http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 12 (16): 3897-3902, 2013
ISSN 1812-5638 / DOI: 10.3923/itj.2013.3897.3902

Method of Real-time Tracking for Respiratory Motion Based on SIFT Feature Matching

Dong Feng, Sun Li-Ning and Ru Chang-Hai
Research Center of Robotics and Micro System and Collaborative Innovation Center of Suzhou
Nano Science and Technology, Soochow University, Suzhou, 215021, Jiangsu, China

Abstract: In this study, a new non-contact method of cancer detection based on SIFT feature matching is proposed to track real-time the respiratory motion. And a new camera calibration method based on planar and 3D checkerboard templates is proposed to get high accuracy calibration parameters. Experiment results show that cameras calibration accuracy is 0.0365 mm and the biggest movement scope of the marker is not lager than 1.5 mm between the vision measurement and practical measurement in a breathing cycle and can achieve the real-time tracking for respiratory motion.

Key words: Precise radiotherapy, binocular vision, camera calibration, sift feature matching, real-time tracking, respiratory motion

INTRODUCTION

The goal of the precise radiotherapy is to kill maximally tumor cells and protect effectively the surrounding normal tissues and vital organs. Therefore, the radiation dose is usually divided into several parts to complete the treatment and it is important to positioning precisely the location of tumor target (Price *et al*., 2009; Peng *et al*., 2011; Wu *et al*., 2005).

There are many technologies about the precise location of tumor target such as CT technology (Garcia-Ramirez *et al*., 2002), ultrasonic technology (US) (Fung *et al*., 2005), Magnetic Resonance Imaging (MRI) (Lathow *et al*., 2004), digital radiology image reconstruction technology (DRRS) (Bollet *et al*., 2003), Positron Emission Tomography (PET) (Eriksson *et al*., 2005), 4DCT technology (Korreman *et al*., 2012; Li, *et al*., 2012; Noel and Parikh, 2011; Sarker *et al*., 2010) etc. However, the tumors and surrounding tissues of the chest and abdomen happen to move with the influence of patient breathing and other physiological factors and the detection error of the tumor location is about 0.5~2.0 cm in the radiation therapy (Zhang *et al*., 2009). X-ray CT technique can't be used for of the whole process of radiotherapy.

And the ultrasonic technology is a contact method. But two methods aren't applied to the real-time detection of tumor location. Although, MRI, DRRS, PET, 4D-CT technologies can be applied to the real-time detection of tumor location, the operation procedures are very difficult, expensive and time-consuming.

The key of tracking real-time tumor is to calculate the specific location of tumor with time changing. In this study, a new camera calibration method based on planar and 3D checkerboard templates is proposed to get high accuracy calibration parameters. And a new non-contact method of cancer detection based on SIFT feature matching is proposed to track real-time the respiratory motion. The SIFT feature matching has strong robustness and accurate matching performance in tracking tumor process. In image matching process, the dynamic choice matching image and local searching method are used to reduce time of calculation and matching. The measurement method can improve the accuracy and speed of image matching and achive the real-time tracking for respiratory motion.

METHOD BASED ON SIFT FEATURE MATCHING

The image-based target tracking procedure had many complex factors such as losing information, image noise, target complex motion, target partial occlusions, complex target shape, light change and real-time requirements etc. (Lowe, 2004; Yilmaz *et al*., 2006). Therefore, this study proposes the real-time tracking method based on SIFT (Scale Invariant Feature Transform) feature matching.

The algorithm of SIFT feature matching includes the generation and matching of SIFT feature vector.

Generation of SIFT feature vector: In order to detect stable key point at scale space, Difference of Gaussian scale-space (DOG) is used which is defined as follows:

Corresponding Author: Ru Chang-hai, Research Center of Robotics and Micro System and Collaborative Innovation Center of Suzhou Nano Science and Technology, Soochow University, Suzhou, 215021, Jiangsu, China

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (1)$$

where, k is a scale factor, s is the scale space factor, L is an image scale space, (x, y) is an image pixel position, G(s, y, s) is the Gaussian function.

The least squares fit method is used to locate the position and scale of key points accurately and exclude the instability edge points in order to enhance the stability of matching and noise immunity.

In order to keep the SIFT operators rotational invariance, the equation of the modulus m(x, y) and direction q(x, y) in the gradient direction of (x, y) position are defined as follows:

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2}$$
$$\theta(x, y) = \arctan((L(x, y+1) - L(x, y-1)) / (L(x+1, y) - L(x-1, y))) \quad (2)$$

The 8×8 window is selected in the center of the key point, the accumulated value of its 8 directions is calculated in each 4×4 sub-region and the gradient histogram of each gradient direction is drew to form a seed point. Therefore, a total of 16 seeds and 128 SIFT feature vectors are generated. Meanwhile, SIFT feature vectors have eliminated some geometric distortion factors such as scale changes, rotation, the influence of illumination change, etc.

Matching of SIFT feature vectors: After the SIFT feature vectors of two images are generated, the Euclidean distance of the feature vectors are taken as the similarity measurement of two images which is expressed as follows:

$$d_s = |S_i - S_j| = \sqrt{\sum_k \left[S_i(k) - S_j(k)\right]^2} \; (k = 1, \cdots, 128) \quad (3)$$

The nearest distance and the near distance of the feature points are compared to eliminate the mismatching:

$$\frac{d_{smin}}{d_{s1}} < R(0 < R \le 1) \quad (4)$$

where, d_{smin} is the nearest distance, d_{s1} is the near distance. When their ratio is less than the percentage threshold value R, the result is the correct matching, or the error matching.

Optimization of SIFT algorithm: Although, the SIFT algorithm has a high robustness, it needs a lot of matching time. In order to improve the feature reliability and the matching speed, the optimization methods are as follows:

- The SIFT algorithm is used mainly to build scale space and generate feature vector. The larger the image scale is, the more the generated feature vectors are. Therefore, after first finishing the image matching between the marker and two videos, the last image position of the marker is taken as the center and the matching area is selected dynamically to calculate the next calibration in the two images; then, the exact position of the image is mapped to the original image. The efficiency of video matching is up to the size of the selected area
- Since, the matching image of the marker is not in a complex environment, the ratio R can be reduced to improve the feature reliability and the matching speed

EXPERIMENT

Design of the real-time tracking system based on binocular vision: The process of the real-time tracking system based on binocular vision is shown in Fig. 1. The system includes the camera calibration module, the video calibration module, the SIFT feature extraction module, the stereo matching module, the 3D coordinate calculation module and the 3D motion calculation module:

- **Camera calibration module:** According to the Tsai camera model and the influence of lens distortion, the study propose a new camera calibration method based on the planar and 3D checkerboard templates to get the internal and external parameters of cameras and distortion parameters

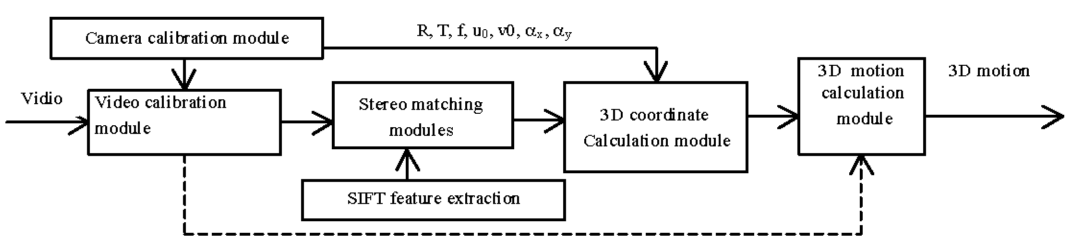

Fig. 1: Process of the real-time tracking system based on binocular vision

- **Video calibration module:** The matrix of the internal parameters is used to calibrate each video image
- **SIFT feature extraction and the stereo matching modules:** The matching points between marker image and two cameras images are got by their same characteristics factors. The center coordinate of marker is got by Eq. 5:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \tag{5}$$

- **3D coordinate calculation module:** The 3D coordinate of marker can be solved by the principle of parallel binocular vision
- **3D motion calculation module:** The movement distance of marker is expressed as follows:

$$S_{kl} = \sqrt{(x_k - x_l)^2 + (y_k - y_l)^2 + (z_k - z_l)^2} \tag{6}$$

The average velocity of the marker is expressed as follows:

$$v_{kl} = \frac{S_{kl}}{t_{kl}} = \frac{\sqrt{(x_k - x_l)^2 + (y_k - y_l)^2 + (z_k - z_l)^2}}{|l-k|\Delta t} \tag{7}$$

The speeds of marker in the x, y, z-axis is expressed as follows:

$$v_x = \frac{(x_k - x_l)^2}{t_{kl}}, v_y = \frac{(y_k - y_l)^2}{t_{kl}}, v_z = \frac{(z_k - z_l)^2}{t_{kl}} \tag{8}$$

- **Image matching selection of subsequent pixel:** According to the above optimization method of SIFT algorithm, the selected matching area is expressed as follows:

$$\begin{cases} M_{px} = v_x t + x + 1.5\ w \\ M_{py} = v_y t + y + 1.5\ w \end{cases} \tag{9}$$

where, w is the image width of the marker.

Experiment: The resolution of cameras is 1280×1024 and the sampling frequency is 15 Hz. The flat checkerboard template is shown in Fig. 2 and each size is 20×20 mm. The cube calibration template is shown in Fig. 3, where O is the center point of the treatment area.

The camera accuracy is expressed as the mean square error between the camera calculated coordinates of the feature points (X_w, Y_w, Z_w) and the actual coordinates of the feature points (X_{w0}, Y_{w0}, Z_{w0}) which is as follows:

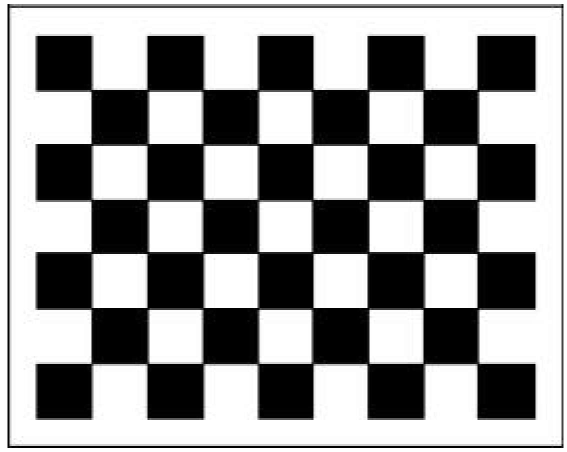

Fig. 2: Flat checkerboard template

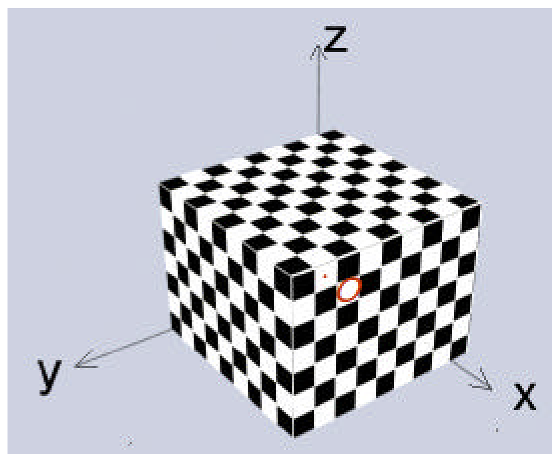

Fig. 3: Cube calibration template

$$\delta = \sqrt{\frac{(X_w - X_{w0})^2 + (Y_w - Y_{w0})^2 + (Z_w - Z_{w0})^2}{3}} \tag{10}$$

The camera calibration results are shown in Table 1. The measurement accuracy of the camera is 0.0365mm, so the calibration method can get higher accuracy calibration parameters.

A cylindrical rubber fixed in the abdomen of the body is selected as the marker which is shown in Fig. 4.

The experimental result of the real-time tracking system is shown in Table 2.

Experimental discussion

Accuracy of matching and time: When the SIFT feature vectors of the first frame image are extracted, the image feature numbers of two cameras are 485 and 568 and the times are 298.25 and 358.65 m sec. It takes about 1 sec to complete the 3D coordinate calculation, so it can not meet the requirement of the real-time tracking. However, when the above optimization method is used from the second frame image, the image feature numbers of two cameras image features are 31 and 30 and the times are 12.01 and 12.12 m sec. It takes about 60 m sec to complete the 3D coordinate calculation.

Table 1: Camera calibration result (unit: mm)

Name	Left camera	Right camera
U_0	162.692982	152.722817
V_0	110.270260	126.393028
K_1	-0.159408	-0.152761
K_2	0.161214	0.142905
S_x	1.000000	1.000000
S_y	1.000000	1.000000
C_x	306.526995	323.916494
C_y	286.809716	264.655358
g	0.000000	0.000000
F	511.363092	521.225593

The internal parameters of Left camera. T_x = -88.915959; T_y = -0.157128; T_z = 4.306600, R_x = 0.000422, R_y = 0.005006, R_z = 0.002866, d = 0.0365 mm

Meanwhile, although the marker moves with the surface motion of the stomach and the feature vectors number of the left and right images (eg 15, 16) is smaller than the feature vectors number of the marker (25), the system still tracks the marker¡äs position. Therefore, the algorithm of SIFT feature matching can meet the accuracy of time and image matching.

Accuracy analysis: The Position-Time curves of the tracking are shown in Fig. 5. The coordinate of the medical robot positioning system is defined as follows: X means the left and right movement direction of the body, Y means the head and foot movement direction of the body and Z means the up and down movement direction of the body. The deviation of X axis is 3.1 mm, the deviation of Y axis is 3.8 mm and the deviation of Z axis is 15.4 mm from the experimental results in Table 2. The actual measurement distance of Z axis is 16.8 mm, so its deviation is 1.4 mm. The marker is placed on the surface of the abdomen, so the deviation happens mainly in the Z-axis direction.

Table 2: Experimental result of the real-time tracking system

Pixel	SIFT of target	SIFT of left image	Extraction time (m sec^{-1})	Coordinate of left image (x, y)	SIFT of right image	Extraction time (m sec^{-1})	Coordinates of right image/(x, y)	Three-dimensional coordinates of marker/(x, y, z)/(mm)
1	25	485	298.25	(507,290)	568	358.65	(301,380)	(292.1,178.6,469.0)
2	25	31	12.01	(505,291)	30	12.12	(302,380)	(292.2,178.1,469.5)
3	25	32	12.63	(504,291)	31	13.29	(300,381)	(292.3,178.0,470.1)
4	25	33	12.78	(505,290)	34	13.82	(294,385)	(292.4,177.5,471.6)
5	25	35	14.03	(501,292)	32	12.65	(290,384)	(292.5,177.1,472.9)
6	25	33	13.22	(500,291)	32	12.61	(287,384)	(292.7,177.0,473.0)
7	25	31	12.15	(498,291)	32	12.59	(283,384)	(292.9,176.9,474.2)
8	25	34	13.68	(497,291)	34	13.49	(282,385)	(293.0,176.7,475.6)
9	25	32	12.65	(495,290)	32	12.57	(278,385)	(293.1,176.4,477.3)
10	25	15	9.36	(494,290)	15	9.67	(276,385)	(293.2,176.3,478.1)
11	25	16	10.13	(492,291)	17	11.25	(274,386)	(293.4,176.3,479.5)
12	25	23	11.61	(490,291)	22	11.54	(272,386)	(293.5,176.2,481.3)
13	25	26	11.68	(488,292)	24	11.51	(270,387)	(293.6,176.1,482.0)
14	25	35	13.90	(487,292)	22	12.07	(269,387)	(293.8,176.0,482.6)
15	25	34	13.19	(486,293)	32	12.71	(276,387)	(294.0,175.8,483.0)

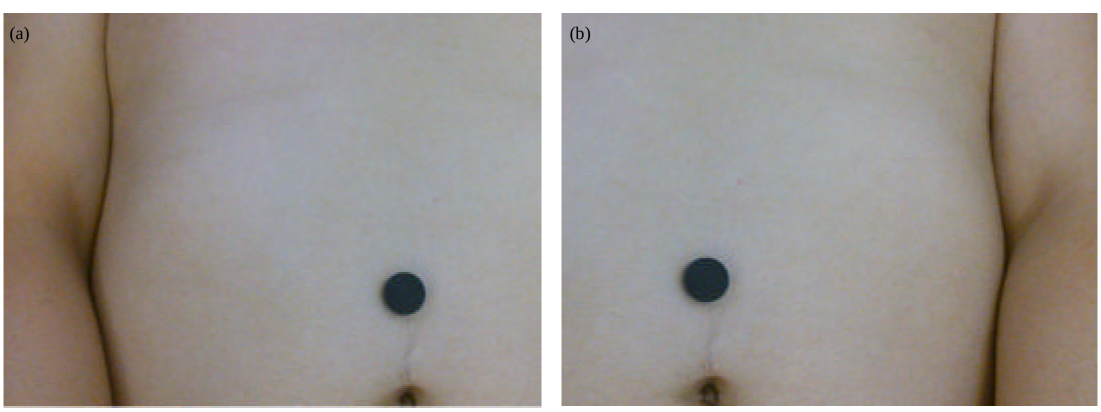

Fig. 4(a-b): Tracking image of marker in the two cameras (a) Right camera and (b) Left camera

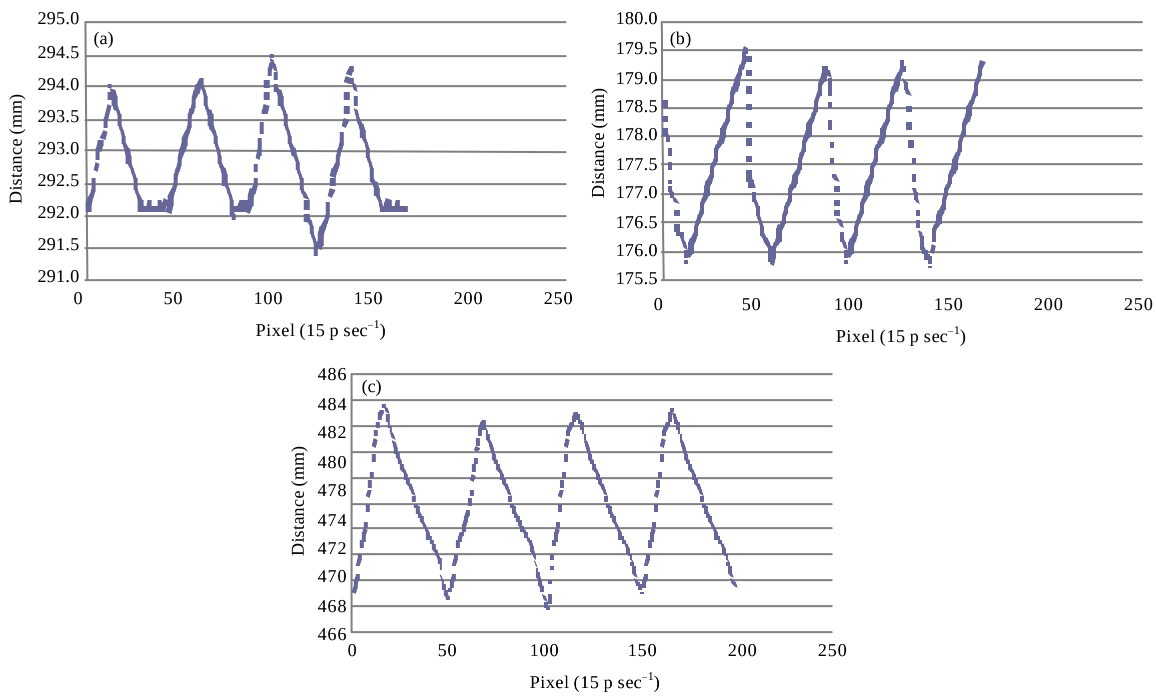

Fig. 5(a-c): Position-time curve of the racking (a) X coordinate axis for the left and right direction of the body, (b) Y coordinate axis for the head and foot direction of the body and (c) Z-coordinate axis for the up and down direction of the body

The errors of the calibration process are primarily the two factors: First, it is the error of measurement system such as the camera lens distortion, resolution and sampling precision etc., secondly, it is the overlay error between the central coordinate of the cube calibration block and the coordinate of the ISO-center.

CONCLUSION

In this study, a new non-contact method of cancer detection based on SIFT feature matching is proposed to track real-time the respiratory motion. And a new camera calibration method based on planar and 3D checkerboard templates is proposed to get high accuracy calibration parameters. Experiment results show that cameras calibration accuracy is 0.0365 mm and the biggest sports scope of the marker is not lager than 1.5 mm between the vision measurement and the practical measurement in a breathing cycle.Therefor, the proposed method can improve the accuracy and speed of image matching and achive the real-time tracking for respiratory motion.

However, the method is used mainly in the motion tracking for the single point of the abdominal surface and the movement of the marker is assumed to be the respiratory movement of the tumor target, so there are still some errors. In radiotherapy, the tumors may grow in different position, have different sizes and have different movement patterns in different stages, etc. Therefore, the tracking multi-points method of respiratory motion needs further study for radiotherapy patients.

ACKNOWLEDGMENTS

The authors are grateful for the support of the National High Technology Research and Development Program of China (863 Program, No. 2012AA040507), the National Natural Science Foundation of China (General Program, No. 61174087) and Jiangsu Natural Science Founds for Distinguished Young Scholar (No. BK2012005).

REFERENCES

Bollet, M.A., H.A. McNair, V.N. Hansen, A. Norman and U. O'Doherty *et al*., 2003. Can digitally reconstructed radiographs (DRRS) replace simulation films in prostate cancer conformal radiotherapy? Int. J. Radiat. Oncol. Biol. Phys., 57: 1122-1130.

Eriksson, B., H. Orlefors, K. Oberg, A. Sundin, M. Bergstrom and B. Langstrom, 2005. Developments in PET for the detection of endocrine tumours. Best Pract. Res. Clin. Endocrinol. Metab., 19: 311-324.

Fung, A.Y., C.A. Enke, K.M. Ayyangar, N.V. Raman and W. Zhen *et al*., 2005. Prostate motion and isocenter adjustment from ultrasound-based localization during delivery of radiation therapy. Int. J. Radiat. Oncol. Biol. Phys., 61: 984-992.

Garcia-Ramirez, J.L., S. Mutic, J.F. Dempsey, D.A. Low and J.A. Purdy, 2002. Performance evaluation of an 85-cm-bore X-ray computed tomography scanner designed for radiation oncology and comparison with current diagnostic CT scanners. Int. J. Radiat. Oncol. Biol. Phys., 52: 1123-1131.

Korreman, S., G. Persson, D. Nygaard, C. Brink and T. Juhler-Nottrup, 2012. Respiration-correlated image guidance is the most important radiotherapy motion management strategy for most lung cancer patients. Int. J. Radiat. Onc. Biol. Phys, 83: 1338-1343.

Lathow, C., S. Ley, C. Fink, M. Puderbach and W. Hosch *et al*., 2004. Analysis of intrathoracic tumor mobility during whole breathing cycle by dynamic MRI. Int. J. Radiat. Oncol. Biol. Phys., 59: 952-959.

Li, H., C. Noel, J. Garcia-Ramirez, D. Low and J. Bradley *et al*., 2012. Clinical evaluations of an amplitude-based binning algorithm for 4DCT reconstruction in radiation therapy. Med. Phys., 39: 922-932.

Lowe, D.G., 2004. Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vision, 60: 91-110.

Noel, C.E. and P.J. Parikh, 2011. Effect of mid-scan breathing changes on quality of 4DCT using a commercial phase-based sorting algorithm. Med. Phys., 38: 2430-2438.

Peng, Y., S. Vedam, J.Y. Chang, S. Gao, R. Sadagopan, M. Bues and P. Balter, 2011. Implementation of feedback-guided voluntary breath-hold gating for cone beam Ct-based stereotactic body radiotherapy. Int. J. Radiat. Oncol. Biol. Phys., 80: 909-917.

Price, G.J., P.J. Sharrock, T.E. Marchant, J.M. Parkhurst and D. Burton *et al*., 2009. An analysis of breast motion using high-frequency, dense surface points captured by an optical sensor during radiotherapy treatment delivery. Phys. Med. Biol., 54: 6515-6533.

Sarker, J., A. Chu, K. Mui, J.A. Wolfgang, A.E. Hirsch, G.T. Chen and G.C. Sharp, 2010. Variations in tumor size and position due to irregular breathing in 4D-CT: A simulation study. Med. Phys., 37: 1254-1260.

Wu, Y.C., G.L. Li and S.X. Tao, 2005. ARTS research status and development of precision radiotherapy system. Chinese J. Med. Phys., 6: 683-690.

Yilmaz, A., O. Javed and M. Shah, 2006. Object tracking: A survey. ACM Comput. Surv., Vol. 38. 10.1145/1177352.1177355

Zhang, D.D., H.Z. Zhang, W. Han, Y. Wang and Y. Zhang *et al*., 2009. Influence of respiratory motion on target dose distribution in radiotherapy for lung tumors. Chinese J. Radiat. Oncol., 3: 191-196.