Recognition of Partially Occluded Objects through Fuzzy Invariant Indexing

Thorsten Graf *, Alois Knoll, and André Wolfram
University of Bielefeld, Faculty of Technology
Postfach 10 01 31, D-33501 Bielefeld, Germany
{graf,knoll,andre}@techfak.uni-bielefeld.de

Abstract

We present a new approach to the recognition of partially occluded objects employing fuzzy invariant values and fuzzy if-then-rules, called fuzzy invariant indexing (FII).

Compared with traditional invariant indexing, the fuzzy method proposed here offers the following advantages: firstly, as shown in the experimental results of this paper, the recognition quality may be considerably increased in the case of similar objects; secondly, the ability is provided to control the recognition process during the hypothesis evaluation stage, and thirdly, a FII-based recognition system can be simply extended in a closed form, i.e. new attributes may be added to the fuzzy classification rules resulting in only minor changes to the original structure of the system. We demonstrate the recognition performance of the new FII-technique for partially occluded (quasi-)planar objects in real image scenes taken from different camera viewpoints and conclude the paper with a discussion of the potential of the method and directions of possible future research.

1. Introduction

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 is a promising and powerful approach to tackle this problem.

Mathematically, invariants are functions of geometric configurations remaining unaffected under particular classes of transformations (for good introductory papers see [2, 6]), e.g. the class of projective transformations modelling the camera mappings of a vision system. Since these invariants are independent of the viewpoint of the camera, the measured projective invariant values of an object can be used efficiently in the hypothesis generation as an index into an object-lookup table (see Sect. 2). This technique is called invariant indexing.

In the recent past, several recognition systems based on invariant theory have been developed, e.g. an early system based on the geometric hashing technique [3], the LEWIS-system [7] or the MORSE-system [4]. All these systems have to cope with the unavoidable fluctuation of invariant values caused mainly by noisy imaging hardware and/or inaccurate feature extraction. This fluctuation causes the invariant values (to be used as indices) to be "smeared" over a certain interval. Hence, if the invariant values of several object models are close to one another, it may become very difficult to discriminate between observed objects, i.e. to establish an unambiguous mapping between the observation and the correct object model.

We present a new invariant indexing method, called fuzzy invariant indexing (FII), that uses fuzzy invariant values, i.e. fuzzified values of invariants, and fuzzy if-then-rules.

The paper is arranged as follows. In Section 2 we explain in detail the proposed FII-technique for generating object hypotheses. It is shown how FII-classification rules can be generated automatically from real image data. In Section 3 we describe the construction of a FII-based recognition system. The recognition performance of the new FII-technique is demonstrated in Section 4 for partially occluded (quasi-)planar objects in real image scenes taken from different camera viewpoints. Finally, in Section 5 our conclusions and directions of possible future research are presented.

2. Object recognition through FII

2.1. Motivation

Since images taken from real-world scenes (and using real-world equipment) are generally discrete, cluttered, and noisy, the observed projective invariant values fluctuate between different perspective views of an object.

* T. Graf's contribution to this work was in part funded by the Deutsche Forschungsgemeinschaft (DFG).

This problem must be handled within the indexing stage of every recognition system based on invariants. In our context indexing means to assign image features to adequate model features and therefore to generate object hypotheses.

Usually, invariant indexing hashes into a discrete index space, where all points belonging to an object are marked. For indexing, a hashing function is evaluated for the measured independent invariant values of a geometric configuration part of an object. The number of the independent invariants depends on the underlying geometric configuration [2, 6]. For example, a pair of coplanar conics has two independent projective invariants (see Sect. 2.2). To overcome the fluctuation not only a single point of the index space is marked but also the neighbouring ones. So invariant values of a certain neighbourhood are mapped to the same object with equal weight.

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

$$\begin{aligned} &\text{IF } i^k_{m1} = \widetilde{I}^k_{m1} \text{ AND} \cdots \text{AND } i^k_{mN^k_m} = \widetilde{I}^k_{mN^k_m} \\ &\qquad\qquad \text{THEN } o^k_m = \widetilde{O}^k \end{aligned} \tag{1}$$

$$\text{with} \quad \begin{aligned} k &= 1, 2, \ldots, K \\ m &= 1, 2, \ldots, M^k \\ n &= 1, 2, \ldots, N^k_m \end{aligned}$$

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 (generated automatically, 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 the fluctuations can be adequately approximated by bell-shaped membership functions $\mu_{\widetilde{I}^k_{mn}}$:

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

where the parameters $\alpha^k_{mn}, \beta^k_{mn}$ determining the shape of the function are chosen as follows:

- The parameter α^k_{mn} determines the position of the maximum of the bell-shaped function (2). Therefore this parameter should be the mean of the fluctuating invariant values: $\alpha^k_{mn} = \frac{1}{N}\sum_l I_l$, where I_l, $1 \le l \le N$ are the invariant values for an object taken in N different images.
- The parameter β^k_{mn} determines the position of the inflexions of (2), which are located at $\alpha \pm \beta$. This parameter should be the standard deviation of the invariant values: $\beta^k_{mn} = \left(\frac{1}{N}\sum_l (I_l - \alpha^k_{mn})^2\right)^{\frac{1}{2}}$.

For example, consider the two well-known and independent projective invariants of a pair of coplanar conics [6]:

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

and

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

where $\mathbf{C_1}, \mathbf{C_2}$ are the conic coefficient matrices and $|\cdot|$ denotes the determinant. Figure 1a shows the invariant values for an object 'rim' (such as shown in Fig. 3) measured in 80 images. The invariant values on the x-axis are calculated using Eq. (3), the invariant values on the y-axis are calculated using Eq. (4). The distributions of these invariant values are depicted in the histograms in Figures 1b+c. For the example (Figure 1a) we get the values $\alpha^k_{11} = 3.766$, $\beta^k_{11} = 0.038$ for the parameters of the first fuzzy invariant value and $\alpha^k_{12} = 4.094$, $\beta^k_{12} = 0.062$ for the second. This leads to the fuzzy invariant values shown in Figures 1d+e.

2.3. Generation of object hypotheses

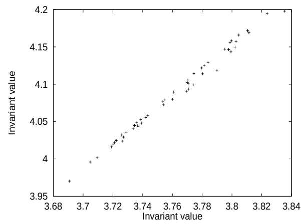
The object hypotheses are generated through the FII-technique by inferring the fuzzy rules:

$$\mu_{o^k_m} := \min_{1 \le n \le N^k_m} \mu_{\widetilde{I}^k_{mn}}(i^k_{mn}) \tag{5}$$

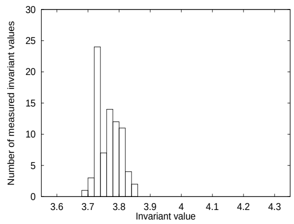
where $\mu_{o^k_m}$ is the output of m-th subrule for object k and i^k_{mn} are the measured invariant values. These subresults are combined disjunctively:

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

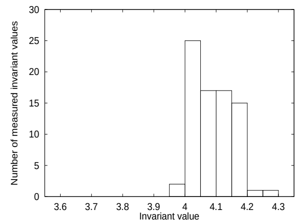
The final result is the indexed k-th object model with the measured credibility μ_{o^k}.



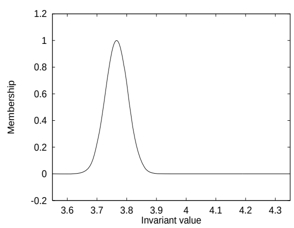
(a) Invariant values (using Eq. (3),(4))



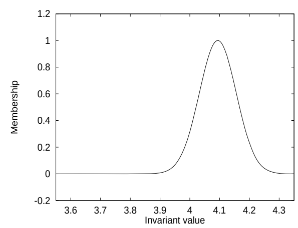
(b) Histogram of invariant values (using Eq. (3))



(c) Histogram of invariant values (using Eq. (4))



(d) Fuzzy invariant value 1



(e) Fuzzy invariant value 2

Figure 1. Example for the generation of fuzzy invariant values (test object 'rim', Fig. 3)

3. The implemented FII-recognition system

Now we apply the proposed fuzzy invariant indexing technique to an object recognition system for partially occluded (quasi-)planar objects. The structure of this FII-recognition system, (see Figure 2), is similar to other systems like [3, 7] but differs in the fuzzy rule base. 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.

The system consists of the following modules:

1. *Edge detection:*
   The first stage of the recognition system is the edge detection. In the implemented system we use the Canny edge detector [1], which takes a greyscale image as input, and generates as its output linked edge points.

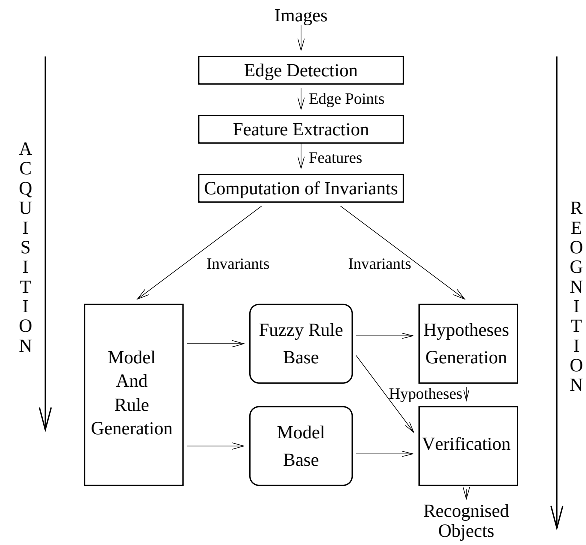


Figure 2. FII-recognition system

2. *Feature extraction:*
   In the feature extraction geometric primitives are fitted to the extracted edge points, where the primitives that are used depend on the objects to be recognized. For the object domain shown in Figure 3 straight lines and ellipses are suitable.

3. *Computation of invariants:*
   As invariants can be computed for different geometric configurations, this module consists of two steps:

   Firstly, the extracted features are grouped into configurations for which the invariants can be computed. In the implemented system we use the invariants of two geometric configurations: For the invariants of a pair of conics (see Eq. (3),(4)) all combinations of two ellipses are calculated and for the invariants consisting of three straight lines and a conic [5] all combinations of three adjoining line segments plus one ellipse are computed.

   Secondly, the invariant values are calculated for the grouped features.

4. *Model and rule generation:*
   In the model and rule generation stage new objects are learned automatically: a new object model is generated that consists of the object name, the extracted features and the computed invariant values. The fuzzy rules are generated as described in Section 2.

5. *Hypothesis generation:*
   In this module the measured invariant values are used to evaluate the rules of the fuzzy rule base. This is

done as described in Section 2.3. If the resulting credibility μ_{o^k} of an indexed object is above a threshold, a new object hypothesis is generated. This hypothesis consists of the object name, the credibility and the features used to compute the invariant values.

6. *Verification:*
   The last stage of the recognition process is the verification of the generated object hypotheses. This is done as usually: The hypothesized object model is mapped into the image and verified against the extracted features.

Although this system is for recognizing (quasi-)planar objects only, this is no principle limitation for the FII-technique. In principle, the system as described so far might also be implemented as a classical system using multivariate pdf's. However, the aforementioned fluctuations (Sect. 1) are very hard to model by mere distribution functions because they may result from systematic (yet unknown) errors, e.g. in the line extraction algorithm. Moreover, the addition of further attributes (e.g. colour values) would entail the recomputation of the pdf.

4. Experimental results

The FII-technique has been tested on several real images with our object recognition system (see Sect. 3). In the sequel we first describe the performance of the system based on two examples. Then we show first experimental results obtained through the use of additional colour attributes. Finally, we compare the crisp indexing method with the proposed FII-technique.

4.1. Performance of the FII-recognition system

The performance of the FII-recognition system is tested for an (originally coloured) object domain of wooden toy objects, such as rims, tyres, nuts and slats (see Fig. 3).

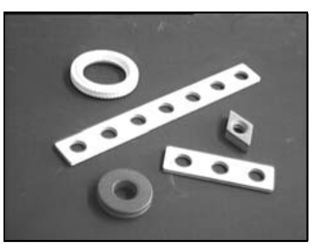
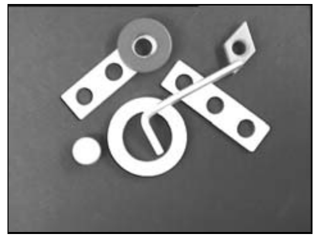


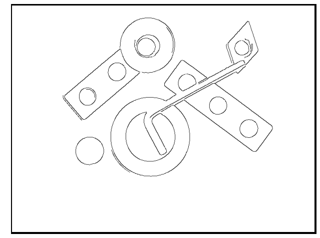
Figure 3. Test objects

All of these objects are quasi-planar, i.e. the depth of the objects is small.

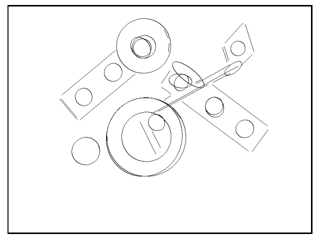
The first scene, Figure 4a, is taken perpendicular to the planar object surfaces. It consists of two three-hole-slats, a nut, a rim, a tyre and two unknown objects, which overlap each other. Since the detected edge points (Figure 4b) as well as the fitted features (Figure 4c) provide a reliable image description, the system recognizes all of the known objects displayed in Figure 4d.



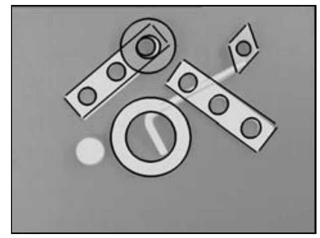
(a) Original image



(b) Extracted edge points



(c) Fitted features



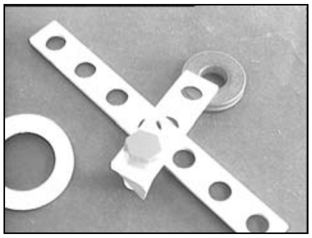
(d) Result

Figure 4. Recognition of scene I.

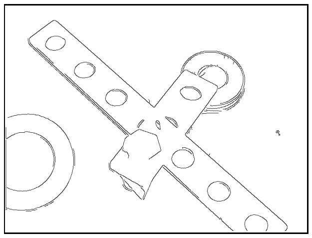
In the second scene, Figure 5a, a three-hole-slat, a seven-hole-slat, a rim and a tyre are used. This scene is taken at an angle of about 25 degrees. Figure 5b shows the detected edge points and Figure 5c the fitted features. In this scene the system detects all of the known objects except for the three-hole-slat (see Fig. 5d). The problem here is a consequence of an inaccurate feature extraction, so the system fails to extract the topology of the three-hole-slat correctly. In this case a single conic is fitted to edge points coming from the top and the bottom of the three-hole-slat. Hence, the calculated invariant values differ too much from the desired values and no object hypothesis is generated.

To demonstrate how the fuzzy rules can easily be extended by adding further attributes, we modify our recognition system by integrating colour attributes to the fuzzy rules learned before. For this we measure the RGB colour information of an object along the underlying geometric structures of the fuzzy rules and transform it into the HSV colour space. Depending on the saturation of the object colour we use the hue or the intensity for generating and evaluating the fuzzy rules, e.g. the rule for the rim in Fig. 3 looks like:

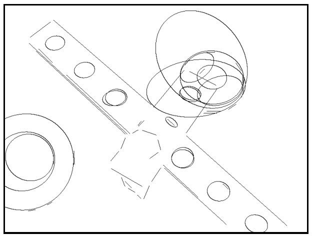
IF (inv1 $\approx$ 3.8) AND (inv2 $\approx$ 4.1) AND (hue is RED)
THEN (object is RIM)



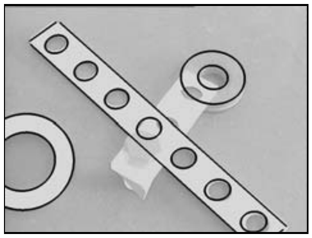
(a) Original image



(b) Extracted edge points



(c) Fitted features



(d) Result

Figure 5. Recognition of scene II.

First experimental results show that this extension reduces the number of generated object hypotheses by 34%, where mainly false positives are suppressed. For example, the extended fuzzy rules decrease the number of hypotheses for Fig. 4a from 124 to 81 and for Fig. 5a from 192 to 137. 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.

4.2. Comparison between crisp indexing and FII

In the following we compare the proposed FII-technique with the usual indexing method. Therefore we implement a "crisp version" (CII) of our recognition system, in which we use intervals instead of fuzzy invariant values. To emphasize the advantages of the FII-technique we apply the system to the difficult case of very similar objects. We use seven different rims with a constant exterior diameter of 50 mm but varying interior diameters of 22 mm to 38 mm (see Fig. 6).

For these objects we get the fuzzy invariant values in Figure 7. Figure 7a shows the fuzzy invariant values computed for the first invariant of a pair of conics (3) and Figure 7b the second invariant (4). From right to left the membership functions represent the objects rim22, rim25, rim28, rim30, rim32, rim35 and rim38, where the numbers denote the interior diameters.

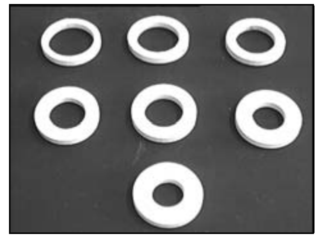
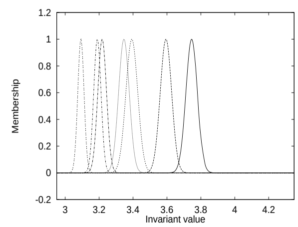
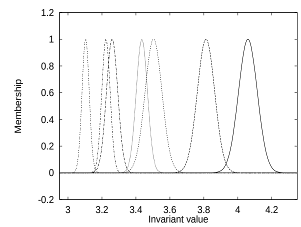


Figure 6. Object domain of seven similar rims



(a) Fuzzy invariant values of first invariant (Eq. (3))



(b) Fuzzy invariant values of second invariant (Eq. (4))

Figure 7. Fuzzy invariant values of Fig. 6

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 property 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 Table 1.

Table 1. Recognition results: comparison between crisp and fuzzy invariant indexing

	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%

It turns out that the recognition system based on the FII-technique provides a better discrimination between the objects than the crisp system; the recognition rate is generally higher. Only for one object, rim22, we get an equivalent rate. As expected, we achieved 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%.

5. Conclusions and future research

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 method, called fuzzy invariant indexing (FII), enhances the usual invariant indexing technique in three ways:

- Since the FII-technique generally produces object hypotheses with different associated *credibilities*, a better discrimination between similar objects is achieved;
- The credibilities of the object hypotheses provide the ability to control the recognition process, e.g. by exploring the most credible hypothesis first;
- The recognition system can be extended in a closed form, i.e. new attributes (also non-invariant attributes like colour, energy, etc.) may be added to the fuzzy classification rules resulting in only minor changes to the original structure of the system.

In the near future we will carry out the following extensions:

- Improve the object discrimination by using multidimensional fuzzy invariant values for modelling the different invariants of a single geometric configuration.
- Use more complex rules and extend the fuzzy rules through further attributes.
- Extend the implemented FII-recognition system for recognizing partially occluded 3D objects.

References

[1] J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679–698, Nov. 1986.

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

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

[4] J. Mundy, C. Huang, J. Liu, W. Hoffman, D. Forsyth, C. 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.

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

[6] J. Mundy and A. Zisserman. Introduction - towards a new framework for vision. In [5], pages 1–39. MIT Press, 1992.

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