A Template Matching Approach Robust to Rotation based on Cross-Correlation

Jose Bernal* and Maria Trujillo†

*Universidad del Valle, Colombia, jose.bernal@correounivalle.edu.co
†Universidad del Valle, Colombia, maria.trujillo@correounivalle.edu.co

Keywords: Template matching, robust to rotation, cross-correlation.

Abstract

Template matching is a widely used technique for finding objects on images. Although there are several approaches, some of them are not robust to some issues that are commonly faced. In this paper, an approach that considers rotation is presented in the template matching field based on cross-correlation. Also the approach is evaluated to observe the behavior in different situations and compare computational complexities of algorithms. Results show improvement with respect to the rotation and that there is no a considerable variation in brightness, but the approach weakens the algorithms in presence of noise or if the target is shifted some pixels.

1 Introduction

Research and development in computer vision has been increasing over the last years. Nowadays, character recognition, medical diagnosis, target detection, and remote sensing are extensively studied fields. However, there is a bad acceptance of commercial vision systems due to this kind of processes demand low cost, high processing speed, result accuracy, and flexibility [2].

Template matching on image processing consists of taking an image named template and search all the occurrences of it in another image that is called input image (see Fig. 1). This process takes into account activities like finding a set of features in the template for matching them in the input image. This problem can be formulated as global optimisation problem where the target function consists of achieving great reliability, but it is a computationally expensive task.

Template matching approaches have to take into account some elements. One of these elements is that the technique has to search the target in the input image and, hence it is necessary to find optimal search techniques. Another is that it is needed to identify if two images are similar. This process is done using a similarity function. The sum of absolute differences (SAD), mean squared error (MSE) and cross-correlation functions are widely used for this task. However, each approach has advantages and disadvantages depending on the problem [4].

Normalised cross-correlation functions are classical area based methods. The procedure consists of taking window pairs from the two compared images, compute the similarity measure, and search for the maximum. The maximum window pair are corresponding ones [4]. However, these are functions that use corresponding pair of pixels. Thus, there are cases which cannot be addressed in this way as follows:

- The template is in the image but it is rotated
- The template is in the image but it is scaled

For overcoming the rotation issue many algorithms were proposed, such as:

- Tsai et al.[10] proposed a template matching based on the wavelet transform. The method considers wavelet decomposition and ring projection (a 2D to 1D transformation) representation to try to decrease the computational complexity of common template matching algorithms.
- Yu et al.[6] describe another proposal in the template matching field named orientation codes. Orientation codes are obtained from the discretisation of gradient angle of each pixel in the subimage and the template. Then, histogram is generated for each image and similarity between histograms is calculated.

The first algorithm losses information when converting 2D into 1D and second algorithm calculates gradient for each pixel.
and generates histograms for each subimage and, hence, implies higher computational cost. In this paper, an algorithm is presented in order to overcome the loss of information without adding considerable computational cost. The paper is organised as follows: In Section 2, common described template matching issues are shown. Then, in Section 3, the approach for template matching algorithms based on cross-correlation functions is exposed. In Section 4, considered cross-correlation techniques are presented. The tests explanation and the analysis of results are in Section 5. Finally, Section 6 contains conclusions of the developed work.

2 Template matching issues

There are some issues that are faced in the template matching field, for example, when the template is in the input image but it is 45 degrees rotated or it is partially occluded by another object. Thus, it is necessary to consider some of the most remarkable template matching issues [3, 7, 11] shown in Fig. 2 for exploring a new approach for template matching tasks. Those issues are described below:

- **Brightness**: The target is in the input image but it is brighter/darker compared to the input image.
- **Measure value**: Cross-correlation measure value is, in most cases, between -1 and 1 (the bigger value represents that compared images are completely similar while the lesser completely different). Due to noise presence in the images and some other factors, 1 is not commonly achieved. Thus, it is important to recognise whether the measure value represent or not a hit even if it is not the biggest value.
- **Motion blur**: Although target is in the input image, motion blur difficult the process.
- **Oclusion**: The target is partially occluded by some objects in the input image.
- **Perspective**: The viewpoint of the taken image is different to the one of the target image.
- **Processing time**: Computer vision applications demand not only result accuracy but also high processing speed.
- **Noise**: The image present random variation of brightness caused by sensors, circuitry, among others.
- **Rotation**: The image is in the input image but rotated some degrees.
- **Scaling**: The image is in the input image but scaled some units.

3 Template Matching Approach Robust to Rotation

Data preprocessing based on the gradient direction is proposed due to rotation is one of the important issue that has to be faced in the process of template matching. The preprocessing takes into account the gradient orientation before comparing the template and the explored window. The gradient orientation was discretised to multiples of π/4 degrees in [0, 2π) interval.

A brief example is used to illustrate the data preprocessing. Giving a 3x3 template with a gradient orientation of 0 and the 3x3 explored window of the input image with a gradient orientation of π/2, the data preprocessing consists in rotating the template window π/2 anticlockwise to align the gradient directions, as shown in Fig. 3.

![Figure 3: Corresponding pixels for comparison. The red arrows represent the gradient orientation, the left matrix corresponds to template with gradient orientation of 0 at the central pixel position and the right matrix corresponds to an explored window with a gradient orientation of π/2 at the central pixel position.](image1)

In other words, pixels in a template and the explored window are aligned based on the gradient direction before using a similarity function.

The proposed approach is described as follows:

**Pixel indexing**

Let

\[ P = \{(x, y) \mid \left(x - \frac{N}{2}\right)^2 + \left(y - \frac{N}{2}\right)^2 \leq \left(\frac{N}{2}\right)^2 \}, \]
be the set of pixels in the circular window, $V$ be a map from each $(x, y)$ pair in $P$ to a number that depends on the position when the pair was added and $\theta$ the gradient direction at the central pixel position. Then, pairs in $P$ are treated as coordinates and are rotated $\theta$ degrees forming another set $P'$. This new set represents the coordinates of the considered pixels. The whole process can be seen in Fig. 4.

Figure 4: Indexing process for pixels located in the circular window taking into account the rotation angle. The pixels inside circular window are indexed from left to right and top to bottom and, then, rotated certain angle.

**Searching**

An exhaustive search of the template in the input image is done. In this way, all possible locations in a explored window are compared, using a similarity function, to the template. The explored window and the template have the same dimension. The process is described in the following algorithm:

**Data:** $t$, $f$, $\text{sim}$  
**Result:** pair $(x, y)$  

$x = 0$;  
$y = 0$;  
$best = -1$;  

for $i = 1$ to $M$ do  
  for $j = 1$ to $M$ do  
    $s = \text{sim}(t, f, i, i+N, j, j+N)$;  
    if $s > best$ then  
      $x = i$;  
      $y = j$;  
      $best = s$;  
    end  
  end  

return $(x, y)$  

**Algorithm 1:** Exhaustive search algorithm

where $t$ represents a $N \times N$ template, $f$ represents a $M \times M$ input image, $\text{sim}$ represents a similarity function, and the output of the algorithm, $(x, y)$, represents the coordinate where the left corner of the explored window was located minimising the difference with respect to the template.

**Comparison**

The gradient is calculated at the central pixel of both, the template and the explored window, using the Sobel mask, as in Fig. 5, for aligning and comparing the explored window and the template.

Figure 5: Illustration of the gradient calculation at the central pixel. This process is applied to compared images for getting corresponding sets of indexed pixels.

Then, the sets and the number assignations for the template and the explored window are obtained taking into account the gradient orientation. The intersection between the number assigned for each pair in $P$ contained in the $V$ maps is used for knowing the pixels that will be used in the measure estimation.

**4 Cross-correlation**

The cross-correlation functions estimate the degree of similarity or distance between the template and each explored window in the image. It is used when sliding the template over the search area and for each candidate the correlation score is calculated[8]. Process is formal defined by the following expression[5]:

$$
\gamma(u, v) = \sum_{x, y} f(x, y) t(x - u, y - v).
$$

In equation 2, $t$ represents a $N \times N$ template, $f$ represents a $M \times M$ input image, $(u, v)$ represents the position where explored window is positioned.

Several algorithms based on cross-correlation functions were considered. In particular, Normalized Cross-Correlation (NCC)[1], Zero-mean Normalized Cross-Correlation (ZNCC)[1], Image Normalized Cross-Correlation (IMNCC)[7] and Image Zero mean Normalized Cross-Correlation (IMZNCC)[7], and the Median Correlation (MC) [9] are presented in Table 1.

In Table 1, variables $\tilde{x}$, $\tilde{y}$ represent the median of the template and the explored window, respectively, and $g$ is a matrix that is related to the following expression:

$$
g_{ij} = \frac{1}{2\pi\sigma^2} e^{-\frac{dist(P_i, P_j)^2}{2\sigma^2}},
$$

where $dist(P_i, P_j)$ function computes the distance between pixels $i$ and $j$. 


Table 1: Normalised Cross-correlation functions using $N \times N$ windows

<table>
<thead>
<tr>
<th>Function</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMNCC</td>
<td>$\sum_{i=1}^{N^2} \sum_{j=1}^{N^2} g_{ij} x_i y_j / \sum_{i=1}^{N^2} \sum_{j=1}^{N^2} g_{ij}$</td>
</tr>
<tr>
<td>IMZNCC</td>
<td>$\sum_{i=1}^{N^2} \sum_{j=1}^{N^2} g_{ij} (x_i - \bar{x}) (y_j - \bar{y}) / \sqrt{\sum_{i=1}^{N^2} (x_i - \bar{x})^2 \sum_{j=1}^{N^2} (y_j - \bar{y})^2}$</td>
</tr>
<tr>
<td>MC</td>
<td>$\sum_{i=1}^{N^2} \sum_{j=1}^{N^2} g_{ij} x_i y_j / \sqrt{\sum_{i=1}^{N^2} (x_i - \bar{x})^2 \sum_{j=1}^{N^2} (y_j - \bar{y})^2}$</td>
</tr>
<tr>
<td>NCC</td>
<td>$\sum_{i=1}^{N^2} x_i y_i / \sum_{i=1}^{N^2} x_i \sum_{i=1}^{N^2} y_i$</td>
</tr>
<tr>
<td>ZNCC</td>
<td>$\sum_{i=1}^{N^2} g_{ij} (x_i - \bar{x}) (y_j - \bar{y}) / \sqrt{\sum_{i=1}^{N^2} (x_i - \bar{x})^2 \sum_{j=1}^{N^2} (y_j - \bar{y})^2}$</td>
</tr>
</tbody>
</table>

5 Results and Analysis

Tests based on tracking and image recognition issues, presented by Nakhmani et al.[7], in paper "A new distance measure based on generalized image normalized cross-correlation for robust video tracking and image recognition", are recreated, in order to compare the behaviour of the proposal with results obtained with the distance measure proposed by Nakhmani et al. Thus, a $15 \times 15$ subimage (see Fig. 6) taken from cameraman image, (https://sites.google.com/site/nakhmania/image, cameraman.tif) is used as template. Tests evaluating translation, rotation, brightness and noise tolerance are described in next subsections. The measures were computed with the following expression:

$$M = 1 - \text{MAX}(0, m)$$

where $m$ means the calculated value for the analysed function. The termination in $R$ represents that data was preprocessed before applying the similarity function.

5.1 Translation Test

Translation test consider vertical shifts of the selected area inside the red rectangle shown in Fig. 6. As can be seen in Fig. 7, the preprocessed IMZNCC (IMZNCCR) is the most translation-robust similarity function compared to the others.

However, in general, the data preprocessing weakens the translation robustness, probably because the center is changing while shifting and, hence, gradient orientation is affected.

5.2 Rotation Test

Rotation test consists of taking the cameraman subimage and rotate it from $-90$ to $90$ degrees. The results are shown in Fig. 8.

It can be observed that in most of cases the preprocessed similarity functions are better than the unmodified ones. However, there are some peaks of them when approaching to -90, -45,
45, and 90 degrees. That is because rotation angles were discretised to angles multiple of 45.

5.3 Brightness Test

Brightness test uses the cameraman subimage with different brightness values in the $[-50,50]$ range. Fig. 9 shows that there is no considerable change on the brightness robustness of all algorithms when considering the approach.

![Figure 9: Behaviour of the algorithms when original cameraman subimage is compared to subimage with a brightness variation from -50 to 50](image)

5.4 Noise Test

Noise test considered consists in adding Gaussian noise with standard deviations between 0 and 90 to the 15x15 subimage. Results are shown in Fig. 10. Thought the trend for most algorithms is to increase when increasing noise in the image, modified algorithms present several peaks along the curve showing that considering the gradient orientation of the central pixel can reduce the robustness to noise.

![Figure 10: Behaviour of the algorithms when Gaussian noise with $\sigma$ in the interval [0,90] is applied to the cameraman subimage and compared to the original subimage](image)

5.5 Complexity Comparison

Table 2 shows the computational complexity for searching the best match with the considered correlation functions. Assume that the template and the input image are a $N \times N$ image and a $M \times L$ image, respectively. Due to pixels indexes are pre-calculated for multiples of $\pi/4$ in $[0,2\pi)$, there is no a big difference between modified and unmodified algorithms.

Table 2 shows that IMNCC and IMZNCC algorithms are the most expensive algorithms among the evaluated algorithms, and that MCR and MC are comparable to ZNCC and ZNCCR algorithms in terms of big-O notation.

<table>
<thead>
<tr>
<th>Function</th>
<th>Computational complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMNCC</td>
<td>$O(ML \cdot N^4)$</td>
</tr>
<tr>
<td>IMNCCR</td>
<td>$O(ML \cdot N^4)$</td>
</tr>
<tr>
<td>IMZNCC</td>
<td>$O(ML \cdot (N^4 + N^2))$</td>
</tr>
<tr>
<td>IMZNCCR</td>
<td>$O(ML \cdot (N^4 + N^2))$</td>
</tr>
<tr>
<td>MC</td>
<td>$O(ML \cdot (2N^2 + N \log(n)))$</td>
</tr>
<tr>
<td>MCR</td>
<td>$O(ML \cdot (2N^2 + N \log(n)))$</td>
</tr>
<tr>
<td>NCC</td>
<td>$O(ML \cdot N^4)$</td>
</tr>
<tr>
<td>NCCR</td>
<td>$O(ML \cdot N^4)$</td>
</tr>
<tr>
<td>ZNCC</td>
<td>$O(ML \cdot 2N^2)$</td>
</tr>
<tr>
<td>ZNCCR</td>
<td>$O(ML \cdot 2N^2)$</td>
</tr>
</tbody>
</table>

Table 2: Computational Complexity of Algorithms

5.6 General remarks

The approach improves the algorithms in rotation robustness, but weakens in translation and noise robustness. That is a consequence of the gradient orientation calculation. In the case of brightness test, there is no considerable difference between modified and unmodified algorithms. The IMZNCCR algorithm behaves better than the others, but it has the highest computational cost. On the other hand, the lowest computational cost algorithms (NCC and NCCR) have also a good behaviour among the tests.

6 Conclusions

Template matching is a technique consisting of taking a template and search for all the occurrences of it on a input image. There are several approaches in this field for optimising the response time considering translation, rotation or scaling. A template matching approach robust to rotation is presented. It takes advantage of the gradient orientation for comparison. For evaluating the effectiveness of this algorithm, some test
were run and it was found that algorithm has some difficulties when template is shifted some pixels or when noise is added to the input image. That issue is, probably, a consequence of considering the gradient orientation of the central pixel for comparing the images. However, it has some strengths when target is found rotated some degrees or brighter/darker than template. Another special feature of the approach is that there is no a considerable addition of computational cost. Results of the tests also shown that IMZNCCR algorithm behaves better among tests compared to other algorithms, but it has the highest computational cost, while NCC and NCCR algorithms that takes less time and are close to the IMZNCCR results. Although there are several algorithms based on cross-correlation functions, there is no a standard test for evaluating algorithm robustness and that allows the comparison among authors’ results.

References


