Published in final edited form as:
Proc SPIE. 2013 March 13; 8669: . doi:10.1117/12.2000569.

Involuntary motion tracking for medical dynamic infrared thermography using a template-based algorithm

Tze-Yuan Cheng and **Cila Herman**
Department of Mechanical Engineering, Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD, USA 21218-2682

Abstract

In medical applications, Dynamic Infrared (IR) Thermography is used to detect the temporal variation of the skin temperature. Dynamic Infrared Imaging first introduces a thermal challenge such as cooling on the human skin, and then a sequence of hundreds of consecutive frames is acquired after the removal of the thermal challenge. As a result, by analyzing the temporal variation of the skin temperature over the image sequence, the thermal signature of skin abnormality can be examined. However, during the acquisition of dynamic IR imaging, the involuntary movements of patients are unavoidable, and such movements will undermine the accuracy of diagnosis. In this study, based on the template-based algorithm, a tracking approach is proposed to compensate the motion artifact. The affine warping model is adopted to estimate the motion parameter of the image template, and then the Lucas-Kanade algorithm is applied to search for the optimized parameters of the warping function. In addition, the weighting mask is also incorporated in the computation to ensure the robustness of the algorithm. To evaluate the performance of the approach, two sets of IR image sequences of a subject's hand are analyzed: the steady-state image sequence, in which the skin temperature is in equilibrium with the environment, and the thermal recovery image sequence, which is acquired after cooling is applied on the skin for 60 seconds. By selecting the target region in the first frame as the template, satisfactory tracking results were obtained in both experimental trials, and the robustness of the approach can be effectively ensured in the recovery trial.

Keywords

dynamic infrared imaging; motion tracking; template-based algorithm

1. INTRODUCTION

1.1 Background

Since the 1990s, following the development of IR camera technology, the novel algorithms of image processing, and the progression in IR sensors, medical IR thermography, a totally non-ionizing and non-invasive imaging modality, is regaining more interest in clinical medicine.

As compared with conventional passive IR imaging, which examines the spatial temperature information on a single image frame, dynamic IR imaging detects not only spatial but also temporal information regarding skin temperature. Dynamic IR imaging first introduces a thermal challenge such as cooling or heating on the human skin, and then a sequence of hundreds of consecutive frames is acquired after the removal of the thermal challenge. As a

Correspondence to: Tze-Yuan Cheng.

result, by analyzing the temporal variation of the skin temperature in the range of hundreds millikelvins, abnormalities such as the malignancy of the skin can be examined [1].

However, during the acquisition period of dynamic IR imaging, involuntary movements of patients are unavoidable, and such movements will produce motion artifacts, which hinder the accurate thermal comparison between two specific points on the skin, therefore undermining the validity of the local temperature analysis. Therefore, in order to screen the abnormalities of tissue accurately in dynamic IR imaging, it is crucial to track the target motion during the acquisition period.

1.2 Literature review

Moving target tracking in the video sequence is generally an active topic in civil and military applications. Due to the lower cost and fast improvement of infrared (IR) technology, object tracking has also been widely used in IR thermography, such as pedestrian detection for surveillance purposes [2–4].

In medical applications, however, only a few researchers have tackled the tracking issue specifically for dynamic IR imaging. To allow the wide use of dynamic IR imaging in breast cancer screening [5–7], based on marker-based image registration, the motion artifact reduction approach using image sequence realignment was proposed [8–10]. In this approach, the registration of each IR image frame is achieved by aligning 5 to 18 markers, which are visible to serve as control points in IR sequence, and the approach can achieve accurate tracking performance in terms of SNR evaluation. However, despite the fact that algorithms are designed for marker localization and control points matching in the approach [8], the registration error due to the mismatch of control points is still not totally eliminated.

Therefore, in order to achieve robust tracking in dynamic IR imaging, based on the brightness constancy of the marker image and the surrounding skin over the entire IR sequence, a template-based tracking approach is presented in this study to correct the motion artifacts in clinical environments.

2. METHOD

2.1 Template-based tracking algorithm

Template-based tracking is achieved by estimating the warp of a template image in an image sequence, and in this paper, the method is described using the notation of Mathews et al[11]. Given that a template image ***T(x)*** is the region of interest, we want to track in a target image ***I(x)***where ***x*** represents the pixel coordinates vector, ***W(x;p)*** is the warp function parameterized in terms of parameters $\boldsymbol{p} = (p^1,\ldots p^n)^T$. In the algorithm, all pixels ***x*** in ***I(x)*** will be warped by the function ***W(x;p)*** to be compared with ***T(x)***, and the optimized parameters ***p*** will be determined such that:

$$\boldsymbol{I}(\boldsymbol{W}(\boldsymbol{x};\boldsymbol{p})) \approx \boldsymbol{T}(\boldsymbol{x}) \quad (1)$$

Since the template image patch is moving in 3D, the affine warp ***W(x;p)*** is utilized for the template-based tracking:

$$\boldsymbol{W}(\boldsymbol{x};\boldsymbol{p}) = \begin{pmatrix} \mathbf{1}+\boldsymbol{p_1} & \boldsymbol{p_3} & \boldsymbol{p_5} \\ \boldsymbol{p_2} & \mathbf{1}+\boldsymbol{p_4} & \boldsymbol{p_6} \end{pmatrix} \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \\ \mathbf{1} \end{pmatrix} \quad (2)$$

Where $\boldsymbol{p}=(p_1, p_2, p_3, p_4, p_5, p_6,)^T$ are the 6 parameters for affine transformation.

2.2 Lucas-Kanade approach

Based on the Lucas-Kanade algorithm [12], in an attempt to search for the optimized parameters ***p*** for warping function, the problem can be formulated as the minimization of the sum of squared errors between the warped image ***I*** and the template patch ***T***:

$$\Sigma_{\boldsymbol{x}}[\boldsymbol{I}(\boldsymbol{W}(\boldsymbol{x};\boldsymbol{p})) - \boldsymbol{T}(\boldsymbol{x})]^2 \quad (3)$$

This minimization is a non-linear optimization problem, and based on a predicted current estimate of ***p***, the Lucas-Kanade algorithm iteratively solves for an increment ***Δp*** to update the current estimate ***p*** by:

$$\boldsymbol{p} \leftarrow \boldsymbol{p}+\boldsymbol{\Delta p} \quad (4)$$

and the expression (3) can be re-written as:

$$\Sigma_{\boldsymbol{x}}[\boldsymbol{I}(\boldsymbol{W}(\boldsymbol{x};\boldsymbol{p}+\boldsymbol{\Delta p})) - \boldsymbol{T}(\boldsymbol{x})]^2 \quad (5)$$

As a result, the iteration will continue until the updated estimate of ***p*** converges, which serves as the optimized set of parameters ***p*** for ***W(x;p)*** at the end of the computation. Furthermore, to linearize the error function (5), the Lucas-Kanade algorithm adopts the first-order Taylor expansion to approximate ***I(W(x;p+Δp))***:

$$\boldsymbol{I}(\boldsymbol{W}(\boldsymbol{x};\boldsymbol{p}+\boldsymbol{\Delta p}))=\boldsymbol{I}(\boldsymbol{W}(\boldsymbol{x};\boldsymbol{p}))+\nabla \boldsymbol{I}\frac{\partial \boldsymbol{W}}{\partial \boldsymbol{p}}\boldsymbol{\Delta p} \quad (6)$$

Therefore, by substituting ***I(W(x;p+Δp))*** into (5), the error function can be written as[13]:

$$\Sigma_{\boldsymbol{x}}\left[\boldsymbol{I}(\boldsymbol{W}(\boldsymbol{x};\boldsymbol{p}))+\nabla \boldsymbol{I}\frac{\partial \boldsymbol{W}}{\partial \boldsymbol{p}}\boldsymbol{\Delta p} - \boldsymbol{T}(\boldsymbol{x})\right]^2 \quad (7)$$

where **Δ*I*** is the gradient of image ***I*** evaluated when the current estimate warping function ***W(x;p)*** is applied, and $\frac{\partial \boldsymbol{W}}{\partial \boldsymbol{p}}$ represents the Jacobian of the warp. For affine transformation, ***W(x;p)*** can be decomposed as : $\boldsymbol{W(x;p)=(W_x, W_y)^T}$and the Jacobian of ***W(x;p)*** is represented as:

$$\frac{\partial \boldsymbol{W}}{\partial \boldsymbol{p}}=\begin{pmatrix} \frac{\partial \boldsymbol{W_x}}{\partial \boldsymbol{p_1}} \cdots \frac{\partial \boldsymbol{W_x}}{\partial \boldsymbol{p_6}} \\ \frac{\partial \boldsymbol{W_y}}{\partial \boldsymbol{p_1}} \cdots \frac{\partial \boldsymbol{W_y}}{\partial \boldsymbol{p_6}} \end{pmatrix}=\begin{pmatrix} \boldsymbol{x0y010} \\ \boldsymbol{0x0y01} \end{pmatrix} \quad (8)$$

Based on equation (8), to minimize the error function (7), we can first get its partial derivative with respect to ***Δp***

$$\boldsymbol{2\Sigma_x}\left[\nabla \boldsymbol{I}\frac{\partial \boldsymbol{W}}{\partial \boldsymbol{p}}\right]^{\boldsymbol{T}}\left[\boldsymbol{I}(\boldsymbol{W}(\boldsymbol{x};\boldsymbol{p}))+\nabla \boldsymbol{I}\frac{\partial \boldsymbol{W}}{\partial \boldsymbol{p}}\boldsymbol{\Delta p} - \boldsymbol{T}(\boldsymbol{x})\right] \quad (9)$$

By setting the expression (9) to zero, the closed-form solution for the ***Δp*** can be obtained [13]:

$$\Delta p = H^{-1} \Sigma_x \left[\nabla I \frac{\partial W}{\partial p} \right]^T [T(x) - I(W(x;p))] \quad (10)$$

where H is the $n \times n$ Hessian Matrix, and the Gauss-Newton approximation is adopted in Lucas-Kanade algorithm for Hessian matrix:

$$H = \Sigma_x [\nabla I \frac{\partial W}{\partial p}]^T [\nabla I \frac{\partial W}{\partial p}] \quad (11)$$

2.3 Infrared image acquisition

The simulated target attached to the skin surface is imaged using the Merlin midwave (3–5μm) infrared camera (MWIR) as shown in Fig. 1(a). The temperature resolution of the camera is 0.025 °C, and a 320 by 256 pixel focal plane array (FPA) is used to acquire 16 bit raw data at frame rate of 60 Hz. Each image frame has the field of view (FOV) of 22 × 16 degrees. In our experimental trials, we recorded two sets of IR image sequence on a subject's hand: the first one is the steady-state image sequence, and the other one is the recovery image sequence, which is acquired after gel-pack cooling is applied on the skin for 60 seconds. In both sets of the IR sequences, the subject's hand movement is deliberately incorporated.

2.4 Tracking marker and simulated lesion

In general, abnormal regions on the skin are not directly identifiable in the IR image. To test the performance of the tracking method proposed in this study, we created two paper markers which are both visible in the IR image (Fig .1 (c) (d)): one is the rectangular marker of size: 5 cm × 5 cm, which serves as the template image as shown in Fig.1 (d), and the other is the round maker used to simulate the location of lesion target on the skin.

2.5 Weighting mask of the template image

The appearance constancy over the entire image sequence is the primary assumption for template-based tracking, whereas this assumption is not feasible in many circumstances where partial occlusions and brightness variations are unavoidable. The assumption might roughly hold for steady-state IR image, however, for dynamic IR image sequences, in which the time-dependent temperature variation are critical for our analysis, this assumption will be violated. Therefore, the robust template tracking algorithm proposed in [14, 15] is adopted to tackle the temperature changes in dynamic IR image data. The robust algorithm incorporates a weighting mask, which selectively suppresses the region with brightness variation in the template image. In our application, according to the highest and lowest temperature value, the template image intensity is normalized to (0, 1) and serves as the weighting mask. The two weighting masks for both steady-state and recovery IR sequence are shown in Figure 2. In this situation, the skin surface will be highly weighted in the tracking computation, while the region receiving cooling, which is also the region with highest variation rate in the recovery IR sequence, will be suppressed in the tracking computation.

2.6 Determining the location of simulated lesion for comparison purposes

For the purpose of evaluating the tracking performance of the proposed method, we need to know the actual position of the simulated lesion as the reference point. After the optimized tracking parameters are determined in the following IR frames respectively, the centroid coordinate of in the first IR frame will be predicted at each subsequent IR frame, and this

predicted location and the actual centroid location will be compared to analyze the tracking errors. To determine the actual location of simulated lesion, a contour segmentation algorithm, random walker [16], is adopted in steady-stage IR sequence. For the recovery IR sequence, in which the temperature changes dynamically, an IR-IR registration technique is applied.

The random walker algorithm [16] is first implemented on the steady-stage IR image sequence, in which the skin temperature is in equilibrium with the room temperature around 27 °C, and the simulated lesion contour is then identifiable over the entire sequence. Therefore, the location of the simulated lesion can be segmented in each IR image frame as shown in Figure 3. After the lesion region is delineated, its centroid is chosen as the reference point to calculate the Euclidean distance with respect to the predicted lesion position from the tracking algorithm.

In our previous clinical application [17], the recovery IR images after cooling reveal critical information to detect the malignancy of melanoma, thus it is of interest to test the proposed tracking method on the recovery IR sequence after cooling. Since the simulated lesion will be covered by the cooling spot and become invisible in IR image during recovery phase, we have to register the simulated lesion outlines via the location of four corners of the square marker, which is visible in both steady-state and recovery IR image sequences.

As shown in Figure 4 (a), the four corners of square marker and the segmented lesion contour are first identified in the steady-state IR image; subsequently, the corresponding points are then manually selected in the recovery IR image (Figure 4 (b)). Based on the four pairs of coordinate correspondences, a 2D projective transformation matrix is solved [18] for the registration between the steady-state and recovery IR image frame, and therefore the simulated lesion contour segmented in steady-state IR image can be mapped to recovery IR image. By applying the registration process from the steady-state IR image to each frame of recovery sequence consecutively, the actual location of simulated lesion location can be determined for the tracking error analysis.

3. RESULTS

3.1 Tracking performance in the steady-state IR sequence

We tested 23 frames of steady-state IR images with random in-plane motion incorporated, and four representative frames are shown in Figure 5. As we compare the location of predicted (circle) with actual (cross) centroids of the simulated lesion, it can be seen that despite the random in-plane motion, the two locations of circle and cross coincide with only slight differences.

The tracking error is evaluated quantitatively: the Euclidean distance between the predicted and actual location of target centroids in units of pixel number is calculated at each frame, and the results are recorded in Figure 6. As evident in Figure 6, the distance between the predicted and actual centroid locations is within 3 pixels. From the actual size of marker, the 30 mm length is approximately equivalent to 86 pixels in in the IR image, which means the distance of 1 pixel in the image approximately quals to 0.35 mm on skin surface.

3.2 Tracking performance for the dynamic IR sequence after cooling

The tracking performance for the dynamic IR sequence recovered from cooling application is illustrated in Figure 7. The frame rate is 1 frame/second in this trial. In the first trial, only the first 23 frames are tested. It can be observed that despite the temperature variation over the sequence, the centorid location of the simulated lesion can still be tracked with reasonable accuracy up to 23 seconds into recovery phase. As shown in Figure 8, the

quantitative analysis was also applied to the 23 frames of the IR sequence, and it is evident that the highest tracking errors are less than 4 pixels. The results suggest that by adopting the first frame after cooling as the template image, along with its normalized image as weighting mask, the template-based tracking algorithm is capable of tackling the transient temperature change in the first 23 seconds in recovery phase. However, as we extended the duration of the testing over 30 frames in the trial, as shown in Figure 9, it appears that the tracking errors begin to grow significantly. It can be seen from the error analysis in Figure 10, a remarkable growth of tracking errors is detected after 30 frames. The error growth can be explained by the fact that after 30 seconds in recovery phase, the thermal contour of the target becomes substantially different from the template image taken at the first frame. As indicated from the Figure 10, the tracking errors will exceed 10 pixels after 40 frames. Therefore, adjustments of the tracking method are needed when significant changes occur in the image sequence.

4. CONCLUSION

In this study, we demonstrated the feasibility of template-based algortihm applied to involuntary motion compensation in dynamic infrared imaging. Our aim was to test if the algorithm can track the target region in the IR image in the presence of involuntary movements. From the experimental trials using steady-state with random in-plane motion, satisfying results were obtained by selecting the target region at the first frame as template image. In the recovery IR image trials, it can be inferred that the weighting mask converted from its normalized image can effectively ensure the robustness for 30 frames; however, its robustness would collapse if the recovery time lasts longer than 40 seoncds, when the temperature contour is significantly different from the template image taken at the initial frame right after cooling.

5. DISCUSSION

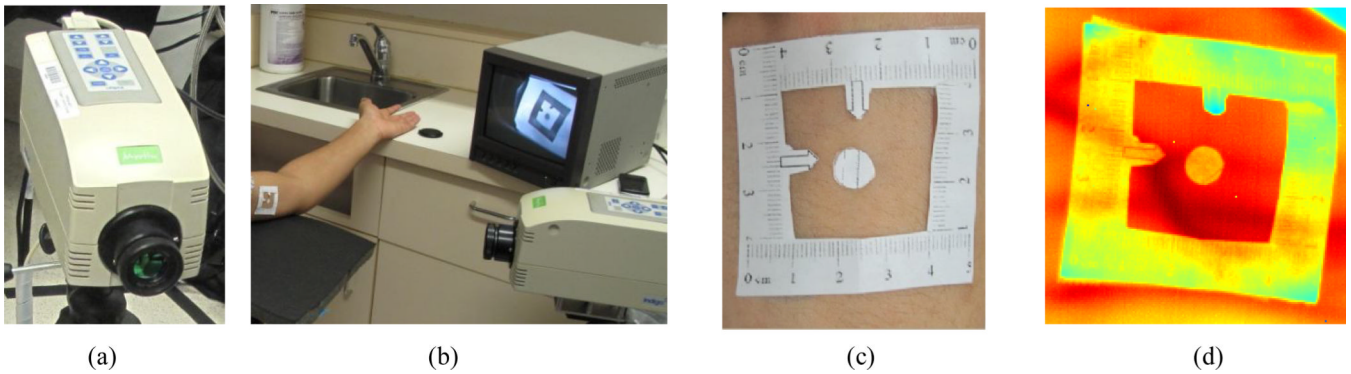
From the error analysis for the three trials, we can conclude that the weighing mask converted from normalized initial frame can stabilize the motion tracking in steady-state IR sequence. For the recovery sequence with transient temperature change, although the intensity changes dramatically, the algorithm can still function effectively for the first 30 seconds. During this period the temperature change is at highest rate. It is not surprising to see the fast error growth after 40 seconds, since then the skin temperature is substantially recovered, and the thermal contour appears to be more similar to the steady state than the template image. Therefore, it is speculated that if the template image can be updated using the image frame at 30 seconds, which has the thermal contour closer to the following frames, the tracking algorithm is expected to function with reasonable improvement. On the other hand, for verification purposes, the lesion location is made to be visible using simulated marker in this study; in the clinical environment, however, since the real lesion is not necessarily visible in IR image, a pin-point technique is desired to identify lesion location before implementing the method.

Acknowledgments

This research was funded by the National Institutes of Health NCI Grant No. 5R01CA161265-02, the National Science Foundation Grant No. 0651981 and the Alexander and Margaret Stewart Trust through the Cancer Center of the Johns Hopkins University. The author would also like to thank Akanksha Bhargava for the assistance in the imaging experiments.

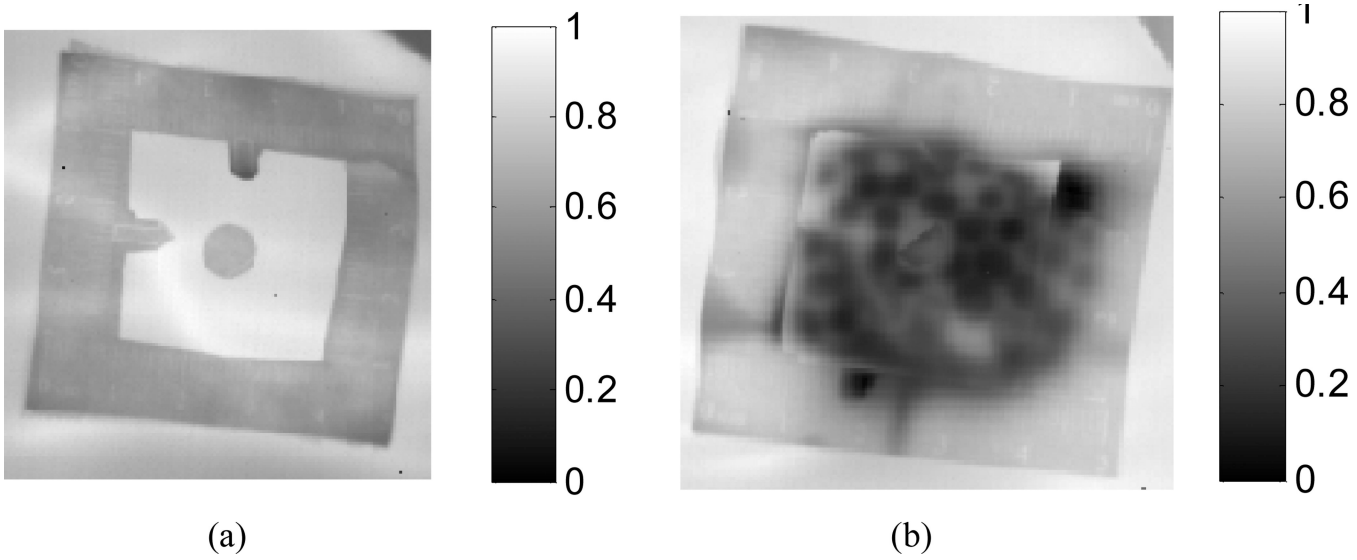
REFERENCES

1. Button TM, Li H, Fisher P, et al. Dynamic infrared imaging for the detection of malignancy. Phys. Med. Biol. 2004; 49(14):3105–3116. [PubMed: 15357184]
2. Dai C, Zheng Y, Li X. Layered representation for pedestrian detection and tracking in infrared imagery. Proc. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR). 2005
3. Yasuno M, Yasuda N, Aoki M. Pedestrian detection and tracking in far infrared images. Proc. CVPRW. 2004; 27((02)):125–125.
4. Zhou J, Hoang J. Real-time robust human detection and tracking system. Proc. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR). 2005
5. Anbar M, Brown C, Milescu L, et al. The potential of dynamic area telethermometry in assessing breast cancer - Emphasizing physiological information in the diagnosis and management of breast cancer. IEEE Eng. Med. Biol. Mag. 2000; 19(3):58–62. [PubMed: 10834117]
6. Anbar M, Milescu L, Naumov A, et al. Detection of cancerous breasts by dynamic area telethermometry. IEEE Eng. Med. Biol. Mag. 2001; 20(5):80–91. [PubMed: 11668900]
7. Anbar M. Assessment of physiologic and pathologic radiative heat dissipation using dynamic infrared imaging. Visualization and Imaging in Transport Phenomena (Ann. Ny. Acad. Sci.). 2002; 972:111–118.
8. Agostini V, Delsanto S, Molinari F, et al. Evaluation of feature-based registration in dynamic infrared imaging for breast cancer diagnosis. Conf. Proc. IEEE Eng. Med. Biol. Soc. 2006; 1:953–956. [PubMed: 17946428]
9. Agostini V, Delsanto S, Knaflitz M, et al. Noise estimation in infrared image sequences: a tool for the quantitative evaluation of the effectiveness of registration algorithms. IEEE Trans. Biomed. Eng. 2008; 55(7):1917–1920. [PubMed: 18595811]
10. Agostini V, Knaflitz M, Molinari F. Motion artifact reduction in breast dynamic infrared imaging. IEEE Trans. Biomed. Eng. 2009; 56(3):903–906. [PubMed: 19272933]
11. Matthews I, Ishikawa T, Baker S. The template update problem. IEEE Trans. Pattern Anal. Mach. Intell. 2004; 26(6):810–815. [PubMed: 18579941]
12. Lucas B, Kanade T. An iterative image registration technique with an application to stereo vision. Proc. 7th IJCAI. 1981; 2:674–679.
13. Baker S, Matthews I. Lucas-Kanade 20 years on: A unifying framework. Int. J. Comput. Vis. 2004; 56(3):221–255.
14. Hager GD, Belhumeur PN. Efficient region tracking with parametric models of geometry and illumination. IEEE Trans. Pattern Anal. Mach. Intell. 1998; 20(10):1025–1039.
15. Ishikawa T, Matthews I, Baker S. Efficient image alignment with outlier rejection. technical report CMU-RI-TR-02-27, Robotics Institute, Carnegie Mellon University. 2002
16. Grady L. Random walks for image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 2006; 28(11):1768–1783. [PubMed: 17063682]
17. Cetingul MP, Herman C. Quantification of the thermal signature of a melanoma lesion. Int. J. Therm. Sci. 2011; 50(4):421–431.
18. Hartley, R.; Zisserman, A. [Multiple View Geometry in computer vision]. Cambridge University Press; 2004. p. 32-33.



(a) (b) (c) (d)

Figure 1.
(a) Merlin midwave infrared camera used for IR image acquisition with (b) monitor; (c) the two paper adhesive markers; the square one serves as the tracking template, and the round one represents the simulated lesion for error analysis purposes; (d) the template image used for motion-tracking computations



(a) (b)

Figure 2.
(a) Weighting mask of template image for steady-state IR image sequence (b) weighting mask of template image for the recovery IR image sequence (after cooling is applied)

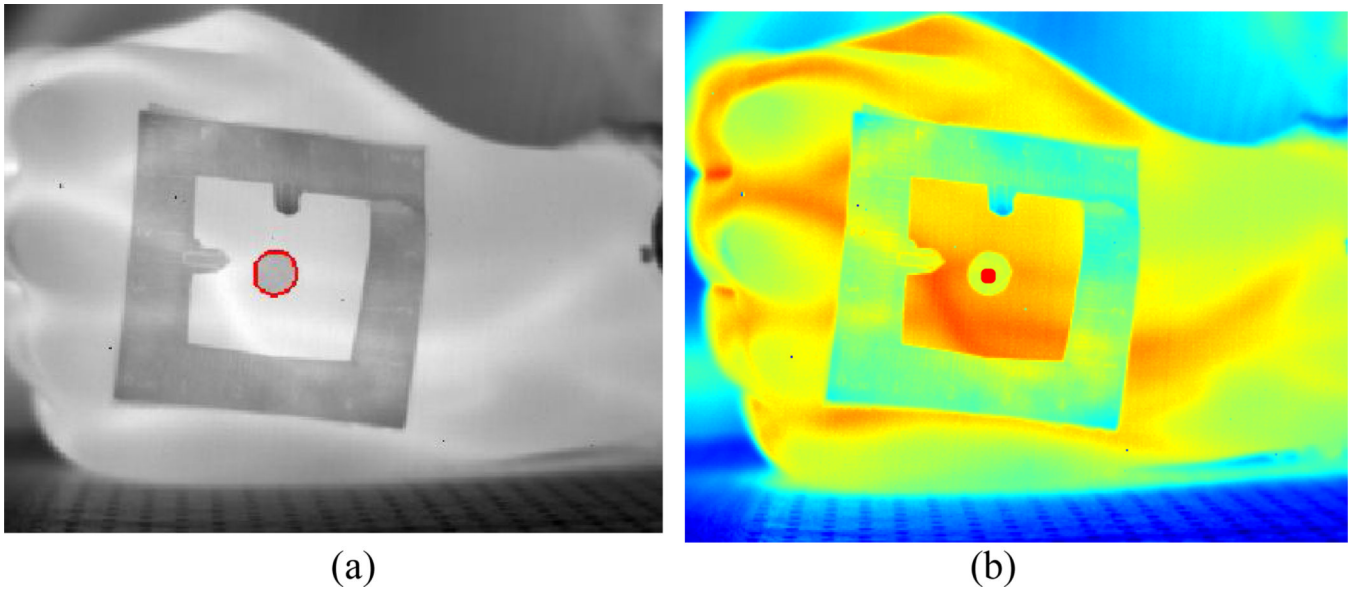


Figure 3.
(a) Adhesive markers and the segmented outline of the simulated lesion (red circle); (b) red point: the centroid location of the simulated lesion

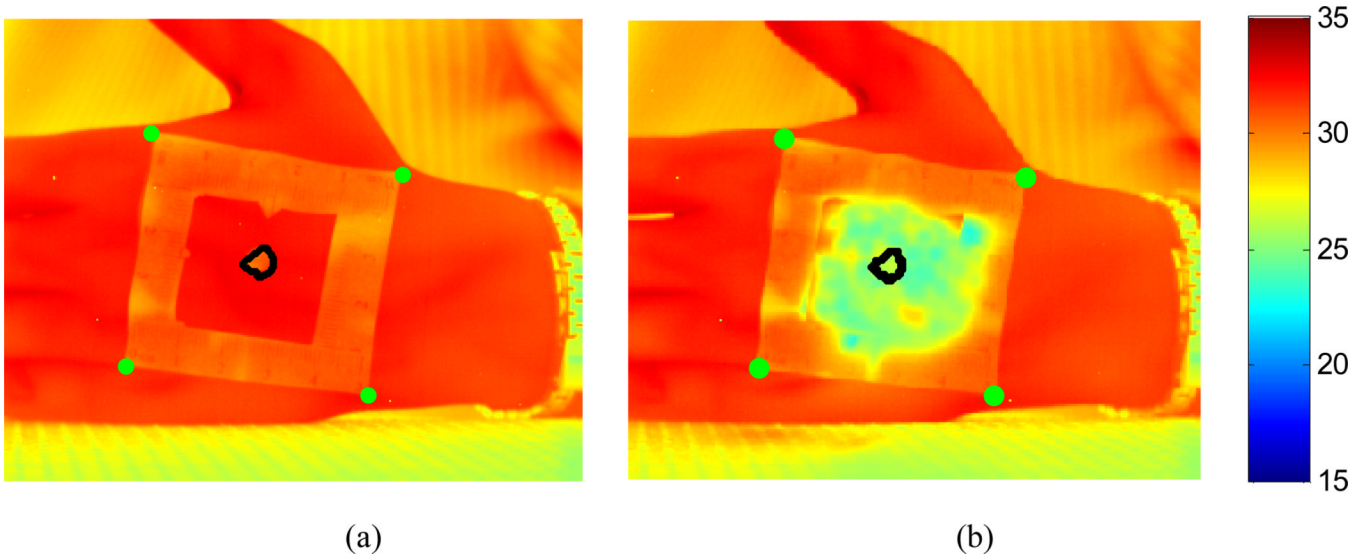


Figure 4.
The registration of the simulated lesion from the steady-state IR image (a) to the recovery IR image (b)

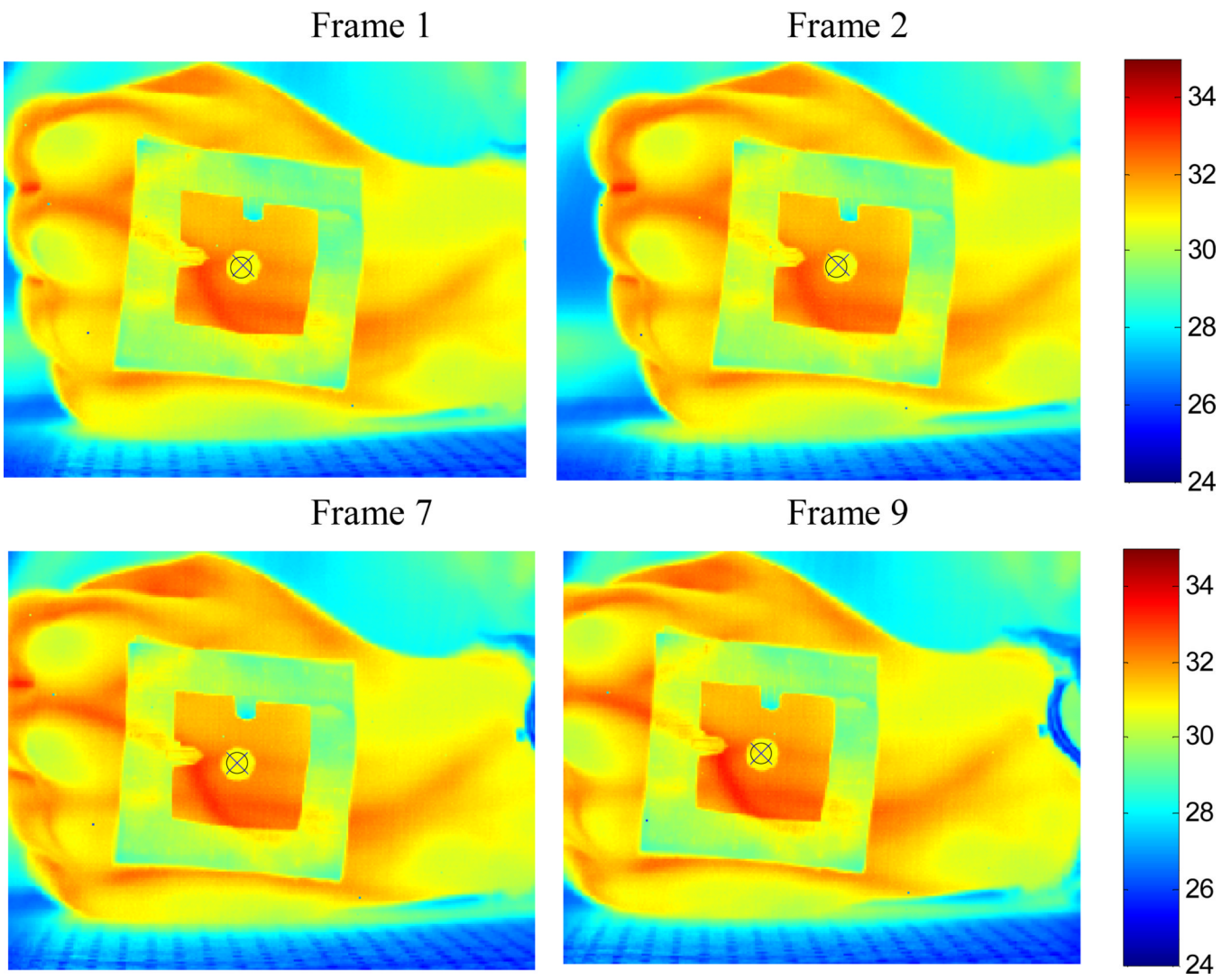


Figure 5.
The tracking performance in steady-state IR image sequence (circle: the centroid location predicted by the tracking algorithm; cross: the actual centroid location of the simulated lesion)

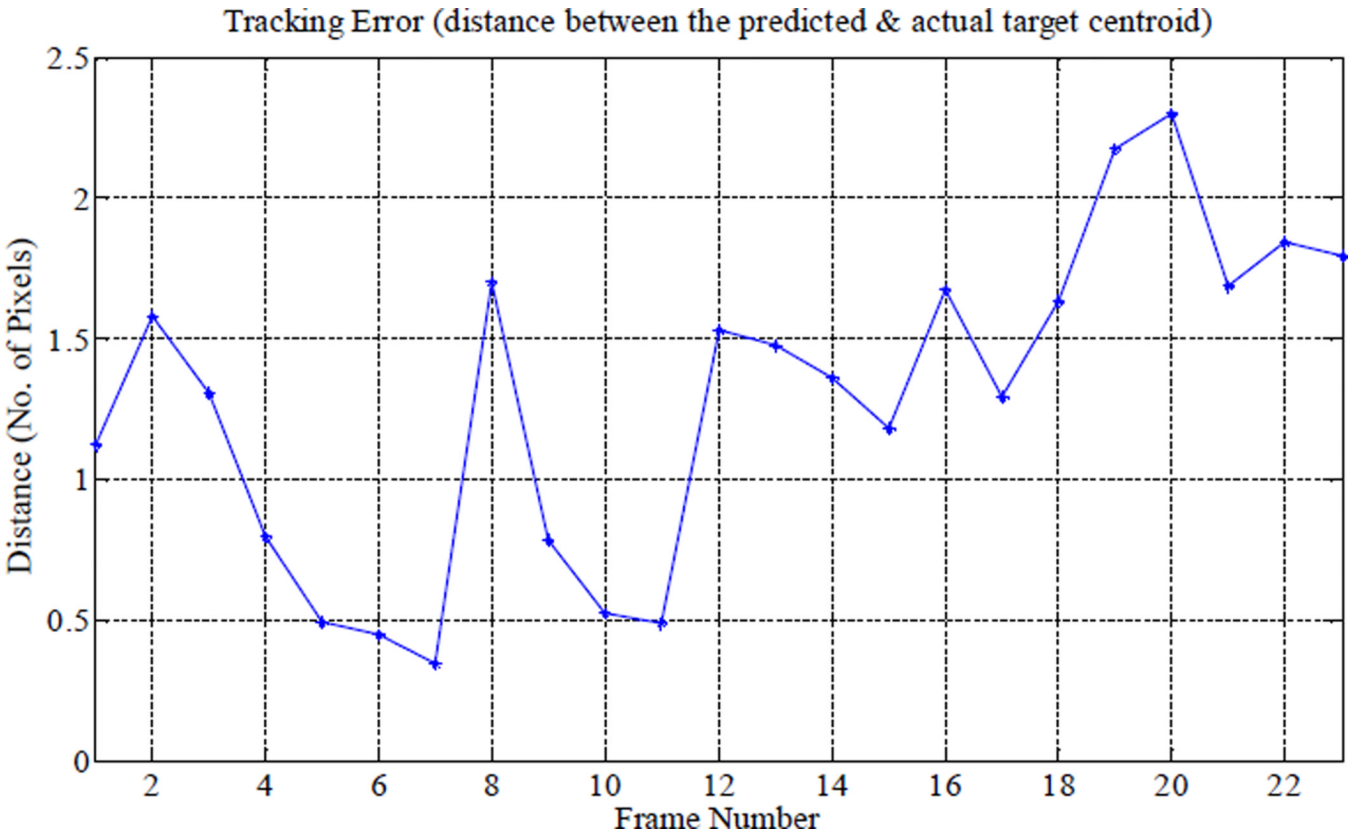


Figure 6.
Tracking error analysis: the distance between the predicted and actual location of simulated lesion centroid in 23 consecutive steady-state IR image frames

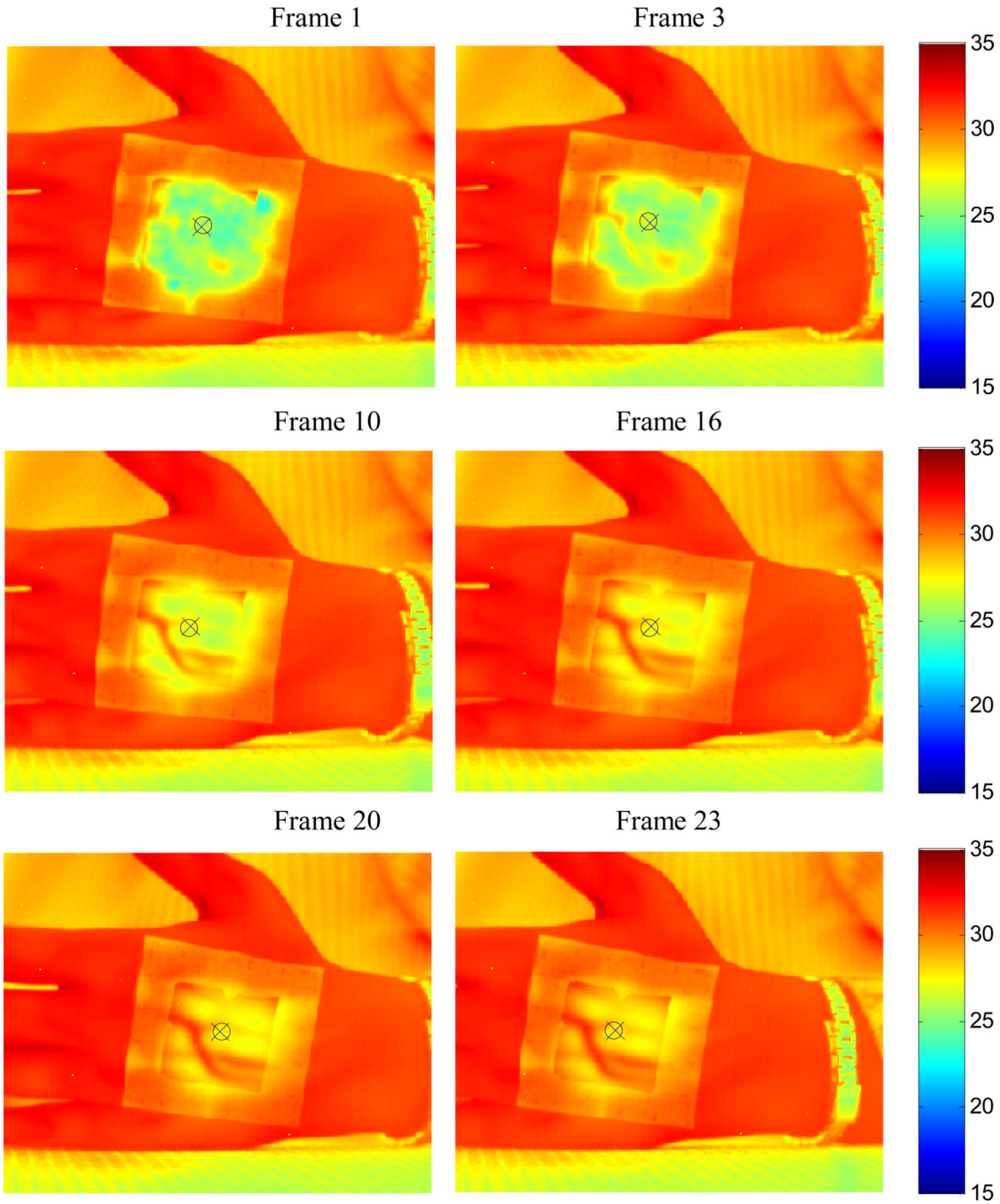


Figure 7.
The tracking performance in the first 23 frames in the recovery IR image sequence (circle: the centroid location predicted by the tracking algorithm; cross: the actual centroid location of the simulated lesion)

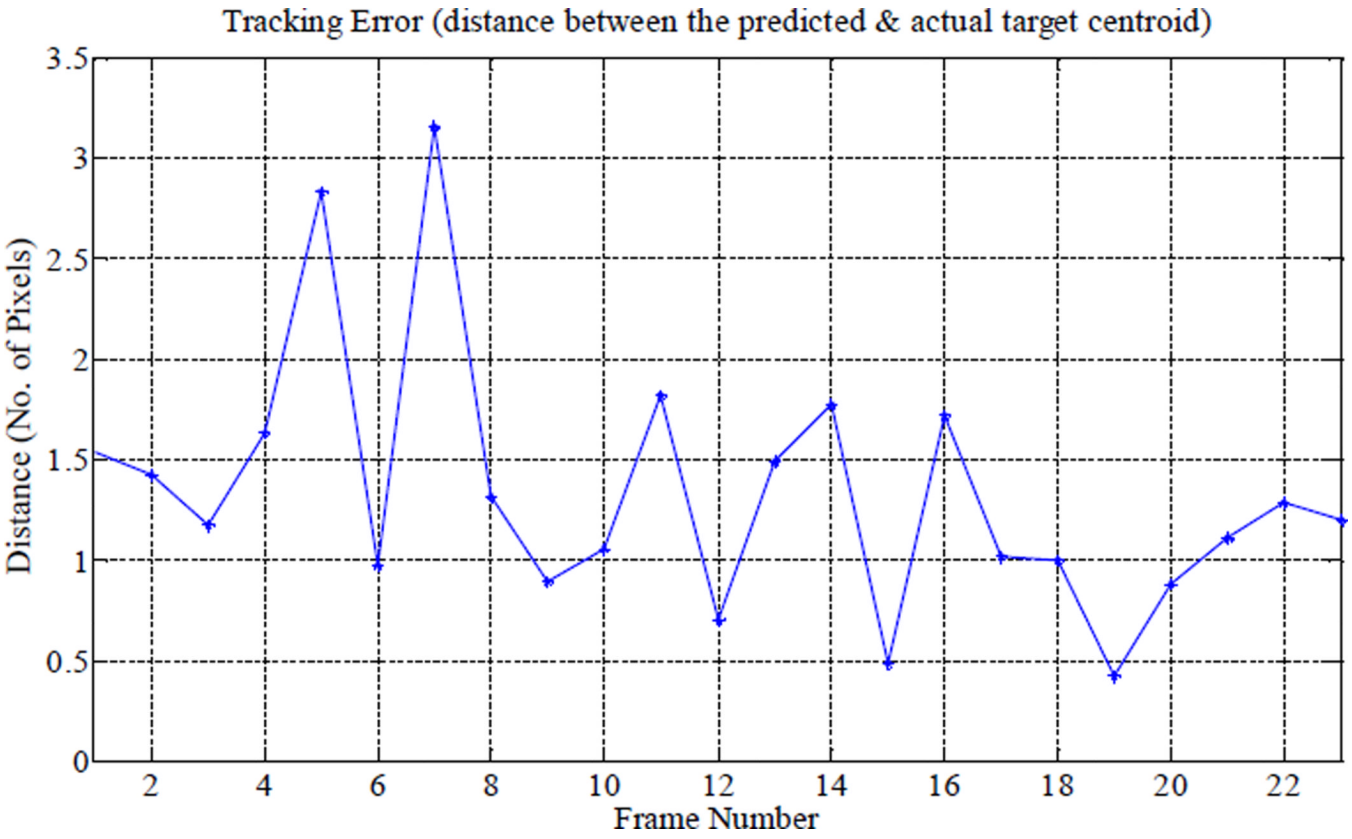


Figure 8.
Tracking error analysis: the distance between the predicted and actual location of simulated lesion centroid from Frame 1 to Frame 23 in the recovery IR image sequence

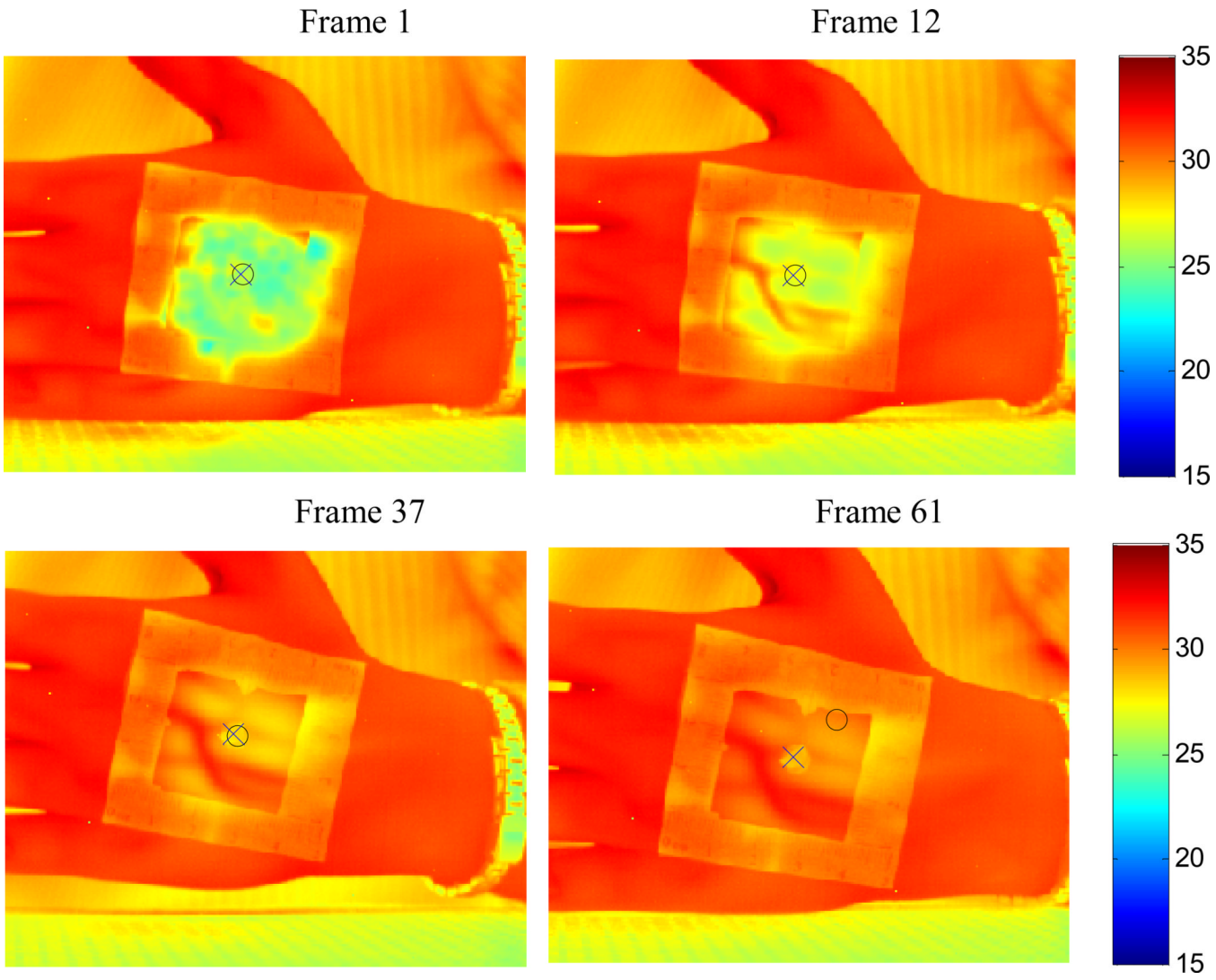


Figure 9.
The tracking performance from Frame 1 to Frame 61 in the recovery IR image sequence (circle: the centroid location predicted by the tracking algorithm; cross: the actual centroid location of the simulated lesion)

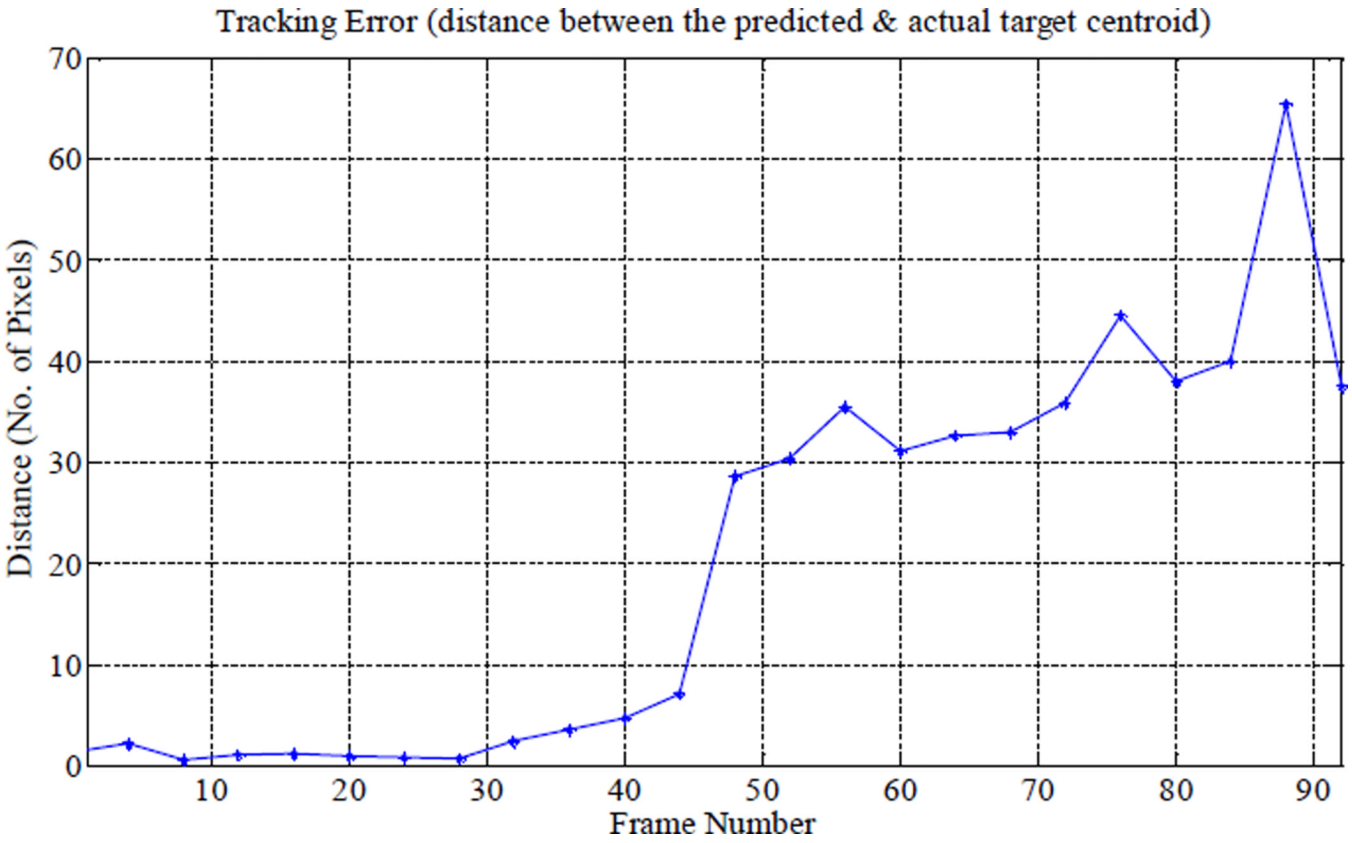


Figure 10.
Tracking error analysis: the distance between the predicted and actual location of simulated lesion centroid from Frame 1 to Frame 92 in the recovery IR image sequence