An Accumulating Interpreter for Cognitive Vision Production Systems

E. Michaelsen, L. Doktorski, K. Luetjen

Fraunhofer Institute of Optronics, System Technologies and Image Exploitation
Gutleuthausstr. 176275 Ettlingen, Germany
eckart.michaelsen@iosb.fraunhofer.de

Knowledge-based recognition and analysis of high dimensional data such as aerial images often has high computational complexity. For most applications time and computational resources such as memory are limited. Therefore approximately correct interpreters with any-time capability are proposed. In this contribution a special software architecture is published, which can handle the administration of complex knowledge-based recognition and analysis in a tractable manner.

Introduction

Two principle categories of automatic recognition and analysis from high dimensional data (such as remote sensing images) can be distinguished: 1) Learning appearances; 2) utilizing machine interpretable knowledge. Emphasis today is on the learning approach. It yields optimal performance given the corpus of data for training is representative and the laws of decision theory are obeyed. The down-side is clear: The labor of labeling the training data will usually be cumbersome, and it always remains questionable whether they are really representative. E.g. in recognition of man-made objects from aerial images very large training sets must be labeled by hand, and still surprising new variants will occur with every new image.

Utilizing machine interpretable knowledge can in principle get along even without a single training image. Most existing work on this topic emphasizes logical correctness, consistency, and even completeness. Such approaches inevitably scale badly with rising numbers of instances in the image and knowledge rules. Computation time and effort can hardly be predicted for deep automatic analysis. The goal of this work is to provide a software package that can keep the semantic richness while it emphasizes practical applicability for time critical tasks.

Related Work: Syntactic methods are among the first options discussed for automatic image understanding [18]. A still valid reference for knowledge based automatic recognition and analysis in general but with focus on semantic nets is [15]. Most internationally well known production system approach for remote sensing data has been Schema (or KBV) [1]. SIGMA of Matsuyama & Hwang [6] also was pioneering work. Contemporary work on syntactic recognition from aerial images can be found in [2]. Our own references are given below. Most of them are also available on http://publica.fraunhofer.de/starweb/pub08/en/

Generation, Reduction, and Accumulation

Knowledge-based Recognition by Production Systems: Context-free constrained multi-set grammars are discussed in [5] particularly with regard to graphical languages and computer interfaces. The basic idea is generalizing the generative string grammars by replacing the concatenation constraint by a more general constraint. Next to their symbolic name the instances have attributes such as locations, orientations, etc., on which the constraints are defined. It is known that such systems can solve the satisfiability problem of propositional logic and therefore are in the general case NP-complete [7]. Such systems can work in both directions – generative and reductive. Generative means that a root instance is given and then, by successive application of productions left to right (where a random generator picks attributes fulfilling the constraints), the objects are replaced until only primitives are left. In this way an image is rendered which is member of the language. Reductive means that from the images primitives are segmented and than by successively testing the constraints all possible right to left replacements are explored, until possibly a root instance is reached.

Recognition Using the Approximate Any-time Interpreter: Precise formal language definitions for constrained multi-set grammars (there called coordinate grammars in accordance with [18]) are given in [7, 13]. This includes accumulative parsing, where - during the right to left application of a production – the right-hand side objects are not removed from the database. It follows that by such accumulative parsing derivations can be made that are not valid in the reducing sense (because of double use of objects). However, if this is a rare exception – due to the constraints – accumulative parsing can be a good option. We see it as an approximate solution saving a huge amount of combinatory administration.

Cluster analysis: Often successive application of the same production has some meaning in a clustering or Hough like estimation sense (e.g. a long contour attached to a short contour segment results again in a long contour). For these situations there is a short-cut production accumulating larger sets in one step [12].

User Independent Software Architecture

The BPI System: [4] proposed the BPI system as user-independent solution for accumulative interpretation of such production systems following the blackboard rationale. Such systems use a dispatcher assigning working hypotheses to computational resources. Such a hypothesis is called **WorkingElement** in Figure 1. It consists of a triggering object instance (called **ImageObject**) and entries from a corresponding production rule (namely left-hand side, i.e. **HypoType**, partners in the right-hand side **PartnerType** – and, optionally, context). The dispatcher module gets the production system as input. Thus, if a **WorkingElement** has no hypothesis attached yet it will form admissible clones, else it will call the appropriate methods searching for partners testing constraints, and if those hold new **ImageObject** instances will be produced. From each newly produced instance (and from primitives segmented from the input image) new **WorkingElement** instances are formed with no hypothesis attached yet.

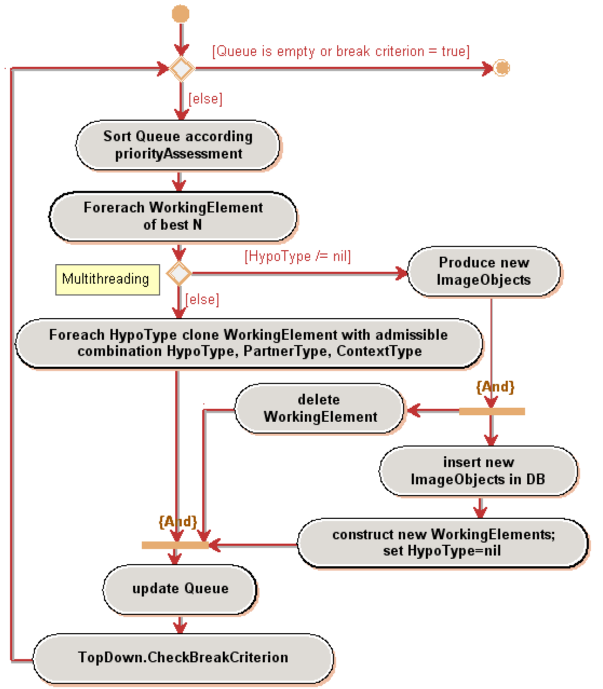

Fig. 1. UML-activity diagram for accumulating dispatch.

This cycle can be repeated until, either all hypotheses have been processed, or the object of interest has been instantiated, or other break criteria (such as maximal admissible time) are meat. The set of **WorkingElement** instances is organized as **Queue** which is ordered according to an assessment value. Such value is by default given through a *quality* measure for the triggering image object (data-driven search). Many systems have an additional assessment component – the *importance*. This is achieved by weight factors on the *quality*. Given a particular state of the search **WorkingElement** instances gain different *importance* for the task at hand – particular **HypoType**s will be of more interest, instances in particular image regions may be of higher or lower relevance. Such use of *top-down* importance for focusing the search is described in detail in [13]. Both assessment components (*quality* and *importance*) have to be provided by the user.

The BPI System was used for many years and many ambitious 2D and 3D recognition and analysis problems (see Table 1). It was implemented using assembler code under VAX-VMS, featured a PASCAL-like syntax for its user language, and a special graphical interface for knowledge acquisition and

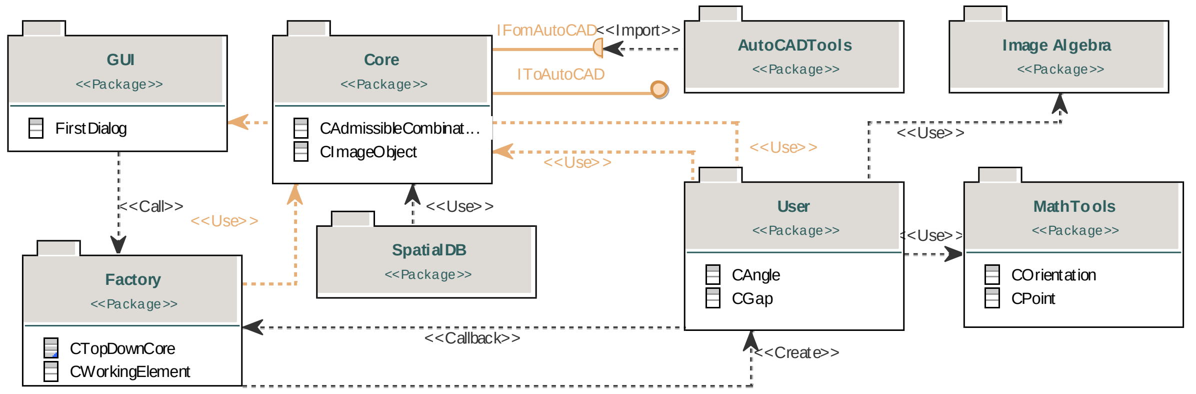

Fig. 2. UML-package diagram for cognitive vision system.

explanation called WEEK (Wissens Erwebs- und Erklärungs Komponente). Emphasis was on swiftness using parallelization and, in particular, associative access to possible **PartnerType** instances using hash mechanisms and also special hardware [8]. BPI activities ended around 2004 mainly due to the restriction to VMS operating system. Intermediately we used a provisory Matlab implementation (see Table 1).

The COGVIS System is a newly implemented variant using .NET functionalities. The architecture emphasizes object oriented programming and modular structure. Figure 2 displays its packages. The dispatcher is implemented in package Factory. In particular this contains the queue handling and handles for re-assessing elements. It uses the most abstract declaration **ImageObject** in the package Core. Main entry to the System is usually interactively provided by the package GUI, where one can choose appropriate productions, input data, and running parameters. But the system can of course also be called from other systems such as in a navigation control loop [11]. GUI also contains run-time visualization threads for the current statistics of accumulated objects and used resources and for a graphical visualization drawing the objects e.g. as overlay to an image. MathTools provides geometrical classes and methods that are commonly used by many constraints and object constructors in the User packages. SpatialDB is meant to provide associative access to possible **PartnerType** instances in the style of data-banking (yet under construction). ImageAlgebra contains some of the usual convolution and morphological filter procedures, thresholding and other segmentation methods constructing primitives from input images. AutoCADInterface replaces the WEEK functionality of BPI. It can isolate instances from an achieved result – together with their derivation tree, measure distances, select smaller input object sets, and construct new object instances interactively.

The user is invited to define his/her own classes as specializations of **ImageObject** and plug them in together with a table of admissible combinations (production system) as a package on its own. Figure 3 shows a class diagram of an example system used for visual landmark based UAV-navigation.

Potential and Intended Applications

Hardly any restriction on possible User packages can be imagined. Table 1 gives a rather incomplete list of published applications so far.

Table 1. Some applications so far

Ref.	System	Task, Domain
[4]	BPI	UAV landmarks
[8]	BPI	3d Vehicle, Ground-based
[16]	BPI	3d Buildings, Aerial im.
[3]	BPI	ATR, Fusion IR/RADAR
[9]	BPI	Vehicle, Aspect based
[16]	Matlab	3d Buildings, LIDAR
[10]	Matlab	Geometric estimation
[14]	CogVis	Buildings, Airborne SAR
[11]	CogVis	UAV landmarks

Exemplarily, we display in Figure 3 a class diagram of the user package used in [11].

Major highways are salient landmarks for UAV navigation. In order to stably achieve also the position along the highways preferably bridges over them are used. They appear as CTCrossing. The diagram shows left to right in red color the part-of relations. From top to bottom in blue color the inheritance hierarchy is shown.

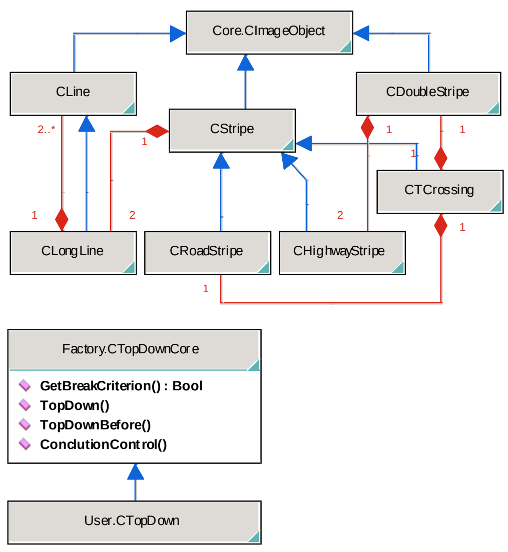

Figure 3. UML-class diagram for an example user package for recognition of bridges.

Conclusion

Knowledge based automatic recognition and analysis of high dimensional data is a long term endeavor being perused for decades. Therefore, occasional re-implementations of the key algorithms and procedures are necessary. The one presented here emphasizes modern object orientation and parallelization.

References

1. Draper, B., Collins, R., Brolio, J., Hanson, A., Riseman, E.: The Schema System. IJCV. – 1989 - Vol. 2.-P. 209-250.
2. Guo, C.-E., Zhu, S.C., Wu, Y.N.:. Modelling visual patterns by integrating descriptive and generative methods. International Journal on Computer Vision, 53 (1), 5-29, 2003.
3. Jäger, K., Bers, K.-H., Jurkiewicz, K.: Three-dimensional intra- and inter-frame target detection in IR/mmW image sequences for missile seekers. In: Andresen. B., Fulop. G., Strojnik. M. (eds): Infrared Technology and Applications XXVI, Proc. SPIE, Vol. 4130, 2000, 679-687.
4. Lütjen, K.: BPI: ein Blackboard-basiertes Produktionssystem für die automatische Bildauswertung // Mustererkennung (DAGM) 1986, Informatik Fachberichte 125, Springer, Berlin, P. 164-168.
5. Marroitt, K., Meyer, B. (eds.): Visual Language Theory. Springer, Berlin 1998.
6. Matsuyama, T., Hwang, V. S.-S.: Sigma a Knowledge-based Image Understanding System. Plenum Press, New York. 1990.
7. Michaelsen, E.: Über Koordinaten Grammatiken zur Bildverarbeitung und Szenenanalyse. Diss. Univ. of Erlangen, 1998, online available under http://www.exemichaelsen.de/Michaelsen_Diss.pdf.
8. Michaelsen, E., Lütjen, K., Stilla, U.: Associative access and special hardware for production nets. Pattern Recognition and Image Analysis, 9(4), 1999, pp.662-666.
9. Michaelsen, E., Stilla, U.: Probabilistic decisions in production nets - an example from vehicle recognition. In: Caelli, T. (ed.): Proceedings / SSPR and SPR, Berlin: Springer, (LNCS 2396), 2002, pp.225-233
10. Michaelsen, E., von Hansen, W. Kirchhof, M., Meidow, J., Stilla, U.: Estimating the essential matrix: GOODSAC versus RANSAC. In: Förstner, W. (ed.): ISPRS Commission III, , PCV 2006: The Intern. Arch. of Photogr., 36, Part 3, pp.161-166
11. Michaelsen, E., Jaeger, K.: A GOOGLE-Earth Based Test Bed for Structural Image-based UAV Navigation. IEEE: FUSION 2009, 12th International Conference on Information Fusion: ISBN: 978-0-9824438-0-4, 2009, pp.340-346.
12. Michaelsen, E, Doktorski, L., Arens, M.: Shortcuts in production-systems - A way to include clustering in structural pattern recognition. (PRIA-9-2008), Nischnij Nowgorod: Lobachevsky State University, ISBN: 978-5-902390-14-5, 2008, pp.30-38
13. Michaelsen, E., Arens, M., Doktorski, L.: Interaction of Control and Knowledge in a Structural Recognition System. In: Mertsching, B.(ed.): KI 2009, Berlin: Springer, (LNCS 5803), 2009, pp.73-80.
14. Michaelsen, E., Stilla, U., Soergel, U., Doktorski, L.: Extraction of Building Polygons from SAR Images: Grouping and Decision-Level in the GESTALT System. Pattern Recognition Letters, 31: 2010, pp. 1071-1076,
15. Niemann, H.: Pattern Analysis and Understanding. Springer, Berlin 1989.
16. Stilla, U., Michaelsen, E., 1997. Semantic modelling of man-made objects by production nets. In: Gruen, A., Baltsavias, E.P., Henricsson, O. (Eds.), Automatic Extraction of Man-made Objects from Aerial and Space Images (II). Birkhäuser, Basel, pp. 43-52.
17. von Hansen, W., Michaelsen, E., Thoennessen, U., 2006. Cluster Analysis and Priority Sorting in Huge Point Clouds for Building Reconstruction. In: Proceedings of 18th International Conference on Pattern Recognition, ICPR 2006, Hong-Kong, vol. 1, pp. 23-26.
18. Rosenfeld, A.: Picture Languages. Academic Press, New York, 1979.