COMPLEMENTARITY OF FEATURE POINT DETECTORS

Guillaume Gales, Alain Crouzil
Institut de Recherche en Informatique de Toulouse, Université P. Sabatier
118 rte de Narbonne, 31062 Toulouse cedex 9, France

Sylvie Chambon
Laboratoire Central des Ponts et Chaussées, rte de Pornic, BP 4129, 44314 Bouguenais cedex, France

Keywords: Feature points, Complementarity, Repeatability, Spatial distribution.

Abstract: The goal of this paper is to provide a study on complementarity of feature point detectors. Many studies have been proposed on these detectors but none deals with complementarity in details. We introduce an evaluation of eleven well-known detectors based on new criteria used to characterize complementarity. The complementarity is computed with spatial distribution and contribution measures as well as repeatability and distribution gains of the association of two detectors.

1 INTRODUCTION

Many applications in computer vision rely on some characteristic points called feature points (or point of interest). Their detectors are designed to select the most distinctive points in an image so they are easy to match without ambiguities. We asked ourselves the following question: if different feature point detectors select the most distinctive points in an image, do they return the same points? In other words, we propose to study the complementary of several feature point detectors. Previous evaluations studied detection criteria such as repeatability or information content and matching criteria such as recall and precision rates. We propose to study the complementarity of feature point detectors based on new criteria: spatial distribution and complementarity measures. The idea is to find out which are the most complementary detectors in order to combine them when needed. Initially, this evaluation is proposed on small-baseline stereo image pairs since feature points may be used for instance in application such as stereo matching (Lhuillier and Quan, 2002) or fundamental matrix estimation (Hartley and Zisserman, 2004).

First, previous work on feature point detectors is described. Second, the new criteria to measure complementarity are introduced. Finally, the results of our experimentations are discussed before the conclusion.

2 PREVIOUS WORK

2.1 Feature Point Detection

Feature points are distinctive points within an image. Therefore, they are used in many applications such as tracking, indexation or stereo matching. There are different families of detectors: contour based, intensity based, parametric model based methods (Schmid et al., 2000). We focus on intensity based methods since they are widely used. They are based on the following steps: (i) computation for each pixel of a response value based on local grey level variations; (ii) non maxima suppression; (iii) post processing (edge response elimination, subset selection, subpixel localisation). In this section we describe briefly the first step for the well known detectors: Moravec, Harris, Kitchen-Rosenfeld, Beaudet, SUSAN, FAST, SIFT, SURF, Harris-Laplace, Hessian-Laplace and Kadir.

Fixed Scale Detectors. The response is computed using a fixed window size.

Moravec (MO). The response is based on an auto-correlation measure based on four directions (Moravec, 1977).

Harris (HA). The response is based on a generalisation of the auto-correlation measure on every shift.
When two images of a same scene are taken under different conditions are submitted to a feature point detector, it is desirable that the returned points are repeated: the two projections of a scene point are detected in the two images. In other words, a feature point detector, it is desirable that the returned points must also be maxima in the scale space of the determinant of the Hessian (Bay et al., 2006).

Repeatability. When two images of a same scene are large (rotation, scale change, perspective change, light change), this detector is robust to that transformation (Schmid et al., 2000; Mikolajczyk and Schmid, 2004; Gil et al., 2009).

Information Content. This value represents how much the feature points of an image are distinct to each other. The more distinctive the feature points are, the less ambiguous they are to match. A descriptor is computed for each feature point and the Mahalanobis distance is used to normalize each component (Schmid et al., 2000).

Complementarity. The complementarity is measured in (Mikolajczyk et al., 2005) in an object recognition context. A clustering algorithm is applied to the feature points, then the number of points from each detector is computed for each class. Ideally, each class must contain points from the same detector only.

3.1 Repeatability

Let \( p^i_{j,i} \) be the pixel at the \( i \)th row and \( j \)th column in the image \( I \). Let \( d^i_{j,j} \) be the disparity vector such as \( p^i_{j,j} = p^i_{j,j} + d^i_{j,j} \) where \( p^i_{j,j} \) is the pixel in the right image which is the projection of the same scene entity as \( p^i_{j,j} \). Let \( \varepsilon \) be a tolerance margin in pixels. We use the following definitions:

- repeated point – a feature point \( p^i_{j,j} \) in the image \( I_1 \) is repeated if a feature point \( p^i_{j,j} \) has been detected in the image \( I_2 \) within a distance lower than \( \varepsilon \) from its theoretical correspondent, i.e. \( ||(p^i_{j,j} + d^i_{j,j}) - p^i_{j,j}|| < \varepsilon \).
3.2 Spatial Distribution

**Distribution in Depth Discontinuity Areas.** Depth discontinuity areas are located near the boundaries between homogeneous depth regions. They result in grey level variations. However, the grey level of the background may be locally different for two correspondents which makes such pixels harder to match than pixels in other areas. Feature point detectors tend to select points in areas where high grey level variations occur and therefore may return points in depth discontinuity areas. Thus, we compute the following measure:

\[
DA = \frac{\text{# feature points in depth discontinuity areas}}{\text{# feature points}}
\]

**Region Wise Distribution.** The regions are extracted using a color segmentation algorithm. The idea is based on the hypothesis that pixels with more or less the same color belong to the same object part (and therefore have more or less the same disparity value). For a detector \( D \), the region wise distribution score is:

\[
RD_{DS} = \frac{\text{# regions holding at least one feature point}}{\text{# regions}}
\]

where \( S \) is a segmentation map. A low ratio reveals a lack of feature points or a bad distribution over the different regions of the image.

3.3 Complementarity

**Contribution Measure.** Let \( \mathcal{P}_{D_1} \) be the set of the feature points returned by a detector \( D_1 \) and \( \mathcal{P}_{D_2} \) the set of the feature point returned by a detector \( D_2 \). The contribution of \( D_2 \) over \( D_1 \) is given by:

\[
\text{contribution}_{D_2|D_1} = \frac{\text{card} \{ \mathcal{P}_{D_2} \} - \text{card} \{ \mathcal{P}_{D_1} \cap \mathcal{P}_{D_2} \}}{\text{card} \{ \mathcal{P}_{D_1} \}}
\]

where the intersection \( \mathcal{P}_{D_1} \cap \mathcal{P}_{D_2} \) is computed considering two points \( p_{i,j} \) and \( p'_{i,j} \) of an image \( I \) as the same point when \( ||p_{i,j} - p'_{i,j}|| < \epsilon \). The higher this value is with a high \( \epsilon \), the larger the differences are between the two sets \( \mathcal{P}_{D_1} \) and \( \mathcal{P}_{D_2} \). This measure is illustrated by figure 1.

This contribution measure is not symmetric since the cardinalities of each detector are different, therefore in order to get an objective idea of the complementarity of the union of two detectors \( \{D_1, D_2\} \), we compute:

\[
\text{contribution}(D_1, D_2) = \text{contribution}(D_2, D_1) = \min(\text{contribution}_{D_2|D_1}, \text{contribution}_{D_1|D_2})
\]

**Repeatability Gain.** The repeatability is computed independently for the detectors \( D_1 \) and \( D_2 \). It is then computed with \( D_1 \& D_2 \) which represents the union \( \mathcal{P}_{D_1} \cup \mathcal{P}_{D_2} \). The gain of repeatability is given by:

\[
\text{gain}_{\text{rep}_{D_1\&D_2}} = \text{rep}_{D_1\&D_2} - \max(\text{rep}_{D_1}, \text{rep}_{D_2})
\]

**Distribution Gain.** The spatial distribution is computed independently for the detectors \( D_1 \) and \( D_2 \). It is then computed with \( D_1 \& D_2 \) which represents the union \( \mathcal{P}_{D_1} \cup \mathcal{P}_{D_2} \). The gain of distribution is given by:

\[
\text{gain}_{\text{RD}_{D_1\&D_2}} = \text{RD}_{D_1\&D_2} - \max(\text{RD}_{D_1}, \text{RD}_{D_2})
\]

where \( S \) is a segmentation map.

4 DATA SET

4.1 Stereo Pairs

Stereo pairs from the Middlebury data set\(^1\) are used for this experimentation\(^2\). They present issues such as

\(^1\)vision.middlebury.edu/stereo/data/
\(^2\)We use the pairs named aloe, art, bowling1, cloth1, cloth4, cones, dolls, mid2, moebius, plastic and teddy.
occlusions and depth discontinuities. They are epipolar rectified and the ground truth is known. However, deformations between two images is not as important as it could occur in applications such as indexation.

4.2 Implementation Details

The programs provided by the detector authors is used when available: HAL\textsuperscript{3}, HEL\textsuperscript{4}, SU\textsuperscript{5}, FA\textsuperscript{6}, SI\textsuperscript{7}, SR\textsuperscript{8} and KA\textsuperscript{9}. Our own implementation is used for the other detectors. For each detector, the following parameters can have an influence on the result:

- **Response Window Size** (for fixed scale detectors) – the smaller this size is, the more the detector is able to select small structures but the greater the noise sensitivity is. On the other hand, the greater this size is, the larger the smoothing effect is and the greater localization errors are.
- **Non Maxima Suppression Window Size** – the smaller this size is, the larger the number of detected feature points is but these feature points may be close to each other.
- **Response Threshold** – a threshold on the response value is often used in order to get rid of false responses. The higher this value is, the smaller the number of feature points is.

It would be interesting to study the influence of each parameter on the result but this exhaustive evaluation is difficult to analyze. Moreover, for the detectors implemented by their authors, we do not always have the possibility to change all the settings. A normalization would also be necessary on all the different values in order to make them comparable. Facing this issue, we decided to fix these parameters once for all for each detector. For the detectors we implemented, we used parameter values that give good results with our experimentations. For the other detectors, we used the default values given by their authors.

5 RESULTS

5.1 Repeatability

The repeatability \( r_{repD} \) of a detector \( D \) is given by computing the mean value of the repeatability rates found over the data set. The repeatability is measured by taking \( \epsilon \in \{0, 1.5, 3\} \). The results for repeatability are shown in figure 2. The detectors with the best repeatability are HA, FA and SI. For \( \epsilon = 0 \) the repeatabilities are very low. Therefore, a “feature points to feature points” matching strategy is not advisable when high precision is required. We recommend a “feature points to neighbours of feature points” matching strategy to settle this issue.

5.2 Spatial Distribution

The results for the depth discontinuity distribution are shown in table 1. The depth discontinuity areas are computed using the provided ground truth disparity maps, see § 4.1. The results for the region wise distribution are shown in table 1. The EDISON\textsuperscript{10} program is used to compute two segmentation maps: (i) an “under” segmentation \( S_1 \) giving about 100 regions for each image ; (ii) a “medium” segmentation \( S_2 \) giving about 500 regions for each image (see figure 1).

The most significant criteria, where the biggest differences are obtained, are \( R_{D,S_2} \) and Card. HAL detector returns very few points and consequently gives also the worst \( R_{D,S_2} \). On the other hand, HA detector seems to obtain the best compromise between cardinality, DA and \( R_{D,S_2} \). According to these results, the best compromises is obtained by the detectors HA, FA and SI.

5.3 Complementarity

**Contribution Measure.** This measure is computed over all the feature points and over the repeated points with \( \epsilon = 3 \) (we consider this value as the minimum

\[ \text{COMPLEMENTARITY OF FEATURE POINT DETECTORS} \]
distance between two different points) in one hand and \(\epsilon = 12\) (large enough value to see which detectors return complementary feature points away from each other) on the other hand. The results are shown in table 2 (the measures are computed over all the feature points) and in table 3 (the measures are computed over the repeated feature points). According to these results, the most complementary detectors are KA+SU, KA+MO and KR+HEL. By reading these results, we can see for instance that by taking the union HA+SI, we add in the worst case 67% of new repeated feature points within a distance of \(\epsilon = 3\) pixels of each other and 2% of new repeated feature points within a distance of \(\epsilon = 12\) (i.e. at least 12 pixels farther than the already repeated feature points).

**Repeatability Gain.** The results are shown in table 4. They show which detectors are the most complementary in terms of repeatability. First, it shows the good complementarity of the following detectors KR+HEL, HEL+SU and KR+SI. Second, it shows which detector to use in combination to another in order to improve the repeatability. The best repeatability was given by FA, see § 5.1. Therefore, adding the feature points from the detector HEL can improve by 2.12% the repeatability of FA.

**Distribution Gain.** The segmentation S2 is used in our experimentation, see § 5.2. The results are shown in table 5. They show which detectors are the most complementary in terms of region based distribution. First, it shows the good complementarity of the following detectors KA+MO, KA+SU and KA+SR. Second, it shows which detector to use in combination to another in order to improve the spatial distribution. The best region based distribution was given by HA, see § 5.2. By reading table 5, we can see that adding the feature points from the detector BE can improve the distribution by 6.14% which gives a score of 86+6.14=92.14% for the union HA+BE.

**Analysis.** These results can be read in two ways: (i) for each detector they give which detector is the most complementary in terms of contribution, repeatability and spatial distribution; (ii) they give the most complementary detectors between them. The best compromises between repeatability and distribution are given by: Harris, FAST and SIFT. Table 6 summa-
Table 4: Repeatability gain. For each couple of detectors $D_1$&$D_2$, we show gain$_{D_1D_2}$ ($\varepsilon = 3$). To determine which detector is the most complementary in terms of repeatability with HA, for instance, look at the HA row and column (here in blue), it shows that it is HEL. For each detector, the best result appears in bold.

<table>
<thead>
<tr>
<th></th>
<th>BE</th>
<th>FA</th>
<th>HAL</th>
<th>HA</th>
<th>HEL</th>
<th>KA</th>
<th>KR</th>
<th>MO</th>
<th>SI</th>
<th>SR</th>
</tr>
</thead>
<tbody>
<tr>
<td>FA</td>
<td>1.57</td>
<td>0</td>
<td>10.36</td>
<td>4.28</td>
<td>4.20</td>
<td>0</td>
<td>3.40</td>
<td>1.48</td>
<td>-0.69</td>
<td>2.60</td>
</tr>
<tr>
<td>HAL</td>
<td>10.75</td>
<td>0</td>
<td>10.36</td>
<td>4.28</td>
<td>4.20</td>
<td>0</td>
<td>3.40</td>
<td>1.48</td>
<td>-0.69</td>
<td>2.60</td>
</tr>
<tr>
<td>HA</td>
<td>3.77</td>
<td>0.87</td>
<td>0</td>
<td>3.88</td>
<td>5.75</td>
<td>1.81</td>
<td>0</td>
<td>3.40</td>
<td>1.48</td>
<td>-0.69</td>
</tr>
<tr>
<td>HEL</td>
<td>4.48</td>
<td>2.12</td>
<td>0.36</td>
<td>3.11</td>
<td>4.47</td>
<td>1.54</td>
<td>2.83</td>
<td>0</td>
<td>3.40</td>
<td>1.48</td>
</tr>
<tr>
<td>KA</td>
<td>0.17</td>
<td>1.69</td>
<td>0.39</td>
<td>3.88</td>
<td>5.75</td>
<td>1.81</td>
<td>0</td>
<td>3.40</td>
<td>1.48</td>
<td>-0.69</td>
</tr>
<tr>
<td>KR</td>
<td>4.71</td>
<td>0.22</td>
<td>0.90</td>
<td>3.88</td>
<td>5.75</td>
<td>1.81</td>
<td>0</td>
<td>3.40</td>
<td>1.48</td>
<td>-0.69</td>
</tr>
<tr>
<td>MO</td>
<td>3.78</td>
<td>0.69</td>
<td>0.62</td>
<td>3.11</td>
<td>4.47</td>
<td>1.54</td>
<td>2.83</td>
<td>0</td>
<td>3.40</td>
<td>1.48</td>
</tr>
<tr>
<td>SI</td>
<td>3.57</td>
<td>1.19</td>
<td>0.76</td>
<td>3.40</td>
<td>2.37</td>
<td>0.82</td>
<td>4.98</td>
<td>3.00</td>
<td>0</td>
<td>3.40</td>
</tr>
<tr>
<td>SR</td>
<td>1.01</td>
<td>0.38</td>
<td>0.09</td>
<td>0.66</td>
<td>2.06</td>
<td>1.48</td>
<td>1.70</td>
<td>0.82</td>
<td>0.19</td>
<td>0</td>
</tr>
<tr>
<td>SU</td>
<td>3.53</td>
<td>1.21</td>
<td>1.65</td>
<td>2.24</td>
<td>5.00</td>
<td>2.60</td>
<td>3.50</td>
<td>2.25</td>
<td>2.46</td>
<td>2.63</td>
</tr>
</tbody>
</table>

Table 5: Distribution gain. For each couple of detectors $D_1$&$D_2$, we show gain$_{D_1D_2}$ ($\varepsilon = 3$) (see 4 for an example of how to read the table).

<table>
<thead>
<tr>
<th></th>
<th>BE</th>
<th>FA</th>
<th>HAL</th>
<th>HA</th>
<th>HEL</th>
<th>KA</th>
<th>KR</th>
<th>MO</th>
<th>SI</th>
<th>SR</th>
</tr>
</thead>
<tbody>
<tr>
<td>FA</td>
<td>3.96</td>
<td>0</td>
<td>3.00</td>
<td>0</td>
<td>3.00</td>
<td>0</td>
<td>3.00</td>
<td>0</td>
<td>3.00</td>
<td>0</td>
</tr>
<tr>
<td>HAL</td>
<td>2.23</td>
<td>1.30</td>
<td>0</td>
<td>2.23</td>
<td>1.30</td>
<td>0</td>
<td>2.23</td>
<td>1.30</td>
<td>0</td>
<td>2.23</td>
</tr>
<tr>
<td>HA</td>
<td>6.14</td>
<td>5.11</td>
<td>1.19</td>
<td>0</td>
<td>6.14</td>
<td>5.11</td>
<td>1.19</td>
<td>0</td>
<td>6.14</td>
<td>5.11</td>
</tr>
<tr>
<td>HEL</td>
<td>8.13</td>
<td>7.80</td>
<td>1.30</td>
<td>6.09</td>
<td>0</td>
<td>8.13</td>
<td>7.80</td>
<td>1.30</td>
<td>6.09</td>
<td>0</td>
</tr>
<tr>
<td>KA</td>
<td>8.53</td>
<td>6.85</td>
<td>7.09</td>
<td>5.54</td>
<td>9.27</td>
<td>0</td>
<td>8.53</td>
<td>6.85</td>
<td>7.09</td>
<td>5.54</td>
</tr>
<tr>
<td>KR</td>
<td>6.47</td>
<td>4.23</td>
<td>5.28</td>
<td>5.86</td>
<td>7.98</td>
<td>6.08</td>
<td>0</td>
<td>6.47</td>
<td>4.23</td>
<td>5.28</td>
</tr>
<tr>
<td>MO</td>
<td>5.82</td>
<td>0.02</td>
<td>5.46</td>
<td>5.26</td>
<td>8.62</td>
<td>12.07</td>
<td>2.30</td>
<td>0</td>
<td>5.82</td>
<td>0.02</td>
</tr>
<tr>
<td>SI</td>
<td>5.32</td>
<td>4.90</td>
<td>1.52</td>
<td>5.59</td>
<td>5.39</td>
<td>6.88</td>
<td>6.12</td>
<td>2.92</td>
<td>0</td>
<td>5.32</td>
</tr>
<tr>
<td>SR</td>
<td>8.13</td>
<td>5.79</td>
<td>4.89</td>
<td>5.67</td>
<td>10.51</td>
<td>11.38</td>
<td>5.79</td>
<td>4.89</td>
<td>5.67</td>
<td>10.51</td>
</tr>
<tr>
<td>SU</td>
<td>6.16</td>
<td>1.01</td>
<td>1.30</td>
<td>4.18</td>
<td>9.72</td>
<td>10.17</td>
<td>3.20</td>
<td>3.27</td>
<td>4.58</td>
<td>1.12</td>
</tr>
</tbody>
</table>

Table 6: This table summarizes the most complementary detectors to the detectors in terms of contribution, repeatability and region-wise distribution. The most complementary detectors between them are Kadir and SUSAN, Kitchen and Rosenfeld and Hessian-Laplace, Moravec and Kadir i.e. they return the most distinct sets of feature points.

<table>
<thead>
<tr>
<th></th>
<th>D</th>
<th>Cont</th>
<th>R</th>
<th>RD$_{D_1D_2}$</th>
<th>S2</th>
</tr>
</thead>
<tbody>
<tr>
<td>HA</td>
<td>HEL</td>
<td>HEL</td>
<td>BE</td>
<td>HEL</td>
<td>HEL</td>
</tr>
<tr>
<td>SI</td>
<td>HEL</td>
<td>KR</td>
<td>KA</td>
<td>HEL</td>
<td>HEL</td>
</tr>
</tbody>
</table>

6 CONCLUSIONS

We proposed an evaluation and a comparison of eleven well-known feature point detectors based on new criteria used to characterize spatial distribution and complementarity. This study aims to be helpful for any applications that need feature points well distributed in the image. It also helps to select the most complementary detectors in terms of region based distribution. This work will be extended on larger transformations between the images.

REFERENCES


