Position-invariant surface recognition and localization using infrared sensors

Billur Barshan
Tayfun Aytac
Bilkent University
Department of Electrical Engineering
TR-06800 Bilkent, Ankara, Turkey
E-mail: billur@ee.bilkent.edu.tr

Abstract. Low-cost infrared emitters and detectors are used for the recognition of surfaces with different properties in a location-invariant manner. The intensity readings obtained with such devices are highly dependent on the location and properties of the surface in a way that cannot be represented in a simple manner, complicating the recognition and localization process. We propose the use of angular intensity scans and present an algorithm to process them. This approach can distinguish different surfaces independently of their positions. Once the surface is identified, its position can also be estimated. The method is verified experimentally with the surfaces aluminum, white painted wall, brown kraft paper, and polystyrene foam packaging material. A correct differentiation rate of 87% is achieved, and the surfaces are localized within absolute range and azimuth errors of 1.2 cm and 1.0 deg, respectively. The method demonstrated shows that simple infrared sensors, when coupled with appropriate processing, can be used to extract a significantly greater amount of information than they are commonly employed for. © 2003 Society of Photo-Optical Instrumentation Engineers.

Subject terms: pattern recognition and feature extraction; surface recognition; position estimation; infrared sensors; optical sensing.

Paper 030092 received Feb. 20, 2003; revised manuscript received May 27, 2003; accepted for publication May 28, 2003.

1 Introduction

In this work, we consider the use of a simple infrared sensing system consisting of one emitter and one detector, for the purpose of surface recognition and localization. This paper complements earlier work where we considered the differentiation and localization of objects with different geometries such as plane, corner, edge, and cylinder.
have been determined using the Phong illumination model, and using this information, the infrared sensor employed has been modeled as an accurate rangefinder for surfaces at short ranges. References 18–20 deal with optical determination of depth information. Reference 21 describes a passive infrared sensing system that identifies the locations of the people in a room. Infrared sensors have also been used for automated sorting of waste objects made of different materials. However, to the best of our knowledge, no attempt has been made to simultaneously differentiate and estimate the position of several kinds of surfaces using a small number of simple, low-cost, point sensors. In this paper, we show that by appropriate processing and application of pattern recognition techniques, it is possible to achieve these objectives. Our results show that it is possible to extract a significantly greater amount of information from simple optical sensors than in their usual applications (e.g., the emitter-detector pair employed in this study is marketed as a simple proximity switch).

This paper is organized as follows: In Sec. 2, we describe the surface differentiation and localization process. Two alternative approaches are employed, which are discussed in Secs. 2.1 and 2.2. In Sec. 2.3, a description of how saturated scans are processed is given. Section 3 provides experimental verification of the approaches presented in this paper. Concluding remarks are made in the last section.

2 Surface Recognition and Localization

The infrared sensor used in this study consists of an emitter and detector, works with 20-to-28-V dc input voltage, and provides analog output voltage proportional to the measured intensity. The detector window is covered with an infrared filter to minimize the effect of ambient light on the intensity measurements. Indeed, when the emitter is turned off, the detector reading is essentially zero. The sensitivity of the device can be adjusted with a potentiometer to set the operating range of the system.

The surfaces employed in this study are aluminum, white painted wall, brown kraft paper, and polystyrene foam packaging material. Our method is based on angularly scanning the surfaces over a certain angular range. The infrared sensor is mounted on a 12-in. rotary table (Fig. 1) to obtain angular scans $I(\alpha)$ from the surfaces. Reference data sets are collected for each surface type with 2.5-cm distance increments, ranging from 12.5 to 57.5 cm, at $\theta = 0$ deg. The intensity signal is processed using an 8-bit microprocessor-compatible analog-to-digital converter chip having a conversion time of 100 $\mu$s.

The resulting reference scans for the four surfaces are shown in Figs. 2(a)–2(d). Notice that the scans are peaked at around $\theta = 0$ deg since both specular and diffuse reflections decrease with increasing $\theta$. The intensity scans are $\theta$-invariant but not $r$-invariant; changes in $r$ do not result in any simple scaling. As is shown below, these scans contain sufficient information to identify and localize different surfaces with a good degree of accuracy. Notice that the return signal intensities saturate at an intensity corresponding to about 11-V output voltage.

We now describe how to recognize and determine the position of an arbitrarily located surface whose intensity scan has been observed. First, we check whether the observed scan $I(\alpha)$ exhibits saturation or not. This situation is treated separately as explained in Sec. 2.3.

We start by identifying the surface. Unfortunately, direct comparison with the corresponding curves in Figs. 2(a)–2(d) is not possible, since we do not yet know the distance to the surface, and comparing with all the curves at all distances would be computationally very expensive. Therefore, we exploit the fact that the successive curves in Figs. 2(a)–2(d) exhibit a monotonic dependence on distance. Furthermore, when an observed scan is compared with the several successive curves in any of Figs. 2(a)–2(d), the two measures of difference between them described in Secs. 2.1 and 2.2 below also exhibit a monotonic fall and rise around a single minimum. Therefore, we are assured that we will not be settling at a suboptimal point if we compare the observed scan, not with all scans at all distances, but only with the four scans (one for each surface type) whose central intensities are closest to that of the observed scan. Therefore, for unsaturated scans, only four comparisons need to be made. This remains the case even if the 2.5-cm increments are reduced to smaller values. This has the advantage that the accuracy of the system can be increased without increasing the cost of computation (although a greater number of scans do have to be stored). As a test, we also ran a version of the method where eight comparisons were made, using the scans with the nearest central intensities both above and below the observed central intensity, and also using all of the scans shown in Figs. 2(a)–2(d). These computationally more expensive approaches (the latter one exceedingly more so) did not improve the results over those of comparison with only four scans. In fact, since the systematic elimination of a priori suboptimal scans eliminates the small possibility that they will mistakenly be chosen as the best matching scan due to noise and other errors, results obtained by using all scans are found to be inferior to those obtained by using four scans.

Two alternative approaches are employed in performing the four comparisons. These are discussed below in the following two subsections.

2.1 Least-Squares Approach

First, we estimate the angular position $\theta$ of the surface as follows: Assuming the observed scan pattern is not satu-
rated, we find the angular location of its maximum and the corresponding intensity value. This angular value, denoted \( \theta_{\text{MAX}} \), can be directly taken as an estimate of the angular position of the plane. Alternatively, the angular position can be estimated by finding the center-of-gravity (COG) of the scan as follows:

\[
\theta_{\text{COG}} = \frac{\sum_{i=1}^{n} a_i I(a_i)}{\sum_{i=1}^{n} I(a_i)}, \tag{1}
\]

where \( n \) is the number of samples in the angular scan. Ideally, these estimates would be equal, but in practice they differ by a small amount. They would be equal under ideal conditions because the scans are symmetric and peaked at their center of symmetry. Symmetry follows from the symmetry of the data acquisition configuration, and the maximum value being at the center follows from the decrease of reflections with increasing \( |\theta| \). We consider the use of both alternatives when tabulating our results. From now on, we will refer to either estimate as the center angle of the scan.

Plots of the intensity at the center angle of each scan in Figs. 2(a)–2(d), as a function of the distance at which that scan was obtained, play an important role in our method. Figure 3 shows these plots for the maximum-intensity case.

In this approach, we compare the intensity scan of the observed surface with the four reference scans by computing their least-squares differences after aligning their centers with each other. The mean square difference between the observed scan and the four reference scans, one for each possible surface, is computed as follows:

\[
\mathcal{E}_j = \sum_{i=1}^{n} \left[ I(a_i - \alpha_{\text{align}}) - I_j(a_i) \right]^2, \tag{2}
\]

where \( I_j, j=1,2,3,4 \), denote the four reference scans. Here, \( \alpha_{\text{align}} \) is the angular shift that is necessary to align the two
patterns. The reference scan resulting in the smallest value of $E$ is declared as the observed surface. Once the type of the surface is determined, the range can be estimated by using Fig. 3. We use the set of points associated with the determined surface type and employ linear interpolation between the points at which reference scans are available to determine a distance estimate from the observed intensity value. For instance, if the surface is determined to be a white wall, and the intensity is observed to be 6 V, we use linear interpolation to estimate the distance as approximately 43.5 cm. Note that, this way, the accuracy of the method is not limited by the 2.5-cm spacing used in collecting the reference scans.

2.2 Matched Filtering Approach

As an alternative, we have also considered the use of matched filtering\(^{26}\) to compare the observed and reference scans. The output of the matched filter is the cross-correlation between the observed intensity pattern and the $j$\textsuperscript{th} reference scan normalized by the square root of its total energy:

$$
y_j(l) = \frac{\sum_k I(a_k) I_j(a_k-l)}{\left(\sum_k I(a_k)^2\right)^{1/2}}. \tag{3}
$$

The surface corresponding to the maximum cross-correlation peak is declared as the observed surface type, and the angular position of the correlation peak directly provides an estimate of the azimuth angle of the surface. Then, the distance is estimated by using linear interpolation in Fig. 3 with the intensity value at the azimuth estimate.

2.3 Saturated Scans

If saturation is detected in the observed scan, special treatment is necessary. In the least-squares approach, the mean square differences between the aligned observed scan and all the saturated reference scans are computed and the reference scan with the minimum mean square difference is chosen. The range estimate of the surface is taken as the distance corresponding to the scan resulting in the minimum mean square difference. Similarly, for the matched filter, correlation between the observed scan and all the stored saturated reference scans is computed, and the reference scan resulting in the highest correlation peak is selected. The range estimate is again taken as that of the best-matching scan.

It should be noted that, in the saturated case, range estimation accuracy is limited by the 2.5-cm interval at which the reference scans were taken. If this accuracy is not satisfactory, it can be improved by reducing the intervals. We underline that the 2.5-cm interval does not limit the range estimation accuracy in the unsaturated case, where accurate interpolation is possible from Fig. 3.

In the unsaturated case, the azimuth could be estimated by taking the angular value corresponding to either the maximum value of the intensity curve or its COG. In the case of saturated scans, a single maximum may not be observed, but the COG can still be used to reliably estimate the azimuth. Even when the maximum intensity is used for the unsaturated scans, the COG approach is used for the saturated scans.

3 Experimental Verification and Discussion

In this section, we experimentally verify the proposed method by locating the surfaces at randomly selected distances $r$ and azimuth angles $\theta$ and collecting a total of 100 test scans. The surfaces are randomly located at ranges from 12.5 to 57.5 cm and azimuths from $-45$ to $45$ deg.

The results of least-squares-based surface differentiation are displayed in Tables 1 and 2 in the form of surface confusion matrices. Table 1 gives the results obtained using the maximum intensity values, and Table 2 gives those obtained using the intensity value at the COG of the scans.

<table>
<thead>
<tr>
<th>Surface</th>
<th>AL</th>
<th>WW</th>
<th>BP</th>
<th>PF</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AL</td>
<td>25</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>25</td>
</tr>
<tr>
<td>WW</td>
<td>—</td>
<td>20</td>
<td>3</td>
<td>2</td>
<td>25</td>
</tr>
<tr>
<td>BP</td>
<td>—</td>
<td>4</td>
<td>18</td>
<td>3</td>
<td>25</td>
</tr>
<tr>
<td>PF</td>
<td>—</td>
<td>—</td>
<td>6</td>
<td>19</td>
<td>25</td>
</tr>
<tr>
<td>Total</td>
<td>25</td>
<td>24</td>
<td>27</td>
<td>24</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 2. Surface confusion matrix: least-squares-based recognition (COG variation).

<table>
<thead>
<tr>
<th>Surface</th>
<th>AL</th>
<th>WW</th>
<th>BP</th>
<th>PF</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AL</td>
<td>25</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>25</td>
</tr>
<tr>
<td>WW</td>
<td>—</td>
<td>20</td>
<td>3</td>
<td>2</td>
<td>25</td>
</tr>
<tr>
<td>BP</td>
<td>—</td>
<td>4</td>
<td>18</td>
<td>3</td>
<td>25</td>
</tr>
<tr>
<td>PF</td>
<td>—</td>
<td>—</td>
<td>6</td>
<td>19</td>
<td>25</td>
</tr>
<tr>
<td>Total</td>
<td>25</td>
<td>24</td>
<td>27</td>
<td>24</td>
<td>100</td>
</tr>
</tbody>
</table>
The average accuracy over all surface types can be found by summing the correct decisions given along the diagonal of the confusion matrix and dividing this sum by the total number of test trials (100). The average correct classification rates obtained by using the maximum intensity and the COG variations of the least-squares approach are 81% and 82%, respectively.

Matched-filter differentiation results are presented in Table 3. The average accuracy of differentiation over all surfaces is 87%, which is better than that obtained with the least-squares approach. In Ref. 1, where we dealt with the differentiation of targets with different geometries as opposed to the different surfaces treated here, the least-squares approach resulted in differentiation accuracies of 93% and 89%, and the matched-filtering approach resulted in an accuracy of 97%. Based on these results, we conclude that differentiating targets with different surfaces is considerably more difficult than differentiating targets with different geometries.

As shown in the tables, aluminum is always correctly identified regardless of which method is used, due to its distinctive signature. The remaining surfaces are comparable in their correct identification percentages. Brown kraft paper is the surface most confused with others, especially polystyrene foam. Although the intensity scans of these two surfaces do not resemble each other in the unsaturated region, their saturated scans are similar, contributing to the misclassification rate. Nearly all misclassified surfaces are located at nearby ranges where the return signal intensities are saturated. This means that the misclassification rate can be reduced by increasing the lower limit of the range interval at the cost of reducing the operating range.

The average absolute range and azimuth estimation errors for the different approaches are presented in Table 4 for all surface types. As seen in the table, using the maximum-intensity and COG variations of the least-squares approach, the surface ranges are estimated with average absolute range error of 1.5 cm in both cases. Matched filtering results in an average absolute range error of 1.2 cm, which is better than that obtained with the least-squares approach. The greatest contribution to the range errors comes from surfaces that are incorrectly recognized. If we average over only correctly recognized surfaces, the average absolute range errors become 1.0, 1.1, and 1.2 cm for the maximum intensity and COG variations of least-squares and the matched filter approaches, respectively. Since these three numbers are closer than the corresponding numbers in Table 4, we may conclude that the superior range accuracy of matched filtering is mostly a consequence of its superior differentiation accuracy.

The major contribution to range errors comes from saturated scans, where linear interpolation from Fig. 3 cannot be employed to obtain better range estimates. Consequently, surfaces for which saturation occurs over a greater portion of the operating range exhibit greater range estimation errors, with aluminum being the worst.

As for azimuth estimation, matched filtering results in an average absolute estimation error of 1.0 deg, which is the best among the approaches compared. Averaging the azimuth errors over only correctly differentiated surfaces does not result in significant changes. This is because azimuth estimation is not dependent on correct differentiation. The COG variation is, on the average, better than the maximum intensity variation in azimuth estimation, because COG-based calculations average out the noise in the return signal intensities.

We have also considered expanding the range of operation of the system. As an example, changing the operating range from [12.5 cm, 57.5 cm] to [5 cm, 60 cm], results in a reduction of the correct differentiation percentage from 87% to 80%. This reduction in performance is mostly a consequence of highly saturated scans and scans with very low intensities, both of which are prone to greater errors.

Light reflected from a surface consists of specular and diffuse components. The specular component is concentrated where the reflection angle equals the incidence angle, whereas the diffuse component is spread in all directions with a cosine factor. For different types of surfaces, the contribution of these two components and the rate of decrease of intensity with the scan angle $\alpha$ is different. It is this difference which results in a characteristic intensity scan pattern (signature) for each surface, enabling us to distinguish them without knowing their positions. In contrast, a system relying only on reflected energy could not distinguish between a highly reflecting distant object and a less reflecting nearby one. Occasionally, two very distinct surfaces may have intensity scans with very similar dependence on $\alpha$, in which case they cannot be reliably differentiated with the present method.

### Table 3 Surface confusion matrix: matched-filter-based recognition.

<table>
<thead>
<tr>
<th>Surface</th>
<th>AL</th>
<th>WW</th>
<th>BP</th>
<th>PF</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AL</td>
<td>25</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>25</td>
</tr>
<tr>
<td>WW</td>
<td>—</td>
<td>21</td>
<td>3</td>
<td>1</td>
<td>25</td>
</tr>
<tr>
<td>BP</td>
<td>—</td>
<td>1</td>
<td>21</td>
<td>3</td>
<td>25</td>
</tr>
<tr>
<td>PF</td>
<td>—</td>
<td>—</td>
<td>5</td>
<td>20</td>
<td>25</td>
</tr>
<tr>
<td>Total</td>
<td>25</td>
<td>22</td>
<td>29</td>
<td>24</td>
<td>100</td>
</tr>
</tbody>
</table>

### Table 4 Absolute range and azimuth estimation errors for all surfaces.

<table>
<thead>
<tr>
<th>Method</th>
<th>Quantity</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AL</td>
<td>WW</td>
</tr>
<tr>
<td>Least squares (max) $r$ (cm)</td>
<td>2.4</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td>$\theta$ (deg)</td>
<td>0.8</td>
</tr>
<tr>
<td>Least squares (COG) $r$ (cm)</td>
<td>2.4</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td>$\theta$ (deg)</td>
<td>0.8</td>
</tr>
<tr>
<td>Matched filter $r$ (cm)</td>
<td>1.7</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>$\theta$ (deg)</td>
<td>0.8</td>
</tr>
</tbody>
</table>

4 Conclusion

In this study, differentiation and localization of four types of surfaces is achieved using an inexpensive infrared emitter-and-detector pair. Different approaches are com-
pared with respect to correct differentiation and to range
and azimuth estimation accuracy. One advantage of our
system is that it does not greatly depend on environmental
conditions, since we employ an active sensing modality. A
typical use of the demonstrated system would be in mobile
robotics in surveying an unknown environment composed of
several different types of surfaces, or in industrial appli-
cations where different materials must be identified and
separated.

The main accomplishment of this study is that even
though the intensity patterns are highly dependent on sur-
face location and properties, and this dependence cannot
be represented by a simple relationship, we achieve position-

invariant differentiation of different types of surfaces. A
correct differentiation rate of 87% over all surface types is
achieved, and surfaces are localized within absolute range
and azimuth errors of 1.2 cm and 1.0 deg, respectively.

In earlier work, we had considered differentiation and
localization of objects having different geometries such as
plane, corner, edge, and cylinder, as opposed to the differ-
entiation and localization of different surfaces considered in
this paper. In that work, 97% correct differentiation was
achieved. Comparing this with the 87% correct differen-
tiation reported in this paper, we conclude that specular and
diffuse reflection characteristics are not as distinctive as
gerometric reflection characteristics. Current work investi-
gates the deduction of both the surface type and the geo-
metry of the target from its intensity scan without knowing its
location. Preliminary results indicate that the method of the
present paper can be applied to this case by treating the
combination of a particular geometry and particular surface
as a generalized target type.27

Acknowledgments
This research was supported by TUBITAK under BDIP and
197E051 grants. The authors would like to thank the De-
partment of Engineering Science of the University of Ox-
ford for donating the infrared sensors.

References
1. T. Aytac and B. Barshan, “Differentiation and localization of targets
2. T. Aytac and B. Barshan, “Rule-based target differentiation and posi-
42(6), 1766–1771 (2003).
3. F. T. S. Yu and S. Yin, Eds., Selected Papers on Optical Pattern
4. C. D. Psaltis, “Scale invariant optical correlation using Mellin trans-
5. M. McDonnell, “Clarification on use of Mellin transform in optical
6. H. H. Arsenault, Y. N. Hsu, and K. Chalasinska-Macukow, Rotation-
7. G. Gheen, “Design considerations for low-clutter, distortion invariant
8. C. Gu, J. Hong, and S. Campbell, “2-D shift invariant volume holo-
10. C. Gu, J. Hong, and S. Campbell, “2-D shift invariant volume holog-