Target Tracking in Medical Images Using Template Matching

著者	Zhang Xiaoyong, Homma Noriyasu
journal or publication title	東北大学医学部保健学科紀要
volume	23
number	1
page range	9-15
year	2014-01-31
URL	http://hdl.handle.net/10097/57147

東北大医保健学科紀要　**23**(1)：9〜15，2014

Review Article

Target Tracking in Medical Images Using Template Matching

Xiaoyong ZHANG and Noriyasu HOMMA

Department of Radiological Imaging and Informatics, Tohoku University Graduate School of Medicine
(*E-mail : xiaoyong@ieee.org*)

テンプレートマッチングを用いた医用画像の目標追跡

張　　暁 勇，本 間 経 康

東北大学大学院医学系研究科　医用画像工学分野

Key words : Visual tracking, template matching, mean squared error, correlation coefficient, correlation ratio, mutual information, radiation therapy, fluoroscopic image

Tracking or recognizing an object in medical image is one of the most critical tasks in many medical image-related applications, such as image-guided radiation therapy (IGRT), computer-aided diagnosis (CAD), etc. Template matching, as one of the common-used techniques in computer vision, has widely been applied to these applications due to its simplicity and flexibility. The goal of this note is to introduce the principles of the template matching technique, explain several similarity matrices used in the template matching, and give an example of the template matching for tracking a tumor target in fluoroscopic image sequence for IGRT. In addition, we will also give some experimental results to evaluate the performance of the template matching based on different similarity in terms of the accuracy and computation time.

1. Introduction

With the development of medical imaging technology, medical image processing has been involved in many clinical applications, ranging from diagnosis, treatment to patient follow-up. In these applications, automatic tracking or recognizing an object in the medical images is an important task. For example, in image-guided radiation therapy (IGRT), a tumor tracking system, which is capable of tracking the respiration-induced tumor motion, allows clinician to deliver high-dose treatment beam to a moving target for improving the treatment efficiency and reducing the risk to healthy tissues around the tumor [1, 2].

As one of the common-used techniques, template matching has widely been applied to various applications due to its simplicity and flexibility. Generally, the template matching is a technique in digital image processing for finding small parts of an observed image which match a template image [3].

In this note, we will introduce the principles of the template matching technique, explain different similarity metrics used in the template matching, and give an example of the template matching for tracking a tumor target in fluoroscopic image sequence for IGRT. We will evaluate the performance of the template matching based on different similarity metrics in terms of the accuracy and computation cost.

2. Template Matching

Let $I(x, y)$ be an observed image of size $M \times N$, where $x=0,1,\ldots, M-1$ and $y=0, 1, \ldots, N-1$ are the coordinates, and I denotes the image intensity at the coordinates of (x, y), and $w(x, y)$ be a template (target) of size $m \times n$. In general, the size of the template is less that that of the observed image.

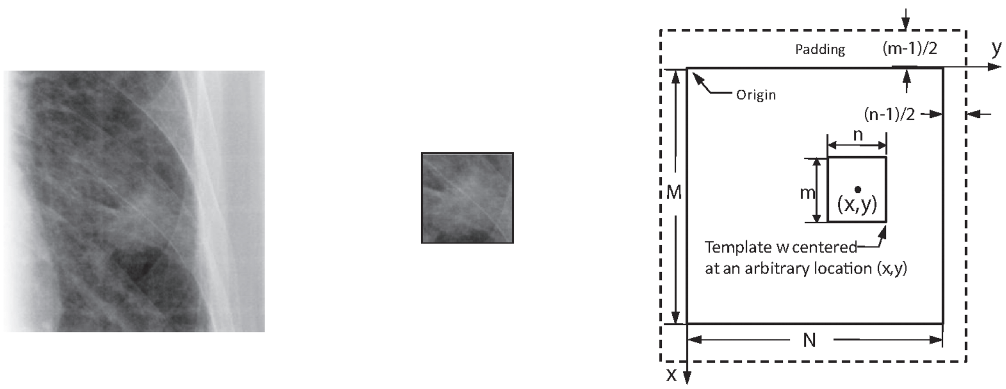
As an example, figures 1(a) and 1(b) show a fluoroscopic image as the observed image $I(x, y)$ and a pre-defined tumor target $w(x, y)$ as the template, respectively. The template matching is to find a sub-image in $I(x, y)$ which is similar to the template.

Figure 1(c) illustrates the mechanism of the template matching in which a template whose center is at an arbitrary location (x, y). The border around $I(x, y)$ is a padding necessary to provide for the situation when the center of $w(x, y)$ is on the border of $I(x, y)$. As usual, we consider that the templates are of odd size for notational convenience. We shift the template from the origin of image $I(x, y)$ pixel by pixel, and compute a pre-defined measure of match between the template and a sub-image (target candidate) overlapped by the template. The measure of match is considered to be a metric that indicates the degree of similarity or dissimilarity between the target and the target candidate. After a full scanning, a matching measure map of size $(M+m, N+n)$ is obtained, and we crop it to the size of the image $I(x, y)$. The template is said to be a match at a particular position in the observed image if a distinct peak in the measure of match is found at that position.

A number of measures of match have been proposed in the fields of the medical image and computer vision technologies. There is no best measures of match but a set of measures that are appropriated for particular applications. In this note, we will introduce several typical measures of match used in template matching.

2.1 Intensity Difference Measures

The intensity difference is the simplest measure that is based on the pixel-by-pixel intensity differences between the target and target candidates. Typical intensity difference measures include the mean squared error (MSE), sum of absolute difference (SAD), to name a few. The lower value of intensity difference indicates the higher similarity, while the higher value of intensity difference indi-



(a) (b) (c)

Fig. 1. An example of the template matching for tracking a tumor in fluoroscopic image. (a) A fluoroscopic image as an observed image $I(x, y)$. (b) A pre-defined tumor target $w(x, y)$ as the template. (c) Mechanism of the template matching.

cates the less similarity.

The MSE is a common measure of mismatch between a template $w(x, y)$ and the candidates in $I(x, y)$, given by

$$\mathrm{MSE}(x,y)=\frac{1}{mn}\sum_{s=-a}^{a}\sum_{t=-b}^{b}[I(x+s,y+t)-w(s,t)]^2 \quad (1)$$

where $a=(m-1)/2$ and $b=(n-1)/2$ are the padding size described previously. Similarly, the SAD is defined as

$$\mathrm{SAD}(x,y)=\sum_{s=-a}^{a}\sum_{t=-b}^{b}|I(x+s,y+t)-w(s,t)| \quad (2)$$

Figure 2 shows an MSE map between the template and the observed image as shown in Figs. 1(a) and 1(b). The match position is located at the coordinates (164, 170) where the MSE reaches its minimum. The intensity difference measures can be considered as the distance between two intensities vectors representing the the target and the target candidate. Therefore, these measures are sensitive to the image intensity variation and geometrical deformation, such as luminance shift, contrast changing, scaling, and so on. An additional reading on MSE may be found in [6].

2.2 Correlation Coefficient (CC)

The correlation coefficient (also called the normalized cross-correlation) of two images is a similarity measure of two images. For the template matching, the CC map between the template and the observed image is computed by

$$\mathrm{CC}(x,y)=\frac{\sum_{s,t}[w(s,t)-\bar{w}][I(x+s,y+t)-\bar{I}_{xy}]}{\sqrt{\sum_{s,t}[w(s,t)-\bar{w}]^2\sum_{s,t}[I(x+s,y+t)-\bar{I}_{xy}]^2}} \quad (3)$$

where $\bar{w}$ is the average intensity of the template, and $\bar{I}_{xy}$ is the average intensity of image in the region where I and w overlap. The value of $\mathrm{CC}(x, y)$ ranges from -1 to 1. A high value for $\mathrm{CC}(x, y)$ indicates a good match between the template and the image, when the template is centered at the coordinates (x, y).

Figure 3 shows a CC map between the template and the observed image as shown in Figs. 1(a) and 1(b). The match position is located at the coordinates (164, 170) where the CC reaches its maximum. A geometric interpretation of correlation coefficient is that the correlation coefficient can be viewed as the cosine of an angle between two vectors that represent the intensities of two images [7]. Therefore, the CC is robust against the image intensity varia-

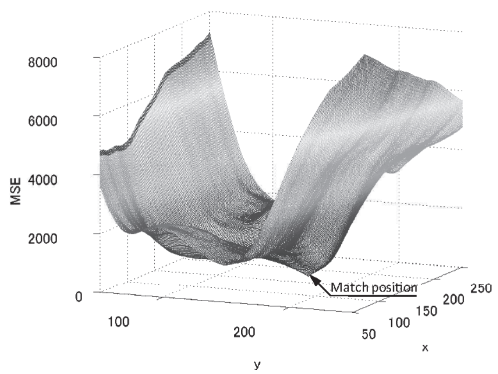


Fig. 2. Template matching based on mean squared error (MSE).

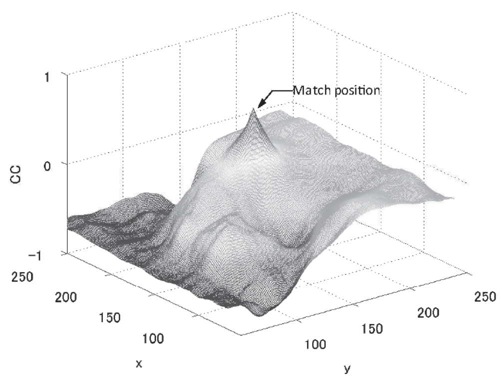


Fig. 3. Template matching based on correlation coefficient (CC).

tion. However, the CC is also sensitive to the geometrical deformation.

2.3 Correlation Ratio (CR)

In probability theory, correlation ratio measures the functional dependence between two random variables [9]. Given two images A and B, the correlation ratio of A and B is defined as

$$C(A, B) = \frac{\sigma_A^2 + \sigma_B^2}{2\sigma_{A \cup B}^2} \tag{4}$$

where σ_A^2 and σ_B^2 are the intensity variances of image A and B, respectively, and $\sigma_{A\cup B}^2$ is the total variance of images A and B. The CR takes on values between 0 and 1 that indicate the dissimilarity and similarity, respectively.

For the template matching, a similarity map based on the CR is computed by

$$\mathrm{CR}(x, y) = \sum_{s=-a}^{a} \sum_{t=-b}^{b} C[I(x+s, y+t), w(s, t)] \tag{5}$$

where $C(\cdot)$ denotes the CR defined by Eq. (4). Figure 4 shows a CR map between the template and the observed image as shown in Figs. 1(a) and 1(b). The match position is located at the coordinates (164, 170) where the CR reaches its maximum.

2.4 Mutual Information (MI)

In probability theory and information theory, the mutual information (MI) of two random variables is a measure of the mutual dependence of the two random variables. The MI has been one of the most popular similarity measures for multi-modal image registration, such as registration of CT and MR images [10].

Given two images A and B, the MI of the two images is defined as

$$M(A, B) = \sum_{i,j} P(i, j) \log_2 \frac{P(i, j)}{P_A(i) P_B(j)} \tag{6}$$

where i and j are intensities of images A and B (For example, $i, j = 0, 1, \ldots, 255$ if A and B are 8-bit images), $P(i, j)$ is the normalized two-dimensional (2-D) histogram (also called joint probability density function) of A and B, $P_A(i)$ and $P_B(j)$ are the normalized 1-D histograms (marginal probability density functions) of A and B, respectively. The MI measures the similarity between A and B, i.e., a high value of MI indicates high similarity, while low value of intensity difference indicates dissimilarity. The detail about the 1-D and 2-D histograms can be found in [8, 10].

Given an observed image and a template, the similarity map based on the MI is computed by

$$\mathrm{MI}(x, y) = \sum_{s=-a}^{a} \sum_{t=-b}^{b} M[I(x+s, y+t), w(s, t)] \tag{7}$$

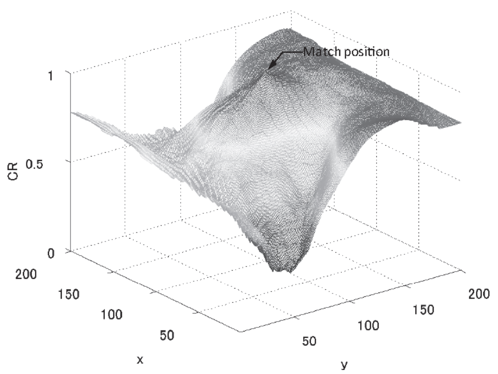


Fig. 4. Template matching based on correlation ratio (CR).

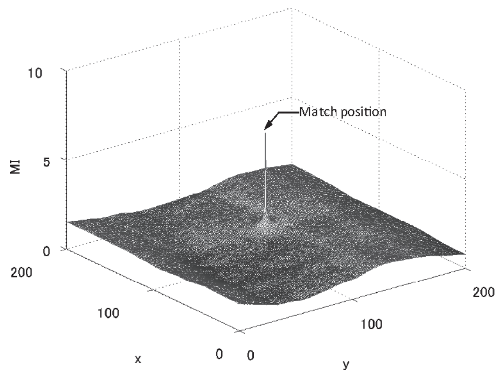


Fig. 5. Template matching based on mutual information (MI).

where $M(\cdot)$ denotes the MI defined by Eq. (6). Figure 5 shows an MI map between the template and the observed image as shown in Figs. 1(a) and 1(b). The match position is located at the coordinates (164,170) at where a distinct peak is located.

3. Tracking Tumor Motion for IGRT

In this section, we will present an example of template matching used for tracking tumor target in fluoroscopic image sequence.

In radiation therapy, as we known, respiration-induced tumor motion limits the efficiency of radiation delivery, especially for abdominal and thoracic tumor. One of the solution for improving the efficiency of radiation delivery is to utilizes a fluoroscopic imaging system to monitor the tumor motion during the treatment. On the base of the imaging system, a tumor tracking system, which is capable of automatically tracking the tumor motion, allows the treatment device to deliver the treatment beam to a moving tumor target. Figure 6 shows a fluoroscopic imaging system mounted on a radiation therapy device and a template matching-based tracking system for tracking the tumor in the fluoroscopic image sequence.

Here, we present several experimental results of the tumor tracking by using different similarity metrics. Experimental data is a fluoroscopic image sequence which consists of 100 frames of size 300×300 pixels obtained by an On-Board Imaging (OBI) system (Varian Medical Systems, Palo Alto, CA). A tumor target is delineated by a rectangle of size 81×81 pixels in the first frame. For the subsequent frames, we utilize the template matching based on four similarity metrics (MSE, CC, CR, and MI) to locate the tumor position in each frame. Figures 1(b) and 1(a) show the first frame of the image sequence and the template used for the template tracking, respectively. The tracking algorithms are implemented by using MATLAB on an Intel Core i7 3.3 GHz computer with Windows 7 OS.

Figures 7(a)-7(d) shows the tumor tracking results along the superior-inferior (SI) and left-right (LR) directions. In these figures, we also plot the ground truths of the tumor motion that are generated by manual localization. Table 1 summaries the root of mean squared errors (RMSEs) (mm) between the tracking results and the ground truths. In this table, we also list the computational times (s/frame) of each tracking method. From this table, we can see that the accuracy of the CC-based template matching is higher than those of MSE-, CR-, and MI-based methods. On the other hand, the computational times show that the lowest computational cost is the MSE-based method.

In the experiments, the template matching is performed by using an exhaustive searching strategy (full scan). In practice, some efficient search strategies, such as coarse-to-fine template matching, can be used to reduce the computational cost to meet

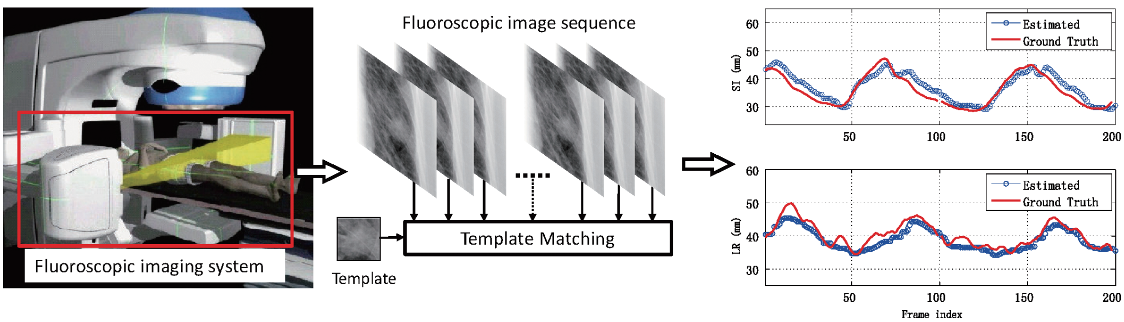


Fig. 6. Tracking tumor motion in fluoroscopic image sequence for IGRT.

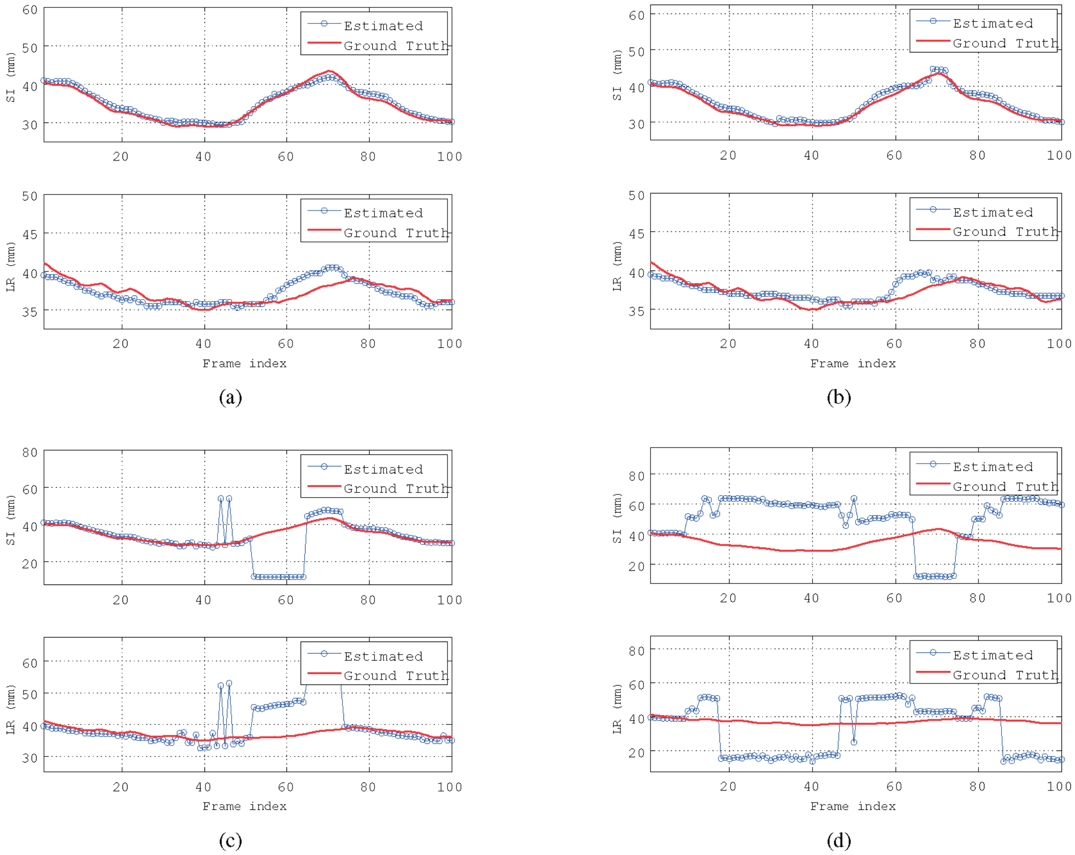


Fig. 7. Experimental results of template matching based on different similarity metrics. (a) Mean squared error (MSE). (b) Correlation coefficient (CC). (c) Correlation ratio (CR). (d) Mutual information (MI).

Table 1. Root of mean squared errors (RMSEs) (mm) in tumor motion estimation

	MSE-based	CC-based	CR-based	MI-based
RMSE (mm)	1.10	1.05	8.40	20.73
Computation time (s/frame)	1.2	2.2	2.6	5.8

the requirement of real-time tracking. Further reading about the efficient search strategies can be found in [11, 12].

4. Conclusion

In this note, we present the fundamental of template matching technique and introduced four common-used similarity metrics used for the template matching : MSE, CC, CR, and MI. By using an experiment of tumor tracking, we also evaluated the performances of these similarity metrics in terms of the accuracy and computational cost.

In computer vision technology, the template matching is only one of the object tracking methods. Extensive surveys of object tracking methods can be found in [4, 5].

References

[1] Xing, L., Thorndyke, B., Schreibmann, E., Yang, Y., Li, T., Kim, G., Luxton, G., Koong, A.: Overview of Image-Guided Radiation Therapy, Medical Dosimetry, **31**(2), 99-112, 2005

[2] Dawson, L.A., Jaffray, D.A.: Advances in Image-Guided Radiation Therapy, Journal of Clinical Oncology, **25**(8), 838-946, 2007

[3] Brunelli, R.: Template Matching Techniques in Computer Vision: Theory and Practice, Wiley, 2009

[4] Yilmaz, A., et al.: Object Tracking: A Survey, ACM Computing Surveys, **38**(4), 1-45, 2006

[5] Trucco, E., Plaks, K.: Video Tracking: A Concise Survey, IEEE Journal of Oceanic Engineering, **31**(2), 520-529, 2006

[6] Zhou, W., et al.: Mean squared error: Love it or leave it? A new look at Signal Fidelity Measures, IEEE Signal Processing Magazine, **26**(1), 98-117, 2009

[7] Rodgers, J.L., Nicewander, W.A.: Thirteen Ways to Look at the Correlation Coefficient, The American Statistician, **42**(1), 59-66, 1988

[8] Gonzalez, R.C., Woods, R.E.: Digital Image Processing, Prentice Hall Press, 2002

[9] Roche, A., et al.: The correlation ratio as a new similarity measure for multimodal image registration, MICCAI, **98**, 1115-1124, 1998

[10] Maes, F., et al.: Multimodality Image Registration by Maximization of Mutual Information, IEEE Transactions on Medical Imaging, **16**(2), 187-198, 1997

[11] Vanderbrug, G.J., Rosenfeld, A.: Two-stage template matching, IEEE Transactions on Computers, **26**, 384-393, 1977

[12] Goshtasby, A., et al.: A two-stage cross correlation approach to template matching, IEEE Transactions on PAMI, **6**, 374-378, 1984