A General Cognitive System Architecture Based on Dynamic Vision for Motion Control

Ernst D. DICKMANNS
Aero-Space Engineering (LRT), Universitaet der Bundeswehr Munich (UBM),
Institut fuer Systemdynamik und Flugmechanik (ISF)
D-85577 Neubiberg, Germany

ABSTRACT

Animation of spatio-temporal generic models for 3-D shape and motion of objects and subjects, based on feature sets evaluated in parallel from several image streams, is considered to be the core of dynamic vision. Subjects are a special kind of objects capable of sensing environmental parameters and of initiating own actions in combination with stored knowledge. Object / subject recognition and scene understanding are achieved on different levels and scales. Multiple objects are tracked individually in the image streams for perceiving their actual state ('here and now'). By analyzing motion of all relevant objects / subjects over a larger time scale on the level of state variables in the 'scene tree representation' known from computer graphics, the situation with respect to decision taking is assessed.

Behavioral capabilities of subjects are represented explicitly on an abstract level for characterizing their potential behaviors. These are generated by stereotypical feed-forward and feedback control applications on a separate systems dynamics level with corresponding methods close to the actuator hardware. This dual representation on an abstract level (for decision making) and on the implementation level allows for flexibility and easy adaptation or extension. Results are shown for road vehicle guidance based on three cameras on a gaze control platform.

Keywords: Active vision, model-based vision, vision system architecture, autonomous vehicles, mobile robots.

1. INTRODUCTION

The third-generation dynamic vision system based on spatio-temporal modeling, developed at UBM, has been dubbed '**EMS-Vision**' according to its main properties: a) **E**xpectation-based image sequence interpretation and behavior decision, b) **M**ulti-focal camera arrangement on a gaze control platform, and c) **S**accadic viewing direction control and perception capabilities [1-4]. Though there are many vision systems for (autonomous) vehicles under development at present, no other one seems to have the above-mentioned properties of vertebrate-type vision with its efficiency, flexibility and growth potential. A good survey may be obtained from the proceedings of the yearly International Symposium on 'Intelligent Vehicles'xx' [5] (started in 1992, xx designating the last to digits of the year).

An approach of similar scope as EMS-vision but with fundamental differences in the methods for vision and knowledge representation is [6]. While spatial representations are done there with the help of grids of different scales, EMS-vision takes advantage of the standard object-oriented methods used in computer graphics exploiting **h**omogeneous **c**oordinate **t**ransformations (HCT) and scale factors directly integrated in the HCT-framework by 4x4 transformation matrices. These also provide for seamless inclusion of perspective projection. The big challenge in vision (as opposed to computer graphics) is that many of the entries into the transformation matrices are not known beforehand, but are the unknowns of the vision process. The 4-D approach to dynamic vision has solved this problem by extending the Kalman filter approach to perspective imaging introducing the adaptive fit of generic spatio-temporal models through recursive iteration [1, 7].

2. BASIC PROCESSING STAGES FOR VISUAL PERCEPTION AND CONTROL

In the latest stage of development, visual perception is organized as a three-stage process, each requiring different image access, data processing methods and attention control:

1. Schematic detection of features in a wide field of view (~ 100° by 40°), covered by two wide-angle cameras with coplanar divergent optical axes; this first stage asks the question: 'Is there something indicating an object of interest to mission performance?' (VQ1). With specific sets of characteristic features for object detection set by the higher system levels (or the human operator), this stage can work without reference to temporal aspects on raw data (lower left in figure 1).

2. Groupings of features in images are hypothesized as elements in perspective maps of stationary or moving 3-D objects under certain aspect conditions. In the near range, this may be done in the wide-angle images directly; for the far range, a fast gaze shift brings features of interest into the field of view of the high-resolution camera (tele or zoom). These objects are then tracked over time by prediction error feedback using recursive estimation techniques well known in systems dynamics [1, 7]. Here, the question is answered: 'What type of object is it, and what is its relative state in 3-D space (including relative velocity components)'? (VQ2). At video-rate, the best estimates for the state variables and parameters of each object are sent to a '**D**ynamic **O**bject data**B**ase (DOB)' serving as information exchange platform in the system (see second (4-D) and third layer (left) from bottom in figure 1).

3. For objects of special interest, especially for those capable of perceiving part of the environment and of initiating action or motion on their own (called 'subjects'), the data in the DOB are observed on a larger time scale. This allows recognizing typical 'maneuvers' (like lane changing or turning-off onto

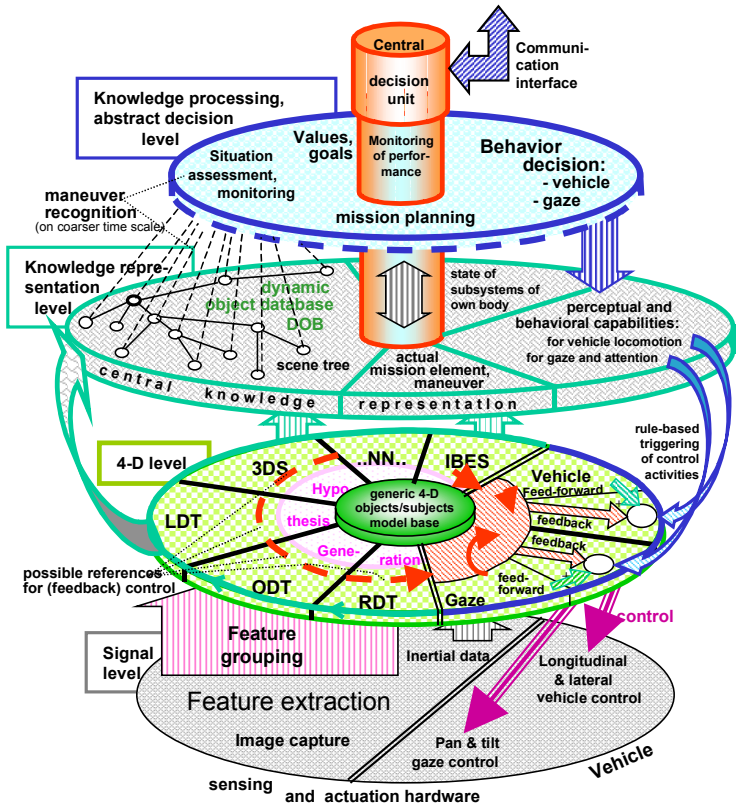


Figure 1: Overall cognitive system architecture in EMS-vision (4 layers)
RDT = Road Detection and Tracking; ODT = Obstacle Detection and Tracking
LDT = Landmark Detection and Tracking; 3DS = 3D Surface Recognition
IBES = Inertially-Based Ego State; NN = Future additional capabilities

crossroads) and behavioral skills of other subjects (like 'lane following' or 'convoy driving'). On this level the question to be answered is: 'What is this guy doing, and what are (possibly) his or her intentions? (VQ3). [Note that this third level does not rely directly on image data any more, but makes use of symbolic descriptions of 'subjects' and their attributes as well as the time histories of their states stored in the DOB.]

Sections 3-5 give some more details on these perception stages; in section 6, the problem of situation assessment is addressed. Section 7 treats behavior decision on an abstract level while in section 8 behavior implementation with systems dynamics methods is discussed. Section 9 gives a brief view on mission performance and section 10 on system integration. After discussing experimental results in section 11, conclusions are drawn in section 12.

3. FEATURE DETECTION AND GENERATION OF OBJECT HYPOTHESES

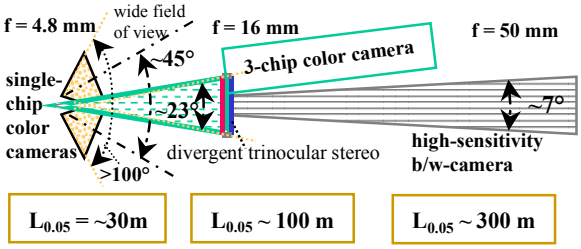
This topic is the most data-intense part of the system. Three cameras of standard resolution (~ 400 000 pixels per frame), two of which are black-and-white, while one is a 3-chip-color camera, yield a video data flow rate of 50 MB/s (at 40 ms cycle time, 25 Hz, see figure 2). A single microprocessor and communication link of today's standard cannot handle this rate continuously beside data processing. Therefore at present, distributed processing with separate video data input is mandatory. In our system, two to three dual-processor systems are devoted to basic feature extraction ***and*** recognition of special classes of objects in video streams (a combination of levels 1 and 2). This keeps communication needs on the feature level low. However, it is not optimal from a methodological point of view. With increased computing power for feature extraction, storage capacity for rich feature data bases, and communication bandwidth for sharing results with little delay time among all potential users on the next higher level, bottom-up feature and object detection could be exhaustively concentrated on this level. For area-based features like color and texture, much more computing power is needed than presently available.

This has to be an area of development for future systems; pattern recognition methods and corresponding hardware for iconic image processing may find a lasting field of application on this level right before the transition to object-oriented methods with strongly increased top-down components is done. Future hardware development will shift the pattern of practical solutions in the transition region between the vision levels 1 and 2 mentioned above.

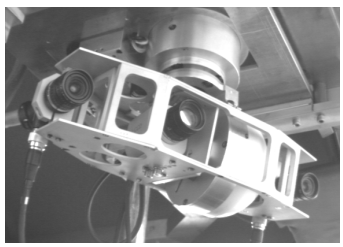
4. OBJECT TRACKING AND RELATVE STATE ESTIMATION

On this level, quite a gain in computing efficiency can be achieved by resorting to spatio-temporal models as pioneered in the 4-D approach [7]. Historical developments of microprocessors and applications in the field of autonomous ground vehicles have led to the object classes shown in the left part of the 4-D level in figure 1 (second from bottom; for the acronyms see the legend).

Road detection and tracking (**RDT**) can now be done on a single standard microprocessor including all feature extraction necessary. General 3-D surface recognition (**3DS**, including



a) Fields of view and viewing ranges of MarVEye.



b) Realization of MarVEye4 with 3 CCD-cameras on a two-axis platform.

Figure 2: MarVEye system parameters (a), camera set in VaMoRs on yaw and pitch platform, large stereo base and one mild tele-camera (b).

real-time detection of negative obstacles like ditches) requires powerful special processors like the ACADIA-board (Pyramid Vision Technology) with up to 80 billion operations per second as a plug-in card for PC's [8, 9].

For driving on rough surface with large perturbations in angular orientation of the own vehicle, a combination of inertial and visual data interpretation has proven advantageous. Inertial sensing includes the effects of perturbations on the acceleration and angular rate level, that is with lead components in systems engineering term. Since these signals are easily interpreted and integrated to pose changes, the own bodily orientation (except for drift disturbances) is more easily derived from inertial data than from vision. Delay times in vision of several video cycles (> 0.1 s) favor this approach, also found in vestibular-ocular interaction in vertebrates. The process IBES (**I**nertially - **B**ased **E**go **S**tate estimation) in the lower back part of figure 1 realizes this feature of EMS-vision. Drift problems with very low corner frequencies (as usual in inertial state estimation) are solved visually by observing features from objects far away (horizon, sky line).

The best estimates for object states of all units observed are stored in the **D**ynamic **O**bject data**B**ase (DOB, 3rd level from bottom in figure 1). This is done with a scene tree in which each node represents a physical object or a frame of reference; the edges of the tree are HCT's, linking features in 3-D space to projected features in images [10]. This is the world of imagination in the 'mental' upper part of figure 1 (left side on level 3). Note that this framework allows computing the Jacobian matrix of first-order derivatives (which is very rich in information for spatial interpretation) between features in the image and their corresponding ones in 3-D space. Prediction error feedback in **E**xtended **K**alman **F**iltering (EKF) thus substitutes perspective inversion on the 4-D level for answering VQ2 (see section 2).

5. MANEUVER AND INTENT RECOGNITION

Handling VQ3 requires a different knowledge base. Here, the notion of typical behavioral capabilities on an extended time scale has to be available. Since this is the same background knowledge as needed for own decision making, these purely 'mental' activities requiring no direct image data input are organized on the uppermost level 4 in figure 1. Typically, information about several different objects / subjects has to be interpreted in conjunction, like one or more other vehicles and their orientation relative to lanes. For deeper understanding of what is happening in the environment, the notion of maneuvers (like lane changing, transition to convoy driving, turning off onto a crossroad) or behavioral modes (like lane keeping, convoy driving) has to be available. The former ones are generated by typical feed-forward time histories for some control variables; the latter ones result from linking control output to some state variables (applying feedback control laws).

In the latest step towards cognitive (EMS-) vision, these behavioral capabilities are represented on two levels in the system: 1. For actual implementation, they are coded on the 4-D level (right part of level 2 in figure 1); this is done in the conventional control engineering approach. 2. The effect of the behavioral capability is represented on the decision level (top) as a state transition taking a certain amount of time; however, no dynamics are involved here. This corresponds to the usual AI-approach, but the lower level being triggered from the higher ones (in general) takes care of symbol grounding. Subjects may be characterized by their behavioral capabilities, both for perception, for decision making and for locomotion. Observing the onset of a typical maneuver, the system can now infer the likely continuation of it and can adjust its own behavior to it (defensive style of driving).

6. SITUATION ASSESSMENT

The integrated view of all objects of relevance in connection with the own mission plan yields the situation. This process is called 'situation assessment' and may have different aspects for the purpose of viewing direction (gaze) control [2f), 11] or locomotion control [2e), 2g), 13 - 15]. Situation aspects are assigned in a fuzzy manner [12] so that their combination allows good rules for behavior decision. Figure 3 symbolizes this approach; the center and left parts have been realized while the right part is an extension for future more versatile assistance systems (the left branch BDL then is passive).

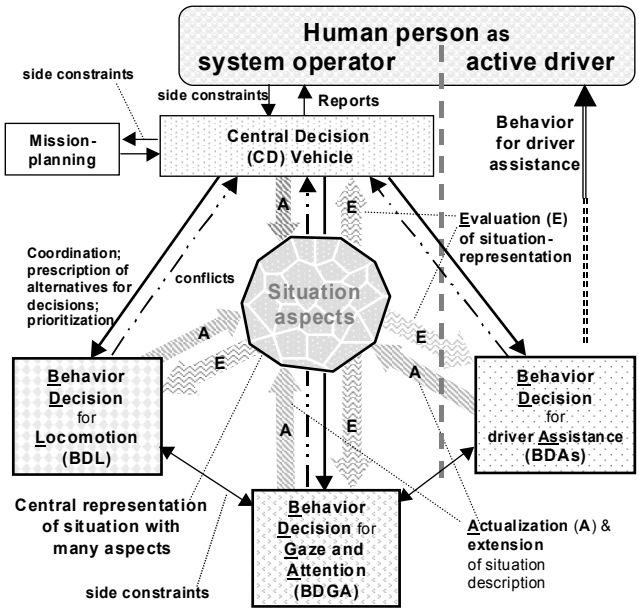


Figure 3: Hierarchically structured, distributed decision agents based on competencies: For visual perception (lower center), locomotion (lower left), driver assistance (lower right), and for overall systems aspects (center top); a common representation of the situation is desirable [after 12].

7. BEHAVIOR DECISION

A distinction is made between tactical and strategic decision making in order to avoid too large a single knowledge base and too heavy communication loads. There are a central unit for strategic decisions ('**C**entral **D**ecision' CD) concerning mission aspects, and two local decision units (three in the future).

Behavior Decision for Gaze and Attention (BDGA)

The purpose of gaze control is to achieve maximal information gain over time by splitting and sequencing attention according to the actual collection of objects / subjects of interest. BDGA collects all requirements with respect to vision from the perception specialists and figures out the best gaze program for achieving maximal information gain over time from all cameras together. This may include periodic saccades with time-slices of

attention on different objects or on different parts of a large object. The interpretation is done in a unified framework. When no new visual input is available (usually for a few tenths of a second during fast saccadic motion) predicted states are filled in exploiting spatio-temporal models. Demand of attention (potential information gain) is computed from the significance of the object / subject for mission performance and from its uncertainty in state [12]. The latter one increases over time during prediction phases.

Behavior Decision for Locomotion (BDL)

The goal of locomotion control is to safely achieve the mission by following the mission plan as good as possible and by avoiding obstacles whenever they occur. BDL performs all tactical decisions during mission performance, like transition into convoy driving mode when running up to a slower vehicle in front, or evading an obstacle by a local maneuver [2g), 8, 15]. When drawbacks for mission performance are likely or when conflicting requirements from BDGA and BDL occur, Central Decision is called and has to come up with harmonizing solutions [14].

Central Decision (CD)

CD is the agent with direct access to overall mission representation. It initiates mission planning and re-planning when the nominal list of mission elements, which is the reference for mission performance by BDL, has become outdated due to some unforeseen event. Monitoring of mission progress and of proper functioning of all system components as well as serving as interface to the human operator are the main tasks of CD (see central cylindrical top in figure 1).

8. REALIZATION OF BEHAVIORS

All behavioral capabilities of the system are represented internally on two separate levels. The 'Vehicle Computer' close to the actuators does the implementation on a lower level. Here, control - engineering methods predominate with control feed-forward and feedback loops. Parallel to this, the effects of these activities and all possible transitions depending on specific events are represented on the higher decision levels in explicit form (Harel state charts) [14]. This allows a clean separation between (quasi-static) 'Artificial Intelligence'- and (dynamic) control-engineering methods. Behavioral capabilities are grouped in a network according to a) the actuators involved and b) their level of sophistication [12, 13, 15].

Simple **skills** (bottom line in dotted area of figure 4), usually realized by parameterized control time histories or feedback control laws, form the basic layer. Stereotypical basic behaviors (central horizontal line in dotted area of figure 4) may consist of parallel or properly time-triggered sequences of simple skills. Before a behavior is activated, the decision level is informed about the actual state of availability of all components required. After activation it can monitor the progress of the maneuver instantiated and take corrective action when expectation and the real trajectory differ by more than a threshold value specified. This high-level supervision is an essential step towards reliable mission performance. The downward looking arrows in the upper right corner of figure 1 symbolize this realization of behavior decision and behavior implementation detailed in figure 4. One of the advantages of this separation is that high data rates to and from the upper levels can be avoided. Handling of usual perturbations is thus removed from the decision level.

9. MISSION PERFORMANCE

When a task is given, a specialist process taking the actually available behavioral capabilities into account generates alternative mission plans, triggered by CD, each consisting of a sequence of mission elements. CD then is in charge of selecting

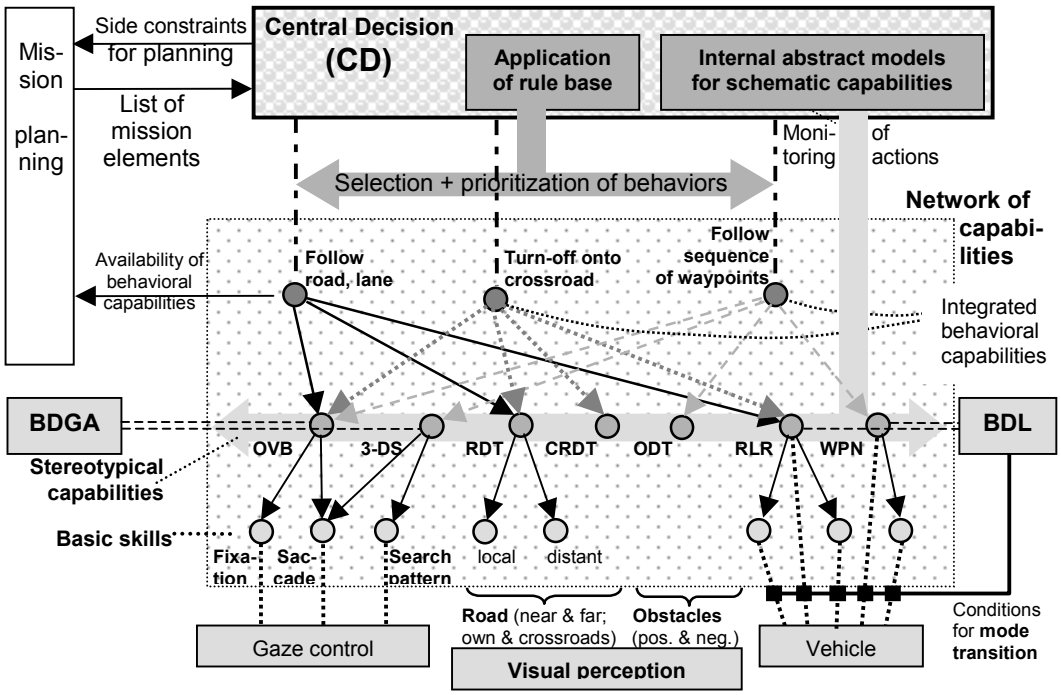


Figure 4: Activation, prioritization and monitoring of some stereotypical capabilities by Central Decision (CD) exploiting the capability network and special decision units for Gaze and Attention (BDGA) and for locomotion (BDL). Capabilities for perceiving roads and obstacles from the image stream delivered are provided by specialist processes for certain object classes (not detailed here).
Legend: **OVB** = Optimization of Viewing Behavior; **3-DS** = Search in 3-D space; **RDT** = Road (or lane) Detection and Tracking, this is achieved with different algorithms (basic skills) for the near and the far range; **CRDT** = CrossRoad Detection and Tracking; **ODT** = Obstacle Detection and Tracking, both stationary and moving, above the ground (positive) and missing support for the wheels (deep potholes and ditches = negative obstacles); **RLR** = Road (or Lane) Running (lateral guidance with appropriate speed; **WPN** = WayPoint Navigation when driving off the road (based on GPS). {For a complex maneuvering capability like turning-off onto a crossroad see [12].}

one according to the performance measures preferred. BDGA and BDL start implementing the plan after being triggered by CD.

Figure 4 shows the internal dependencies of complex behavioral capabilities on more basic ones (arrows) and their activation as they occur during dynamic mission performance. Note that optimization of complex maneuvers does not require individual re-coding but adjusting parameters and activation rules (triggering).

10. SYSTEM INTEGRATION

Figure 1 gives a coarse visualization of the overall system architecture. The dashed curved arrow from the left on the 4-D level (running through all specialist processes for object recognition towards control implementation) is intended to demonstrate the flexibility of the architecture with respect to behavior generation. The arrow is meant to indicate that any object of these classes can be assigned by 'Behavior Decision' as reference for feedback control loops in gaze or locomotion on the lower level. This allows very flexible behaviors like view fixation on an object in gaze control, road tracking and lane keeping, range estimation and distance keeping to the vehicle in front (convoy driving) or crossroad recognition and turning off onto it in locomotion control.

Systems with three to four cameras for vision have been implemented on a cluster of four DualPentium PC. Beside Ethernet for booting they were linked by a 'Scalable Coherent Interface' (SCI) with an effective data rate of close to 100 MB/s (upper bar in figure 5 across all four PC's). PC's 1 to 3 are devoted to image processing and visual perception from specific video data streams. The fourth PC (labeled 'behavior' PC to the right in figure 5) is connected to two subsystems running under hard real-time constraints. The 'gaze' subsystem implements inertial gaze stabilization (at 500 Hz) and active gaze control (at video rate of 25 Hz). The 'vehicle' subsystem is the interface computer to conventional sensors and actuators. In addition, a GPS-receiver is connected to the 'behavior'-PC which also serves as the Human-Machine Interface (HMI) for the operator. All other PC's of the system are handled through the embedded PC demon process EPC in each system.

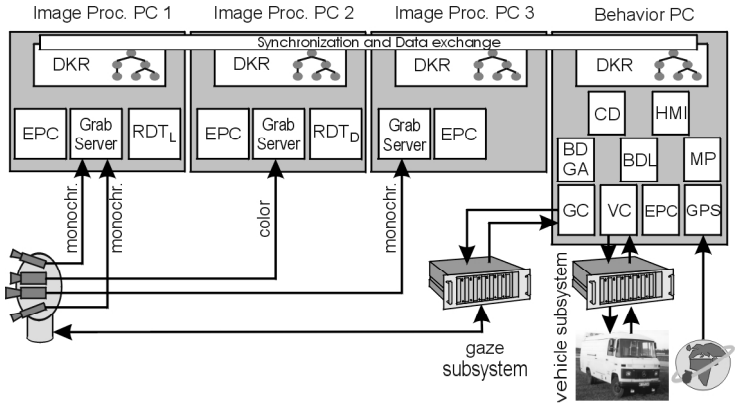


Figure 5: Hardware realization on a cluster of PC's plus two subsystems for hardware interfacing (gaze and vehicle control).

11. EXPERIMENTAL RESULTS

Experimental results in road vehicle guidance have been achieved with the two test vehicles **VaMoRs** (5-ton van) and **VaMP** (Mercedes 500 SEL) of UBM. Detecting a crossroad in a network of minor roads without lane markings and turning off onto it as well as convoy driving at normal highway speeds have been demonstrated.

Figure 6 shows test results from detecting, tracking and turning-off onto a crossroad. Figure 6a shows the yaw (pan) angle time history of the platform during this maneuver. From second ~ 90 to 110, a saccading maneuver in gaze is performed in order to alternatively collect data on the crossing (distance and speed of approach) and on the geometry of the crossroad (width and angle of intersection). In 6b the saccade bit is shown telling the

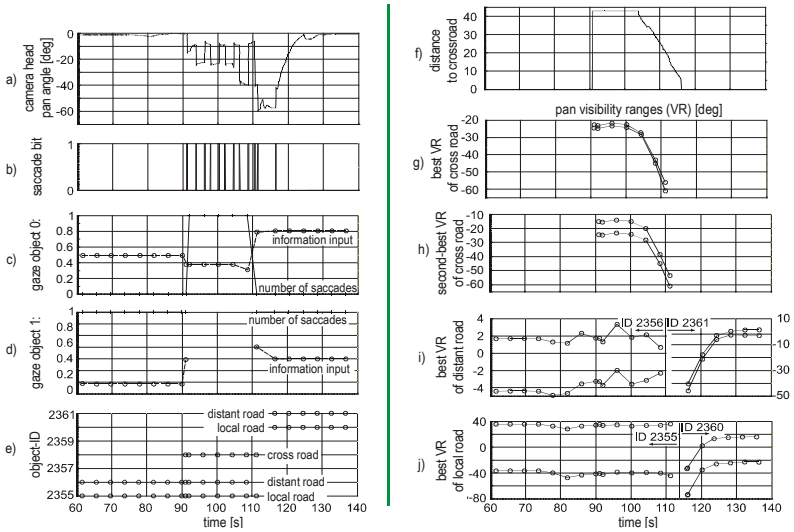


Figure 6: Detection, recognition and turn-off onto a crossroad with **VaMoRs** and **E**xpectation-based, **M**ulti-focal, **S**accadic (EMS-) vision.

image evaluation processes whether it makes sense to process images or whether the images are blurred and they should stick to prediction with spatio-temporal models (when the bit is up). From sub-figures 6c - e the objects observed can be seen. The crossroad is inserted into the scene tree at around 90 seconds and becomes the new reference road (split into local (near) and distant) at around 115 seconds (s). During approach of the intersection (6f), the gaze angle in yaw increases up to ~ 60° (6g, 6h), telling that the vehicle turns its viewing direction 'over the shoulder' while driving still straight ahead on the old road. At ~ 116 s reorganization of the scene tree is finished, with the old crossroad now being the new reference road. The gaze angle of MarVEye is constantly in direction of this new reference while the vehicle turns underneath it until at ~ 130 s gaze is almost in direction of the body longitudinal axis again (6i, 6j). (These figures show the best viewing ranges (VR) as evaluated by BDGA. Small offsets from zero may stem from the fact that the system is preparing for leaving the road and turning onto the grass surface; since only one boundary of the road can be tracked by the tele-camera at the range specified, the gaze angle selected is 2° (6i, right).

As final demonstration of the development contract, in October 2001 a coherent mission consisting of many arcs has been performed for an international audience:

1) crossroad recognition and turning off onto the left-hand side,
2) leaving the driveway to the right and entering grassy terrain for cross-country driving along a route fixed by several (virtual) way-points (GPS-coordinates),
3) while driving on uneven grassy ground, recognize a sealed road under ~ a right angle,
4) entering this road with an appropriate 90° turn to the left,
5) road following and crossing of another road,
6) turning off to the left and entering grassy terrain again,
7) detect a ditch (~ 0.8 m wide) sufficiently early for
8) stopping in front of it.

In the meantime, autonomous dodging of the ditch and driving around its near corner with gaze fixation of the corner has been demonstrated [9].

12. CONCLUSIONS

The general cognitive system architecture based on dynamic vision for motion control has proven efficient in an implementation on a cluster of standard DualPentium PC. Its implementation in C++ has brought about easy portability. Multi-focal scene recognition with active (including saccadic) vision has been realized for the first time in road vehicles.
Visual perception is partitioned into three stages: 1. for detection of visual features indicative of objects of interest, 2. hypothesis generation, tracking and relative state estimation for single objects (for n of those in parallel), and 3. recognition of maneuvers performed by subjects and inference of their intention(s). The results of stage 2 are collected in a dynamic object database (DOB) exploiting a scene tree representation with homogeneous coordinate transformations. This dynamic knowledge base underlying imagination of the actual situation is shared among all cognitive processes by a wide-bandwidth communication network (SCI).
A situation assessment process for central decisions, and other ones specialized for gaze and locomotion control, analyze the motion state including maneuvers under performance for all objects / subjects of relevance. This final visual perception stage does not rely on image data directly but looks at state variable time histories on a larger scale. Only with special type of knowledge on behavioral capabilities of subjects can cognitive entities as maneuvers or intentions be recognized. This is why situation assessment and own behavior decision are grouped together on this higher level relying on a similar knowledge base. Explicit representation of perceptual and behavioral capabilities has been introduced for this purpose. Behavioral capabilities have a dual representation with quite different aspects. The abstract (quasi-static) transitions effected by the capability serve as base for decision making on the higher level. Real-world implementation is done on the lower level with direct access to a broad data stream exploiting control-engineering methods (both feed-forward and feedback control components). The approach has been verified, among others, with the test vehicle **VaMoRs** in a small but complex mission.

13. REFRENCES

[1] Dickmanns E.D., Wuensche H.-J.: Dynamic Vision for Perception and Control of Motion. In: B. Jaehne, H. Haußenecker and P. Geißler (eds.) Handbook of Computer Vision and Applications, Vol. 3, Academic Press, 1999, pp 569-620

[2] Proceedings of Symposium on 'Intelligent Vehicles'00'. Dearborn, MI, USA, Oct. 2000, with the following contributions on EMS-Vision:
a) R. Gregor, M. Lützeler, M. Pellkofer, Siedersberger K.H., E. D. Dickmanns.: EMS-Vision: A Perceptual System for Autonomous Vehicles. pp. 52 – 57.
b) R. Gregor, E. D. Dickmanns.: EMS-Vision: Mission Performance on Road Networks. pp. 140 – 145.
c) U. Hofmann, A. Rieder, E.D. Dickmanns: EMS-Vision: An Application to Intelligent Cruise Control for High Speed Roads. pp. 468 – 473.
d) M. Lützeler, E.D. Dickmanns: EMS-Vision: Recognition of Intersections on Unmarked Road Networks. 302 – 307.
e) M. Maurer: Knowledge Representation for Flexible Automation of Land Vehicles. pp. 575 – 580.
f) M. Pellkofer, E.D. Dickmanns: EMS-Vision: Gaze Control in Autonomous Vehicles. pp. 296 – 301.
g) K.- H. Siedersberger, E. D. Dickmanns: EMS-Vision: Enhanced Abilities for Locomotion. pp. 146 – 151.

[3] Gregor, R., Lützeler, M., Pellkofer, M., Siedersberger, K.H. and Dickmanns, E.D.: EMS-Vision: A Perceptual System for Autonomous Vehicles. IEEE Trans. on Intelligent Transportation Systems, Vol.3, No.1, March 2002, pp. 48-59

[4] A. Rieder: Fahrzeuge sehen. PhD dissertation, UniBwM, LRT, 2001

[5] Proc. of the International Symposium on 'Intelligent Vehicles' starting 1992, (organized yearly by I. Masaki and various institutions)

[6] Albus J.S., Meystel A. M.: Engineering of Mind. – An introduction to the science of intelligent systems. J. Wiley & Sons Publication, New York, 2001, 411 pages.

[7] Dickmanns E.D.; Graefe V.: a) Dynamic monocular machine vision. Machine Vision and Appl., Springer International, Vol. 1, 1988, pp 223-240. b) Applications of dynamic monocular machine vision. (ibid), pp. 241-261.

[8] Siedersberger K.-H.; Pellkofer M., Lützeler M., Dickmanns E.D., Rieder A., Mandelbaum R., Bogoni I.: Combining EMS-Vision and Horopter Stereo for Obstacle Avoidance of Autonomous Vehicles. Proc. ICVS, Vancouver, July 2001

[9] Pellkofer M., Hofmann U., Dickmanns E.D.: Autonomous Cross Country Driving Using Active Vision. SPIE-Aero-Sense, Proc. 'Unmanned Ground Vehicles', Orlando, April 2003

[10] Dirk Dickmanns: Rahmensystem für visuelle Wahrnehmung veränderlicher Szenen durch Computer. Diss., UniBw Munich, Informatik, 1997.

[11] Pellkofer M., Lützeler M., Dickmanns E.D.: Interaction of Perception and Gaze Control in Autonomous Vehicles. Proc. SPIE: Intelligent Robots and Computer Vision XX; Oct. 2001, Newton, USA, pp 1-12

[12] M. Pellkofer: Verhaltensentscheidung für autonome Fahrzeuge mit Blickrichtungssteuerung. PhD. dissertation, UniBwM, LRT, 2003

[13] Pellkofer M., Dickmanns E.D.: Behavior Decision in Autonomous Vehicles. Proc. of the Int. Symp. on ‚Intell. Veh.'02', Versailles, June 2002

[14] M. Maurer: Flexible Automatisierung von Straßenfahrzeugen mit Rechnersehen. Fortschrittsberichte VDI Reihe 12, Nr. 443, 2001

[15] K.-H. Siedersberger: Komponenten zur automatischen Fahrzeugführung in sehenden (semi-) autonomen Fahrzeugen. PhD. dissertation, UniBwM, LRT, 2003