

# Fuzzy matching of visual cues in an unmanned airborne vehicle

Thord Andersson

Silvia Coradeschi

Alessandro Saffiotti

Linköping University Electronic Press  
Linköping, Sweden

<http://www.ep.liu.se/ea/cis/1999/008/>

*Published on August 16, 1999 by  
Linköping University Electronic Press  
581 83 Linköping, Sweden*

**Linköping Electronic Articles in  
Computer and Information Science**  
*ISSN 1401-9841  
Series editor: Erik Sandewall*

©1999 Thord Andersson, Silvia Coradeschi, and Alessandro Saffiotti  
*Typeset by the authors using TeX  
Formatted using étendu style*

**Recommended citation:**

*<Authors>. <Title>. Linköping electronic articles  
in computer and information science, Vol. 4(1999): nr 8.  
<http://www.ep.liu.se/ea/cis/1999/008/>. August 16, 1999.*

*This URL will also contain a link to the authors' home pages.*

*The publishers will keep this article on-line on the Internet  
(or its possible replacement network in the future)  
for a period of 25 years from the date of publication,  
barring exceptional circumstances as described separately.*

*The on-line availability of the article implies  
a permanent permission for anyone to read the article on-line,  
to print out single copies of it, and to use it unchanged  
for any non-commercial research and educational purpose,  
including making copies for classroom use.*

*This permission can not be revoked by subsequent  
transfers of copyright. All other uses of the article are  
conditional on the consent of the copyright owners.*

*The publication of the article on the date stated above  
included also the production of a limited number of copies  
on paper, which were archived in Swedish university libraries  
like all other written works published in Sweden.  
The publisher has taken technical and administrative measures  
to assure that the on-line version of the article will be  
permanently accessible using the URL stated above,  
unchanged, and permanently equal to the archived printed copies  
at least until the expiration of the publication period.*

*For additional information about the Linköping University  
Electronic Press and its procedures for publication and for  
assurance of document integrity, please refer to  
its WWW home page: <http://www.ep.liu.se/>  
or by conventional mail to the address stated above.*

## Abstract

Computer vision systems used in autonomous mobile vehicles are typically linked to higher-level deliberation processes. One important aspect of this link is how to connect, or *anchor*, the symbols used at the higher level to the objects in the vision system that these symbols refer to. Anchoring is complicated by the fact that the vision data are inherently affected by uncertainty. We propose an anchoring technique that uses fuzzy sets to represent the uncertainty in the perceptual data. We show examples where this technique allows a deliberative system to reason about the objects (cars) detected by a vision system embarked in an unmanned helicopter, in the framework of the WITAS project.

### Authors' affiliations:

**Thord Andersson:** Department of Electrical Engineering, Linköping University, Sweden.

**Silvia Coradeschi:** Department of Computer and Information Science, Linköping University, Sweden.

**Alessandro Saffiotti:** Applied Autonomous Sensor Systems, Örebro University, Sweden.

# 1 Introduction

Autonomous mobile vehicles need to use computer vision capabilities in order to perceive the physical world and to act intelligently in it. These systems also need the ability to perform high-level, abstract reasoning in order to operate reliably in a dynamic and uncertain world without the need for human assistance. For example, a mail delivery robot faced with a closed door should decide whether it is better to plan an alternative way to achieve its goal, or to reschedule its activities and try again this delivery later on. In general, autonomous vehicles need to incorporate a decision-making system that uses the perceptual data to guide the activity of the vehicle toward the achievement of the intended task, and also guides the activity of the perceptual subsystem according to the priorities of this task.

An important aspect in integrating the decision-making and the computer vision systems is the connection between the abstract representations used by the symbolic decision-making system to denote a specific physical object, and the data in the computer vision system that correspond to that object. Following [5], we call *anchoring* the process of establishing this connection. We assume that the decision-making process associates each object to a set of properties that (non-univocally) describe that object. Anchoring this object then means to use the vision apparatus to find an object whose observed features match the properties in this description. For example, suppose that the symbolic system has an object named ‘car-3’ with the description ‘small red Mercedes on Road-61.’ Anchoring this object means to: (i) find an object in the image that matches this description; and (ii) update the description of ‘car-3’ by using the observed features, so that the same object can later be re-identified.

One of the difficulties in the anchoring problem is that the data provided by the vision system are inherently affected by a large amount of uncertainty. This may result in errors and ambiguities when trying to match these data to the high-level description of an intended object. In order to improve the reliability of the anchoring process, this uncertainty has to be taken into account in the proper way. In this work, we propose to use techniques based on fuzzy logic to define a *degree of matching* between a perceptual signature and an object description. The possibility to distinguish between objects that match a given description at different degrees is pivotal to the ability to discriminate perceptually similar objects under poor observation conditions. Moreover, these degrees allow us to consider several possible anchors, ranked by their degree of matching. Finally, these degrees can be used to reason about the quality of an anchor in the decision making process; for example, we can decide to engage in some active perception in order to get a better view of a candidate anchor.

In the rest of this paper, we deal with the anchoring problem in the context of an architecture for unmanned airborne vehicles. This

architecture, outlined in the next section, integrates several subsystems, including a vision system and an autonomous decision making system. In section 3, we discuss how we represent the inexact data provided by the vision system with fuzzy sets. In section 4, we show how we compute the degrees of matching between these data and the intended descriptions. Section 5 illustrates the use of these degrees by going through a couple of examples, run in simulation. Finally, section 6 discusses the results and traces future directions.

## 2 The WITAS project

The WITAS project, initiated in January 1997, is devoted to research on information technology for autonomous systems, and more precisely to unmanned airborne vehicles (UAVs) used for traffic surveillance.

The general architecture of the system is a standard three-layered agent architecture consisting of a deliberative, a reactive, and a process layer. The deliberative layer generates at run-time probabilistic high-level predictions of the behaviors of agents in their environment, and uses these predictions to generate conditional plans. The reactive layer performs situation-driven task execution, including tasks relating to the plans generated by the deliberative layer. The reactive layer has access to a library of task and behavior descriptions, which can be executed by the reactive executor. The process layer contains image processing and flight control, and can be reconfigured from the reactive layer by means of switching on and off groups of processes. Besides vision, the sensors and knowledge sources of the system include: a global positioning system (GPS) that gives the position of the vehicle, a geographical information system (GIS) covering the relevant area of operation, and standard sensors for speed, heading and altitude.

The system is fully implemented in its current version. Because of the nature of the work, most of the testing is being made using simulated UAVs in simulated environments, even though real image data has been used to test the vision module. In a second phase of the project, however, the testing will be made using real UAVs. More information about the project can be found at [7].

Of particular interest for this presentation is the interaction between the reactive layer and the image processing in the process layer. This is done by means of a specialized component for task specific sensor control and interpretation, called the The Scene Information Manager (SIM). This system, illustrated in figure 1, is part of the reactive layer and it manages sensor resources: it reconfigures the vision module on the basis of the requests of information coming from the reactive executor, it anchors symbolic identifiers to image elements (points, regions), and it handles simple vision failures, in particular temporary occlusion and errors in car re-identification.

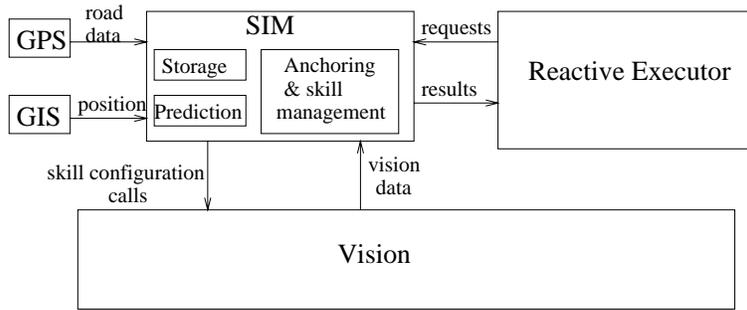


Figure 1: Overview of the Scene Information Manager and its interaction with the Vision module and the Reactive Executor.

Two of the main aspects of anchoring implemented in the SIM are identification of objects on the basis of a visual signature expressed in terms of concepts, and re-identification of objects that have been previously seen, but have then been out of the image or occluded for a short period.

For identification and re-identification the SIM uses the visual signature of the object, typically color and geometrical description, and the expected positions of the object. For instance if the SIM has the task to look for a red, small Mercedes near a specified crossing, it provides the vision module with the coordinates of the crossing, the Hue, Saturation and Value defining “red” and the length, width and area of a small Mercedes. The measurements done in the vision module have a degree of inaccuracy, and the SIM also provides the vision module with the intervals inside which the measurement of each of the features is acceptable. The size of the interval depends on how discriminating one wants to be in the selection of the objects and also, in the case of re-identification of an object, on how accurate previous measurements on the object were.

The vision module receives the position where to look for an object and the visual signature of the object and it is then responsible for performing the processing required to find the objects in the image whose measures are in the acceptability range and report the information about the objects to the SIM. The vision module moves the camera toward the requested position and for each object in the image and for each requested feature of the object, it calculates an interval containing the real value. If the generated interval intersects with the interval of acceptability provided in the visual signature for the feature, the feature is considered to be in the acceptability range. The vision module reports information about color, shape, position, and velocity of each object whose features are all in the acceptability range to the SIM.

Intersection of intervals is a simple, but not very discriminating method to identify an object. As a consequence, several objects that are somehow similar to the intended one can be sent back by the

vision module to the SIM. The SIM then needs to apply some criteria in order to perform a further selection of the best matching object between those reported by the vision module. The selection of the best matching object should depend on how well the objects match the different aspects of the signature, but also on the accuracy of the measurements performed by the vision and their reliability. In what follows, we show how we perform this selection using fuzzy signature matching.

### 3 Fuzzy-set representation of visual cues

Cues obtained from the vision system, e.g., color, shape, position and velocity, are affected by uncertainty and imprecision in several ways. In this work, we propose to explicitly represent the inexactness which is inherent to these data, and to take this inexactness into account when performing signature matching. In order to justify our representation, we need to analyze the way in which we extract the needed parameters from the image.

Consider the measurement of the shape parameters (length, width and area) of an observed car. Roughly, the measurement starts with a segmented and labeled binary image containing our candidate cars. This binary image is created by combining and thresholding the feature images produced by the different feature channels available, e.g., orientation, color, IR and velocity (currently, we only use the color channels). For each object in the labeled image, we then compute the moment of inertia matrix. From this  $2 \times 2$  matrix, we calculate the two eigenvalues which correspond to the largest and smallest moment of inertia, respectively, and convert them into the length and width of the object under the assumption that our objects (cars) are rectangular. We also measure the area by counting the pixels that belong to the same object. The length, width and area measures are then converted to metric measures through multiplication by a scale factor describing the meter per pixel ratio. This ratio is computed from the field-of-view angle and from the position and angles of the camera.

There are a number of factors that influence the correctness of the values measured by the above procedure. First, in the segmentation phase, the discretization of the image limits the precision of the measure. Second, continuing the segmentation phase, we apply some binary operations (e.g., “fill” and “close”) on the binary image in order to try to connect and “bind” segmented pixels into objects. This operation slightly alters the shape, thus limiting the precision. The above two factors together produce a *segmentation error*, denoted by  $\epsilon_s$ . Third, the measurement model may be inaccurate, thus introducing an error, the *model error*, denoted by  $\epsilon_m$ ; for example the above assumption that cars are rectangular is almost never completely true. Note that the impact of the  $\epsilon_s$  and  $\epsilon_m$  errors on the quality of the

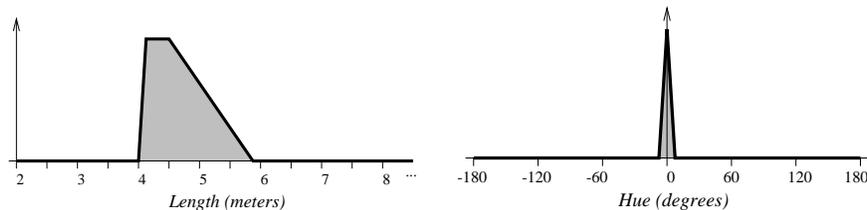


Figure 2: Fuzzy sets for the measured length (left) and hue (right).

measurements depends on the size of the car in the image, which in turn depends on its distance from the camera and on the focal length of the camera. A fourth factor that affects the measurement is the perspective distortion due to the angle  $\alpha$  between the normal of the car plane and the optical axis: if the car plane is not perpendicular to the optical axis, the projection of the 3D-car on the image plane will be shorter. We denote this *perspective error* by  $\epsilon_\alpha$ . Finally, all the geometric parameters needed to compute the length may themselves be affected by errors and imprecision. For example, the distance from the camera depends on the relative position of the helicopter and the car; and the  $\alpha$  angle depends on the slope of the road; both these values may be difficult to evaluate. We summarize the impact of these factors on our measurement in a *geometric error* term, denoted by  $\epsilon_g$ .<sup>1</sup>

The above discussion reveals that there is a great amount of uncertainty that affects the measured value, for example, the length of an object; and that this uncertainty is very difficult to precisely quantify — in other words, we do not have a *model* of the uncertainty that affects our measures. Similar observations can be made for other features measured by the vision system: for example, the measurement of the color of an object is influenced by the spectral characteristics of the light that illuminates that object. Given this difficult nature of the uncertainty in the data coming from the vision system, we have chosen to represent these data using fuzzy sets [8]. Fuzzy sets offer a convenient way to represent inexact data whose uncertainty cannot be characterized by a precise, stochastic model — but for which we have some *heuristic* knowledge. For example, Fig 2 (left) shows the fuzzy set that represents a given length measurement. For each value  $x$ , the value of this fuzzy set at  $x$  is a number in the  $[0, 1]$  interval that can be read as “the degree by which  $x$  can be the actual length of the object given our measurement.” (See [9] for this possibilistic reading of a fuzzy set.)

In our work, we use trapezoidal fuzzy sets, both for computational

---

<sup>1</sup>There are more sources of errors in this process. For example, when  $\alpha$  increases, the car projection may seem longer due to the fact that the sides of the car will appear. Also, the measured value can be totally invalid if there has been an error in the segmentation and/or labeling phases; for instance, if the car has been merged with its shadow, or with another car in front of it. Accounting for these possibilities is part of our current work.

reasons and for ease of construction. The possibilistic semantics give us some simple guidelines on how to build a trapezoidal fuzzy set to represent an inexact measurement. The flat top part of the fuzzy set (its *core*) identifies those values  $x$  that can be fully regarded as the actual length value given our measurement. In the example in Fig 2 (left), these values are spread over an interval rather than concentrated in a point because of the segmentation effect: our measurement cannot tell us more than what is allowed by the pixel size. The base of the fuzzy set (its *support*) identifies those values  $x$  that can possibly be regarded as the actual length value: given the errors that may affect our measurement, the actual length may be anywhere in the support interval — but under no circumstances it can be outside this interval. Put differently, the support constitutes a sort of worst case estimate: however big the error is, the actual value must lie somewhere in this interval. While the core constitutes a best case estimate: even when there is no error in our measurement, we cannot be more precise than this.

Let us now discuss in detail how we have built the fuzzy set in Fig 2. The vision system has calculated the length to 29.9 pixels, which correspond to  $l = 4.23$  meters. The segmentation error  $\epsilon_s$  is estimated to a constant  $\pm 1$  pixel, which with a scale factor of  $s = 0.14$  meter/pixel gives us  $\epsilon_s = 0.14$  meter. This segmentation error is inherent to our measurement process, no matter how good our models and computations are, and it thus defines the core of the trapezoid in the picture, given by the interval  $[l - \epsilon_s, l + \epsilon_s] = [4.09, 4.37]$ .

Our estimates for the other errors are all collected in the support of the trapezoid. The model error  $\epsilon_m$  is estimated in a coarse but simple way by comparing the measured area  $a_m$  with the computed area  $a_c = wl$  (where  $w$  is the calculated width). The difference between these areas defines  $\epsilon_m$  such that  $a_m$  will lie in the interval  $[(w - \epsilon_m)(l - \epsilon_m), (w + \epsilon_m)(l + \epsilon_m)]$ . If, for example,  $a_m$  is greater than  $a_c$ ,  $\epsilon_m$  becomes: (As a simplification we have assumed that  $\epsilon_m$  is the same for both the width and for the length.)

$$a_m - (w + \epsilon_m)(l + \epsilon_m) = 0 \implies \epsilon_m = -\frac{(w+l)}{2} + \sqrt{\frac{(w+l)^2}{4} + (a_m - a_c)} \quad (1)$$

which in our case gives us  $\epsilon_m = 0.04$  m. As for the perspective error  $\epsilon_\alpha$ , in our case we have  $\alpha = 40.3^\circ$ . If we assume that we measure the projected length as  $l \cos \alpha$ , then the worst case error due to  $\alpha$  becomes  $\epsilon_\alpha = l_{\text{MAX}}(1 - \cos \alpha)$ , where  $l_{\text{MAX}}$  is the estimation of the maximum object length. If we set  $l_{\text{MAX}} = 6$  m we get  $\epsilon_\alpha = 1.42$  m. Since the support of our fuzzy set must include all the values which are possible in a worst case error situation, we include all the above errors in it.<sup>2</sup> This gives us the interval  $[l - \epsilon_s - \epsilon_m, l + \epsilon_s + \epsilon_m + \epsilon_\alpha] = [4.05, 5.83]$  for the base of our trapezoid. Note that  $\epsilon_\alpha$  only affect the upper

---

<sup>2</sup>In our current experiments in the simulated environment, we have  $\epsilon_g = 0$  since the helicopter pose, the camera parameters, and the road geometry are all perfectly known.

bound of the interval, i.e., the car may seem smaller in the image when  $\alpha$  increases. The correct length in our example was 4.42 m.

The construction of the fuzzy sets for the other features follow similar guidelines. For example, Fig 2 (right) shows the fuzzy set that represents the observed Hue value. (At the current stage of development, however, we have mainly focused on the shape parameters.) Although the definitions of these fuzzy sets are mostly heuristic, they have resulted in good performance in our experiments.

## 4 Fuzzy signature matching

We now focus on the problem of anchoring a high-level description coming from the symbolic system (reactive executor) to the data coming from the vision module. As an example, consider the case in which a task needs to refer to ‘a small red Mercedes.’ The SIM system has to link two types of data: on one side, the description containing the symbols ‘red,’ ‘small’ and ‘Mercedes’ received from the symbolic system; and on the other side, the measurable parameters of the observed objects which are sent by the vision system. Anchoring implies to convert these representations to a common frame, and to find the car that best matches the description. In our case, we have chosen to convert symbols to the universe of the measurable parameters.

In general, symbolic descriptions contain linguistic terms like ‘red’ and ‘small’ that do not denote a unique numerical value. Sticking to a common practice [8, 3], we have chosen to map each linguistic term of this kind to a fuzzy set over the relevant frame. For example, we associate the term ‘red’ with the fuzzy set shown in Fig 3 (left): for each possible value  $h$ , the value of  $red(h)$  measures, on a  $[0, 1]$  scale, how much  $h$  can be regarded as ‘red.’<sup>3</sup> As a second example, Fig 3 (right) shows how we represent the length associated to the linguistic term ‘small-Mercedes’ by a fuzzy set over the space of possible lengths. In our system, we use a database that associates each car type to its typical length, size, and area, represented by fuzzy sets. Cars of unknown types are associated with generic fuzzy sets, like the ‘small’ (car) set in the picture. Once again, we only use trapezoidal fuzzy sets for computational reasons.

Once we have represented both the desired description and the observed data by fuzzy sets, we can compute their *degree of matching* using fuzzy set operations. This choice is justified in our case since fuzzy sets can be given a semantic characterization in terms of degrees of similarity [4]. Consider two fuzzy sets  $A$  and  $B$  over a common domain  $X$  which respectively represent the observed data and the target description. The degree of matching of  $A$  to  $B$ , denoted by  $match(A, B)$ , is the degree by which the observed value  $A$  can be one

---

<sup>3</sup>In practice, we associate this term to three sets for hue, saturation, and value, respectively.

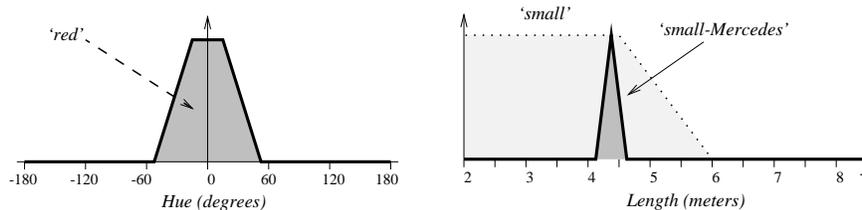


Figure 3: Fuzzy sets associated to the symbols ‘small-Mercedes’ and ‘red.’

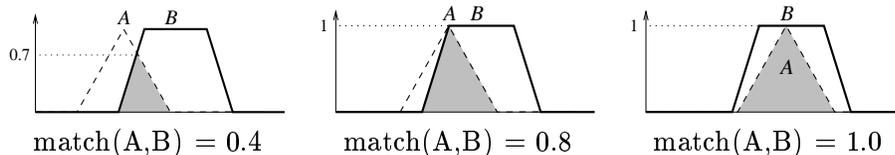


Figure 4: Three examples of partial matching.

of those that satisfy our criterion  $B$ . In the experiments presented in this note, we use the following measure:

$$\text{match}(A, B) = \frac{\int_{x \in X} \min\{A(x), B(x)\} dx}{\int_{x \in X} B(x) dx} \quad (2)$$

Intuitively, (2) measures the degree by which  $A$  is a (fuzzy) subset of  $B$  by looking at how much of  $A$  is contained in  $B$ . (See, e.g., [2] for different measures.) The behavior of this measure is graphically illustrated in Fig. 4. To ensure an efficient computation, we approximate (2) by the ratio between the area of the inner trapezoidal envelope of  $A \cap B$  and the area of  $B$ . These areas can be computed very easily when  $A$  and  $B$  are trapezoidal fuzzy sets.

Once we have computed a degree of matching for each individual feature, we need to combine all these degrees together in order to obtain an overall degree of matching between the intended description and a given percept. In our case, we need to combine the degrees of matching of the length, width, area, hue, saturation, and value criteria into one summarized degree of matching. The simplest way to combine our degrees is by using a *conjunctive* type of combination, where we require that each one of the features matches the corresponding part in the description. Conjunctive combination is typically done in fuzzy set theory by T-norm operators [6, 3], whose most used instances are min, product, and the Łukasiewicz T-norm  $\max(x+y-1, 0)$ . In our experiments, we have noticed that the latter operator provides the best results. (See [1] for an overview of the use of alternative operators with applications to image processing.)

The overall degree of matching is used by the SIM to select the best anchor among the candidate objects provided by the vision module. For each candidate, the SIM first computes its degree of matching to the intended description; then it ranks these candidates by their

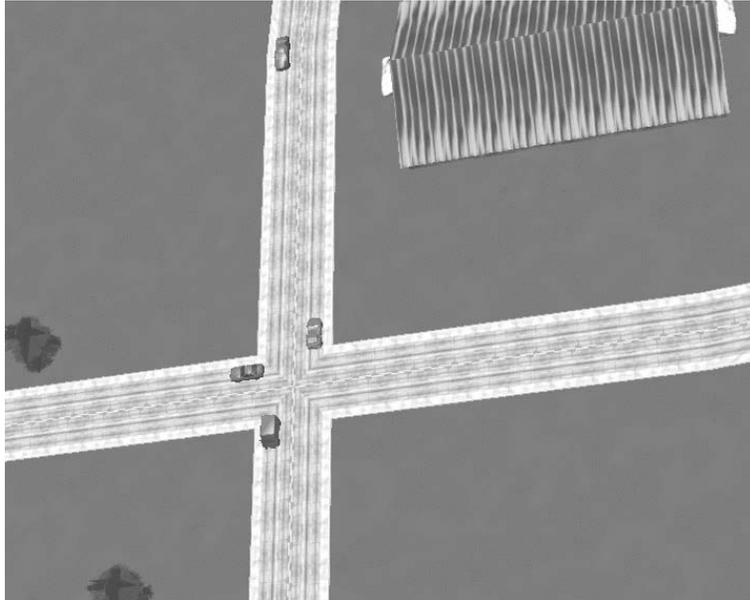


Figure 5: The simulated scenario for our examples.

degree; and finally returns the full ordered list to the reactive executor. Having a list of candidates is convenient if the currently best one later turns out not to be the one we wanted. Also, it is useful to know how much the best matching candidate is better than the other ones: if the two top candidates have similar degrees of matching, we may decide to engage in further exploratory actions in order to disambiguate the situation before committing to one of them — for instance, we may give the vision system the task to zoom on each candidate in turn in the hope to get more precise data.

## 5 Fuzzy signature matching at work

We illustrate the use of the fuzzy signature matching by two examples on a scenario taken from the WITAS project. In this scenario, the deliberative system is interested in a red car of a specified model in the vicinity of a given crossing. Four cars are situated around that crossing, moving in different directions. The cars are all red, but of different models: a small van, a big Mercedes, a small Mercedes, and a Lotus. In the first example the helicopter is above the cars. In the second example discriminating between the cars is made more difficult by the fact that the helicopter views the crossing at an inclination of about 30 degrees (see Fig. 5): this results in some perspective distortions, thus introducing more uncertainty in the extraction of geometrical features.

In our first example, the deliberative system decides to follow ‘Van-B’, which is described as a red van. The SIM sends the proto-

typical signature of a red van to the vision module. Since all the four cars in the image are red, and they have fairly similar shapes, the vision module returns the observed signatures of all the four cars to the SIM. These signatures are then matched against the desired signature by our routines, resulting in the following degrees of matching:

ID	Color	Shape	Overall
66	1.0	0.58	0.58
67	1.0	0.38	0.38
68	1.0	1.0	1.0
69	1.0	0.0	0.0

The ID is a label assigned by the vision system to each car found in the image. The degree of matching for the color is obtained by combining the individual degrees of Hue, Saturation, and Value; in our case, this will be 1.0 for all the cars as they are all red. The degree of matching for the shape is the combination of the individual degrees of matching of length, width, and area. The overall degree is the Łukasiewicz combination of the color and shape degrees. In this case, car 68 is correctly<sup>4</sup> identified as the best candidate, and an anchor to that car is thus returned to the deliberation system.

In the second example, the deliberative system is interested in ‘Car-D’, a red small Mercedes. The SIM sends the corresponding prototypical signature to the vision module, and again gets the signatures of all the four cars in the image as an answer. In this case however, the helicopter is at a long distance from the crossing and it views the crossing at an inclination of about 30 degrees. By applying our fuzzy signature matching routine, we obtain the following degrees:

ID	Color	Shape	Overall
66	1.0	0.65	0.65
67	1.0	0.84	0.84
68	1.0	0.0	0.0
69	1.0	0.97	0.97

Cars 66, 67 and 69 match the desired description to some degree, while car 68 can safely be excluded. The SIM decides that these degrees are too close to allow a safe discrimination, and it tries to improve the quality of the data by asking the vision module to zoom on each one of cars 66, 67, and 69 in turn. Using the observed signatures after zooming, the SIM then obtains the new degrees of matching:

ID	Color	Shape	Overall
66	1.0	0.30	0.30
67	1.0	0.70	0.70
69	1.0	0.21	0.21

---

<sup>4</sup>This verification was done manually off-line by analyzing some additional information, like the road on which a car is.

The closer view results in a smaller segmentation error, since the scale factor is smaller, and hence in more narrow fuzzy sets. As a consequence, all the degrees of matching have decreased with respect to the previous observation. What matters here, however, is the relative magnitude of the degrees obtained from comparable observations, that is, those which are collected in the above table. These degrees allow the SIM to select car 67 as the best candidate.

The SIM now also has the option to try to further improve its choice by commanding the helicopter to fly over car 67 and take another measurement from above the car — the best observation conditions for the vision system. If we do this, we finally obtain a degree of matching of 1.00 for car 67. Note that this degree could as well have dropped, thus indicating that car 67 was not really the car that we wanted. In this case, the SIM might have used the partial match information and go back to cars 66 and 69 to get more accurate views.

## 6 Conclusions

Anchoring symbols to the physical objects they are meant to denote requires the ability to integrate symbolic and numeric data under uncertainty. In this paper, we have considered an instance of the anchoring problem in which we link the car identifiers used at the decision-making level to the perceptual data provided by a vision system. Our experimental results indicate that our technique is adequate to handle the ambiguities that arise when integrating uncertain perceptual data and symbolic representations. In particular, fuzzy signature matching improves our ability to discriminate among perceptually similar objects in difficult situations (e.g., perspective distortion). Moreover, degrees of matching allow us to exclude unlikely candidates, and to rank the likely ones by their similarity to the intended description. Finally, these degrees can help in decision making; for example, if these degrees indicate a large amount of anchoring ambiguity, the system may decide to engage in active information gathering such as zooming or getting closer to the object in order to obtain better information.

The work reported in this paper is still in progress, and many aspects need to be further developed. First, the treatment of perspective distortions presented here is rather primitive; in our next experiments, we shall use different models of each car viewed from different observation angles. Second, we need to account for still more sources of errors, including the possibility that the detected object is not a car. Third, we need to study more sophisticated forms of aggregation of the individual degrees of matching of different features into an overall degree. For example, in some situations some of the features are more critical than others, and we would like their degree of matching to have a stronger impact on the overall degree. Fourth, we

plan to include features of a different nature in the matching process, like the observed position and velocity of the cars. Finally, until now we have only performed experiments in simulation. At the current stage of development of the WITAS project, the vision system takes the video frames produced by a 3D simulator as input. Although this configuration results in some amount of noise and uncertainty in the extracted features, we are aware that a real validation of our technique will only be possible when we have access to the real data from an embarked camera.

## Acknowledgments

Discussions with Dimiter Driankov, and Klas Nordberg have been greatly beneficial. This work is part of the WITAS project, which is fully sponsored by the Knut and Alice Wallenberg Foundation.

## References

- [1] I. Bloch. Information combination operators for data fusion: A comparative review with classification. *IEEE Trans. on Systems, Man, and Cybernetics*, A-26(1):52–67, 1996.
- [2] S. Gottwald. *Fuzzy sets and fuzzy logic*. Vieweg, Braunschweig, Germany, 1993.
- [3] G. Klir and T. Folger. *Fuzzy sets, uncertainty, and information*. Prentice-Hall, 1988.
- [4] E. H. Ruspini. On the semantics of fuzzy logic. *Int. J. of Approximate Reasoning*, 5:45–88, 1991.
- [5] A. Saffiotti. Pick-up what? In C. Bäckström and E. Sandewall, editors, *Current Trends in AI Planning — Proc. of EWSP '93*, pages 166–177. IOS Press, Amsterdam, NL, 1994.
- [6] S. Weber. A general concept of fuzzy connectives, negations and implications based on t-norms and t-conorms. *Fuzzy sets and systems*, 11:115–134, 1983.
- [7] WITAS web page: <http://www.ida.liu.se/ext/witas/>.
- [8] L. A. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.
- [9] L. A. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1:3–28, 1978.