Improving the Reliability of Object Recognition Based On Template Matching

Dr. A.S.M. Khedher
Technical Institute of Mosul

Dr. A. M. Alkababji
Computer Engg. Dept.
College Of Engineering\University of Mosul

O. A. Hadi
Computer Engg. Dept.

Abstract

Object recognition in computer vision is the task of finding a given object in an image or video sequence. During the last decades it has received increasing attention from the computer vision community for a variety of reasons, ranging from counting objects for industrial application to the development of practical biometric systems and interactive, emotion-aware and capable human–machine interfaces. There are variety of approaches for object recognition problem, depending on the type of object, the degree of freedom of the object and the target application. Template matching is the most advanced and intensively developed areas of computer vision and has been a classical approach to the problems of locating and recognizing of an object in the image. The object of this paper is to improve the reliability of object recognition by describing a modified method for template matching based on the Sum of Squared Differences (SSD) equation, that gives the highest margin between other template matching methods, the main advantage is that the high margin resulting from it can be considered as more safe to avoid wrongly detecting /recognizing an object.

Keywords: object recognition, template matching, Sum of Squared Differences (SSD).

تحسين موثوقية تميز الكائن بالاعتماد على المطابقة المتغيرة

عمر أنمار هادي
قسم هندسة الحاسوب

د. أحمد مأمون الكبابجي
قسم هندسة الحاسوب
كلية الهندسة/جامعة الموصل

د. عبد الستار محمد خضر
المعهد التقني/الموصل

الخلاصة

تمييز الكائن في الرؤيا الحاسوبية يعني ايجاد كائن معلوم في صورة او تسلسل فيديو. خلال العقود الاخيرة يتلقى تمييز الكائن اهتماما متزايدا من رابطة الرؤيا الحاسوبية لعدة ضروريات، منها عد الكائنات في التطبيقات الصناعية لتطوير انظمة تدقيق حقيقية وتطوير تواصل تفاعلي، مدرك للانفعالات بين الانسان والالة. توجد عدة طرق لحل مشكلة تمييز الكائن اذ يعتمد ذلك على طبيعة الكائن ودرجة حريته والتطبيق المطلوب. تعتبر المطابقة المتغيرة من اكثر المجالات تقدما وتطورا في الرؤيا الحاسوبية ولاتزال طريقة تقليدية لحل مشاكل ايجاد وتمييز الكائن في الصور. يهدف هذا البحث الى تحسين موثوقية تمييز الكائن من خلال طريقة معدلة للمطابقة المتغيرة معتمدة على معادلة مجموع تربيع الفروق(SSD) ،والتي تعطي الهامش الاعظم مقارنة بالطرق الاخرى للمطابقة المتغيرة والذي يعتبر اكثر امانا لتجنب الايجاد والتمييز الخاطئ للكائن.

الكلمات الدالة: تمييز الكائن، المطابقة المتغيرة، مجموع تربيع الفروق

1. Introduction

Computer vision is a wide research field that aims at creating machines that see. Not in the limited meaning that they are able to sense the world by optical means, but in the more general meaning that they are able to understand its perceivable structure. Template matching techniques, as now available, have proven to be a very useful tool for this intelligent perception process and have led machines to super human performance in tasks such as face recognition [1].

Humans recognize a large amount of objects in images with little effort, regardless of the fact that the image of the objects may differ some what in different viewpoints, in several different sizes /scale or even when they are translated or rotated [2].

Object recognition has many areas of application, including industrial application (here often the term "machine vision" is used), security and tracking application as well as searching and detection application [3].

Each application imposes different requirement and constraints on the object recognition task, a few categories are mentioned below:

- Evaluation time: especially in industrial application, the data has to be processed in real time. Evaluation time mainly depends on the number of pixels covered by the object as well as the size of the image area to be examined.
- Accuracy: in some application the object position has to be determined very accurately, a trade-off between fast and accurate recognition has to be found during the design of a vision system.
- Invariance: the vision system has to be insensitive to some kind of variance of the object to be recognized, variance can be introduced by the image acquisition process as well as the object itself, depending on the application it should achieve invariance with respect to: illumination, scale, rotation, background clutter, partial occlusion and viewpoints change.
- Recognition reliability: all methods of object recognition try to reduce the rate of "false alarms", in which correct objects classified as defect and "false positives" in which wrong objects classified as correct as much as possible. In industrial application this is critical where there is more pressure to prevent misclassifications and thus avoiding costly production errors.

There are so many distinct approaches to object recognition because the different nature of each application, it's specific requirements and constraints. There is no "general-purpose-approach" applicable in all situations [2].

Mainly the object recognition approaches can be classified into: Geometry-based approaches, Appearance-based approaches and Features-based approaches.

One of the global methods for object recognition is the template matching, in this paper we tried to improve one of its main requirements "recognition reliability" by increasing the safe margin between the correct and false templates matches.

2. Methodology

In computer vision, Template Matching is a technique used to categorize objects; a template is a small image (sub-image) and could be used to find occurrences of this template in a larger image.

The basic template matching algorithm consist of calculating the distortion function at each position of the image under examination to measures the degree of similarity between the template and the image. Then, the minimum distortion, or maximum correlation, position is taken to locate the template into the examined image.

Number of techniques has been developed and still developing to measure the match between input/reference image and the template. Matching accuracy of the measure depends on the type of method or algorithm selected kind of problem which is to be solved, and type of template and for what application it has to be done [4].

Template matching technique, especially in two dimensional cases, has many applications in object detection image registration ,object recognition and other computer vision applications[5,6,7], Even now, it is a fundamental technique to solve them.

2.1 Matching methods

There are several matching approaches for template matching, which are the Sum of Squared Difference (SSD), Cross Correlation (CC) and Coefficient Correlation method. As well as normalized versions of all methods which are more robust to the light differences between template and image but are much slower. Depending on its matching algorithm, the matching result may be slightly different. These methods are the most popular mostly due to their computational efficiency [8].

The template matching start by sliding the template one pixel at a time (left to right and up to down), at each location a metric is calculated so it represent how "good" or "bad" the match at that location

After matching all possible locations with the template, the results are stored in a resultant matrix R, which stores the coefficient value for each matched location in pixel.

With the size of the source image is $W \times H$ where W and H representing the width and height, respectively, and $w \times h$ is the product of width and height of the template image, the matrix size is given by $(W - w + 1) \times (H - h + 1)$. The equations for the three methods are shown below with their normalized methods:

1. Sum of Squared Difference (SSD)

This method matches the squared difference, there for a perfect match will be zero and bad one will be greater, it's also faster than other methods but less accurate.

$$R_{SSD}(x, y) = \sum_{x',y'} \left(T(x',y') - I(x + x', y + y') \right)^2 \qquad \dots \dots \dots \dots \dots ..(1)$$

2. Normalized Sum of Squared Difference

$$R_{NSSD}(x, y) = \frac{R_{SSD}(x,y)}{\sqrt{\sum_{x',y'} T(x'y')^2 . \sum_{x',y'} I(x+x', y+y')^2}} \qquad \dots \dots \dots \dots \dots ..(2)$$

3. Cross Correlation

This method multiplicatively match the template against the image, so a perfect match will be large and bad matches will be small or zero, it's slower than the Sum of Squared Difference but more accurate and robust.

$$R_{CC}(x, y) = \sum_{x'y'}\left(T(x', y').\, I(x + x', y + y')\right) \qquad \dots\dots\dots\dots\dots..(3)$$

4. Normalized Cross Correlation

$$R_{NCRC}(x, y) = \frac{R_{CRC}(x,y)}{\sqrt{\sum_{x',y'} T(x'y')^2.\, \sum_{x',y'} I(x+x', y+y')^2}} \qquad \dots\dots\dots\dots\dots (4)$$

5. Coefficient Correlation

This method match a template relative to its mean against the image relative to its mean, so a perfect match will be 1 and a perfect mismatch will be –1 while value of zero simply means that there is no correlation between the template and the image which is an additional result gained from this method.

$$R_{COC}(x, y) = \sum_{x'y'}\left(T'(x', y').\, I\,(x + x', y + y')\right) \qquad \dots\dots\dots\dots\dots..(5)$$

Where
$$T'(x', y') = T(x', y') - \frac{1}{(w.h)}.\, \sum_{x'',y''} T(x'', y'') \qquad \dots\dots\dots\dots\dots..(6)$$

$$I'(x+ x', y+ y') = I(x+ x', y+ y') - \frac{1}{(w.h)}.\, \sum_{x'',y''} I(x + x'', y + y'') \qquad \dots\dots\dots\dots\dots..(7)$$

6. Normalized Coefficient Correlation

$$R_{NCOC}(x, y) = \frac{R_{COC}(x,y)}{\sqrt{\sum_{x',y'} T'(x'y')^2.\, \sum_{x',y'} I'(x+x', y+y')^2}} \qquad \dots\dots\dots\dots\dots..(8)$$

Where

x' = 0 … w – 1 and y' = 0 … h– 1 for all three methods.

Considering T is the template image and I is the input source image. R(x,y) is numerical index of the used Method in the range [0,1] at position (x,y) after matching. Since the interpretation of SSD will be opposite of the other two methods, i.e., 0 shows the best while 1 shows the worst, we invert the result for SSD by subtracting it from 1 so that the results we observe will be consistent [4].

3. Improving procedure

Because equation(1) score is proportional to the template size the value could be divided by the template size to approch to some standardization, on the other hand the exprision inside the summation is powred by 2 so it could be rised to the power ½ and the output would be in the same scale or the range of the image and the template.

So the modified Sum of Squared Difference will become

$$R_{SSD}(x,y)' = \frac{R_{SSD}(x,y)}{\text{template size(in pixel)}} \qquad \dots\dots\dots\dots\dots..(9)$$

In case of a color image, template summation in the numerator and each sum in the denominator is done over all of the channels and separate mean values are used for each channel. That is, the function can take a color template and a color image. The result will still be a single-channel image, which is easier to analyze[9].

Experiments are carried out to analyze the performance of our modified method and comparing its results against other methods. The experiment data set used consists of three-colored image with 3 templates from each on with different sizes, which mean we match all the nine templates with every image (shown in Figure 1) and The experiment is performed under C++ environment with Microsoft Visual Studio 2010.

To accomplish the actual template matching, we used the template matching function built into Open CV (Open Source Computer Vision Library) library, where OpenCV is an open-source BSD-licensed library that includes several hundreds of computer vision algorithms.

The OpenCV library is written in C and C++ and runs under Linux, Windows and Mac OS X, there are active developments on interfaces for Python, Ruby, Matlab, and other languages.

| Image 1 | Image2 | Image3 |

Figure (1): The data set images

3.1 Experiments results

Table 1 shows the comparison of each method used for matching by its matching score where the score 1 represents a perfect match and the score zero represents no matching. As a matter of fact the matching score values achieved from the Normalized Correlation Coefficient are higher than other methods even our modified Sum of squared Difference, but it's has the lower margin which mean the cut-off point that we must choose should be >0.855.While our modified Sum of squared Difference gives the largest margin between all the template matching methods So, the cut-off point that we can choose is >0.556.

Table (1): Template matching methods scores

Method	template	Image1	Image2	Image3
Normalized Sum of squared Difference	Temp.1	**0.997849**	0.40926	0.444629
	Temp.2	**0.996507**	0.514168	0.645269
	Temp.3	**0.997925**	0.642539	0.648596
	Temp.4	0.366079	**0.998936**	0.595553
	Temp.5	0.365288	**0.999179**	0.469501
	Temp.6	0.305914	**0.999478**	0.61933
	Temp.7	0.390621	0.691201	**0.999298**
	Temp.8	0.63213	0.517465	**0.999511**
	Temp.9	0.60314	0.610142	**0.999466**
Normalized Cross Correlation	Temp.1	**0.999495**	0.718254	0.744772
	Temp.2	**0.998997**	0.761987	0.830131
	Temp.3	**0.999495**	0.85237	0.855156
	Temp.4	0.688501	**0.999468**	0.807603
	Temp.5	0.708592	**0.99959**	0.808974
	Temp.6	0.663454	**0.999739**	0.8297
	Temp.7	0.709288	0.84632	**0.999649**
	Temp.8	0.81813	0.764987	**0.999755**
	Temp.9	0.807355	0.826013	**0.999733**
Normalized Correlation Coefficient	Temp.1	**0.999051**	0.430951	0.412239
	Temp.2	**0.996909**	0.296739	0.307567
	Temp.3	**0.99901**	0.72701	0.695988
	Temp.4	0.227413	**0.998487**	0.602776
	Temp.5	0.202719	**0.998294**	0.318717
	Temp.6	0.183973	**0.999446**	0.62597
	Temp.7	0.313188	0.68833	**0.999221**
	Temp.8	0.2213	0.288041	**0.997934**
	Temp.9	0.079324	0.51953	**0.998616**
Modified Sum of squared Difference	Temp.1	**0.961737**	0.373854	0.402032
	Temp.2	**0.948554**	0.402412	0.476995
	Temp.3	**0.962894**	0.451658	0.454554
	Temp.4	0.287596	**0.971358**	0.481253
	Temp.5	0.328935	**0.977515**	0.414229
	Temp.6	0.211631	**0.979178**	0.419094
	Temp.7	0.335801	0.556734	**0.979265**
	Temp.8	0.384401	0.319304	**0.978256**
	Temp.9	0.305043	0.339853	**0.973105**

Table 2 shows the calculated margin for all methods, a Hit represents a positive template matching and a miss represents negative template matching and Min represents the minimum score for a hit and Max is the highest score for a miss.

Average I is average of the hit values while Average II is the average of the miss values and the Margin represents the difference between the minimum and maximum for each used method.

Table (2): The margin for templates matching methods

Method	Hit		Miss		Margin
	Min.	Average I	Max.	Average II	
Normalized SSD	0.996507	0.998683	0.691201	0.5091880	0.305306
Normalized Correlation Coefficient	0.997934	0.998552	0.72701	0.396803	0.270924
Normalized Cross Correlation	0.999468	0.999546	0.85237	0.783421	0.147098
Modified normalized SSD	0.948554	0.970206	0.556734	0.385949	**0.39182**

4. Conclusion

The problem of object recognition is still challenging and there are many approaches to deal with it in image or video, Template Matching is one of the main methods for object detection and recognition for many industrial applications, the Sum of Squared Difference method are faster than the correlation method but it's less accurate and robust.

Recognition reliability is a main requirement for object recognition to provide accurate classification especially in industrial application where production errors are costly . The approach proposed in this paper is to use our modified Sum of Squared Difference to increase the cut-off point between the correctly matched template from the other wrongly matched template, A large margin separates the true and false verified dataset reducing the chances of false identified as imposter.

5. References

[1] R. Brunelli " Template Matching Techniques in Computer Vision: Theory and   Practice". John Wiley & Sons, Ltd .ISBN: 978-0-470-51706-2, 2009.

[2] MA. Treiber "An Introduction to Object Recognition" Springer- Verlag London Limited, ISBN 978-1-84996-234-6, 2010.

[3] K. Ahuja and P. Tuli " Object Recognition by Template Matching Using Correlations and Phase Angle Method". International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 3, March 2013.

[4]T. Mahalakshmi, R. Muthaiah, and P. Swaminathan "Review Article: An Overview of Template Matching Technique in Image Processing". Research Journal of Applied Sciences, Engineering and Technology 4(24): 5469-5473, 2012.

[5]D. Kim, D. Lee, H. Myung and H. T. Choi "Object Detection and Tracking for Autonomous Underwater Robots Using Weighted Template Matching". IEEE Trans, 2011.

[6] J. N. Sarvaiya, S. Patnaik and S. Bombaywala "Image registration by template matching using Normalized Cross-Correlation" International Conference on Advances in Computing, Control, and Telecommunication Technologies,2009.

[7] Y. C. Zeng and J. F. Chen "Adaptive Template-matching Method for Recognition of Low-resolution License Number". IEEE Trans, 2011.

[8]W. Ouyang, F. Tombari, S. Mattoccia, L. Di Stefano and W. Cham "Performance evaluation of full search equivalent pattern matching algorithms" PAMI, 2012.

[9]"The Open Source Computer Vision Reference Manual Release 2.4.9.0" 2014.

[10] S. K Lam, C. Y Yeong, C. T. Yew, W. S. Chai and S. A. Suandi "A Study on Similarity Computations in Template Matching Technique for Identity Verification". Lam S K et al. / (IJCSE) International Journal on Computer Science and Engineering Vol. 02, No. 08, 2659-2665, 201