A Method to Eliminate Wrongly Matched Points for Image Matching

Xiao-fei Al*

ABSTRACT

In this paper, we proposed a new method to eliminate wrong point pairs by distance constraint and bidirectional matching. It was aimed at the problems of many wrong pairs and poor matching accuracy in feature-based image matching. It calculated the distance of point pairs obtained by SURF, eliminated part of wrong point pairs by this distance, and then got the best point pairs by bidirectional matching. Used this point pairs, it calculated the homography to mark out the boundaries of reference image. At last single image and video were used to verify the validity. The results indicated that the method can eliminate most wrong matches. The accuracy was improved by 3.07% compared with RANSAC, the speed of single image matching was 0.23s, the frame rate of video matching reached 25.108fps, and the matching speed of SURF was also guaranteed.

INTRODUCTION

Fast and accurate image matching is one of the key issues of image processing. Study on SIFT\(^1\), SURF\(^2\), BRISK\(^3\), ORB\(^4\) and FREAK\(^5\) makes the accuracy of image matching improved. However, no matter what kind of matching algorithm is used, there are always some wrong point pairs. At present, the methods to eliminate wrongly matched points are mainly divided into three categories: based on functional approximation\(^6\), based on statistical model\(^7\), and based on graph\(^8\). These methods can improve the accuracy of matching, but there are some problems such as low accuracy and long time consuming.

In this paper, the SURF will be used to extract the feature points of two images, and these points will be matched by brute force. The minimum value of the coordinate distance between point pairs will be calculated. According to the threshold set from minimum distance, it will eliminate part of point pairs tentatively, and repeat this process by matching in opposite direction. According to the remaining point pairs, it will select four feature points to compute homography, and finally the reference image will be identified in the target image.

THE MODEL OF ELIMINATING WRONG MATCHES

Defining the feature points of reference image as \(P = \{p_1, p_2, p_3, \ldots\}\), the feature points of target image as \(Q = \{q_1, q_2, q_3, \ldots\}\). After feature detecting by SURF and matching by

---

*Corresponding author: 1191256868@qq.com
brute force, the set of point pairs is \( M = \{ m_1, m_2, m_3, \ldots \} \). The distance of each point pairs can be calculated by follow.

\[
d_i = \sqrt{(p_{ix} - q_{ix})^2 + (p_{iy} - q_{iy})^2}
\]

(1)

\( p_{ix}, p_{iy}, \) \( q_{ix}, q_{iy} \) are the \( x, y \) coordinates of each point pairs in \( M \). The minimum of distance can be defined as follow.

\[
d_{\min} = \min\{ d_1, d_2, d_3, \ldots \}
\]

(2)

And then we set the threshold as \( E = a \times d_{\min} \). Comparing \( d_i \) with \( E \), we save the point pairs \( m_i \) when \( d_i < E \) as \( N = \{ n_1, n_2, n_3, \ldots \} \). Exchanging reference image and target image, we can obtain point pairs set \( L = \{ l_1, l_2, l_3, \ldots \} \) by using formula (1), (2). Therefore, the best point pairs set satisfied the following equation.

\[
M = N \cap L
\]

(3)

Therefore, if we get \( P \) and \( Q \), we can find the best point pairs \( M \) by using this measure.

**FIND HOMOGRAPHY**

Since the camera gets the image from different angles, the relationship between the reference image and the target image obtained by camera belongs to the mapping of plane to plane, so the homography can be used to establish the relationship between the two images.

\[
q_i = H \times p_i
\]

(4)

\( p_i = (x_i, y_i, w_i) \) is the homogeneous coordinates of feature point in reference image, \( q_i = (x_2, y_2, w_2) \) is the homogeneous coordinates of feature point in target image, and \( H \) is a 3*3 matrix written as follow.

\[
H = \begin{bmatrix}
    h_{00} & h_{01} & h_{02} \\
    h_{10} & h_{11} & h_{12} \\
    h_{20} & h_{21} & h_{22}
\end{bmatrix}
\]

(5)

Therefore, the relationship can be written as formula (6) by use formula (5) into formula (4).

\[
\begin{bmatrix}
    x_2 \\
    y_2 \\
    w_2
\end{bmatrix} = 
H 
\begin{bmatrix}
    x_1 \\
    y_1 \\
    w_1
\end{bmatrix} = 
\begin{bmatrix}
    h_{00} & h_{01} & h_{02} \\
    h_{10} & h_{11} & h_{12} \\
    h_{20} & h_{21} & h_{22}
\end{bmatrix} 
\begin{bmatrix}
    x_1 \\
    y_1 \\
    w_1
\end{bmatrix}
\]

(6)

If homography \( H \) is given, and each element of it is multiplied by the same number \( a \), the homography \( a \times H \) and \( H \) have the same function. Because the new homography turns the homogeneous point \( p_1 \) into the homogeneous point \( a \times p_2 \), and the point \( a \times p_2 \) is the same as \( p_2 \) in the image. So there are only 8 free elements in a homography \( H \).

Generally, we set the lower right corner of the element \( h_{22} = 1 \) to normalize \( H \). Therefore, we only need 8 equations to solve homography \( H \). In other word, we can get \( H \) through only 4 non-collinear points.

We assume the homogeneous coordinates of the feature points on the two images are \((x_1, x_2, 1)^T\), \((x_2, y_2, 1)^T\), the relationship between the two points can be written as follow.

\[
x_2 = \frac{x_1 h_{00} + y_1 h_{01} + h_{02}}{x_1 h_{20} + y_1 h_{21} + 1}
\]

(7)
\[ y_2 = \frac{x_1 h_{0} + y_1 h_{1} + h_{2}}{x_1 h_{21} + y_1 h_{22} + 1} \]  

(8)

Therefore, if the vertex coordinates of the reference image are known, the point coordinates of the reference image after the transformation \( H \) can be computed by using the formula (7), (8).

**EXPERIMENT AND ANALYSIS**

**Experimental Environment**
- CPU: Intel(R) Core(TM) i5-3210M CPU @2.5 GHz 2.5 GHz
- Graphics card: NVIDIA GeForce 610M
- Camera: USB2.0 UVC HD Web Camera
- Integrated Developing Environment: Visual Studio 2013
- Third part library: OpenCV v2.4.9

**Brute Force Match with SURF Features**

Given a reference image with a resolution of 600*450 and a target image generated by flipping the reference image vertically 180 degrees, paper uses SURF provided by OpenCV to detect feature points, and uses brute force method to match this two images. The result is shown as figure 1. The left is reference image, and right is target image.

![Figure 1](image1.png)

**Eliminating Wrong Matches**

As we can see, there are many wrong matches using brute force matching. Because the reference image is centrosymmetric with the target image, the correct matches should also be centrosymmetric. Therefore, paper uses the result generated by above method to substitute \( M \) in formula (1), then calculates minimum distance by formula (2), and finally gets the best point pairs \( M_i \). The result can be shown as figure 2.

![Figure 2](image2.png)

Comparing figure 2 with figure 1, the wrong point pairs are eliminated basically.
Application Homography

In order to mark the position of the reference image, randomly selected coordinates of four feature points, calculate the homography $H$ by formula (6). Then, setting four vertices coordinates of reference image are as follows: (0, 0), (600, 0), (600,450), (0,450), paper gets the position of reference image in the target image by formula (7), (8). In the program, the reference image is marked with the green line. The result can be shown as figure 3.

![Figure 3. The result of applying homography.](image)

Contrast Test and Video Analysis

In order to illustrate the matching accuracy of this method, the experiment is compared with the RANSAC (RANdom SAmple Consensus) mismatch culling algorithm, and the results are shown in table 1.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Matches</th>
<th>Best Matches</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>RANSAC</td>
<td>423</td>
<td>388</td>
<td>91.73%</td>
</tr>
<tr>
<td>Paper Method</td>
<td>423</td>
<td>401</td>
<td>94.80%</td>
</tr>
</tbody>
</table>

The above experiment is single image feature matching. In some applications, target images are often from video. Therefore, paper uses VideoCapture class provided by OpenCV to obtain video frames and matches with reference image. The result, shown as figure 4, marks the location of the reference image accurately.

![Figure 4. Matching from video.](image)

One of the advantages of the SURF is its fast execution speed. In order to ensure that the method does not affect the speed of SURF matching, the experiment is compared with the SURF with brute force matching. The frame rate of the SURF matching (black line...
representation) and the frame rate when the method is executed (blue line representation), are shown as figure 5.

The result can be calculated: after adding the method, the maximum execution frame rate reaches 25.1058fps, the minimum is 20.6258fps, the average value is 23.4372fps, the variance is 0.398198, slightly better than the SURF frame rate, can ensure high efficiency.

CONCLUSION

We have presented a new method to eliminate wrongly matched points for image matching. This method can eliminate the wrong point pairs effectively, and solve the efficiency and accuracy problems of image matching. Applications range from target recognition to image stitching. However, there are still some limits in this method, such as the real-time problems. For this reason, further studies are still needed.

REFERENCES