1 Decision Support for Object Recognition from Multi-Sensor Data

Alexander Bauer, Jürgen Geisler

Fraunhofer Institute for Information and Data Processing (IITB)
Fraunhoferstr. 1, 76131 Karlsruhe, Germany

Recognition of objects by remote sensing is a crucial task for reconnaissance and surveillance in the defence and security domain. Beyond an effective processing of the sensor data in order to support the perceptibility of objects of interest for human interpreters, those interpreters also need support to cope with the huge amount of relevant objects and their appearances with respect to different sensors.

An object database in conjunction with an efficient way to search objects given their recognition features can obviously improve recognition performance. To recognize an object in optimal time and with maximum accuracy, it becomes moreover crucial for the interpreter to select features which are easily recognized and provide a powerful discrimination at the same time.

Under the title RecceMan® (Reconnaissance Manual) the Fraunhofer IITB has developed a software concept to guide mainly image interpreters in the military reconnaissance domain through the complex world of recognition features using optimal feature selection. RecceMan® is right on the way to become the basic decision support tool for the image interpreters in the airborne IMINT (imagery intelligence) units of the German Bundeswehr. The concept can be translated easily to non-imaging sensors and other application domains such as civil security, especially forensic analysis.

1.1 Introduction

Object recognition describes the process of matching an unknown object with known categories of objects. The categorization allows us to project properties from the recognized category to the unknown object. In primitive times, the ability to recognize an animal as belonging to the category of tigers and the knowledge about tigers being dangerous, has given us the opportunity to escape the predator. To provide our security today, object recognition remains a vital link in the chain from sensor data acquisition to decision making.

Advances in the field of computer vision have led to an automation of object recognition especially in industrial applications. Those algorithms perform well under defined environment conditions, which are easily established in a production line. However, in natural and urban environments, computer vision algorithms are still inferior to the trained human eye.

RecceMan® (Reconnaissance Manual) is a decision support system for object recognition. Decision support systems approach problems by combining the strength of computers and humans. Computers provide lossless storage of a vast amount of data and are able to accurately perform numerical computations in short time, while humans are better at detecting small patterns in a signal and are superior at making judgements [1]. Born and raised in the field of military reconnaissance applications, the concept RecceMan® is applicable to any object recognition task demanding high accuracy and speed from users with limited training.

1.2 Object Recognition Requirements

How is object recognition accomplished? All approaches require at least the following four components [2]:

1. *Object representation*. The relevant characteristics of the object to be categorized must be determined and represented.
2. *Category representations*. Each category must be represented by its characteristics.
3. *Comparison process*. There must be some way in which the object representation is matched or compared against a possible category representation.
4. *Decision process*. There must be a method for deciding, on the basis of results of comparison processes, to which category a given object belongs.

Object and category representations can be realized at different levels of abstraction. At the lowest level, data acquired by a sensor itself (e.g. an image) is directly compared to reference data of the category. Besides the fact that data level comparison is tedious for a human interpreter, the sensor data acquired for a category can vary dramatically depending on the sensor parameters. Higher level representations of objects and categories, suitable for a decision-support system, describe an object in terms which can be easily understood and named by a human interpreter, the so called recognition features.

Typically, recognition features express sensor independent characteristics of an object and its part structure. The object as a whole is roughly described by its shape and dimensions. A more detailed description is achieved by further describing its functional parts (e.g. engine, wing and tail for an airplane) which are characterized by their type (e.g. jet engine), quantity, position and so on. To express the structural dependency the features are arranged in a feature tree.

The structural representation can be further extended by contextual features (e.g. in which country do we expect the object category) and sensor dependent recognition features (e.g. colour, texture).

Using the recognition feature representation, the comparison process is achieved easily by matching the presence of recognition features in both the category and object representation. The process is automatically executed by the decision-support system.

1.3 Optimal feature selection

The number of recognition features necessary to differentiate a large set of categories tends to be very high (e.g. about 500 for airplanes), rendering a complete assessment of all features of the object inefficient. However, assuming the category of the object is known, a few characteristic features are sufficient to confirm or reject the assumed category (e.g. about 10 features for an F-16). As we do not know the category of the object until unique recognition, an iterative approach is necessary to choose an optimal subset of relevant features.

Fig. 1 visualizes the iterative object recognition process. Each step, a recognition feature is assessed. Every feature assessment leads to a reduction of possible object categories, the so called candidates. If only a few candidates remain, a direct comparison of category and object leads to the final decision about the category the object belongs to. If too many candidates remain, relevant features which guarantee a maximum expected reduction of candidates [3] are determined for the next feature assessment.

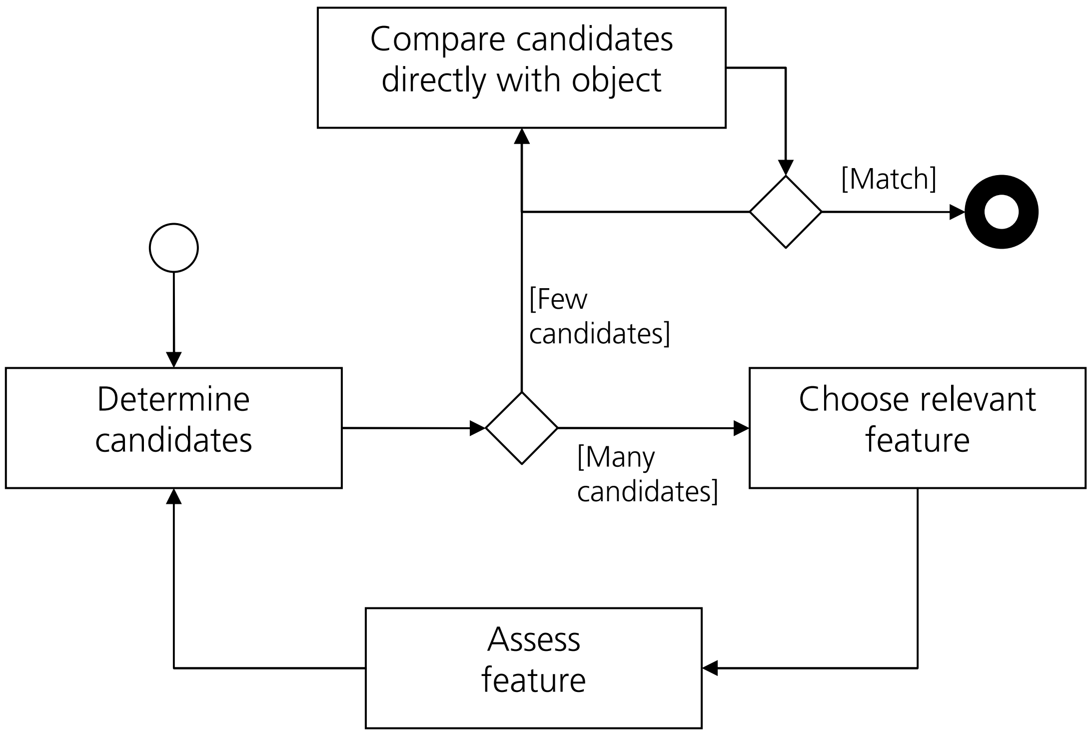

Fig. 1: Iterative object recognition

Exemplarily, the relevancy of a binary feature is defined: Let S be a set of candidates and F a feature which divides S into two sets S_F and $S_{\overline{F}}$, representing the set of candidates which match the feature or do not match the feature respectively. The relevance of the feature F is expressed as the information gain:

$$Gain(S,F) \equiv Entropy(S) - \frac{1}{2}\left(Entropy(S_F) + Entropy(S_{\overline{F}})\right)$$

The entropy of a candidate set S from the set of categories C is defined as

$$Entropy(S) \equiv \sum_{i=1}^{|C|} - p_i \log_2 p_i \ ,$$

given p_i is the probability of a category, which is assumed to be uniformly distributed over the candidate set S. It can be shown that the information gain is maximized by a feature which divides S into two equally sized candidate sets.

Using the information gain theorem, RecceMan® is able to support the process of selecting relevant features by presenting the features in order of their information gain, or by highlighting the most relevant features on the user interface.

1.4 User interface

The user interface of RecceMan® provides easy interaction between the system and the human interpreter. Fig. 2 shows the main screen during the recognition of a ground vehicle. The features are either displayed textually by a feature tree, or in this case, depicted by an icon inside the feature navigation component on the left side. Relevant features are highlighted by an opaque presentation, while non-relevant features are grayed-out. This way, the human interpreter is guided to the most relevant features. Feature assessments are entered by either confirming or rejecting the feature. The right side component displays the candidate list, enhanced by the display of a reference image for each category. The degree of similarity of the displayed category with the assessed features is visualized by different colors, whereas green color represents a perfect match, yellow color stands for a reasonable match and red color expresses a poor match.

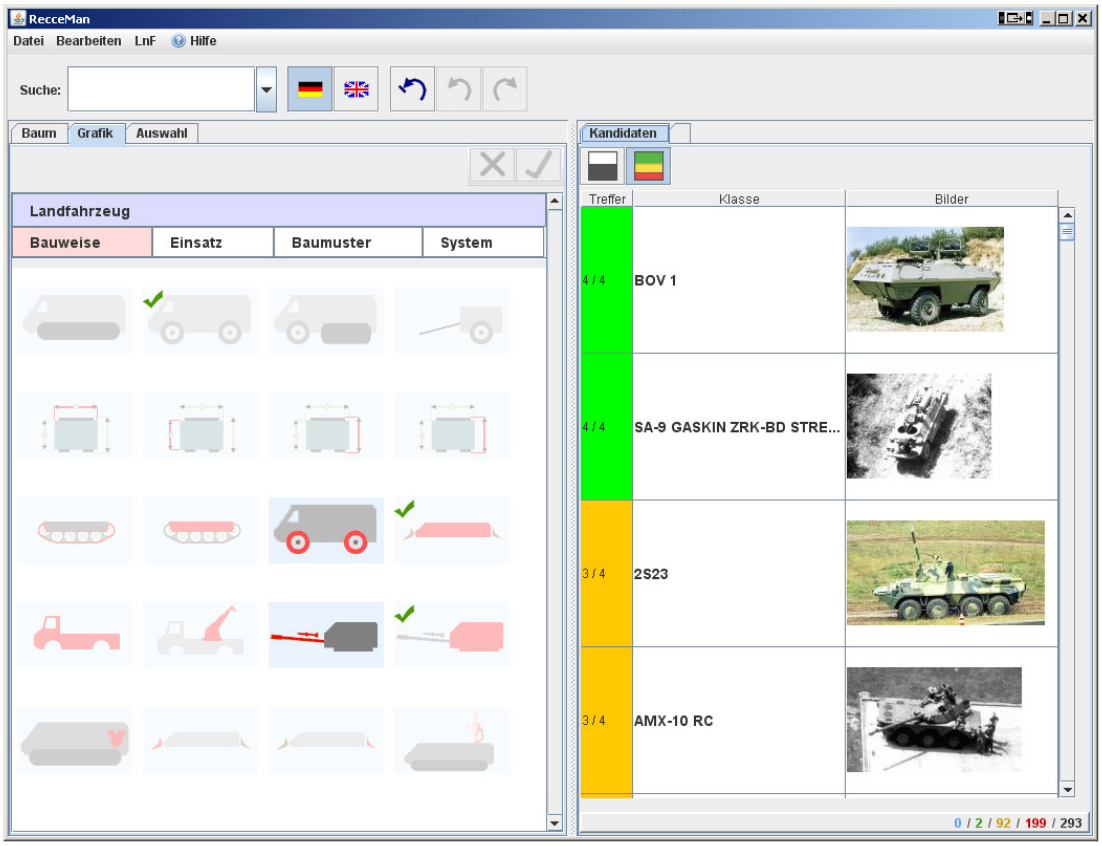

Fig. 2: RecceMan® User Interface

1.5 Conclusion

Time-critical and accurate object recognition is achieved by a recognition process employing iterative feature assessment supported by optimal feature selection. The decision support system RecceMan® interactively guides a human interpreter through the recognition process by automating the comparison process and explicitly highlighting relevant recognition features. The concept is universally applicable to object recognition problems, which frequently occur in the security domain, especially in forensic analysis.

1.6 References

1. Fitts, P.M. (Ed.), Human Engineering for an effective Air-Navigation and Traffic Control System, NRC, Washington, DC (1951).
2. Palmer, S. E., Vision Science: Photons to Phenomology, MIT Press (1999)
3. Mitchell, T., Machine Learning, McGraw-Hill Education (1997)