Object Matching Algorithms Using Robust Hausdorff Distance Measures

Dong-Gyu Sim, Oh-Kyu Kwon, and Rae-Hong Park

Abstract—A Hausdorff distance (HD) is one of commonly used measures for object matching. This work analyzes the conventional HD measures and proposes two robust HD measures based on M-estimation and least trimmed square (LTS) which are more efficient than the conventional HD measures. By computer simulation, the matching performance of the conventional and proposed HD measures is compared with synthetic and real images.

Index Terms—Hausdorff distance, least trimmed square, M-estimation, object matching.

I. INTRODUCTION

Object matching in two-dimensional (2-D) images has been an important topic in computer vision, object recognition, and image analysis [1]–[3]. The performance of the matching method depends on the type of the features used, matching measure criterion, and so on. Low-level matching algorithms, i.e., algorithms using a distance transform (DT) [4]–[6] and a Hausdorff distance (HD) [7]–[9] have been investigated because they are simple and insensitive to changes of image characteristics.

In this paper, we apply the robust statistics of regression analysis [10], [11] to the computation of the HD measures for object matching, resulting in two robust HD measures: M-HD based on M-estimation and least trimmed square-HD (LTS-HD) based on LTS. The two proposed robust approaches yield the correct results, even though the input data are severely corrupted.

The rest of the paper is structured as follows. Section II reviews the conventional HD measures and related robust statistics. In Section III, two proposed robust object matching algorithms based on HD measures are presented [12]–[14]. Experimental results for images with occlusions and noisy images are shown in Section IV, and conclusions are given in Section V.

II. OBJECT MATCHING ALGORITHMS USING THE CONVENTIONAL HD MEASURES

The HD measure computes a distance value between two sets of edge points extracted from the object model and a test image. The classical HD measure [7] between two point sets \( A = \{a_1, \ldots, a_{N_A}\} \) and \( B = \{b_1, \ldots, b_{N_B}\} \) of sizes \( N_A \) and \( N_B \), respectively, is defined as

\[
H(A, B) = \max \{h(A, B), h(B, A)\}
\]

where \( h(A, B) \) and \( h(B, A) \) represent the directed distances between two sets \( A \) and \( B \).

Huttenlocher et al. proposed the partial HD measure in comparing partial portions of images containing severe occlusions or degradation [7]. The directed distance of the partial HD is defined as

\[
h_K(A, B) = R_{a \in A} d_B(a)
\]

where \( d_B(a) \) represents the minimum distance value at point \( a \) to the point set \( B \), and \( R_{a \in A} \) denotes the \( K \)th ranked value of \( d_B(a) \). This HD measure needs one parameter, \( f = K/N_A \), whose range is from 0.0 to 1.0. Depending on the fractional value of \( f \), its performance varies. Experimentally, when \( f \) is about 0.6, good matching results are obtained.

By modifying the HD based on the ranked order statistics, Azencott et al. proposed the CHD measure in comparing binary images [8]. The directed distance of the CHD is defined as

\[
h_{K, I}(A, B) = P_{a \in A} Q_{b \in B} \|a - b\|
\]

where \( P_{a \in A} \) denotes the \( P \)th ranked value of \( Q_{b \in B} \|a - b\| \), with \( Q_{b \in B} \) representing the \( Q \)th ranked value of the Euclidean distance set. The CHD measure requires two parameters: \( p = P/N_A \) and \( q = Q/N_E \), where \( N_E \) denotes the size of the Euclidean distance set. The range of both parameters is from 0.0 to 1.0. Experimentally, when \( p \) is from 0.8 to 0.9 and \( q \) is from 0.01 to 0.05, good matching results are obtained.

Dubuisson and Jain proposed the MHD based on the average distance value in comparing the synthetic images contaminated by four types of noise [9]. The directed distance of the MHD is defined as

\[
h_{MHD}(A, B) = \frac{1}{N_A} \sum_{a \in A} d_B(a).
\]
In the proposed LTS-HD based on the LTS scheme [11], [12], the directed distance \( h_{\text{LTS}}(A, B) \) is defined by a linear combination of order statistics
\[
h_{\text{LTS}}(A, B) = \frac{1}{H} \sum_{i=1}^{H} d_{\text{LTS}}(a)_{(i)}
\]
where \( H \) denotes \( h \times N_A \), as in the partial HD case, and \( d_{\text{LTS}}(x)_{(i)} \) represents the \( i \)th distance value in the sorted sequence \( (d_{\text{LTS}}(x)_{(1)} \leq d_{\text{LTS}}(x)_{(2)} \leq \cdots \leq d_{\text{LTS}}(x)_{(N_A)}) \). The measure \( h_{\text{LTS}}(A, B) \) is minimized by remaining distance values after large distance values are eliminated. So, even if the object is occluded or degraded by noise, this matching scheme yields good results. An optimal fraction \( h \), whose range is from zero to one, depends on the amount of occlusion. If \( h \) is equal to one, this HD measure corresponds to the conventional MHD. In the proposed algorithms, a full search is adopted for finding the optimal matching location, that is, the point that yields a minimum distance is selected because the proposed algorithm are based on a distance measure. The computation time can be reduced by introducing efficient search algorithms such as hierarchical or pyramid schemes.

The object matching algorithms based on two robust HD measures are insensitive to outliers and occlusions, because of employment of the robust estimation in computing the HD measures. Also they yield an efficiency larger than that of the conventional ones, because of the average operation embedded into them, resulting in algorithms relatively insensitive to the change of parameters, e.g., \( \tau \) of the M-HD and \( h \) of the LTS-HD.

The M-HD measure requires comparison and summation operations, whereas the LTS-HD measure requires sorting and summation operations. The partial HD requires a sorting operation as in LTS-HD measure computation, where sorting operations can be computed with linear time and real-time algorithms [16], [17]. As a result, the computational complexity of the two proposed HD measures is almost the same as that of the conventional HD measures such as the partial HD and the MHD. Note that the DT map can be used for fast computation of various HD measures and that the computational complexity of the CHD is much higher than that of other HD measures.

IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

To show the matching performance of the proposed HD measures, we compare the performance of the conventional and proposed HD measures, in terms of the matching position, for synthetic and several real images with various levels of noise, distortion, and occlusions. The matching position is defined by the position, detected by each HD measure, of the object model with respect to the test image. Because the object matching by the CHD is not attractive in terms of the computational load and its matching performance greatly depends on two parameters \( p \) and \( q \), we just select the partial HD and MHD as the conventional HD measures in performance comparison. They have almost the same computational complexity and the same number of parameters as those of the proposed HD measures.

A. Performance Comparison of HD Measures for Matching of Occluded and Degraded Objects

Fig. 1 shows the matching result of a synthetic test image containing occlusion. The 72 \( \times \) 72 object model, a square, is shown in Fig. 1(a), and a 256 \( \times \) 256 test image composed of two quadrangles, a circle, and a triangle is shown in Fig. 1(b). Also the matching result of the proposed LTS-HD, with \( h = 0.6 \), is shown in Fig. 1(c), where the object model is superimposed on the test image. Note that the
Table I shows the comparison of the matching performance in terms of the matching position detected by the conventional and proposed HD measures, for the synthetic image shown in Fig. 1(b), where the correct position is (74, 62). The proposed HD measures yield the position error smaller than the partial HD and MHD. In experiments with the synthetic image, the proposed HD measures give more accurate position than the partial HD, because of the average operation embedded in the proposed HD measures. Also, the proposed robust HD measures produce better results than the MHD by effectively removing outliers caused by occlusions.

As a real test image, we use the 256 × 256 “road” image contaminated by Gaussian noise (standard deviation $\sigma = 30$) [8]. Also about 35% pixels of the target portion of an input image is deleted. The 72 × 72 object model and the test image are shown in Fig. 2(a) and (b), respectively. The matching result obtained by the two proposed HD measures is shown in Fig. 2(c), where the object model is superimposed on the test image.

Table I shows that the proposed HD measures yield better results than the conventional ones for the real image. The matching position estimated by the two proposed HD measures (M-HD with $\tau = 3, 4, 5$ and LTS-HD with $h = 0.6, 0.7, 0.8$), is correct, thus the superimposed matching results are identical, as shown in Fig. 2(c). On the other hand, the matching position detected by

---

**TABLE I**

<table>
<thead>
<tr>
<th>Algorithms</th>
<th>Synthetic image (74,62)</th>
<th>Road image (123,83)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Conventional</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partial HD</td>
<td>$f=0.6$</td>
<td>(73,63)</td>
</tr>
<tr>
<td></td>
<td>$f=0.7$</td>
<td>(72,63)</td>
</tr>
<tr>
<td>M-HD</td>
<td>$f=0.8$</td>
<td>(69,59)</td>
</tr>
<tr>
<td></td>
<td>$f=0.9$</td>
<td>(69,63)</td>
</tr>
<tr>
<td></td>
<td>$f=1.0$</td>
<td>(59,55)</td>
</tr>
<tr>
<td><em>Proposed</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\tau=3$</td>
<td>(73,64)</td>
</tr>
<tr>
<td></td>
<td>$\tau=4$</td>
<td>(73,64)</td>
</tr>
<tr>
<td>LTS-HD</td>
<td>$\tau=5$</td>
<td>(73,64)</td>
</tr>
<tr>
<td></td>
<td>$\tau=6$</td>
<td>(73,64)</td>
</tr>
<tr>
<td></td>
<td>$\tau=7$</td>
<td>(73,64)</td>
</tr>
<tr>
<td></td>
<td>$h=0.6$</td>
<td>(74,64)</td>
</tr>
<tr>
<td></td>
<td>$h=0.7$</td>
<td>(74,63)</td>
</tr>
<tr>
<td></td>
<td>$h=0.8$</td>
<td>(74,63)</td>
</tr>
<tr>
<td></td>
<td>$h=0.9$</td>
<td>(74,63)</td>
</tr>
<tr>
<td></td>
<td>$h=1.0$</td>
<td>(77,63)</td>
</tr>
</tbody>
</table>

---

Fig. 2. Matching result of a real test image. (a) Object model (72 × 72). (b) Test image (256 × 256). (c) Identical matching result of (a) and (b) by the two proposed HD measures.
two conventional HD matching (with $f = 0.6$ for the partial HD, matching results are almost correct) is not correct. According to Table I, a false alarm (123, 158) is detected. This point yields a small distance with a large similarity, whenever the correct matching point is detected.

B. Performance Comparison of HD Measures for Noisy Binary Images

Fig. 3 shows the RMS matching position error for the noisy road image by the conventional and proposed HD measures, as a function of the corresponding parameter of each measure, with the amount of uniform noise level $U$ and fixed line noise ($V = 10$) varying. As the uniform noise $U$ increases, the RMS position error increases.

Fig. 3(a) shows the RMS matching position error by the conventional partial HD, when the partial fraction $f$ is from 0.5 to 1.0. Because the outliers are not effectively removed with respect to ranked order statistics, when $f$ is larger than 0.6, the matching performance is not good. In Fig. 3(b), because there is no data which can measure the similarity between objects, when $\tau$ is equal to 0.0, the matching result is not good. Also for large $\tau$, because the outliers are not effectively removed by the cost function $\rho$, the RMS position error is large. Fig. 3(c) shows that if the parameter $h$ of the LTS-HD is from 0.50 to 0.90, the RMS position error is 0, i.e., the matching results are perfect. If $h$ is larger than 0.90, the LTS-HD yields incorrect results, because the outliers are not effectively removed.

The parameters of the proposed M-HD and LTS-HD can not be accurately determined. Though the parameters of the proposed robust HD measures depend on the amount of noise and various kinds of objects/noise, appropriate $\tau$ and $h$ of M-HD and LTS-HD are about 5 and 0.7, respectively, which are selected experimentally.

V. CONCLUSIONS

This correspondence proposes two efficient HD measures that are robust to outliers and occlusions. They are realized by applying the robust estimators such as M-estimation and LTS methods to the conventional HD algorithm. The effectiveness of the proposed HD measures are tested with various images and noise environments. Further research will focus on the application of the hierarchical structure to two proposed HD measures to reduce the computation time. Also application of other robust estimators to computation of the HD measures is to be investigated.

REFERENCES

Hybrid Estimation of Navigation Parameters from Aerial Image Sequence

Dong-Gyu Sim, Sang-Yong Jeong, Doh-Hyeong Lee, Rae-Hong Park, Rin-Chul Kim, Sang Uk Lee, and In Chul Kim

Abstract—This work presents a hybrid method for navigation parameter estimation using sequential aerial images, where navigation parameters represent the position and velocity information of an aircraft for autonomous navigation. The proposed hybrid system is composed of two parts: relative position estimation and absolute position estimation. Computer simulation with two different sets of real aerial image sequences shows the effectiveness of the proposed hybrid parameter estimation algorithm.

Index Terms—Aerial image, digital elevation model, image matching, navigation, recovered elevation model.

I. INTRODUCTION

Estimation of navigation parameters is important for autonomous navigation and many approaches have been presented [1]–[5]. This paper investigates the estimation of navigation parameters for an aircraft using sequential aerial images. Because only the aerial image sequence is used as an input in our navigation system, the navigation system has advantages in that it is not detected by enemies nor guided by external signals, compared with other active approaches. Also, it can be attached to an aircraft without any special apparatus for compensation of an attitude change. Two test aerial sequences used in this work were acquired from a camera fixed on a light airplane and a helicopter, in which the optical axis of the camera varies according to the aircraft attitude.

This work presents an integrated system for navigation parameter estimation using aerial sequence images. The main objective of the paper is to develop an effective algorithm for real-time implementation. The proposed system is composed of two parts: relative position estimation and absolute position estimation. The former is based on stereo modeling of two successive image frames, whereas the latter is accomplished by image matching with reference images or by using digital elevation model (DEM) information.