



# An Online Learning Approach for Robust Motion Tracking in Liver Ultrasound Sequence

Chunxu Shen<sup>1</sup>, Huabei Shi<sup>1</sup>, Tao Sun<sup>1</sup>, Yibin Huang<sup>3</sup>,  
and Jian Wu<sup>1,2(✉)</sup>

<sup>1</sup> Tsinghua University, Beijing 100084, China  
{scxl6, shbl6, suntl6}@mails.tsinghua.edu.cn,  
wuj@sz.tsinghua.edu.cn

<sup>2</sup> Graduate School at Shenzhen, Tsinghua University, Shenzhen 518055, China

<sup>3</sup> Shenzhen Traditional Chinese Medicine Hospital, Shenzhen 518034, China  
huangyb2004@126.com

**Abstract.** Suffering from respiratory motion and drift, radiotherapy requires real-time and accuracy motion tracking to minimize damage to critical structures and optimize dosage delivery to target. In this paper, we propose a robust tracker to minimize tracking error and enhance the quality of radiotherapy based on two-dimensional ultrasound sequences. We firstly develop a scale adaptive kernel correlation filter to compensate deformation. Then the filter with an improved update rule is utilized to predict target position. Moreover, displacement and appearance constrains are elaborately devised to restrict unreasonable positions. Finally, a weighted displacement is calculated to further improve the robustness. Proposed method has been evaluated on 53 targets, yielding  $1.13 \pm 1.07$  mm mean and 2.31 mm 95%ile tracking error. Extensive experiments are performed between proposed and state-of-the-art algorithms, and results show our algorithm is more competitive. Favorable agreement between automatically and manually tracked displacements proves proposed algorithm has potential for target motion tracking in abdominal radiotherapy.

**Keywords:** Target tracking · Kernel correlation filter · Scale adaptation  
Displacement and appearance constrain · Radiotherapy

## 1 Introduction

Motion in the abdomen is worth accounting for during radiotherapy image guided intervention [1] and focus ultrasound surgery [2]. The motion induced in abdominal organs is mainly due to breathing motion, drift and surgical instruments. Therefore, motion tracking of abdominal target is crucial to minimize the damage to surrounding crucial structure and optimize dosage delivery to target.

Respiratory gating is one of the most conventional approach to deal with abdomen motion, whereas it potentially increases treatment time [3]. Motion modeling like implanting fiducial markers to target region [4] is an alternative method, but it is usually at the expense of healthy tissue. Tracking base on medical image e.g. magnetic

resonance (MR), ultrasound (US) generally becomes a superior to localize abdomen target. De Senneville [5] generates an atlas of motion fields based on magnitude data of temperature-sensitive MR acquisitions. They suppose that motion of target region is periodic and can be estimated in the next moment, so it just recovers deformation caused by periodic component. 4D MR [6] is also introduced to respiratory motion reconstruction, but low signal-to-noise ratio and additional high cost must be considered in clinical practice. US is an appealing choice for abdominal target tracking, by contrast, as it has high temporal resolution and sub-millimeter spatial resolution along the beam direction.

Recently several literatures focus on tracking hepatic landmark and reconstructing liver motion of free breathing. Block matching [7], optical flow [8], particle filter [9], image registration and mechanical simulation [10] are widely investigated. Meanwhile temporal regularization [7] and distance metric [10] are also introduced to reject false tracking results. While some results have achieved a great process, many limitations remain to be discussed like tradeoff between real-time and accuracy, as well as robustness for acoustic shadowing and large deformation due to out-of-plane motion.

Our tracking approach is motivated by kernel correlation filter (KCF) [11], which achieves a fast and high performance on Visual Tracker Benchmark [12]. KCF provides an effective solution for translation, but its performance would degrade because of the scale and deformation of targets. Li et al. [13] suggests an effective scale adaptive scheme. Without discussing update strategy adequately, however, better tracking results cannot be remerged in US sequence. Besides, we integrate intensity feature, namely speckle patten, to proposed tracking frame as it includes much information about anatomical structure. In fact, if all the speckle patterns are stable, target motion can be easily reconstructed. Unluckily, speckle patterns are not identical because of out-of-plane motion and acoustic shadowing [14]. Moreover, similarity metrics is another important ingredient in proposed method. While mutual information (MI) has been suggested to be the most suitable metric for US to US match, high computation limits its usage in real-time target tracking. In this work, normalized cross-correlation (NCC) is chosen as it is easy to implement and effective to perform block matching.

In this work, we propose a real-time, robust tracking algorithm to compensate target motion in abdominal radiotherapy. Our contributions mainly focus on four aspects: first, we propose a scale adaptation strategy to alleviate deformation and scale change. Second, an improved update rule for proximate periodic motion is applied to reducing accumulation error in long-term tracking. Third, we integrate displacement and appearance constrains to proposed method in order to restrict unreasonable target prediction. And fourth, we suggest to use weighted displacement to determine target displacement.

## 2 Method

### 2.1 The KCF Tracker

In KCF tracker, Henriques et al. [11] suppose that the cyclic shifts version of base sample is approximate the dense samples over the base sample. Take one-dimension data  $\mathbf{x} = [x_1, x_2, \dots, x_n]$  for example, a cyclic shift of  $\mathbf{x}$  is defined as

$\mathbf{P}\mathbf{x} = [x_n, x_1, x_2, \dots, x_{n-1}]$ . Therefore, all the cyclic shift samples,  $\{\mathbf{P}^u\mathbf{x}|u = 0, \dots, n - 1\}$ , can be concatenated to form sample matrix  $\mathbf{X}$ , which also called circulant matrix as the matrix is purely generated by the cyclic shifts of  $\mathbf{x}$ . This matrix has a helpful property that all the circulant matrices can be formulated as follows:

$$\mathbf{X} = \mathbf{F}^H \text{diag}(\mathbf{F}\mathbf{x})\mathbf{F} \tag{1}$$

Where,  $\mathbf{F}$  is the Discrete Fourier Transformation (DFT) matrix.  $\mathbf{F}^H$  is the Hermitian transpose of  $\mathbf{F}$ . Benefit from the decomposition of circulant matrix, it can be used to the solution of linear regression. Moreover, the objective function of linear ridge regression can be written as:

$$\min_{\mathbf{w}} \sum_i^n (f(\mathbf{x}_i) - y_i)^2 + \lambda \|\mathbf{w}\| \tag{2}$$

Where,  $f$  is linear combination of basis samples,  $f(\mathbf{x}) = \mathbf{w}^T\mathbf{x}$ . The ridge regression has a close-form solution,  $\mathbf{w} = (\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}^T\mathbf{y}$ . The solution can be rewritten with Eq. 1,  $\hat{\mathbf{w}}^* = \frac{\hat{\mathbf{x}}^* \odot \hat{\mathbf{y}}}{\hat{\mathbf{x}}^* \odot \hat{\mathbf{x}} + \lambda}$ . Where,  $\hat{\mathbf{x}} = \mathbf{F}\mathbf{x}$  donates the DFT of  $\mathbf{x}$ ;  $\hat{\mathbf{x}}^*$  is the complex-conjugate of  $\hat{\mathbf{x}}$ ;  $\odot$  denotes element-wise multiplication. So during the process of extracting patches explicitly and solving a general regression problem, this step can save much computational cost. In order to construct a more powerful classifier in case of non-linear regression, Henriques et al. [11] adopt a kernel tracker,  $f(\mathbf{z}) = \mathbf{w}^T\mathbf{z} = \sum_{i=1}^n \alpha_i \mathcal{K}(\mathbf{z}, \mathbf{x}_i)$ . Then dual space confident  $\alpha$  can be learned as follows:

$$\hat{\alpha}^* = \frac{\hat{\mathbf{y}}}{\hat{\mathbf{k}}^{\mathbf{xx}} + \lambda} \tag{3}$$

$\mathbf{k}^{\mathbf{xx}}$  is defined as kernel correlation. Similar to the linear classifier, the dual coefficients are learned in Fourier domain.  $\mathbf{y}$  is a regression target vector in Fourier domain and has the same size with  $\mathbf{x}$ ;  $\lambda$  is regularization weight in ridge regression. Note that the search window, which is the size of  $\mathbf{x}$ , has 2.5 times the size of the target in the implementation of KCF. In case of Gaussian kernel function, the kernel correlation can be denoted as:

$$\mathbf{k}^{\mathbf{xx}'} = \exp \left( -\frac{1}{\sigma^2} \left( \|\mathbf{x}\|^2 + \|\mathbf{x}'\|^2 \right) - 2\mathbf{F}^{-1}(\hat{\mathbf{x}} \odot \hat{\mathbf{x}}'^*) \right) \tag{4}$$

Where  $\mathbf{F}^{-1}$  denotes inverse Fourier transform.

In detection step, the regression function Eq. 5 is applied to predict the position of target where the maximum regression value locates.

$$\hat{\mathbf{f}}(\mathbf{z}) = \left( \hat{\mathbf{k}}^{\tilde{\mathbf{x}}\mathbf{z}} \right)^* \odot \hat{\alpha} \tag{5}$$

Where  $\tilde{\mathbf{x}}$  denotes basic data template to be learned in the model;  $\mathbf{z}$  is the candidate patch, which has the same size and location with  $\mathbf{x}$  in next frame. When we transform

$\hat{f}(z)$  back into the spatial domain, the translation with respect to the maximum response is considered as the displacement of the tracked target.

### 2.2 Scale Adaptive KCF

Deformations and scale variations of targets is potential to increase the tracking error and reduce robustness, even fail. However, these negative factors are common in abdominal targets. In our clinical practice, there are two situations leading to target deformation. First, with the contraction and relaxation of the diaphragm in free breathing situation [15], the hepatic targets would suffer from deformation. Second, because of free breathing and drift, the appearance of cross section between ultrasound beam and targets would change. In this part, we propose a scale adaptive strategy to compensate these deformations and scale variations.

Suppose that the size of search window sets as  $\mathbf{s}_T = (s_x, s_y)$ , we define a scaling pool  $\boldsymbol{\eta} = \{\eta_1, \eta_2, \dots, \eta_m\}$  to expand search range to different scale space, which can be donated as  $\tilde{\mathbf{s}}_T = \{(\eta_i s_x, \eta_j s_y) | \eta_i, \eta_j \in \boldsymbol{\eta}\}$ . Because the dot-product requires the search window with the fixed size in kernel correlation filter, we resize  $\tilde{\mathbf{s}}_T$  into the fixed size of  $\mathbf{s}_T$  using bilinear-interpolation. Note that our proposed scale adaptive method is different from Li's work [13], which adopts  $\check{\mathbf{s}}_T = \{\eta_i \mathbf{s}_T | \eta_i \in \boldsymbol{\eta}\}$ . Therefore, the response  $\mathbf{R}(\eta_i, \eta_j)$  in difference scale space can be calculated.

$$\mathbf{R}(\eta_i, \eta_j) = \mathbf{F}^{-1} \hat{f}(z(\eta_i, \eta_j)) \tag{6}$$

Where  $z(\eta_i, \eta_j)$  is the sample patch resampled by scaling pool and the size of  $z(\eta_i, \eta_j)$  is  $(\eta_i s_x, \eta_j s_y)$ , which is subsequently resized to the fixed size of  $\mathbf{s}_T$ .

### 2.3 Improved Update Rule for Approximate Periodic Motion

According to Eq. 5, there are two sets of coefficient should be update. One is dual space coefficient  $\boldsymbol{\alpha}$ , another is basic template  $\tilde{\mathbf{x}}$ . Original update rule is realized by combining new filter with old one linearly as Eq. 7 illustrates.

$$\begin{cases} \tilde{\mathbf{x}}_{t+1} = \mu \tilde{\mathbf{x}}_{t+1} + (1 - \mu) \tilde{\mathbf{x}}_t \\ \hat{\boldsymbol{\alpha}}_{t+1} = \mu \hat{\boldsymbol{\alpha}}_{t+1} + (1 - \mu) \hat{\boldsymbol{\alpha}}_t \end{cases} \tag{7}$$

Where  $\mu$  is the linear interpolation factor.

While the update rule above achieves impressive success for nature video tracking, it is so sensitive that cannot support for long-term tracking in our work. An explanation is that Eq. 7 pays more attention to learn new characteristics from a new image. Once ultrasound images suffer from noise severely, like acoustic shadowing and speckle decorrelation, the performance of online classifier could degrade largely. With prior knowledge that motion of liver is approximate periodic in free breathing, the target in first frame would also appear in subsequent sequence. Therefore, an improved update rule for long-term tracking of approximate periodic motion is proposed as Eq. 8 shows:

$$\begin{cases} \tilde{\mathbf{x}}_{t+1} = \beta \tilde{\mathbf{x}}_1 + (1 - \beta - \mu) \tilde{\mathbf{x}}_{t+1} + \mu \tilde{\mathbf{x}}_t \\ \hat{\boldsymbol{\alpha}}_{t+1} = \beta \hat{\boldsymbol{\alpha}}_1 + (1 - \beta - \mu) \hat{\boldsymbol{\alpha}}_{t+1} + \mu \hat{\boldsymbol{\alpha}}_t \end{cases} \quad (8)$$

Where  $\beta$  is recurrence factor.

### 2.4 Restricting Unreasonable Target Prediction

Though NCC has been a popular similarity measure in speckling tracking, it still suffers from acoustic shadowing, speckle decorrelation and other artifacts. Here, in order to alleviate these adverse effect, we provide displacement and appearance constrains to restrict unreasonable target prediction.

**Displacement Constrain.** In clinical ultrasound image guided abdominal radiotherapy, we notice that the target displacement in two consecutive frames is very small (<3 mm, acquisition frequency is 13–23 Hz). So a displacement cost function is employed to restrict unreasonable prediction. Suppose that  $\mathbf{D} = (\Delta\mathbf{x}(\eta_i, \eta_j), \Delta\mathbf{y}(\eta_i, \eta_j))$  is the displacement prediction and  $\mathbf{R}(\mathbf{d}_{ij} | \mathbf{d}_{ij} \in \mathbf{D})$  is corresponding response map, therefore, the response with displacement constrain can be expressed by:

$$\mathbf{R}_{\text{dis}}(\eta_i, \eta_j) = \mathbf{R}(\eta_i, \eta_j) \odot \exp\left(-\frac{\Delta\mathbf{x}^2 + \Delta\mathbf{y}^2}{\sigma_{\text{dis}}}\right) \quad (9)$$

Where  $\sigma_{\text{dis}}$  is the bandwidth of displacement constrain.

**Appearance Constrain.** For alleviating the unreasonable matching from NCC, we also employ a set of confidence response to determine target displacement instead of selecting the displacement that the best response locates. Supposing the threshold of confidence response is  $\theta_{\text{app}}$ , the appearance constrain can be expressed as Eq. 10 shows.

$$\mathbf{R}_{\text{dis}}^{\text{app}}(\eta_i, \eta_j) = \begin{cases} \mathbf{R}_{\text{dis}}(\eta_i, \eta_j), & \text{if } \mathbf{R}_{\text{dis}} \geq \max\{\mathbf{R}_{\text{dis}}\} \cdot \theta_{\text{app}} \\ 0, & \text{others} \end{cases} \quad (10)$$

With constrains of displacement and appearance, the best scale space can be determined by maximize the average response  $\overline{\mathbf{R}_{\text{dis}}^{\text{app}}(\eta_i, \eta_j)}$  in Eq. 10:

$$\arg \max \overline{\mathbf{R}_{\text{dis}}^{\text{app}}(\eta_i, \eta_j)} \quad (11)$$

### 2.5 Weighted Displacement

Motivated by Carletti’s work [9], a weighted displacement is calculated to enhance the robustness of proposed tracking algorithm. The displacements used to calculate

weighted displacement are from Eq. 10, namely  $r_{ij} \in \mathbf{R}_{\text{dis}}^{\text{app}}(\eta_i, \eta_j)$ . Finally the target displacement can be determined in adjacent frames.

$$\bar{\mathbf{d}} = \frac{\sum_{i=1}^M \sum_{j=1}^N r_{ij} \mathbf{d}_{ij}}{\sum_{i=1}^M \sum_{j=1}^N r_{ij}} \quad (12)$$

Note that  $\bar{\mathbf{d}}$  is the displacement in best scale space, we get the real displacement  $\bar{\mathbf{d}}_r$  by performing scale inverse transformation with scale parameters from Eq. 11. Therefore, by combining the target position in last frame  $\mathbf{p}_{\text{old}}$  and displacement  $\bar{\mathbf{d}}_r$ , new target position  $\mathbf{p}_{\text{new}}$  in current frame can be determined.

$$\mathbf{p}_{\text{new}} = \mathbf{p}_{\text{old}} + \bar{\mathbf{d}}_r \quad (13)$$

Finally, the overall algorithm is summarized into Algorithm 1

---

**Algorithm 1.** Overall of proposed tracking algorithm

---

**Require:**

- The template of tracked target,  $\tilde{\mathbf{x}}$  ;
- The dual space coefficient,  $\alpha$  ;
- The newly search window,  $\mathbf{y}$  ;
- The last target position,  $\mathbf{p}_{\text{old}}$  ;

**Ensure:**

- The updated template of tracked target,  $\tilde{\mathbf{x}}$  ;
  - The updated dual space coefficient,  $\alpha$  ;
  - The new target position,  $\mathbf{p}_{\text{new}}$  ;
- 1: **for** every  $(\eta_i s_x, \eta_j s_y)$  in  $\tilde{\mathbf{s}}_{\mathbf{T}}$  **do**
  - 2: Sample the new search window  $\mathbf{z}(\eta_i, \eta_j)$  based on size  $(\eta_i s_x, \eta_j s_y)$  and resize it to  $\mathbf{s}_{\mathbf{T}}$  with bilinear interpolation
  - 3: Calculate the corresponding response  $\mathbf{R}(\eta_i, \eta_j)$  in different scale space with Eqn. 5
  - 4: Apply displacement and appearance constrains with Eqn.9 and Eqn.10
  - 5: Get the best scale space and corresponding scale parameters
  - 6: Calculate the weighted displacement  $\bar{\mathbf{d}}$  in best scale space with Eqn.11
  - 7: **end for**
  - 8: Get the real displacement  $\bar{\mathbf{d}}_r$  with scale parameters
  - 9: Get the new target position  $\mathbf{p}_{\text{new}}$  with Eqn.13.
  - 10: Get the new  $\tilde{\mathbf{x}}$  and  $\alpha$  base on the new position  $\mathbf{p}_{\text{new}}$  with Eqn.3.
  - 11: Update  $\tilde{\mathbf{x}}$  and  $\alpha$  with Equation 8.
  - 12: **return** update  $\tilde{\mathbf{x}}$  and  $\alpha$ ;
-

### 3 Experiments and Results

#### 3.1 Dataset and Parameter Settings

**Datasets and Resource.** Our 2D liver ultrasound sequences are provided by MICCAI 2015 Challenge on Liver Ultrasound Tracking (CLUST) [16] training database, and it consists of five different datasets CIL, ETH, ICR, MED1 and MED2. Each dataset is acquired by different scanner with different image resolution (0.30–0.55 mm) and acquisition frequency (13–23 Hz). Besides, our code is implemented using MATLAB R2017b on an Intel Core i7-4910MQ CPU @ 2.90 GHz.

**Parameter Settings.** The parameters in our algorithm come from two parts. One is from the original KCF tracker and we adopt the default parameters as [11] recommends. The learning rate  $\lambda$  in Eqs. 2 and 3 sets to  $10^{-4}$ ; the  $\sigma$  used in Gaussian function Eq. 4 sets to 0.2; the linear interpolation factor  $\mu$  in Eq. 8 sets to 0.1; and the size of search window is 2.5 times to the size of target. Another part is from our contributions, which is used to ensure proposed tracker more accuracy and robust. We adopt scaling pool with the suggestion from our experienced radiologist  $\eta = \{0.85, 0.90, 0.95, 1.00, 1.05, 1.10, 1.15\}$ . And the recurrence factor  $\beta$  in Eq. 8, bandwidth of displacement  $\sigma_{\text{dis}}$  in Eq. 9 and the threshold of confidence response  $\theta_{\text{app}}$  in Eq. 10 set to 0.15, 10 and 0.95 respectively. Parameters are same for all following experiments.

Note that proposed method needs image patches as initialization. Therefore, we generate a rectangular region manually with the guidance of experienced radiologist in the first frame. During online tracking process, the center of rectangular region is recorded and then used to evaluate tracking performance.

#### 3.2 Tracking Results

We employ Euclidean distance suggested by Organizers of CLUST [16] to evaluate the tracking performance. In our experiments, we compute errors between each manual annotation and the output of proposed algorithm, and then mean, standard deviation (SD), 95%ile and maximum errors are counted. Additionally, processing speed is estimated by counting frames that are tracked per second (FPS).

**Performance Evaluation on CLUST.** Firstly, we evaluate the performance of proposed tracking algorithm using the five datasets of CLUST database. The number of objects means the total objects being tracked in corresponding dataset. The following Table 1 shows the tracking error distribution of each dataset and the total 2D ultrasound sequences respectively.

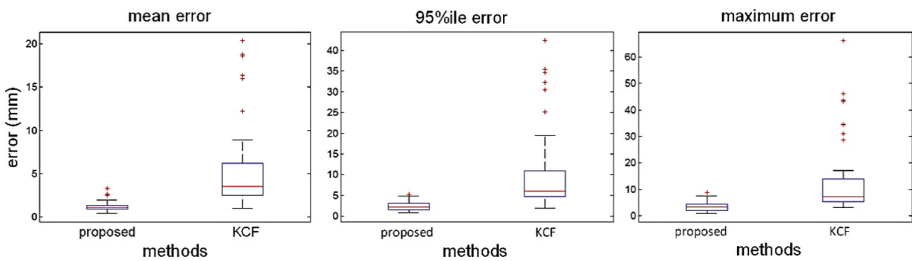
**Comparison Proposed with Baseline Algorithm.** Then a performance comparison experiment is performed between proposed and baseline algorithm, and the results are shown in Fig. 1.

Compared with baseline algorithm, proposed method achieves state-of-the-art results with mean decreasing by 78.8% (from 5.33 mm to 1.13 mm), 95%ile error

decreasing by 77.1% (from 10.08 mm to 2.31 mm) and maximum error decreasing by 82.8% (from 66.10 mm to 11.37 mm) respectively.

**Table 1.** A summary for performance evaluation on CLUST. All tracking errors are in millimeters and processing speed is presented by frames per second.

Dataset	No. objects	Mean	SD	95%ile	Maximum	FPS
CIL	3	0.99	1.16	2.02	3.61	20.33
ETH	16	0.89	0.60	1.73	4.18	23.57
ICR	12	1.00	0.54	2.31	6.23	23.14
MED1	19	1.39	1.62	2.74	11.37	22.21
MED2	3	1.38	2.04	3.01	7.88	31.00
Total	53	1.13	1.07	2.31	11.37	23.22



**Fig. 1.** Tracking errors comparison between proposed and baseline algorithm on CLUST. Left is mean tracking error; middle is 95%ile error; right is maximum error.

**Comparison Proposed with State-of-the-art Algorithms.** Extensive comparison experiments are performed among our tracker and some state-of-the-art trackers. The following Table 2 gives a summary of tracking error distribution. It is worth mentioning that we compare these algorithms whose tracking performance is also evaluated on CLUST training database. Compared with TMG [17], RMTwS [17] and Hybrid [18], proposed algorithm achieves a competitive accuracy with maximum tracking error decreasing by 40.4%–47.8%, which means it would provide a more effective guidance for clinical operation. Experimental results also indicate our tracker is more real time than the existing state-of-the-art trackers.

**Table 2.** Comparison of published results with our tracking results. All tracking errors are in millimeters and processing speed is presented by frames per second.

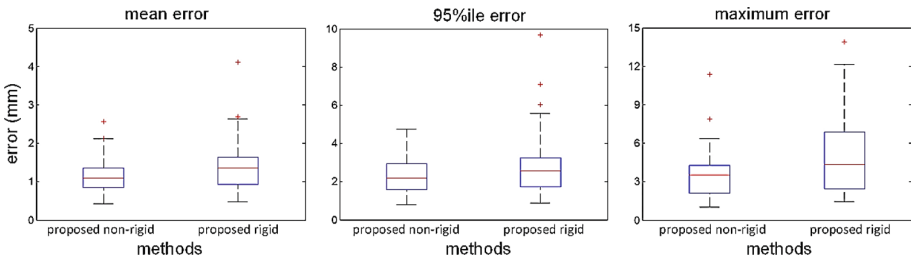
Algorithms	Mean	SD	95%ile	Maximum	FPS
Proposed	1.13	1.07	2.31	11.37	17–34
TMG [17]	1.17	0.89	2.61	21.78	8–23
RMTwS [17]	1.12	0.81	2.19	21.78	3–16
Hybrid [18]	0.80	0.80	1.85	19.08	8–32



### 3.3 Experimental Analysis

In this section, we first perform an ablation analysis to understand the benefit of scale adaptive strategy. Then a detailed parameters analysis are performed to find out the effectiveness of improved update rule (Eq. 8) and appearance/displacement (Eqs. 9 and 10) constraints.

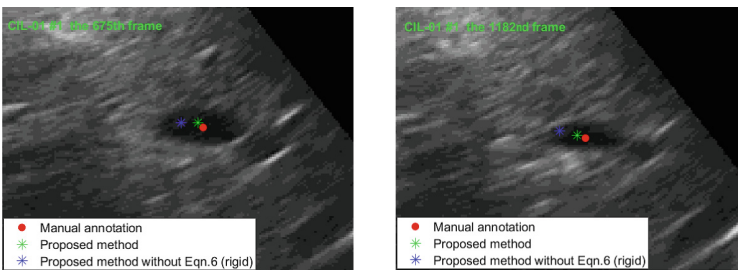
**Ablation Study About Scale Adaptive Strategy.** Deformation is common in liver ultrasound sequence. In this part, we perform a comparison experiment between non-rigid (with Eq. 6) and rigid (without Eq. 6) tracking. Results are shown in Fig. 2.



**Fig. 2.** Tracking error distributions (mm) for proposed non-rigid and rigid tracking method. Left is mean error; middle is 95%ile error, right is maximum error.

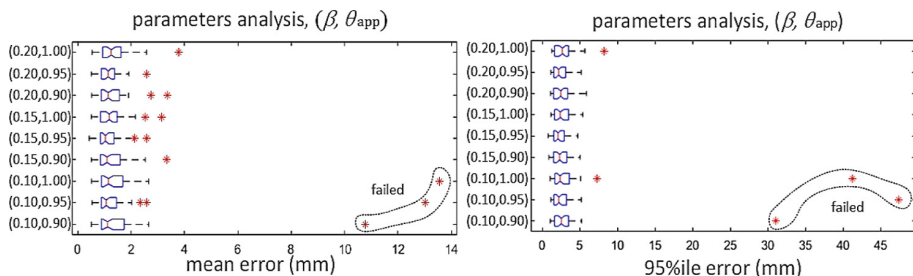
Compared with rigid tracking, non-rigid tracking achieves a better performance with mean decreasing by 20.4% (from 1.42 mm to 1.13 mm), 95%ile error decreasing by 19.5% (from 2.87 mm to 2.31 mm) and maximum error decreasing by 18.1% (from 13.88 mm to 11.37 mm) respectively. That means non-rigid deformation should be considered seriously in precise radiotherapy.

Figure 3 shows an instance to compare the results from non-rigid and rigid tracking. The target position calculated by rigid tracking yields larger deviations, by contrast, the positions from proposed method are more accurate and robust.



**Fig. 3.** An example for showing deviation between non-rigid and rigid tracking. Images are both from CIL-01 #1 in CLUST. Left is the 675th frame and right is the 1182nd frame.

**Parameters Analysis.** There are four parameters,  $[\boldsymbol{\eta}, \sigma_{\text{dis}}, \theta_{\text{app}}, \beta]$ , needing more discussion. Among them, scaling pool  $\boldsymbol{\eta}$  can be designed when the deformation of target is estimated. And we also can determine  $\sigma_{\text{dis}}$  by magnitude of target motion and frequency of image acquisition. However,  $\theta_{\text{app}}$  and  $\beta$  are assigned empirically. In this part, we investigate the effect when we change the threshold of confidence response and recurrence factor. Without loss of generality, we choose  $\theta_{\text{app}} \in [0.90, 0.95, 1.00]$  and  $\beta \in [0.10, 0.15, 0.20]$  to perform parameters analysis on CLUST training database. Here, mean and 95%ile tracking errors, as regardful indicators for our project, are chosen to evaluate the results of parameters analysis. Results are shown in Fig. 4 and Table 3.



**Fig. 4.** The results of parameters analysis on CLUST. Left is results of mean error and right is results of 95%ile error with parameters  $(\beta, \theta_{\text{app}})$  changing.

**Table 3.** Statistic results of mean and 95%ile errors (mm) with  $(\beta, \theta_{\text{app}})$  changing.

$(\beta, \theta_{\text{app}})$	(0.10,0.90)	(0.10,0.95)	(0.10,1.00)	(0.15,0.90)	(0.15,0.95)
mean	1.45	1.43	1.53	1.24	1.13
95%ile	3.23	3.36	3.54	2.54	2.31
$(\beta, \theta_{\text{app}})$	(0.15,1.00)	(0.20,0.90)	(0.20,0.95)	(0.20,1.00)	
mean	1.26	1.24	1.18	1.32	
95%ile	2.65	2.56	2.43	2.78	

Therefore, recurrence factor is a crucial parameter in proposed algorithm. A smaller  $\beta$  has a terrible effect on long-term tracking (like  $\beta = 0.10$ , see Fig. 4). But a larger one would also enlarge tracking error by unduly limiting learning ability for proposed method. Besides, a smaller or larger  $\theta_{\text{app}}$  are not a wise chose, which would potentially introduce more unreasonable position or be not adaptive for artifacts well respectively. Therefore, (0.15, 0.95) is a better combination for accuracy and robust tracking in our project.

## 4 Conclusion and Discussion

In this paper, we present a 2D real-time tracking approach, which consists four steps namely (1) initial target regions selection, (2) tracking with scale adaptive kernel correlation filter, (3) displacement and appearance constrains, and (4) weighted displacement. The initial target regions are generated by our experienced radiologist. Then we train an online classifier to predict targets position. Because deformation of targets can lead to error accumulation in learning phase, we employ adaptive scale strategy to mitigate this adverse effect. Considering US images suffer from acoustic shadowing and speckle decorrelation, NCC is more susceptible to bias. We employ displacement and appearance constrains to constrict unreasonable position prediction by carefully investigating the motion extents of landmarks in liver under free breathing. Furthermore, with prior knowledge that target motion in liver is approximately periodic under free breathing, we revise the update rule by introducing a recurrence factor to improve robustness in long-term tracking. Finally, inspired by success of particle filter in noise circumstance, we obtain new target positions by calculating weighted displacement.

However, we just adopt single feature to realize target tracking. Accuracy and robustness for proposed method may continue to improve by combining other image features like texture and shape, which is a major research direction for future work. Also, similarity metrics is a core ingredient for target tracking. While a large of similarity metrics have been proposed in computer vision community, there are no clear rules about how to select the most suitable one but to try them in different condition.

There are several avenues of future work that would potentially improve proposed method. Integrating texture feature into our tracking method would be helpful to improve accuracy. And adaptive recurrence factor strategy will be investigated to improve robustness for long-time tracking.

In conclusion, we propose an online learning approach for robust and real-time motion tracking in liver ultrasound sequences and evaluate it on five different datasets. Favorable agreement between automatically and manually tracked displacements, along with real-time processing speed prove that proposed algorithm has potential for target motion tracking in abdominal radiotherapy.

**Acknowledgement.** This work is supported in part by Knowledge Innovation Program of Basic Research Projects of Shenzhen under Grant JCYJ20160428182053361, in part by Guangdong Science and Technology Plan under Grant 2017B020210003 and in part by National Natural Science Foundation of China under Grant 81771940, 81427803.

## References

1. Riley, C., Yang, Y., Li, T., Zhang, Y., Heron, D.E., Huq, M.S.: Dosimetric evaluation of the interplay effect in respiratory-gated RapidArc radiation therapy. *Med. Phys.* **41**, 011715 (2014)
2. Jenne, J.W., Preusser, T., Günther, M.: High-intensity focused ultrasound: principles, therapy guidance, simulations and applications. *Zeitschrift Für Medizinische Physik* **22**, 311–322 (2012)

3. Okada, A., et al.: A case of hepatocellular carcinoma treated by MR-guided focused ultrasound ablation with respiratory gating. *Magn. Reson. Med. Sci. Mrms Off. J. Jpn. Soc. Magn. Reson. Med.* **5**, 167 (2006)
4. Kothary, N., Dieterich, S., Louie, J.D., Chang, D.T., Hofmann, L.V., Sze, D.Y.: Percutaneous implantation of fiducial markers for imaging-guided radiation therapy. *AJR Am. J. Roentgenol.* **192**, 1090–1096 (2009)
5. de Senneville, B.D., Mougnot, C., Moonen, C.T.: Real-time adaptive methods for treatment of mobile organs by MRI-controlled high-intensity focused ultrasound. *Magn. Reson. Med.* **57**, 319–330 (2007)
6. Rank, C.M., et al.: 4D respiratory motion-compensated image reconstruction of free-breathing radial MR data with very high undersampling. *Magn. Reson. Med.* **77**, 1170 (2016)
7. De Luca, V., Tschannen, M., Székely, G., Tanner, C.: A learning-based approach for fast and robust vessel tracking in long ultrasound sequences. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (eds.) *MICCAI 2013. LNCS*, vol. 8149, pp. 518–525. Springer, Heidelberg (2013). [https://doi.org/10.1007/978-3-642-40811-3\\_65](https://doi.org/10.1007/978-3-642-40811-3_65)
8. Chuang, B., Hsu, J.H., Kuo, L.C., Jou, I., Su, F.C., Sun, Y.N.: Tendon-motion tracking in an ultrasound image sequence using optical-flow-based block matching. *Biomed. Eng. Online* **16**, 47 (2017)
9. Carletti, M., Dall’Alba, D., Cristani, M., Fiorini, P.: A robust particle filtering approach with spatially-dependent template selection for medical ultrasound tracking applications. In: 11th International Conference on Computer Vision Theory and Applications, pp. 522–531. SCITE Press, Rome (2016)
10. Royer, L., Krupa, A., Dardenne, G., Le, B.A., Marchand, E., Marchal, M.: Real-time target tracking of soft tissues in 3D ultrasound images based on robust visual information and mechanical simulation. *Med. Image Anal.* **35**, 582–598 (2017)
11. Henriques, J.F., Caseiro, R., Martins, P., Batista, J.: High-speed tracking with kernelized correlation filters. *IEEE Trans. Pattern Anal. Mach. Intell.* **37**, 583–596 (2015)
12. Wu, Y., Lim, J., Yang, M.H.: Online object tracking: a benchmark. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2411–2418. IEEE press, Portland (2013)
13. Li, Y., Zhu, J.: A scale adaptive kernel correlation filter tracker with feature integration. In: Agapito, L., Bronstein, Michael M., Rother, C. (eds.) *ECCV 2014. LNCS*, vol. 8926, pp. 254–265. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-16181-5\\_18](https://doi.org/10.1007/978-3-319-16181-5_18)
14. Liang, T., Yung, L., Yu, W.: On feature motion decorrelation in ultrasound speckle tracking. *IEEE Trans. Med. Imaging* **32**, 435–448 (2016)
15. Lei, P., Moeslein, F., Wood, B.J., Shekhar, R.: Real-time tracking of liver motion and deformation using a flexible needle. *Int. J. Comput. Assist. Radiol. Surg.* **6**, 435–446 (2011)
16. Luca, V.D., et al.: The 2014 liver ultrasound tracking benchmark. *Phys. Med. Biol.* **60**, 5571–5599 (2015)
17. Ozkan, E., Tanner, C., Kastelic, M., Mattausch, O., Makhinya, M., Goksel, O.: Robust motion tracking in liver from 2D ultrasound images using supporters. *Int. J. Comput. Assist. Radiol. Surg.* **12**, 941–950 (2017)
18. Williamson, T., Cheung, W., Roberts, S.K., Chauhan, S.: Ultrasound-based liver tracking utilizing a hybrid template/optical flow approach. *Int. J. Comput. Assist. Radiol. Surg.* **13**, 1–11 (2018)