errors.

Table 7. Experimental results for different initial distributions of the CMA-ES based method: initial mTREs vs final mTREs (unit: mm). The pelvis phantom was used as the testing data.

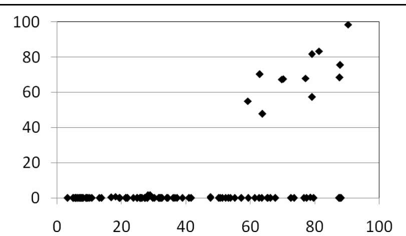
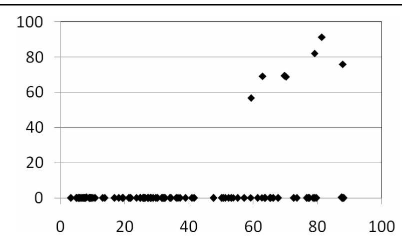
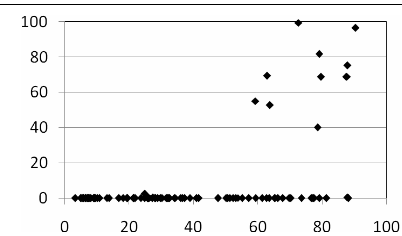
σ=0.5	**σ=1.0**	**σ=2.0**
		

Table 8. Experimental results for different initial distributions of the CMA-ES based method: capture range (mm), accuracy (mm) and computation time (s). The pelvis phantom was used as the testing data.

	σ=0.5	**σ=1.0**	**σ=2.0**
Capture Range	75	80	78
Accuracy	0.12±0.27	0.07±0.07	0.10±0.28
Computation Time	119±33	94±26	115±27

REFERENCES

[1] R. H. Gong, P. Abolmaesumi, J. Stewart, "A robust technique for 2D-3D registration", IEEE Engineering in Medicine and Biology (EMBC), pp. 1433-1436, NY, Aug 2006.

[2] R. H. Gong, J. Stewart, P. Abolmaesumi, "A new method for CT to fluoroscopes registration based on unscented Kalman filter", Medical Image Computing and Computer Assisted Intervention (MICCAI), LNCS 4191, pp. 891-898, Copenhagen, Oct 2006.

[3] C. Bethune, "Adaptive Slice Geometry for Hardware Assisted Volume Rendering", MSc thesis, Queen's University, 2003.

[4] N. Hansen, "The CMA Evolution Strategy: A Tutorial", http://www.bionik.tu-berlin.de/user/niko/ cmatutorial.pdf, accessed Aug 2007.

[5] G. P. Penney, J. Weese, J. A. Little, P. Desmedt, D. L. Hill, D. J. Hawkes, "A comparison of similarity measures for use in 2-D/3-D medical image registration", IEEE Transactions on Medical Imaging, vol. 17, no. 4, pp. 586-595, 1998.

[6] E. A. Wan, R. van der Merwe, “The unscented Kalman filter for nonlinear estimation”, Proc. of Symposium on Adaptive Systems for Signal Processing, Communication and Control, pp. 153-158, 2000.

[7] M. Moghari, P. Abolmaesumi, "A high-order solution for the distribution of the target registration error in rigid-body point-based registration", Medical Image Computing and Computer Assisted Intervention (MICCAI), LNCS 4191, pp. 603-700, Copenhagen, Oct 2006.

[8] S. Lavallee, R. Szeliski, L. Brunie, “Matching 3-D Smooth Surface with their 2-D Projections using 3-D Distance Maps”, SPIE Vol. 1570, Geometric Methods in Computer Vision, pp. 217-238, 1991.