Assisted Interpretation of Infrastructure Facilities from Aerial Imagery

A. Bauer*
Fraunhofer Institute for Information and Data Processing (IITB), 76131 Karlsruhe, GERMANY

ABSTRACT

The evaluation of a country's critical infrastructure requires a detailed analysis of facilities such as airfields, harbors, communication lines and heavy industry. To improve the interpretation process, an interactive support system for the interpretation of infrastructure facilities from aerial imagery is developed. The aim is to facilitate the training phase for beginners, increase the flexibility in the assignment of interpreters and improve the overall quality of the interpretation. An analysis of the approach chosen by professional interpreters has been the basis to identify critical steps which can be effectively supported by a software system. To evaluate the benefit of the system, an experimental setup is proposed.

Keywords: Aerial Image Interpretation, Scene Interpretation, Interactive Support Systems, Probabilistic Inference

1. INTRODUCTION

The assessment of a country's infrastructure requires a detailed analysis of facilities such as airfields, harbors, communication lines and heavy industry. Such facilities usually span larger areas of land and the man-made objects stand of well in an aerial image, so their interpretation is a well-suited application for remote sensing. However, in order to make qualified statements about an infrastructure facility based on an aerial image, a thorough understanding of both the sensor's imaging characteristics and the detailed composition of such facilities is required. Most professional image interpreters therefore focus on the interpretation of a single type of infrastructure and become experts in their particular domain through intensive training and experience.

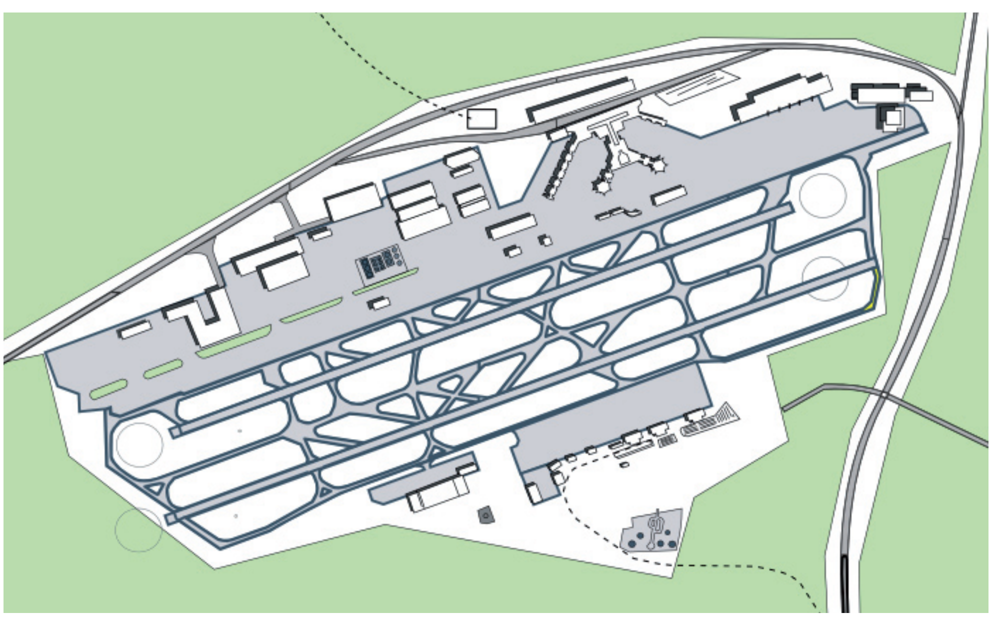

Fig. 1: Sketch of an airfield, depicting the complexity of infrastructure facilities

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

Electro-Optical and Infrared Systems: Technology and Applications VI, edited by David A. Huckridge, Reinhard R. Ebert,
Proc. of SPIE Vol. 7481, 748105 · © 2009 SPIE · CCC code: 0277-786X/09/$18 · doi: 10.1117/12.830298

Under the growing demand for rationalization and the increase in complexity of infrastructure in the globalized world, the image interpreter is and will be forced to cover more diverse assignments, without abundant extension of training periods. Still the quality of the image interpretation result has to be maintained. These conflicting requirements (illustrated in Fig. 2) can only be satisfied by additional support for the image interpreter during the interpretation process.

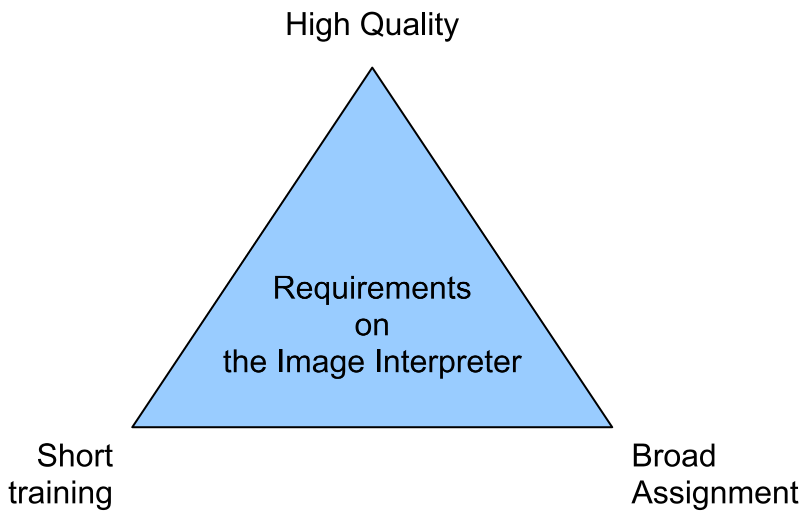

Fig. 2: Conflicting requirements on the image interpreter

Interactive support systems for image interpretation are successfully applied in single object classification such as land, air and sea vehicle recognition, for example using the system RecceMan® (Reconnaissance Manual) developed at Fraunhofer IITB[1]. A similar approach has been used by Nagy et al. for face and flower recognition[2]. These systems help to solve 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 judgments[3]. The interactive classification provided by such systems allows the determination of possible classes of objects, based on features which are assessed by the human interpreter from the image signature. The matching between user assessed and system-stored object class features results in the determination of possible object classes.

As a natural consequence of the application of interactive support systems for single object recognition, an interactive support system for the interpretation of complex scenes (such as infrastructure facilities) from aerial imagery is developed. The aim is to facilitate the training phase for beginners, increase the flexibility in the assignment of interpreters and improve the overall quality of the interpretation. As the interpretation of infrastructure facilities is an application of scene interpretation, *inference methods* from the fields of computer vision and artificial intelligence can be used to extend the functionality for assistance[4-11].

An analysis of the approach chosen by professional interpreters has been the basis to identify critical subtasks which can be effectively supported by software functionality. The current implementation of an interactive support system for infrastructure interpretation implements a subset of the identified functionality. For experiments and evaluation, the software system is implemented for the domain of airfields and harbors.

2. INTERPRETATION OF INFRASTRUCTURE FACILITIES

To design a support system for a complex task such as the interpretation of an infrastructure facility from aerial imagery, the process must be analyzed to identify subtasks on which a support system can provide helpful information and plausible suggestions. In cooperation with active image interpreters of the German Bundeswehr, the procedure applied by the interpreters during the interpretation of several infrastructure facilities was recorded and generalized.

Three levels of abstraction were identified (depicted on Fig. 3), on which an interpreter is developing the scene interpretation. On the *functional level*, the infrastructure is described by its overall function and by the functional composition of physical objects necessary to ensure the operation of the facility. On the *object level*, the objects are described by their general classification which can be extracted from the investigation of the object without taking account for its context. On the image *signature level*, the appearance of the object in the image is analyzed. For example, an image signature of rectangular shape, clearly visible shadow is recognized as a building with flat roof on the object level, and under consideration of the context can be identified as a maintenance hangar, which describes its function inside the facility. The transitions between different levels of abstractions are accomplished by the human interpreter by cognitive processes, which are described in literature as "top-down" and "bottom-up" processes[4].

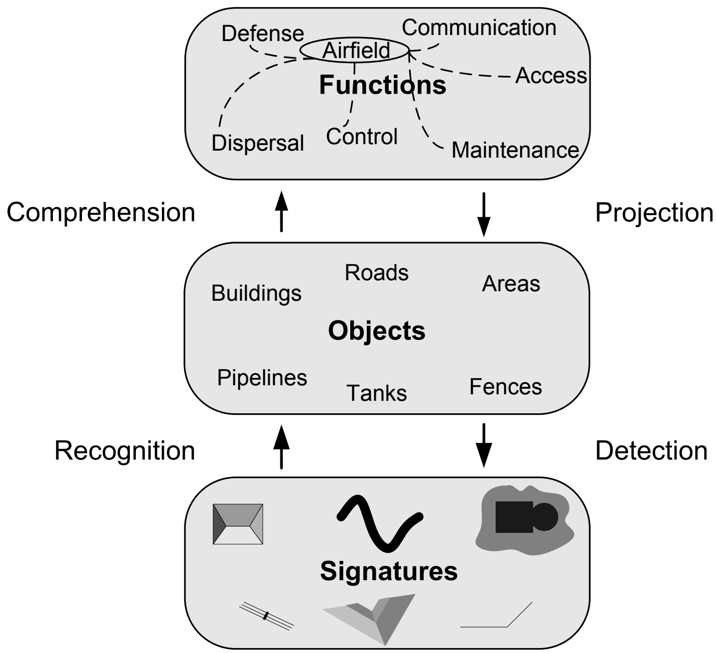

Fig. 3: Levels of abstraction during infrastructure scene interpretation

The transitions are therefore defined as elementary subtasks of the interpretation process:

- ***Recognition*** – 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. If an image of reasonable quality is available most objects can be categorized to the object level without considering spatial context yet.
- ***Comprehension*** – From the composition of objects, their occurrence and spatial arrangement, the interpreter determines possible functions of the infrastructure facility and its parts using the knowledge about the functional structure and its possible implementations by object compositions.
- ***Projection*** – Given hypotheses of the functional structure of the facility has been established from the bottom-up subtasks recognition and comprehension, expectations for objects and their compositions are derived.
- ***Detection*** – The interpreter investigates the image on signatures corresponding to the expectations on the object level.

3. ASSISTING THE INTERPRETATION PROCESS

Based on the definition of interpretation subtasks in section 2, it is possible to derive potential assistance functionality that considers the characteristics (goal, required information) of the respective subtask. In Table 1, a collection of potential assistance functionality for each subtask is presented.

Table 1: Potential assistance functionality by interpretation subtask

Subtask	**Potential assistance functionality**
Recognition	• *Automated object recognition (AOR) methods*, such as building recognition, as far as available for a certain object class and sensor characteristics, are able to classify objects which signature has been selected in the image. • An *interactive classification module* provides hints on probable object classes based on features which have been extracted by the human interpreter, similar to the approach in RecceMan®. • A *reference object catalog* helps to research on possible object classes in the specific facility domain, their description and exemplary images from different sensor types.
Comprehension	• A *bottom-up inference method* suggests possible functions for a composition of objects, based on their occurrence and spatial relations and a respective model of the infrastructure domain. • *Schematic drawings* of characteristic object compositions for a specific function help the interpreter to find tangible hypotheses.
Projection	• A *top-down inference method* suggests possible objects based on functions which are expected at the functional level from the comprehension subtask and a functional model of the infrastructure domain. • *Schematic drawings* help to understand how a specific function can be implemented by objects in the infrastructure facility to derive hypotheses on possible objects.
Detection	• The *top-down inference method* determines regions of interest (ROIs) for the occurrence of specific object classes based on the expected spatial arrangements of objects from a model of the infrastructure domain. • *Automated object recognition (AOR) methods* detect objects of the expected object class applying pattern recognition methods on the image. • A *reference object catalog* supports the interpreter in the detection task by providing description and exemplary images from different sensor types.

Passive methods, which do not take into account the input of the user on the current state of his interpretation of the infrastructure facility, such as the display of references images, schematic drawings and text descriptions on objects are summarized into the basic functionality *electronic manual*. Active methods take into account the image and higher level information extracted from the image such as recognized objects. Automated object recognition methods, which directly work on the image signature, are able to detect objects on the whole image or to classify an object based on its image signature. Interactive methods rely on the cooperation with the interpreter, like the interactive classification approach realized in RecceMan® (see Introduction) and inference methods which are able to derive hypotheses from available high-level scene information like detected objects.

Table 2: Assistance functionality associated to the interpretation subtasks

	Recognition	**Comprehension**	**Projection**	**Detection**
Automated object recognition	✓			✓
Interactive classification	✓			
Electronic manual	✓	✓	✓	✓
Inference methods		✓	✓	✓

To summarize the possible assistance functionality, Table 2 shows the association of different assistance functionality to the subtasks of the interpretation process. An electronic manual supports all subtasks, so it should be the basic functionality of any support system for infrastructure interpretation. The remaining functionality will further improve the assistance performance of the system, but tangible algorithms and methods must be available or developed first, to implement the demanded functionality.

4. INTERACTIVE SUPPORT SYSTEM

Based on the determination of potential functionality in section 3, an interactive support system was designed and implemented. The system implements a subset of the proposed functionality, a combination of the electronic manual and the application of an inference method. The inference method is based on an algorithm which is able to reason about object occurrence in complex scenes[11]. The user interface is depicted in Fig. 4, the database of the system has been populated for the interpretation of airfields. The system main window is divided into three areas. In the top-left component a tree-structure *object catalog* is displayed, providing access to all interpretation-relevant terms, structured from coarse areas of the infrastructure domain to fine details such as single objects. On the right hand side, inference results such as objects expected in the image based on the inference algorithm are ordered according to their probability of occurrence. If a term is selected from the object catalog or the inference results, reference information such as text description and exemplary images and drawings are displayed to the user.

The interpretation process is supported by the supply of reference information accessible from a structured object catalog, but more importantly, the system is able to provide hints on possible object occurrences which have not been detected yet. The inference is based on a scene model which holds the prior knowledge about the occurrence of objects. Each time new evidence is collected, such as observations of objects in the infrastructure, the inference result is updated. Thus, in an iterative process, the interpreter develops the description of the infrastructure in cooperation with the system, which provides him with suggestions how to interpret the infrastructure and on which objects he should focus his attention.

As soon as sufficient objects are classified and their observations confirmed to the system, the inference algorithm is able to create hypothesis about the interpretation of the whole infrastructure facility. For example, in the context of airfields, the determination of the purpose of the airfield (for example for the deployment of military airplanes, as a civil airport or for cargo transport) is based on the occurrence of certain objects. The probability of different interpretations is displayed (see Fig. 5). For each hypothesis, the objects which should be detected to complete the hypothesis are highlighted.

As soon as the interpretation is completed, the confirmed objects collected by the system during the interpretation process can be exported as a basis for the interpretation report.

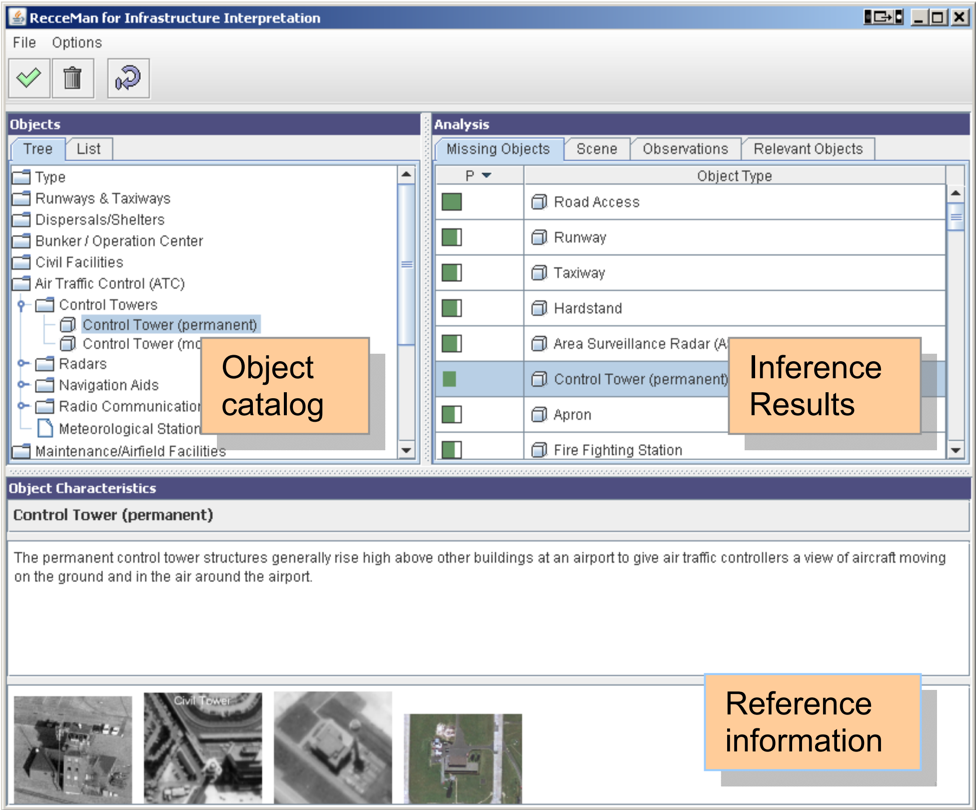

Fig. 4: Graphical user interface of the interactive support system

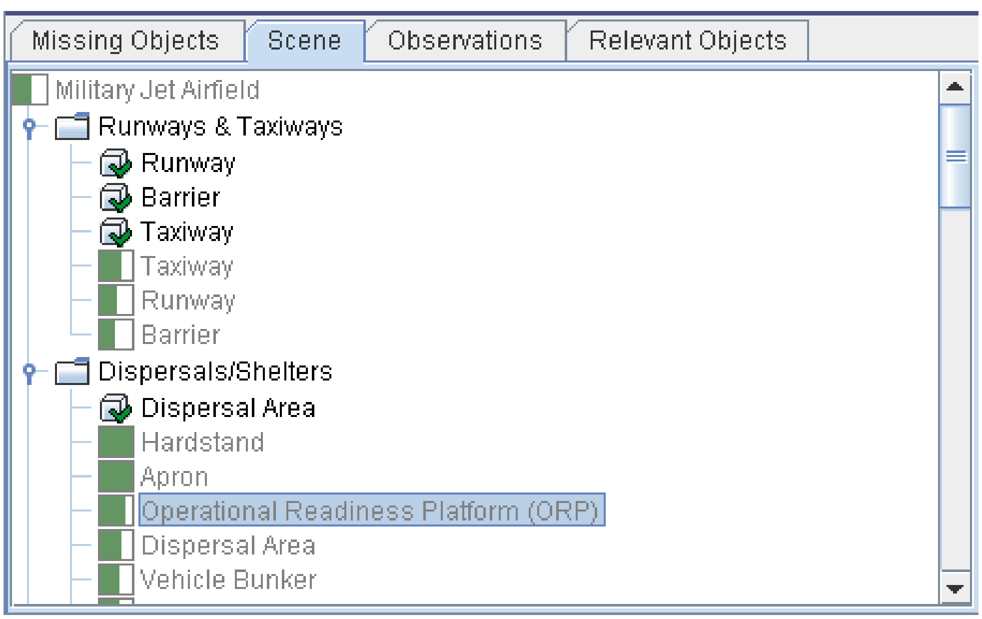

Fig. 5: Interpretation hypothesis determined by the inference algorithm

5. EXPERIMENTAL EVALUATION OUTLINE

A simulation of different strategies for scene interpretation has shown that the application of an inference method can lead to a significant benefit in interpretation performance[11] (depicted in Fig. 6).

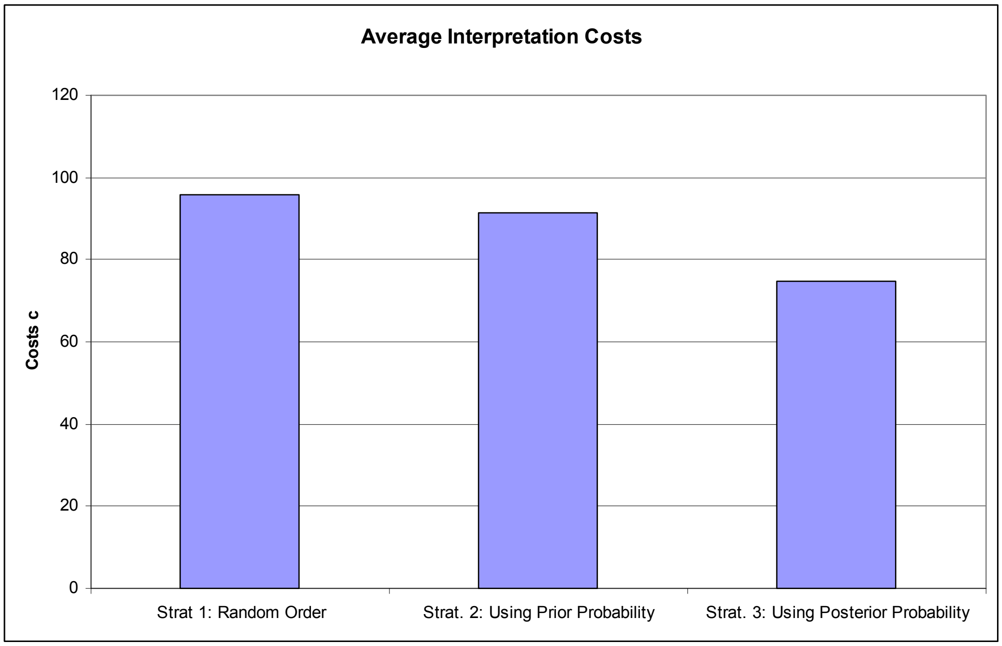

Fig. 6: Simulated average interpretation costs for different image interpretation strategies (taken from the article "Probabilistic reasoning on object occurrence in complex scenes" [11]). The costs represent the effort to create a complete scene description. Strategy 3 uses the proposed inference method to adaptively search for object classes with a high occurrence probability. Strategy 2 merely uses the prior occurrence probability, without incorporating intermediate detections of objects. Strategy 1 randomly selects object classes to search for.

To evaluate the added benefit of the inference method for an interactive support system in addition to electronic manual functionality, an experiment is going to be conducted. The experiment will assess the performance of interpreters, measured as the relation between result accuracy and execution time, for two groups of participants. One group will work using an assistance system which only implements the electronic manual, whereas the second group will be additionally supported by the inference method. To exclude effects of prior knowledge of the participants, an artificial domain of infrastructure facilities will be the developed. For example, in the domain of industry facilities, factory types which produce different imaginary products are defined and their structure (occurring objects and their composition) is modeled. Artificial images of instances of different factory types are generated, and their interpretation is used as the evaluated task for the experiment.

6. CONCLUSION AND OUTLOOK

The need for assistance in the interpretation of infrastructure facilities is evident due the increasing and conflicting requirements on the image interpreter. The design of interactive support systems for the interpretation of complex infrastructure facilities is derived based on an assessment of the approach taken by experienced image interpreters, taking place on different levels of abstraction. Several functionalities supporting the image interpretation process are identified, of which the interactive manual and an inference method are implemented in an interactive support system for infrastructure interpretation. The system is able to suggest possible undetected object occurrences based on recognized objects during the iterative process of image interpretation. The database of the system is populated for harbors and airfields. The implementation of the remaining suggested functionality will be investigated. For evaluation of the system's benefit, the outline of an experimental evaluation is proposed.

REFERENCES

[1] Bauer, A.; Geisler, J., "Decision support for object recognition from multi-sensor data", Thoma, Klaus (Editor), Fraunhofer Verbund für Verteidigungs- und Sicherheitsforschung: Future security: 3rd Security Research Conference Karlsruhe; 10th-11th September 2008: Congress Center Karlsruhe, Germany. Stuttgart: Fraunhofer IRB Verl., 321-326 (2008).

[2] Evans, A. , Sikorski, J., Thomas, P., Cha, S.-H., Tappert, C., Zou, G., Gattani, A., Nagy, G., "Computer Assisted Visual Interactive Recognition (CAVIAR) Technology," 2005 IEEE International Conference on Electro-Information Technology, Lincoln, NE, (2005)

[3] Fitts, P.M. (Ed.), Human Engineering for an effective Air-Navigation and Traffic Control System, NRC, Washington, DC (1951).

[4] Matsuyama, T. and Hwang, V., SIGMA: A Knowledge-Based Aerial Image Understanding System, Plenum Press, (1990).

[5] Russ, T. A., MacGregor, R. M., Salemi, B., Price, K. and Nevatia, R., "VEIL: Combining Semantic Knowledge with Image Understanding," ARPA Image Understanding Workshop, (1996).

[6] Dillon, C., and Caelli, T., "Learning image annotation: the CITE system," Journal of Computer Vision Research, 1(2), 90-121(1998).

[7] Hanson, A., Marengoni, M., Schultz, H., Stolle, F., Riseman, E. and Jaynes, C., "Ascender II: a framework for reconstruction of scenes from aerial images," Workshop Ascona 2001: Automatic Extraction of Man-Made Objects from Aerial and Space Images (III), 25-34 (2001).

[8] Rimey, R. D. and Brown, C. M., "Control of Selective Perception Using Bayes Nets and Decision Theory", International Journal of Computer Vision, Vol 17, 173-109 (1994)

[9] Lueders, P., "Scene Interpretation Using Bayesian Network Fragments", Lecture Notes in Economics and Mathematical Systems, Volume 581, 119-130, doi:10.1007/3-540-35262-7_7, (2006).

[10] Lin, L., Wu, T., Porway, J. and Xu, Z., "A stochastic graph grammar for compositional object representation and recognition", Pattern Recognition, Volume 42, Issue 7, Pages 1297-1307, ISSN 0031-3203, doi: 10.1016/j.patcog.2008.10.033, (2009).

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