Probabilistic Phase Based Sparse Stereo

Ilkay Ulusoy, Ugur Halici
Computer Vision and Intelligent Systems Research Lab., METU, 06531 Ankara, Turkey
Edwin R. Hancock
Department of Computer Science University of York, York, Y01 5DD, UK

Abstract

In this study, a multi-scale phase based sparse disparity algorithm and a probabilistic model for matching are proposed. The disparity algorithm and the probabilistic approach are verified on various stereo image pairs.

1. Introduction and Motivation

In sparse stereo [6], [12], [13], [14], [18], distinctive image features are extracted and corresponding pairs are matched using a feature-based criteria. To be successful for stereo applications, local features must be robust to image deformations such as noise, rotation, scale and brightness changes. Jenkin, Jepson and Fleet [7], [8], [9] and Sanger [19] describe promising methods based on the output phase behavior of band-pass Gabor filters. Recently, Carneiro and Jepson show that the phase information provided by steerable filters is often locally stable with respect to noise, scale and brightness changes [2], and it is also possible to achieve stability under rotation using steerable filters [1].

Experience shows that disparity estimation from local phase-differences is reliable at near edge locations but yields poor results between elsewhere. In [5] a probabilistic lattice structure is proposed to fill such unreliable regions. There are many other probabilistic algorithms which estimate the disparity map between stereo images. Their starting point is usually to maximize the conditional probability of disparity given a pair of stereo images. In [10] the independence and the similarity of neighboring disparity are modeled. Myers, Wilson and Hancock applied their graph edit distance method to uncalibrated stereo correspondence problem [15]. Olson used template based matching [16]. In [21] feature-based and area-based relaxation is presented.

In this study, a multi-scale phase based sparse disparity algorithm and a probabilistic model for matching are proposed.

2. Multi-scale phase based disparity algorithm

Here we follow the population coding method of Ludtke, Wilson and Hancock [11] to represent feature points using the steerable filters as described in [3]. First, the image is filtered with basis filters at different scales and at each scale we interpolate the filtered images at orientations between $0^\circ$ to $180^\circ$ with $10^\circ$ intervals. Then, the orientation estimate for pixel location $x$ is performed as follows

$$\tilde{\theta}(x) = \arg \left\{\max_{n=1}^{S} r_{n,\theta}(x) \right\}$$

where $r_{n,\theta}(x)$ is the response of steerable filtering at pixel location $x = (x, y)$ for scale $n$ and orientation $\theta$ and $S$ is the number of scales. Finally, the sum of response magnitudes over scales is thresholded and feature points are extracted. The main attributes used in this study for feature-point matching are multi-scale phase. We compute phase at each feature point as follows:

$$\phi_i = \arctan \left( \frac{\text{real}(r_{n,\theta}(x_i))}{\text{imag}(r_{n,\theta}(x_i))} \right)$$

where, $n$ and $\theta$ are the scale and orientation of the applied filter, $\text{real}(r_{n,\theta}(x_i))$ and $\text{imag}(r_{n,\theta}(x_i))$ are real and imaginary parts of the filtered image respectively. We use the phase measurements for three filters of different width, i.e. different scales, to locate correspondences. The width of the narrowest filter is six pixels and the largest filter is 18 pixels. Let $\Phi = [\phi_1 \phi_2 \phi_3]$ be a vector of phase estimates obtained using these filters. For each feature point in the right image, we search over a window for feature points of similar phase vector in the left image (3). A weighting matrix $C_{ij}$ is constructed by using the method described in [19]. Each element $C_{ij}$ measures the similarity of the magnitudes of the feature points $i$ and $j$ at scale $n$. Hence, it is the confidence value for the pair $i - j$. In this study, the disparity is computed as the horizontal pixel distance between corresponding feature point locations, i.e. $\delta = x_i - x_j$. 

0-7695-2128-2/04 $20.00 (C) 2004 IEEE
\[
\hat{j} = \arg \{ \max \{ \Phi_j C_{ij} \Phi_j^T \} \}
\]

We find correspondences for 537 out of 980 feature points in the right-hand image of the “Blocks 1” stereo pair (Figure 1). Out of the 537 matched feature points only 62 are in error, hence the success rate is 90%. Most of the errors are for feature points having orientation in the disparity direction.

Figure 1. Disparity result for Blocks 1 stereo pair.

Disparity results for other image pairs are shown in Figure 2. Blocks 1 and Blocks 2 stereo pairs are taken from CMU stereo database. Tsukuba, Sawtooth and Venus pairs are taken from Middlebury Stereo web page. Also, the results are summarized in Table 1. The success rate is calculated as the percentage of the correct matches over the total number of matched pairs. For textured images, the feature points extracted and the multi-scale phase are particularly informative so that the success rate is high.

Table 1. Success rate of the algorithm for different stereo pairs.

<table>
<thead>
<tr>
<th>Image pair</th>
<th>Number of matched pairs</th>
<th>Success rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blocks 2</td>
<td>1265</td>
<td>73</td>
</tr>
<tr>
<td>Venus</td>
<td>1163</td>
<td>94</td>
</tr>
<tr>
<td>Sawtooth</td>
<td>2938</td>
<td>95</td>
</tr>
<tr>
<td>Tsukuba</td>
<td>2089</td>
<td>83</td>
</tr>
</tbody>
</table>

3. Probabilistic Model

We use a mixture of von Mises distributions (4) to fit the phase differences between corresponding feature points of the Venus stereo pair for each scale where \( P(w) \) is the mixing proportion. We use the von Mises distribution (5) because of its appropriateness to angular data.

\[
p(\phi_k) = \sum_{w=1}^{W} p(\phi_k | \kappa, \mu) P(w) \quad (4)
\]

\[
p(\phi_k | \kappa, \mu) = \frac{1}{2\pi I_0(\kappa)} \exp \{ \kappa \cos (\phi_k - \mu) \} \quad (5)
\]

We seek the parameters \( P(w), \mu \) and \( \kappa \) of the distribution that minimise the negative log-likelihood

\[
E = - \ln \ell (\phi_k) = - \sum_{k=1}^{K} \ln p(\phi_k) \quad (6)
\]

Unfortunately, direct estimation of the parameters is not possible due to the non-linear nature of the resulting system of equations. Here we have used the EM (Expectation Maximization) algorithm in order to find the parameters iteratively. The update equation for \( \mu_w \) is

\[
\mu_w^{(n+1)} = \frac{1}{2} \arctan \left[ \frac{\sum_{k} P^{(n)}(w | \phi_k) \sin(2\phi_k)}{\sum_{k} P^{(n)}(w | \phi_k) \cos(2\phi_k)} \right] \quad (7)
\]
where

\[ P(w|\phi_k) = \frac{p(\phi_k|\kappa_w, \mu_w) P(w)}{p(\phi_k)} \quad (8) \]

For \( \kappa_w \), the solution is not obtained directly. Instead we estimate the following quantity which is in the range \([0, 1]\):

\[ R = \frac{I_1(\kappa_w^{(n+1)})}{I_0(\kappa_w^{(n+1)})} = \frac{\sum_k^K P^{(n)}(w|\phi_k) \cos \left(2 \left(\phi_k - \mu_w^{(n)}\right)\right)}{\sum_k^K P^{(n)}(w|\phi_k)} \quad (9) \]

The solution for \( \kappa_w^{n+1} \) can be approximated to \((1/6)R (12 + 6R^2 + 5R^4)\) for \( R \leq 0.5 \) and \( 1/ \left[2 (1 - R) - (1 - R)^2 - (1 - R)^3\right]\) for \( R > 0.5 \). Finally, the updated mixing proportion is given by

\[ P^{(n+1)}(w) = \frac{1}{R} \sum_k^K P^{(n)}(w|\phi_k) \quad (10) \]

With the parameters of the mixture model to hand, we compute the correspondence probability using fitted mixture model and the measured phase difference \( \Delta \phi_{ij} \). The probability that the point \( i \) is in correspondence with point \( j \) is the product of correspondence probabilities computed at different scales, i.e.

\[ P(i = j) = \prod_n P^n(\phi_{ij}) \quad (11) \]

where

\[ P^n(\phi_{ij}) = p^n(\Delta \phi_{ij}) / \sum_j p^n(\Delta \phi_{ij}) \quad (12) \]

The correspondences are taken so as to maximize \( P(i = j) \), i.e.

\[ \hat{j} = \arg \max P(i = j) \quad (13) \]

Using the correspondences located in this way the computed disparities for Venus stereo pair are very similar to those found using the method described in Section 2. Both results are shown in Figure 3.

The same probabilistic model is applied to other stereo pairs and the results are given in Figure 4. The important thing here is that the probabilistic model obtained from Venus stereo pair is used for other image pairs as well. In Table 2 success rates are listed. Although the number of matched pairs for each stereo pair is more in Table 2 we obtain higher accuracy.

### 4. Summary and Conclusion

We have presented a feature based stereo correspondence method and a probabilistic model of it. Instead of calculating disparities using oriented filters and pooling the results over different orientations [4],[17],[19], a single orientation for each feature is obtained prior to disparity computation. By using multi-scale filtering highly informative feature points are extracted even they are at different scales. Although feature points extracted from image pairs are sparse, since they are the points of high contrast edges that define the bounding contours of objects, they still prove to be informative. Correspondences between feature points are located using multi-scale phase information. The confidence weighting is used to augment phase information with information concerning the magnitude of the steerable filtered image to improve the correspondence method. The smallest success rate is obtained for Blocks 2 stereo pair because nearly half of the total number of edges have orientation parallel to the epipolar line and due to the lack of texture, phase contents are very similar for neighboring edges. The next better success rate is for Tsukuba stereo pair. Tsukuba image pair is known to be a hard image pair for stereo algorithms because they have very complicated depth discontinuities and repeated patterns. The best success rates are obtained for Venus and Sawtooth stereo pairs, which are good examples for matching algorithms. In [20] comparison of feature based correspondence algorithms is given. For the Tsukuba image pair the maximum success rate that they could reach was 71.7% whereas we reach 83% success rate with our matching algorithm. The complexity of our algorithm is \( O(NMSRn^2) \) where \( n \) is assumed to be the width of the widest filter, \( S \) and \( R \) are the total num-

![Figure 3. Disparity found by the method in Section 2 (left image) and by the probabilistic model (right image) for Venus.](image)

<table>
<thead>
<tr>
<th>Image pair</th>
<th>Number of matched pairs</th>
<th>Success rate %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blocks 2</td>
<td>1505</td>
<td>86</td>
</tr>
<tr>
<td>Venus</td>
<td>1310</td>
<td>94</td>
</tr>
<tr>
<td>Sawtooth</td>
<td>3079</td>
<td>96</td>
</tr>
<tr>
<td>Tsukuba</td>
<td>2350</td>
<td>83</td>
</tr>
</tbody>
</table>

![Table 2. Success rate of the probabilistic model for different image pairs.](image)
The number of scales and orientations used and $N$ and $M$ are image dimensions. The feature extraction step of our algorithm takes time whereas matching is real time. Also in this study, a probabilistic algorithm for correspondence matching is proposed which has very different roots than many earlier probabilistic models. Mixture of von Mises distributions are used to model phase differences between corresponding pairs of Venus stereo pair and the model is verified on other images. The model provides not only better results especially for Block 2 stereo pair but also flexibility in search region in matching. The important thing here is that although modeling is done by one stereo pair data only, it works fine for many other images. Another important thing is that von Mises being a circular distribution has been used in EM algorithm for the first time.

References