A Fuzzy Invariant Indexing Technique for Object Recognition under Partial Occlusion

Thorsten GRAF, Alois KNOLL, and André WOLFRAM

University of Bielefeld, Faculty of Technology
Postfach 10 01 31, D–33501 Bielefeld, Germany
Phone: +49 (0) 521 106-2952, Fax : +49 (0) 521 106-2962
E-mail:{graf,knoll,andre}@techfak.uni-bielefeld.de

Abstract. In this paper a new approach to the recognition of partially occluded objects employing fuzzy invariant values and fuzzy if-then-rules is presented, called fuzzy invariant indexing (FII). Compared with traditional invariant indexing, the FII-technique offers the following advantages: most importantly, as it generally produces object hypotheses with different credibilities, the recognition quality can be considerably increased, especially in the case of very similar objects; secondly, the ability is provided to control the recognition process during the hypothesis evaluation stage, e.g. by exploring the most credible hypothesis first; and thirdly, a FII-based recognition system can be extended in a closed form, i.e. new attributes may be added to the fuzzy classification rules causing only minor changes to the original structure of the system. The recognition performance of the FII-technique is demonstrated for partially occluded (quasi-)planar objects in real image scenes taken from different camera viewpoints.

Keywords: fuzzy sets, invariant indexing, object recognition, occlusion

1 State of the Art

The recognition of partially occluded objects is undoubtedly one of the most challenging tasks in computer vision. Recent research has indicated that the use of invariants as shape descriptors (i.e. functions of geometric configurations remaining unaffected under particular classes of transformations; see Sect. 2.2) is a promising and powerful approach to tackle this problem [1,2]. Several recognition systems based on invariant theory have been developed, e.g. the early system based on the geometric hashing technique [3], the LEWIS- [4] or the MORSE-system [5].

All these systems have to deal with the unavoidable fluctuation of invariant values caused mainly by a noisy imaging hardware or an inaccurate feature extraction. Hence, contrary to the theoretical expectations, e.g. the projective invariant values observed in real images fluctuate for perspective views of an object. This problem must be handled within the indexing stage in every recognition system based on invariants. Usually the invariant indexing is done by hashing into a discrete index space, where all points belonging to an object are marked. To handle the fluctuation of the invariant values, (overlapping) values within certain environments rather than only a single point are used to index an object model. Unfortunately, this results in the problem of an invariant value possibly indexing several object models. This produces different but equally weighted object hypotheses diminishing the possibility to identify the most likely hypothesis.

2 Object Recognition through Fuzzy Invariant Indexing

2.1 New Aspect

We model the fluctuation of the invariant values of a geometric configuration that can be extracted for an object with fuzzy sets [6]. For performing the indexing the resulting fuzzy invariant values are used in disjunctive connected fuzzy if-then-rules of the following form:

$$\begin{array}{c}\text{IF } i^k_{m1} = \widetilde{I}^k_{m1} \text{ AND } \cdots \text{ AND } i^k_{mN^k_m} = \widetilde{I}^k_{mN^k_m} \\ \text{THEN } o^k_m = \widetilde{O}^k\end{array} \qquad \text{with} \qquad \begin{array}{rcl} k &=& 1,2,\ldots,K \\ m &=& 1,2,\ldots,M^k \\ n &=& 1,2,\ldots,N^k_m \end{array} \tag{1}$$

where i^k_{mn} denotes the n-th input variable of subrule m for the k-th object, $\widetilde{I}^k_{mn}$ the corresponding fuzzy invariant value (automatically generated, see Sect. 2.2), o^k_m the ouput variable of subrule m and $\widetilde{O}^k$ the k-th object class modelled as a fuzzy singleton. The total amount of antecedents (N^k_m) depends on the number of independent invariants of the underlying geometric configuration of subrule m for object k and the total amount of subrules (M^k) depends on the number of different geometric configurations for object k. Since we use fuzzified invariant values and the generation of object hypotheses is done by evaluating the fuzzy rules, we call this approach fuzzy invariant indexing (FII).

2.2 Generation of Fuzzy Invariant Values

The main problem of the fuzzy rule generation is to find appropriate membership functions to model the fuzzy invariant values $\widetilde{I}^k_{mn}$ in (1) for a given object. The investigation of invariant values measured in different perspective views has indicated that these fluctuations can be adequately approximated by bell-shaped membership functions $\mu_{\widetilde{I}^k_{mn}}$:

$$\mu_{\widetilde{I}^k_{mn}}(u) = e^{-\frac{(u-\alpha^k_{mn})^2}{2\beta^k_{mn}{}^2}}, \quad u \in \mathbb{R} \tag{2}$$

The parameters $\alpha^k_{mn}, \beta^k_{mn}$ determining the shape of the function (2) are chosen as follows: $\alpha^k_{mn} = \frac{1}{N}\sum_l I_l$ and $\beta^k_{mn} = \frac{1}{N}\sum_l (I_l - \alpha^k_{mn})^2$, where I_l, $1 \le l \le N$ are the invariant values for an object taken in N different images.

For example, consider the test object in Fig. 1a. It consists of a geometric structure of a pair of coplanar conics for which two well-known and independent projective invariants can be evaluated [2]:

$$I_1(\mathbf{C_1}, \mathbf{C_2}) = \frac{\mathrm{trace}(\mathbf{C_1}^{-1}\mathbf{C_2})\,|\mathbf{C_1}|^{\frac{1}{3}}}{|\mathbf{C_2}|^{\frac{1}{3}}} \tag{3}$$

and

$$I_2(\mathbf{C_1}, \mathbf{C_2}) = \frac{\mathrm{trace}(\mathbf{C_2}^{-1}\mathbf{C_1})\,|\mathbf{C_2}|^{\frac{1}{3}}}{|\mathbf{C_1}|^{\frac{1}{3}}} \tag{4}$$

where $\mathbf{C_1}, \mathbf{C_2}$ are the conic coefficient matrices and $|\cdot|$ denotes the determinant. To generate the membership functions modelling the fuzzy invariant values we take 30 images of the object from different views. The histograms of the calculated invariant values, using Eq. (3),(4), are shown in Fig. 1b. Using Eq. (2) and its associated equations for determining the parameters α^k_{mn} and β^k_{mn}, we obtain the resulting membership functions in Fig. 1c.

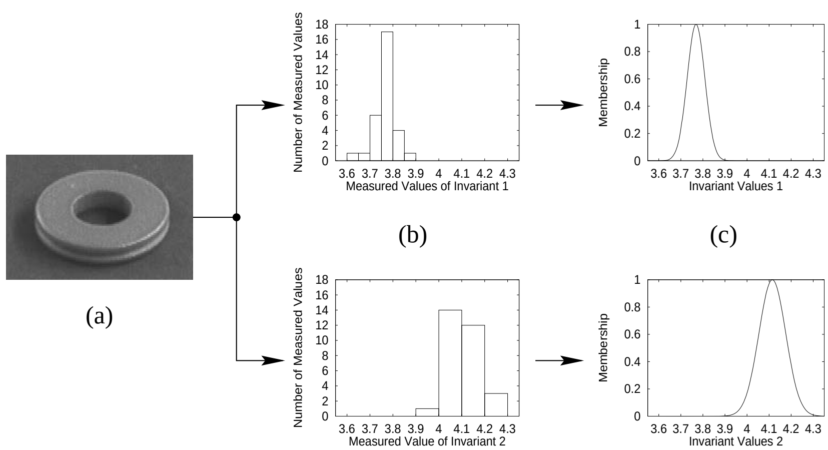


Figure 1: (a) Test Object, (b) Histograms of Invariant Values (from 30 Images) and (c) Fuzzy Invariant Values (see Eq. (2): $\alpha^k_{11} = 3.77, \alpha^k_{12} = 4.11, \beta^k_{11} = 0.04$ and $\beta^k_{12} = 0.06$).

2.3 Generation of Object Hypotheses

The generation of object hypotheses is done by inferring the fuzzy rules:

$$\mu_{o_m^k} := \min_{1 \le n \le N_m^k} \mu_{\widetilde{I}_{mn}^k}(i_{mn}^k) \tag{5}$$

where $\mu_{o_m^k}$ is the output of m-th subrule for object k and i_{mn}^k are the measured invariant values. These subresults are connected disjunctively:

$$\mu_{o^k} = \max_{1 \le m \le M^k} \mu_{o_m^k} \tag{6}$$

The final result is the indexed k-th object model with the measured credibility μ_{o^k}.

3 Experimental Results

3.1 FII-Based Recognition System

We have applied the proposed fuzzy invariant indexing technique to an object recognition system for partially occluded (quasi-)planar objects (Fig. 2).

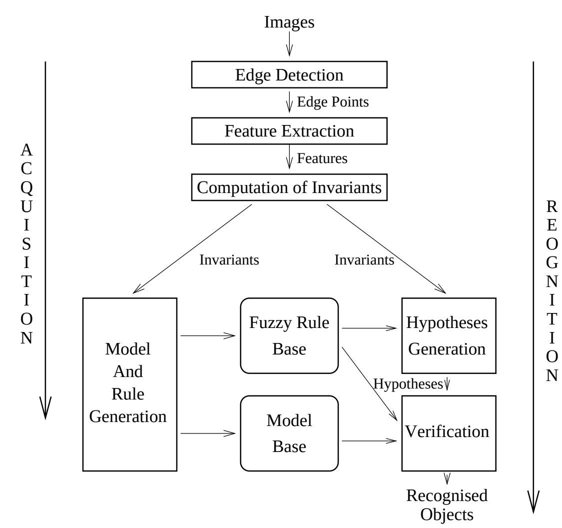
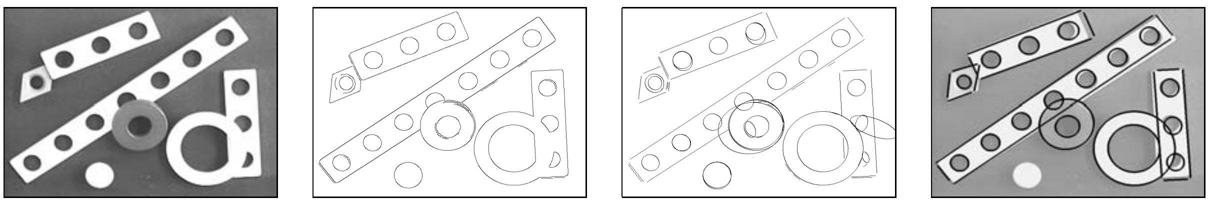


Figure 2: Structure of the FII-Recognition System.

The recognition process consists of: (a) edge detection, (b) feature fitting (lines and ellipses), (c) invariant calculation (for the geometric structures: pair of conics, Eq. (3),(4), and one conic and three lines [4]), (d) hypothesis generation as shown in Sect. 2.3 and (e) verification. As shown, the system is able to learn the fuzzy rules for the recognition process automatically. This is done offline in the model and rule generation. Although this system is intended for recognizing (quasi-)planar objects only, this is no principle limitation for the FII-technique.

3.2 Performance of the FII-Recognition System

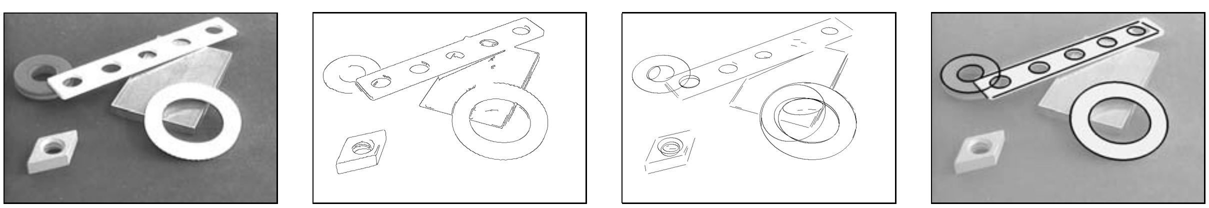
The performance of the FII-recognition system is tested for an object domain of (quasi-) planar, wooden toy objects (rims, tyres, nuts and slats). The first test scene, Fig. 3a,



(a) Original Image. (b) Extracted Edges. (c) Fitted Features. (d) Result.

Figure 3: FII-Recognition Results I.

is taken perpendicular to the planar object surfaces. It consists of two three-hole-slats, one seven-hole-slat, a nut, a rim, a tyre and one unknown object, with varying degree of occlusion. Since the detected edge points (Fig. 3b) as well as the fitted features (Fig. 3c) provide a reliable image description, all of the known objects are recognized correctly (Fig. 3d). In the second test scene, Fig. 4a, a five-hole-slat, a nut, a rim, a tyre and



(a) Original Image. (b) Extracted Edges. (c) Fitted Features. (d) Result.

Figure 4: FII-Recognition Results II.

one unknown object are used. This scene is taken from an angle of about 40 degrees. Fig. 4b shows the detected edge points and Fig. 4c the fitted features. In this scene the system detects all of the known objects except for the nut (Fig. 4d). The reason for this deficiency is a consequence of an inaccurate feature extraction: The relation between the conic size and the line locations does not correspond to the learned object model; hence, the calculated invariant values differ too much from the desired values and no object hypothesis is generated.

3.3 Comparison between Crisp and Fuzzy Invariant Indexing

In the following we compare the proposed FII-technique with the usual invariant indexing method. We implemented a "crisp version" of our recognition system (CII), in which we use intervals instead of fuzzy invariant values. To emphasize the advantages of the FII-technique over the CII-technique we apply the system to the difficult case of very similar objects named rim22, rim25, rim28, rim30, rim32, rim35 and rim38 in Fig. 5a with constant exterior diameter (50mm) but varying interior diameters, where the numbers denote the interior diameters.

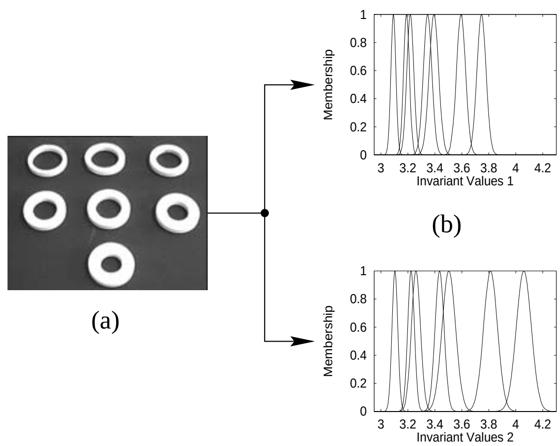


Figure 5: Fuzzy Invariant Values of Seven Similar Rims.

For these objects we get the fuzzy invariant values shown in Fig. 5b, computed for the first (top) and second (bottom) invariant of a pair of conics. From right to left the membership functions represent the objects rim22 to rim38. As shown, the membership functions overlap each other. In two cases, between objects rim28 and rim30 as well as between rim32 and rim35, the overlap is extremely high. For these objects we expect a considerable difference in the discrimination ability of the two implemented systems. The comparison is done through recognizing the objects in 210 different images. The results of this recognition process are summarized in Tab. 1.

It turns out that the recognition system based on the FII-technique provides a better discrimination between the objects than the crisp system, as the recognition rate is gen-

Table 1: CII-/FII-Recognition Results.

	CII			FII		
	corr.	false	rate	corr.	false	rate
rim22	27	3	90.0%	27	3	90.0%
rim25	17	13	56.7%	23	7	76.7%
rim28	19	11	63.3%	24	6	80.0%
rim30	6	24	20.0%	11	19	36.7%
rim32	18	12	60.0%	19	11	63.3%
rim35	2	28	6.7%	19	11	63.3%
rim38	19	11	63.3%	29	1	96.7%
$\sum$	108	102	51.4%	152	58	75.2%

erally higher. Only for one object, rim22, we get an equivalent rate. As expected, we achieve the greatest differences in the recognition rates for the objects with the biggest "clash". The recognition rate of the crisp system for rim30 is only 20% and for rim35 as low as 6.7%. The FII-technique improves these rates to 36.7% and 63.3% respectively. Altogether, the FII-recognition system possesses a recognition rate of about 75.2% while the crisp version of the system only reaches a rate of 51.4%.

3.4 Integration of further Attributes

In this section we demonstrate how the fuzzy rules can easily be extended by adding further attributes. For this we modify our recognition system by integrating the non-invariant attribute of brightness to the fuzzy rules learned before, where the brightness is measured along the underlying geometric structures of the fuzzy rules; e.g. the rule for the rim in Fig. 1a looks like:

$$\text{IF (inv1} \approx 3.77\text{) AND (inv2} \approx 4.11\text{) AND (brightness is HIGH)}$$
$$\text{THEN (object is RIM)}$$

First experimental results show that this extension reduces the number of generated object hypotheses by 14%, where merely false positives are suppressed. For example, the extended fuzzy rules decrease the number of hypotheses for Fig. 3a from 615 to 562 and for Fig. 4a from 154 to 148. As a result the integration of further attributes enhances the performance of the recognition system in two ways: It speeds up the recognition process since fewer object hypotheses must be investigated into in the time consuming verification stage and secondly the robustness of the system is increased, since fewer false positives are established.

4 Conclusions

We have presented a new invariant indexing technique for the hypothesis generation of recognition systems based on fuzzy invariant values and fuzzy if-then-rules. This new approach uses the computed invariant values not only as object clues but also as measure for the credibility of the object hypotheses. As shown, a high discrimination ability between objects is achieved, especially in the case of very similar shape. Due to the use of human readable fuzzy rules the system may be flexibly adapted to new environments.

Future work will concentrate on using more complex rules and extending the fuzzy rules through further attributes, especially color, and furthermore on extending the implemented FII-recognition system for recognizing partially occluded 3D objects.

Literature

[1] D.A. Forsyth, J.L. Mundy, A.P. Zisserman, C. Coelho, A. Heller, and C.A. Rothwell. Invariant descriptors for 3-D object recognition and pose. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(10):971–991, October 1991.

[2] J.L. Mundy and A. Zisserman. Introduction - towards a new framework for vision. In [7], pages 1–39. MIT Press, 1992.

[3] Y. Lamdan, J.T. Schwartz, and H.J. Wolfson. Object recognition by affine invariant matching. In *Proc. Computer Vision and Pattern Recognition*, pages 335–344, 1988.

[4] C.A. Rothwell. *Object recognition through invariant indexing.* Oxford University Pr., 1995.

[5] J.L. Mundy, C. Huang, J. Liu, W. Hoffman, D.A. Forsyth, C.A. Rothwell, A. Zisserman, S. Utcke, and O. Bournez. Morse: A 3D object recognition system based on geometric invariants. In *Proc. ARPA Image Understanding Workshop, Monterey, California, November 13-16*, pages 1393–1402, 1994.

[6] L.A. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.

[7] J.L. Mundy and A. Zisserman, editors. *Geometric invariance in computer vision.* Cambridge, Mass.: MIT Press, 1992.