Exploiting Context for Assisted Aerial Image Interpretation

Alexander Bauer*, Oliver Herschelmann
Fraunhofer-Institute of Optronics, Systems Technologies and Image Exploitation (IOSB),
76131 Karlsruhe, GERMANY

ABSTRACT

The evaluation of a country's critical infrastructure requires a detailed analysis of facilities such as airfields, harbors and heavy industry. To improve the assessment of such facilities, an assistance system for the interpretation of infrastructure facilities from aerial imagery is developed. In this paper we point out recent advances of the system's recommendation function. Besides suggesting the occurrence of undetected objects based on a probabilistic scene model and previously detected objects, the system is now able to suggest the classification of objects based on intrinsic object features and both local context (spatial relations) and global context (overall scene classification). To justify our approach the results of an experimental evaluation of the system for the classification of industrial installations is presented.

Keywords: Image Interpretation, Image Understanding, High-level Vision, Bayesian Inference

1. INTRODUCTION

Infrastructure facilities (such as airfields, harbors and heavy industry) span larger areas of land and the man-made objects they are composed of stand of well in an aerial image. Therefore, the analysis of infrastructure is a prominent application of aerial reconnaissance. Figure 1 shows an aerial image of a harbor. A typical task for an image interpreter is to classify the main function of the infrastructure (military harbor) as well as to describe all relevant objects (such as buildings) in terms of their function (workshop, storage, administration building, etc.).

Figure 1. Aerial image of a harbor. To identify the function of the harbor the interpreter has to analyze the occurring singular objects such as buildings. Features, occurrence frequency and spatial arrangement of objects give valuable hints on their potential function in the infrastructure.

* alexander.bauer@iosb.fraunhofer.de; phone +49 721 6091395; www.iosb.fraunhofer.de

Exploiting context for assisted aerial image interpretation
Alexander Bauer and Oliver Herschelmann, Proc. SPIE 7835, 78350K (2010), DOI:10.1117/12.865119

In the resulting description of the infrastructure, it is not sufficient to merely describe a building in terms of its location size, roof shape, color and height – additionally, the actual function of the building as part of the infrastructure has to be determined. In order to do that, the image interpreter uses his expert knowledge about typical arrangements of objects both on the local level (spatial proximity of objects) as well on the global scene level (occurrence of objects in a specific type of infrastructure).

To compensate for reduced training periods of image interpreters and the increasing complexity and diversity of infrastructure facilities, Fraunhofer IOSB is developing an assistance system for the analysis of complex scenes in aerial images. After a short introduction on the application of inference methods for image interpretation and the resulting development of probabilistic scene models, we point out recent advances of the assistance system's recommendation functions in comparison to the system reported at SPIE Europe Security + Defense 2009 [1]. In order to assess spatial relations without the need for additional input from the interpreter, an image viewer has been embedded into the system and objects can now be annotated in the scene using polygons. Recommendations for the most probable interpretation of an object are then derived using Bayesian inference and graph matching methods [2]. Each time a new object is annotated in the image or new features of the objects are assessed, the recommendations are updated. To highlight the benefit of the recommendation function, results of an experimental evaluation for the classification of industrial installations is presented.

2. INFERENCE METHODS FOR ASSISTING THE IMAGE INTERPRETATION PROCESS

In [1] we presented an analysis of the image interpretation process into the following subtasks:

- ***Recognition*** – In this task, objects in the scene are classified based on their image signature. Low-level characteristics of the image signature are used (shape, color, etc.) to distinguish basic categories of objects such as buildings, roads, fences, etc. The sensor's imaging characteristics massively influence the image signature of an object; therefore the interpreter needs a good understanding of the sensor for this task. If an image of reasonable quality is available, most objects can be categorized on the object level without considering spatial context.
- ***Comprehension*** – In the comprehension task, the interpreter determines possible functions of the infrastructure facility and its parts using the knowledge about the functional structure and possible implementations of that structure by a certain configuration of objects.
- ***Projection*** – Based on the hypothesis derived in the comprehension task, expectations for objects and their compositions are generated.
- ***Detection*** – The interpreter investigates the image on signatures corresponding to the expectations derived in the projection task.

Especially the tasks comprehension and projection require a high level of expertise in the specific domain of infrastructure.

2.1 Probabilistic Scene Model

To support the subtasks of image interpretation using inference methods, a probabilistic scene model has been developed which is able to model object features, object occurrence probability and spatial relations between objects with respect to specific scene classes [2,3]. Figure 2 gives an overview on the dependencies between the notions of scene, image, object observation and probabilistic scene model. A scene is a spatio-temporal section of the world, for example the Airport Berlin-Tegel. The image, a two-dimensional projection of the scene in a limited spectral band is the basis for describing objects in the scene in terms of their features and spatial relations. The goal of the image interpretation is to find a task-relevant description of the objects in the scene, the scene description. In the probabilistic scene model, possible scene descriptions of a specific domain are represented. In order to derive recommendations for the classification of objects, graph matching methods are applied to find a good assignment between object observations and object classes with respect to object features and spatial relations. From the best fitting scene descriptions, recommendations for the classification of the whole scene can be derived. Unmatched objects in the scene descriptions are a sign for undetected objects in the image and can be exploited to support the projection task.

Exploiting context for assisted aerial image interpretation
Alexander Bauer and Oliver Herschelmann, Proc. SPIE 7835, 78350K (2010), DOI:10.1117/12.865119

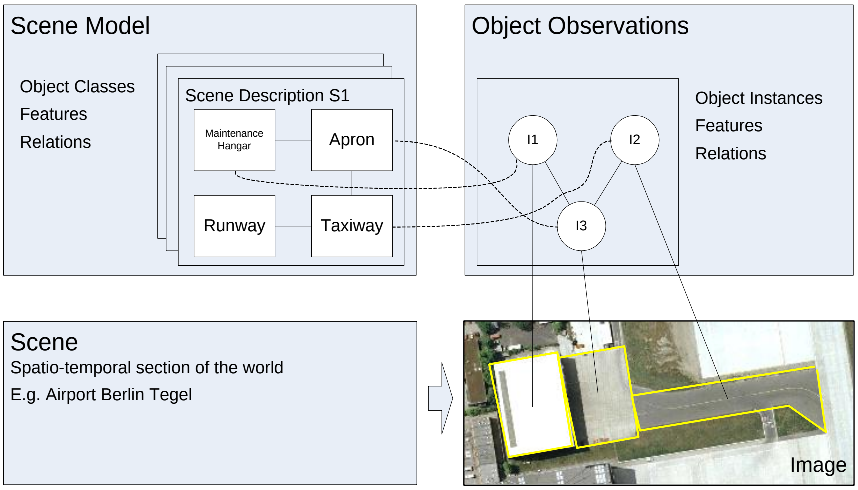

Figure 2. Object observations in an aerial image depicting a real-world scene are matched against a set of scene descriptions. Spatial relations and occurrence probabilities are considered during the matching process using graph matching methods.

2.2 Spatial Relations in Infrastructure Facilities

How are those spatial relations of a specific infrastructure domain identified and modeled? In most processes inside an infrastructure facility, goods, material, energy etc. are transported, transformed or assembled. In facilities related to transport, the major process is to enable people to change between different means of transport in a secure and efficient way. For example at an airport, the departure process starts with the arrival at the airport either by car or by train. Once we have arrived at the car park, we enter the terminal building and check in our luggage. This process goes on with boarding and ends with the plane taxiing to the runway and finally taking off (as illustrated in Figure 3). As there is a high demand for optimizing the runtime of such processes, buildings and other objects involved in the process are arranged to minimize the distance along the process route. This results in a high occurrence of “near” relations between objects involved in successive parts of the process. Starting from the processes occurring in a certain type of facilities, it is easy for an expert to define spatial nearness between objects.

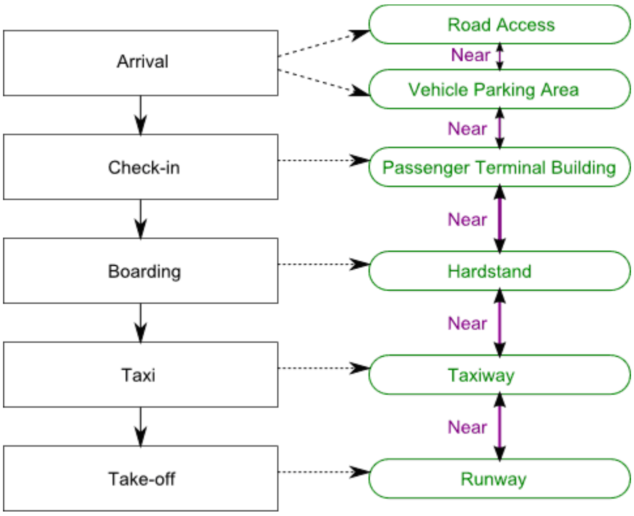

Figure 3: The departure process at an airport. Objects involved in successive process steps are usually arranged in spatial proximity in order to reduce overall process execution time.

Exploiting context for assisted aerial image interpretation
Alexander Bauer and Oliver Herschelmann, Proc. SPIE 7835, 78350K (2010), DOI:10.1117/12.865119

In a similar way, explicit farness between objects may result from regulations about the minimum distance between certain objects for safety reasons. Objects which may endanger other objects (for example the fuel storage and the terminal building) must not be arranged in spatial proximity. These and other relations (e.g., containment) can be represented in the probabilistic scene models and exploited to give recommendations for the image interpreter in the comprehension and projection tasks.

3. SITEANALYST – ASSISTED ANALYSIS OF COMPLEX SCENES

To demonstrate and evaluate the application of the probabilistic scene model for generating recommendations on missing objects, object classification and the scene classification, the software SiteAnalyst has been developed. The system is an extension of the system previously presented in [1].

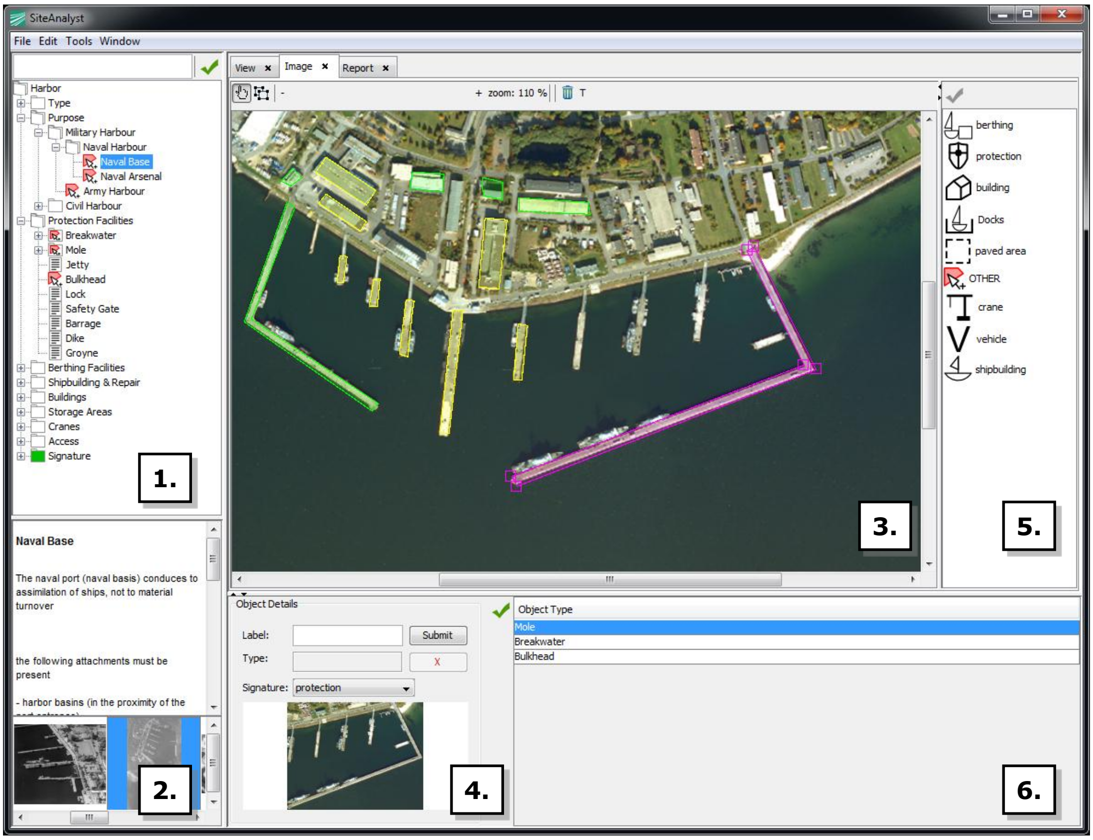

Figure 4: Graphical user interface of SiteAnalyst, with annoted elements: 1. Electronic manual browser, 2. Display for manual entries, 3. Image viewer and object annotation tool, 4. Property view for selected object 5. Toolbar for object signature selection, 6. Display of recommended object classes derived from the scene model

3.1 General Overview

Figure 4 shows a screenshot of the graphical user interface during the interpretation of an aerial image of a harbor. In the left column, a tree navigation element (element 1) enables the user to browse through an electronic manual of vocabulary relevant for the interpretation in the specific domain. For each selected node in the tree, a description and a set of pictures and diagrams are stored in a database. The description, images and drawings of the currently selected vocabulary are displayed in element 2. The central element of the user interface is an image viewer (element 3), which

Exploiting context for assisted aerial image interpretation
Alexander Bauer and Oliver Herschelmann, Proc. SPIE 7835, 78350K (2010), DOI:10.1117/12.865119

displays the image of the current interpretation task. Elements 4, 5 and 6 provide tools to describe objects and to display suggestions produced by the probabilistic scene model.

3.2 Describing Objects

To build the description of the image, the user outlines recognized objects by a polygon shape in the image viewer. If the user is immediately able to classify the actual function of a new object, the respective vocabulary can be selected directly from element 1 and associated to the object by drag and drop or using the confirmation button (green checkmark).

In the case that the user is not able to identify the class of the object directly, the system's recommendation function suggests likely object classes in element 6. To narrow the list of likely object classes, the user can describes the general signature of the object (e.g. building, dock, paved area, crane, vehicle, etc.). Those terms to describe object signatures are shown in a toolbar on the right side of the image viewer (element 5). From there, the object signature can be associated to the object by dragging it from the toolbar. The associated signature information and the spatial arrangement of other objects in the scene are input for the determination of likely object classes from the probabilistic scene model. As output, the most probable object classes are displayed taking into account object features (the signature), local context (specific spatial relations between objects such as nearness) and global context (general occurrences of other object classes in the scene).

3.3 What to look for next?

Infrastructure facilities usually consist of a large set of smaller objects. The most prominent (large hangars, berthing elements, etc.) easily gain attention by the user. As soon as those prominent objects have been described, the user needs a strategy for choosing the next object to describe (driven by the object signature in the image) or for focusing his attention on specific object classes to look for in the image (driven by hypothesis on the overall scene interpretation, as part of the projection task described in Section 2). From the probabilistic scene model, likely occurring but not yet detected objects classes can be determined. Those object classes are displayed in Element 6 if no object is currently selected. For each object class, a detailed description can be displayed in Element 2 by selecting the object class in the list. If based on that, a new object class is detected in the image, the user can start to outline it by dragging the object class to the position in the image.

3.4 Classifying the scene

If enough distinctive objects in the image are described, the probabilistic scene model can also be used to determine the probability distribution of all possible scene classes. In a similar manner as in Section 3.2 and 3.3, a list of suggested scene classes is displayed. In the description details of the scene classes, a list of objects which likely occur in this particular scene class is presented, and checkmarks are shown for those objects classes which have been already found in the image by the user. This way, the user can comprehend how well that scene class corresponds to the model and why it has been suggested by the system.

3.5 Integration with external knowledge bases and manuals

Depending on the level of experience and training, the user may still be overstrained with some explanations and suggestions given by the system. Since it is designed to support the image interpretation under operational conditions, descriptions are intentionally kept brief – they are rather meant to give a reminder on things which have been intensively studied in training. In contrast to that, for example e-learning systems for image interpretation give a detailed explanation on the structure of the object classes in a domain and its relations. To bridge the gap between learning and operational phases, a method for context-aware retrieval of learning units from operational contexts has been developed and integrated in the system. Mareth et al give a detailed description on the integration and the algorithm for retrieval in [4].

4. EXPERIMENTAL EVALUATION

In order to evaluate the potential benefit of the system in operational use, a study was conducted with 10 participants who had no particular experience in infrastructure interpretation. We compared the performance of the participants in interpreting 10 synthetic images of industrial facilities with/without using the functions provided by the probabilistic scene model. Figure 5 shows an example of the synthetic image.

Exploiting context for assisted aerial image interpretation
Alexander Bauer and Oliver Herschelmann, Proc. SPIE 7835, 78350K (2010), DOI:10.1117/12.865119

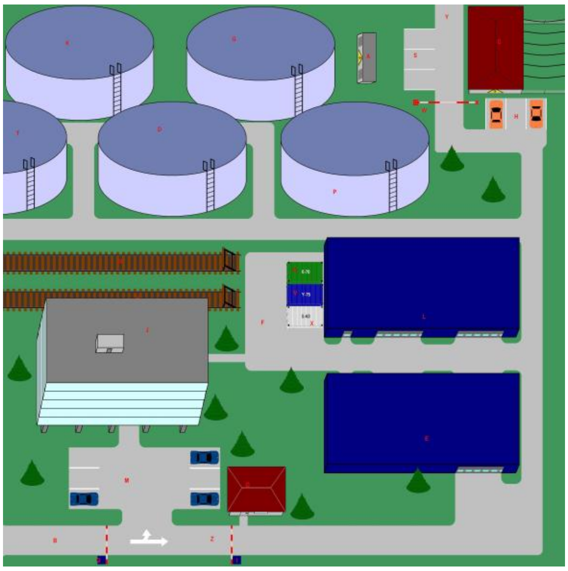

Figure 5: Example of a synthetic image used during the evaluation of the system for industrial facility analysis.

The participants were asked to first describe all objects in the scene and then to classify the whole scene (type of industry) displayed in the image. The parameters of the probabilistic scene model were chosen to give plausible recommendations. After a training phase, the average number of correctly classified objects per time unit was determined. To assess the performance for the scene classification, the number of correct scene classifications per time spent in this phase was determined.

Figure 6 shows all the performance measures determined during the experiment. In the scene classification task, the performance was significantly improved when using the recommendation function for possible scene classes. As evaluating each possible scene class is a very complex and time consuming task, the hint on likely scene classes helped to find the correct scene class faster than by browsing the list of scene classes in alphabetical order. For the object classification task however, there was no significant difference in performance. As in the synthetic images only 15 different object classes occurred, users quickly learned where to find those objects in the tree browser and reached the same performance as choosing from the suggestion list. We expect that in a real domain in which about 100 different object classes are possible, a significant benefit using the recommendation will be observable.

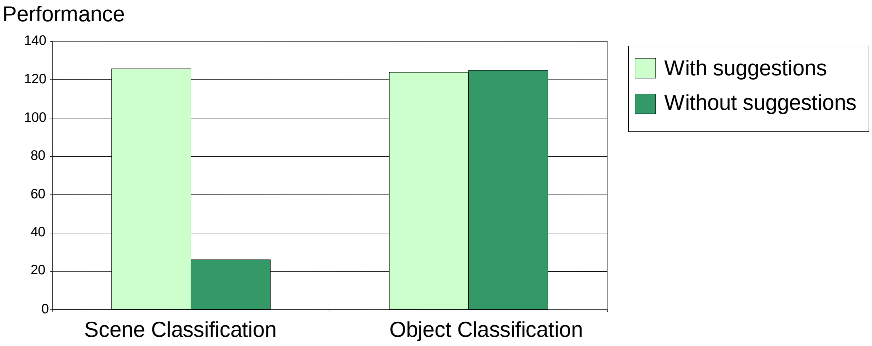

Figure 6: Results of the performance evaluation with/without object and scene class suggestions.

Exploiting context for assisted aerial image interpretation
Alexander Bauer and Oliver Herschelmann, Proc. SPIE 7835, 78350K (2010), DOI:10.1117/12.865119

5. CONCLUSION AND OUTLOOK

Interpreting complex aerial images such as images of infrastructure facilities require a lot of training and experience. If the prior knowledge about the structure of such scenes in terms of object occurrences and spatial configuration can be modeled, useful hints can be given to the interpreter. In this paper, we presented the software SiteAnalyst, an interactive support system with an integrated image viewer and annotation tool and several recommendation functions for object classification, missing object classes and scene classification. The recommendations are based on the signature of the annotated objects and their spatial configuration. In an experimental evaluation we have verified that recommendation functions can improve the performance of inexperienced image interpreters.

To evaluate the performance gain for real aerial images, an experiment is currently prepared for the interpretation of images in the domain of airfields, harbors, oil and metal industry.

REFERENCES

[1] Bauer, A., "Assisted Interpretation of Infrastructure Facilities from Aerial Imagery", Proc. SPIE 7481, 748105 (2009).

[2] Bauer, A., "Probabilistic Scene Models for Image Interpretation", In: Communications in Computer and Information Science, vol. 81, Springer, (2010) .

[3] Bauer, A., "Probabilistic reasoning on object occurrence in complex scenes", Image and Signal Processing for Remote Sensing, Proc. SPIE 7477A, 74770A (2009).

[4] Mareth, N., Streicher, A., Bauer, A., Roller, W., "Context-Aware Retrieval of Learning Units", Proceedings of the IADIS International Conference e-Learning 2010, Vol. 1, IADIS Press, Freiburg, 357-364 (2010).

Exploiting context for assisted aerial image interpretation
Alexander Bauer and Oliver Herschelmann, Proc. SPIE 7835, 78350K (2010), DOI:10.1117/12.865119