Projection registration of X-ray image and CT image

Zhang Hui[1] Luo Limin[1] Shu Huazhong[1] Li Songyi[1] Pascal Haigron[2]

([1] Department of Biological Science and Medical Engineering, Southeast University, Nanjing 210096, China)

([2] Laboratory of Treatment of Signal and Image, University of Rennes I, France 35042)

Abstract: A methodology for alignment of an X-ray image and a CT image, based on the Chamfer 3-4 distance transform and simulated annealing optimization algorithm is presented. Firstly, an initial transformation matrix is constructed. For the convenience of computing, geometric models of the X-ray device to reconstruct the calibration matrix are used. Then, by defining the distance between the 3-D projective and the 2-D object image, we optimize this distance matching problem, using the simulated annealing algorithm. This method is also integrated into medical intra-operation, dealing with the data set acquired from 3-D image workstation and active navigation.

Key words: image registration; calibration matrix; image segmentation; distance transformation; simulated annealing

Clinical diagnosis, as well as planning and evaluation of therapy, is often supported by several imaging modalities. Different modalities usually provide complementary information. Single photon emission computed tomography (SPECT), positron emission tomography (PET), and magnetic resonance spectroscopy (MRS) provide functional information, but delineate anatomy poorly, whereas magnetic resonance imaging (MRI), ultrasound, and functional X-ray imaging, including computed tomography (CT), depict aspects of anatomy, but provide little information. Registration is a useful way to utilize this information as a whole.

Medical image registration is a rapidly evolving field of interest, with its own specific operational conditions[1]. One of the most difficult tasks is to register, if possible in real time, a video or X-ray image of the patient (intra-operative image) with a pre-operative image (MRI or CT). Basically, one must find the 3-D – 2-D projective transformation (composition of a rigid displacement and a perspective projection) which maps a 3-D object onto a 2-D image of this object, unknowing the relative positions of the 3-D object and the 2-D sensor. This problem can be solved with artificial markers visible in both images, but this produces unacceptable constraints (e.g. pre-operative images must be taken on the same day as the intervention; stereo tactic frames are very painful and can prevent free access by the surgeon, etc.).

Recently, computer vision techniques have been proposed to solve this 3-D – 2-D registration problem without artificial markers. Grimson[2] used an intermediate laser range finder, which provided a 3-D description of the patient's surface. This surface is then matched against the surface of the segmented corresponding surface in the medical volume image. Colchester[3] and Kanade[4] also developed reconstruction/rigid registration frameworks. The approach adopted by Schweikard[5] is a 2-D correlation scheme between the radiography and pre-computed simulated radiographies of an MRI image. Finally, Lavallee[6] used the occluding contours of the 3-D object in the 2-D image to perform the 3-D – 2-D registration task.

The method we propose in this paper aims at finding the geometric transformation matrix between the 3-D object structure and the 2-D radiography image. Firstly, we construct an initial transformation matrix. For the convenience of computing, we use geometric models of the X-ray device to reconstruct the calibration matrix. Then, by defining the distance between the 3-D projective and the 2-D object image, we optimize this distance matching problem. This paper focuses on the registration of X-ray image and information of the CT image; the intervention process is used in a computer aided survey.

The whole approach is divided into four steps:

1) Construct a calibration model for initial approximated parameters;

2) Automatically segment object structure from X-ray image;

Received 2002-08-01.

Foundation items: The National Natural Science Foundation of China (60272045) and the Key Project of Ministry of Education of China.

Biographies: Zhang Hui (1976—), male, graduate; Luo Limin (corresponding author), male, doctor, professor, luo.list@seu.edu.cn.

3) Use an optimization algorithm to compute the precise projection matrix;

4) Calculate the 3-D coordinate from the interactive selected point of the X-ray image, using retro-projection.

1 Calibration

The acquisitions of X-ray image are carried on a Philips Integers 2000 System controlled from a remote desk. This digital radiography device is made up mainly of an arch supporting the source-detector set (See Fig.1). Digital images under different incidences are taken by rotating the arch around the table where the patient lies.

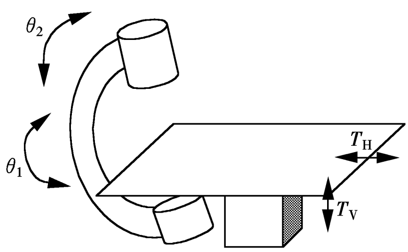


Fig.1 Digital radiography system

To recover 3-D information from several views, the imaging device has to be modeled. The geometric calibration aims at the estimation of the transformation between the 3-D volume coordinate system $R_{3\text{-D}}$ and the 2-D image frame $R_{2\text{-D}}$, included several parameters.

These parameters can be divided into two categories:

1) Extrinsic parameters, which modify the global position of the source-detector set regarding the 3-D coordinate system. Firstly, we must calculate the rotation angle θ_1 around the patient and the tilt angle θ_2 of the source detector set. To provide full function of rotation of the device, we must take into account rotation around another axis. Secondly, T_H and T_V are adjusted on the horizontal and vertical translations of the patient table relative to the source-detector set, respectively.

2) Intrinsic parameters, which affect the scale factors and resolution of image such as translation between pixels, physical distance and the source detector distance (SDD).

Some intrinsic parameters are defined as follows:

x, y, and z indicate coordinates of CT volume (in pixel);

X, Y, and Z indicate coordinates of CT volume (in physical distance);

U and V indicate coordinates of X-ray image (in physical distance);

u and v indicate coordinates of X-ray image (in pixel).

At first, we transform the object coordinate system to the world coordinate system. The 3-D coordinate of CT volume can be generated by firstly rotating the volume, then translating and scaling it.

$$\{X, Y, Z\}^T = \boldsymbol{S}_1(\boldsymbol{R} \cdot \{x, y, z\}^T + \boldsymbol{T}_1)$$

Rotation matrix $\boldsymbol{R}$ is the matrix production of $\boldsymbol{R}_{\theta 1}$, $\boldsymbol{R}_{\theta 2}$ and $\boldsymbol{R}_{\theta 3}$:

$$\boldsymbol{R} = \boldsymbol{R}_{\theta 3}\boldsymbol{R}_{\theta 2}\boldsymbol{R}_{\theta 1} = \boldsymbol{R}_y\boldsymbol{R}_z\boldsymbol{R}_x = \begin{bmatrix} \cos\theta_3 & 0 & \sin\theta_3 \\ 0 & 1 & 0 \\ -\sin\theta_3 & 0 & \cos\theta_3 \end{bmatrix} \cdot \begin{bmatrix} \cos\theta_2 & -\sin\theta_2 & 0 \\ \sin\theta_2 & \cos\theta_2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_1 & -\sin\theta_1 \\ 0 & \sin\theta_1 & \cos\theta_1 \end{bmatrix}$$

where θ_3 is the additional rotation angle, for the implementation of various rotation orientations.

Translation vector $\boldsymbol{T}_1 = \{0, T_H, T_V\}^T$. If we set original point to the source, then $\boldsymbol{T}_1 = \{0, T_H, D - T_V\}^T$.

Scale factor $\boldsymbol{S}_1 = \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & s_z \end{bmatrix}$. Under the condition of interpolated isotropic volume, $s_x = s_y = s_z$.

According to the pin-hole model (perspective projection) indicated in Fig.2,

$$U = X\frac{D}{Z}, \quad V = Y\frac{D}{Z}$$

$$\{U, V\}^T = \boldsymbol{P}\{X, Y, Z\}^T$$

where $\boldsymbol{P}$ is the perspective projection matrix. For convenient computing, we use a normalized matrix instead, that is

$$\{U, V, K\}^T = \boldsymbol{R}\{X, Y, Z, 1\}^T$$

Transformation from physical distance to image pixel is

$$\{u, v\}^T = \boldsymbol{S}_2\{U, V\}^T + \boldsymbol{T}_2$$

where $\boldsymbol{S}_2 = \begin{bmatrix} 1/s_u & 0 \\ 0 & 1/s_v \end{bmatrix}$, s_u and s_v are the scale factors from pixel to physical distance. In the isotropic X-ray image, $s_u = s_v$; $\boldsymbol{T}_2 = \begin{Bmatrix} c_u \\ c_v \end{Bmatrix}$, c_u and c_v are the initial coordinates of the pin-hole model.

From above, the projection matrix comprises ten

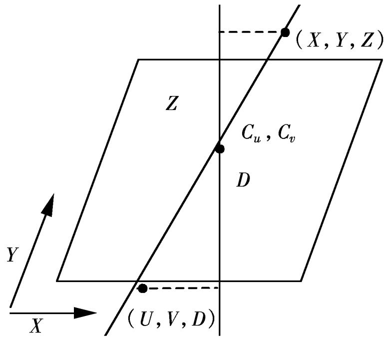


Fig.2 Pin-hole model

parameters: θ_1, θ_2, θ_3, T_V, T_H, D, $s_x(s_y, s_z)$, $s_u(s_v)$, c_u, c_v. These parameters are used in the following optimization step.

2 Segmentation

Segmentation is necessary for the distance transformation-based optimization algorithm. In the procedure of intervention, object structure, such as vascular, is projected onto the 2-D X-ray image. The original image is shown in Fig.3, which is a cube containing three hollow channels, and it is used to simulate a vascular structure in the CT image. The 3-D display of this structure is shown in Fig.4.

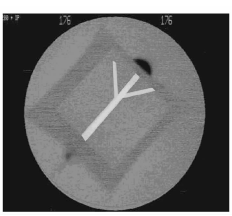


Fig.3 Original X-ray image

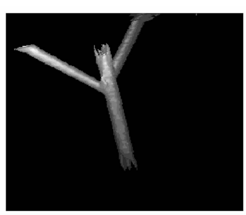


Fig.4 3-D display of object

The approach we undertake to segment is based on multi-step segmentation. Firstly we use open operator (erosion and dilation in morphology), then try to extract the circle area of the whole image, at last we use some automatic threshold algorithms to segment the object out.

1) Use open operator to erase the annotation of the whole image.

2) Extract circle region of the image. We try to depict the image from the ortho-coordinate to polar-coordinate. That is, we try to transform the coordinate from (x, y) to (ρ, θ), then we calculate the average gray level and mean difference along the ρ axis. The corresponding radius can be easily calculated.

3) Segment the object's structure within a circle area. When we exclude the non-circle area, the background effect can be greatly removed. Using the threshold-based segmentation, we extract the structure, as shown in Fig.5.

Through the operations above, the structure is automatically segmented out. Before the optimization step, a thinning operator and a Chamfer 3-4 operation (sometimes a sobel operator in contour points matching) are used. Fig. 6 is the skeleton map and the contour map, Fig. 7 is the distance maps generated respectively.


Fig.5 Segmentation result

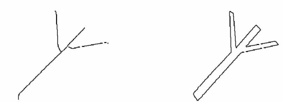


Fig.6 Skeleton and contour map

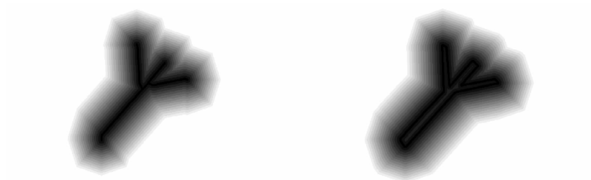


Fig.7 Skeleton distance map and contour distance map

3 Simulated Annealing Optimization

After the initial calibration approach, an approximate projection of a 3-D structure is generated, as shown in Fig.8. Using an interpolation method, this projection is taken into our algorithm input as a point set. The distance value of each projection point depicts the error between the 2-D image object and projection of the 3-D image object. So we propose to deal with this error by using a distance function $E(U)$, which is a real valued function that measures the difference between point set U and the 2-D image, namely another data set V. The specific function we proposed is

$$E(U) = \sum_{u_i \in U} d^2(u_i)$$

where u_i is the point of set U and $d(u_i)$ is the distance value of point u_i in the corresponding distance map. Under the ideal condition of correct projection, $E(U) = 0$.

So optimization is described as below to solve the following problem: Given a segmented object image with its distance map; a set of 3-D points X, whose projection by matrix is a 2-D point set U; and a projection matrix $\boldsymbol{M}$ with ten parameters and its initial approximate value, find the final space transformation matrix $\boldsymbol{M}$ that aligns the 3-D point set X with 2-D image well.

In our energy function, the solution is to find the minimum energy value, which makes the point set U mostly coincide with the segmented image.

P. Besl[7], J. Feldmar[8] and D. Kozinska[9] used some algorithms such as ICP and Marquardt-Levenberg optimization method to solve this problem. These optimization processes need some calculation of gradients and matrix operations, such as the matrix inverse, which make algorithms complex to implement and unstable to produce the correct solution. We propose a new optimization approach instead.

Fig.9 is the sample contour graph of energy function which is simplified to only 2 parameters. Point A is the initial position with parameter (p_{1A}, p_{2A}). Firstly, parameter p_2 is a constraint and parameter p_1 is resampled along the p_1 axis around p_{1A}. We can get the minimum energy point $(p_{1A'}, p_{2A})$. Then we constrain p_1 at $p_{1A'}$ to get another minimum point $(p_{1A'}, p_{2A'})$. At last, the local minimum point B is reached. After that, similar steps are used at point B. If the calculation time is enough, other local minimum value points C and D and others will be reached soon. From these local minimum points, we find the global minimum point, which is the correct optimization solution. The reason why we constrain some parameters is to reduce the calculation time, and make this algorithm's time one-dimensional relative to the number of parameters.

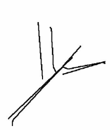


Fig.8 Initial projection

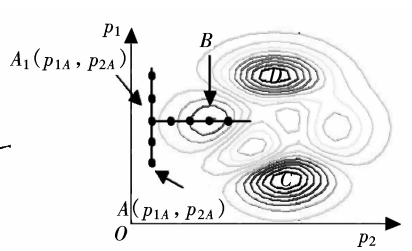


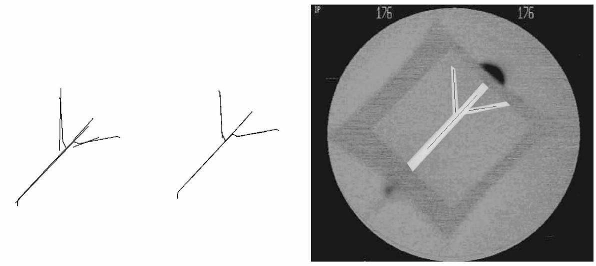
Fig.9 Resample process

The detailed algorithm is described as follows.

```
While iterative time < a threshold
    For every parameter p_i
        Resample p_i around last value, find the minimum energy
        one;
    End for
    If E < E_min
        E_min = E;
        Iterative time = 0;
    Else
        Random jump of all parameters from last found minimum
        value point;

    End if
End while
```

The algorithm result is shown in Fig.10. Fig.10(a) is the projection during the procedure. Fig.10(b) is the final result after optimization. Fig.10(c) is the projection located in the original X-ray image. From these resulting images, we can see that the 3-D skeleton's projection coincides with the original X-ray image well.



(a) (b) (c)

Fig.10 Projection process. (a) Projection during procedure; (b) Final result of registration; (c) Projection to original image

4 Retro-Projection

The use of registration is trying to build a transformation matrix between the 2-D image and the 3-D CT data. It is very useful in intra-operation, such as intervention and virtual endoscope procedure. The CT image can be acquired before the operation and the X-ray image is obtained during operation.

When physicians select a point in the X-ray image by retro-projection, the corresponding 3-D position in the CT volume will be calculated out in real time. With the help of some CT image workstation systems, physicians can see the virtual image from that point along the extracted path and can also be provided the chance to simulate a surgical process.

The retro-projection method used is indicated in Fig.11. When point A in the 2-D image is selected, the distances of that point to all paths are compared. Among them, if the distance is lower than a threshold and point displacement is between the start point and end point of that path, this point belongs to that path. The corresponding space position will be computed out by the ratio where the point is located in the path.

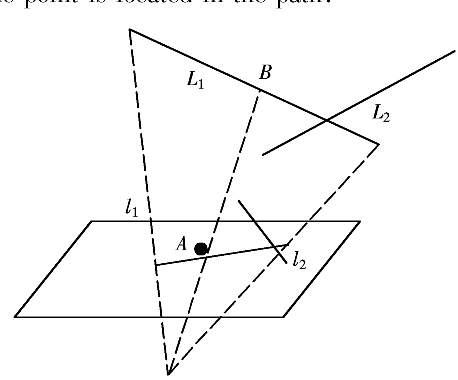
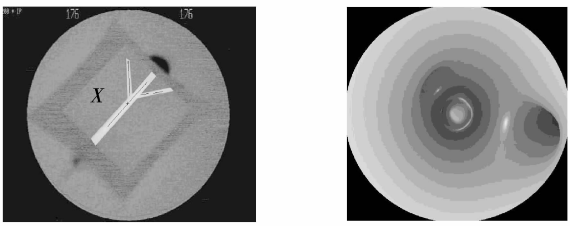


Fig.11 Retro-projection

As to point A, $l_1 = M(L_1), l_2 = M(L_2), A \in l_1$, the position of space point B can be calculated by the ratio A in l_1. In Fig.12(a), point X is the selected path point in the 2-D X-ray image. Fig.12(b) is the simulated image generated by virtual endoscopy from corresponding the 3-D point.



(a) (b)

Fig.12 Retro-projection and rendering result. (a) Select path point X; (b) Rendering by virtual endoscopy

5 Discussion

We have developed a methodology for distance-based registration and demonstrated the effectiveness of this approach in alignment of the X-ray image and CT image. The result shows that the approach is useful in the skeleton projection.

The approach stated in this paper has four key ingredients which are complementary: evaluation of projection matrix, image segmentation, Chamfer 3-4 distance transformation and simulated annealing optimization. The projection matrix with its parameters is used in optimization step. Segmentation limits the accuracy of registration, so our segmentation is based on many image processing operators, trying to precisely extract structure of interest. The Chamfer 3-4 distance is the basis of construction of the energy function, which is followed by an optimization approach. A simulated annealing method can reach a stable solution.

References

[1] Antoine Maintz J B, Viergever M A. A survey of medical image registration[J]. *Medical Image Analysis*, 1998, **2**(1):1 - 36.

[2] Grimson W E L, Lozano-Perez T, Wells W M, et al. An automatic registration method for frameless stereotaxy, image guided surgery, and enhanced reality visualization[J]. *IEEE Trans Medical Imaging*, 1996, **15**(2):129 - 140.

[3] Colchester A C F, Zhao J, Henri C, et al. Craniotomy simulation and guidance using a stereo video based tracking system (vislan)[A]. In: Robb R A, ed. *Proceedings of Visualization in Biomedical Computing*[C]. Rochester, Minnesota, 1994, **2359**:541 - 551.

[4] Simon D, Hebert M, Kanade T. Techniques for fast and accurate intra-surgical registration[J]. *Journal of Image Guided Surgery*, 1995, **1**(1):90 - 97.

[5] Shweikard A, Tombropoulos R, Adler J, et al. Planning for camera-guided robotic radio surgery[J]. *Transactions on Robotics and Automation*, 1998, **14**(6):951 - 962.

[6] Lavallee S, Szeliski R, Brunie L. Matching 3-D smooth surfaces with their 2-D projections using 3-D distance maps[J]. *Geometric Reasoning for Perception and Action*, 1991, **32**(4): 217 - 238.

[7] Besl P, Mckay N. A method for registration of 3-D shapes[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1992, **14**(2):239 - 256.

[8] Feldmar J, Ayache N, Betting F. 3-D - 2-D projective registration of free-form curves and surfaces[J]. *Computer Vision and Image Understanding*, 1997, **65**(3):403 - 424.

[9] Kozinska D, Tretiak O J, Nissanov J, et al. Multi-dimensional alignment using the euclidean distance transform[J]. *Graphical Models and Image Processing*, 1997, **59**(6):373 - 387.

X射线图像和CT图像的投影配准

张 惠[1] 罗立民[1] 舒华忠[1] 李松毅[1] Pascal Haigron[2]

([1] 东南大学生物科学与医学工程系, 南京 210096)

([2] 法国雷恩第一大学信号与图像处理实验室, 法国雷恩 35042)

摘 要 提出了一种基于Chamfer 3-4距离变换和模拟退火优化的方法,用于X射线图像和CT图像间的配准.首先,建立一个初始化的几何变换矩阵.为了便于计算,通过X射线设备的几何模型重新建出变换矩阵.然后,通过定义三维投影和二维物体图像之间的距离,并且通过优化算法来使之达到最小.该方法同样被用于医学介入手术中,处理从三维图像工作站中获取的主动漫游数据.

关键词 图像配准; 校准矩阵; 图像分割; 距离变换; 模拟退火

中图分类号 TP391.1