Using Local Affine Invariants to Improve Image Matching

Daniel Fleck and Zoran Duric
Department of Computer Science, George Mason University, Fairfax VA 22030, USA
{dfleck,zduric}@cs.gmu.edu

Abstract

A method to classify tentative feature matches as inliers or outliers to a transformation model is presented. It is well known that ratios of areas of corresponding shapes are affine invariants [6]. Our algorithm uses consistency of ratios of areas in pairs of images to classify matches as inliers or outliers. The method selects four matches within a region, and generates all possible corresponding triangles. All matches are classified as inliers or outliers based on the variance among the ratio of areas of the triangles. The selected inliers are used to compute a homography transformation. We present experimental results showing significant improvements over the baseline RANSAC algorithm for pairs of images from the Zurich Building Database [5].

1. Introduction

The goal of image matching is to determine if all or part of one image matches all or part of another image. In many applications (e.g. location recognition), after determining a match is present, a registration step is used to align the images so the matching parts overlap precisely.

A recent review of image matching algorithms was conducted by Mikolajczyk and Schmid [9]. The reviewed algorithms for matching image pairs typically have four phases. In the first phase features are detected in both images. In the second phase feature descriptors are computed as the “signatures” of the features. In the third phase the features of the first image are compared to the features of the second image and ordered by similarity. The matches generated in the third phase are considered tentative due to the high percentage of incorrect matches produced at this stage. Thus, a fourth phase is required to filter out incorrect matches. The fourth phase in a typical algorithm attempts to fit a transformation model to the tentative matches. The model is then used to classify each tentative match as an inlier or outlier. The focus of this paper is improving the fourth phase of the matching process.

The remainder of this paper is organized as follows. Section 2 describes the regional affine filtering algorithm. Section 3 describes the experiments performed and presents the results. Section 4 concludes the paper.

2. Regional Affine Filtering

In this work we propose a novel approach to classify matches as inliers or outliers using affine invariants applied to local image regions. Commonly, a transformation model is built by iteratively testing a set of model parameters against the tentative matches. After a sufficient number of iterations the best model is chosen as the transformation between the images. The resulting algorithms are \(O(mn)\) where \(m\) is the number of models tested and \(n\) is the total number of matches. This approach is the general idea behind the very popular Random Sample Consensus (RANSAC) [2] algorithm. Our approach detects inliers using
the property that affine invariant transformations maintain a constant ratio of areas of shapes [6]. To exploit this invariance the technique iteratively samples four matches in a local region. Using the four matches the ratios of areas generated by the four possible triangles are measured. If all corresponding triangles produce a consistent ratio of areas the matches are considered inliers. In this paper we describe this approach and present experimental results showing significant improvements over the traditional RANSAC approach.

An affine transformation can model several changes between pairs of images. The transformation has six degrees of freedom including translation in the X and Y directions, rotation, non-isotropic scaling, and shear. The general affine transformation matrix relating image coordinates is shown in equation 1.

\[
\begin{pmatrix}
  x' \\
y' \\
1
\end{pmatrix} =
\begin{pmatrix}
a_{11} & a_{12} & t_y \\
a_{21} & a_{22} & t_x \\
0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
x \\
y \\
1
\end{pmatrix}
\] (1)

Images related by an affine transformation have invariant properties including parallelism of corresponding lines, ratio of the lengths of corresponding parallel lines and ratios of areas of corresponding shapes [6]. Previous work demonstrated a consistent ratio of areas across an entire image can be used to detect inliers [3]. The underlying assumption was that the relationship between the image pairs could be approximated by an affine transform. It was shown in [4] that as the perspective distortion grew, this assumption broke down and the algorithm generated unstable results.

In this paper we present a new method that chooses matches within a local region to minimize the effect of distortion and produce stable results. The algorithm, shown as Algorithm 1, first computes the Delaunay triangularization [1] of the tentative matches to find nearby around a given match. For each match three other nearby matches are chosen. From these four matches four triangles can be formed. If any of the four triangles are sharp (i.e. have an angle < 10°), new nearby matches are selected and the triangles are checked again. This is because small localization errors cause sharp triangles to have unstable areas. The ratios of areas of all corresponding triangles are computed. The ratios are normalized by dividing them by the maximum ratio. If the normalized minimal ratio is \(< 1 - \varepsilon\) (we used \(\varepsilon = 0.09\)) all matches are discarded as erroneous.

In the second phase all rejected matches are evaluated. Every ‘outlier’ is coupled with three ‘inliers’ to determine if some of the matches previously classified as outliers are actually inliers. These sets of four are then evaluated for affine consistency. If the set is consistent then the previously rejected match is added to the set of inliers. The two phase approach ensures that matches originally misclassified are recovered. A final RANSAC fit is performed on the set of inliers.

```
Input: a set of tentative matches
Output: inliers from the tentative matches
dt ← DelaunayTriangularization

foreach (match in allMatches) do
    // Find local matches
    lm ← LocalMatches(dt, match)
    repeat
        allM ← pickMatches(lm, match)
        aun ← anySharp(allM)
        until (aun == false)
        maxR, minR ← getRatios(allM)
        diff ← 1 - (minR / maxR)
        if diff < threshold then
            inliers += fourMatches
    end
end
```

Algorithm 1: Phase one of the regional affine approach for inlier detection

### 3. Experimental Results

In this section experimental results are presented comparing the regional affine approach to baseline RANSAC [2]. Tests were conducted using 550 image pairs from the Zurich Building Database [5]. Lowe’s SIFT [7] algorithm generated the tentative matches used as input to the al-
gorithm. The resulting inliers were evaluated by applying the normalized direct linear transform algorithm to fit a homography transformation [6]. Using the generated model, the reprojection error was computed to evaluate the transformations using Eq. 2 [8]; \( \tilde{x}^j \) is the model projection of feature \( j \) and \( x^j \) is the detected location of feature \( j \).

\[
R = \sum_j \sqrt{(\tilde{x}^j_1 - x^j_1)^2 + (\tilde{x}^j_2 - x^j_2)^2}
\]  

Fig. 1 shows two sample image pairs and the results from each algorithm. Fig. 2 shows the number of inliers for the model generated by RANSAC versus the regional affine approach for each image pair. Points below the diagonal indicate an image pair where more inliers were found using the regional affine approach. The algorithms are very similar in the number of matches detected. However, the time comparison in Fig. 3 shows a significant improvement in efficiency for regional affine versus RANSAC. Points above the diagonal indicate an image pair where regional affine was more efficient. The figures show that the regional affine approach is much more efficient than RANSAC while maintaining a similar accuracy. Additionally, it is interesting to evaluate the speed and accuracy among different types of images pairs. Table 1 presents results split into image groups based on number of inliers found. These results show that as the number of inliers increases, RANSAC does improve in speed. This is because RANSAC stops when a large percentage of matches is found, and thus can stop evaluating models early in some cases. However, as the images become harder to match RANSAC significantly slows while the regional affine approach maintains the same efficiency. Additionally Table 1 shows that the number of inliers found by both algorithms is very similar.

4. Conclusion

This paper presents a new method to determine inliers and outliers for image matching. The proposed approach uses the affine invariance of ratios of areas of shapes. It detects inliers in local feature constellations quickly and accurately. Results on a large number of images demonstrate the advantages of this approach over the traditional RANSAC algorithm. Future work includes comparing the approach to other model fitting approaches and further increasing the accuracy and efficiency of the algorithm.
Figure 1: Sample matching results from image pairs. Left pairs: Regional Affine. Right pairs: RANSAC

<table>
<thead>
<tr>
<th>Number of Inliers</th>
<th>RANSAC Time</th>
<th>Regional Affine Time</th>
<th>RANSAC Inliers</th>
<th>Regional Affine Inliers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std Dev</td>
<td>Mean</td>
<td>Std Dev</td>
</tr>
<tr>
<td>&lt; 20</td>
<td>16.20</td>
<td>2.52</td>
<td>3.02</td>
<td>1.45</td>
</tr>
<tr>
<td>(20,50]</td>
<td>16.65</td>
<td>1.97</td>
<td>4.66</td>
<td>1.20</td>
</tr>
<tr>
<td>(50,100]</td>
<td>15.55</td>
<td>4.90</td>
<td>5.80</td>
<td>1.74</td>
</tr>
<tr>
<td>&gt; 100</td>
<td>8.63</td>
<td>4.71</td>
<td>5.00</td>
<td>0.56</td>
</tr>
</tbody>
</table>

Table 1: Statistical comparison of inlier ranges between the RANSAC and Regional Affine algorithms.

References