Maneuvers as Knowledge Elements for Vision and Control*

Ernst D. Dickmanns, *Member, GI*

Abstract— **Visual dynamic scene understanding requires recognition of both 3-D objects in motion and of the overall actual situation in the task domain given. This encompasses knowledge about the links between image features, 3-D objects capable of motion in 3-D space, and about situations. Treatment of these three distinct levels in parallel is achieved by early jumps to 4-D hypotheses on moving objects and the situation given. The system architecture resulting is superior to approaches using inverse perspective projection that start from differences in feature positions between two consecutive images and then proceed to objects in motion and to situations. This embedding in a much richer environment and in a closed-loop real-time fashion allows more efficient tests like prediction error feedback. It provides tools for realizing even complex integrated systems capable of handling strong perturbations with moderate computing power needed.**

Stereotypic classes of maneuvers for finite state transitions in appropriate time and corresponding knowledge elements for application under nominal and perturbed conditions are discussed as elements for mission performance. They represent parameterized knowledge about processes in certain task domains with typical environments and objects involved.

I. Introduction

In the early 1960's both digital processing of single images (snapshots) in a field that was to become 'Computer Science / Artificial Intelligence' (CS/AI), and recursive estimation in the field of 'Control Engineering' started in parallel with the introduction of digital microprocessing. Even today many people tend to see image sequence processing as evaluation of single images that allows the transition to understanding temporal processes by comparing results from two or more consecutive images evaluated separately. However, it has been shown that using so called 'dynamical models' combining spatial and temporal degrees of freedom in the form of differential / difference equations allows understanding of motion processes with reduced computer work load [1; 2].

In the field of computer vision the use of this approach with real-world dynamical models in image sequence processing has brought about a quantum jump in real-time performance in the mid 1980's [3; 4; 5; 6]. With less than a dozen conventional 16-bit microprocessors of 5 MHz clock rate the 5-ton test vehicle 'VaMoRs' in 1987 was able to run at speeds up to 60 mph on empty highways fully autonomously (both lateral and longitudinal guidance by computer vision at 12.5 Hz) [6; 7]. Competing vehicles around the world drove at walking speed with image evaluation rates one to two orders of magnitude lower. The key element allowing this level of performance was the use of prediction-error-feedback for inclined edge features in the real world. 3-D perspective projection was part of the nonlinear model for the measurement of edge features in the images of a running sequence. None of the previous images had to be stored. The model parameters of both the curved road observed and of the own state relative to it contained the information about the past needed, extracted previously.

It is essential that the dynamical models used represent objects in the real world, including their behavior over time (and not in some intermediate measurement space like the image plane). Our knowledge about the world is mainly geared to physical objects and object classes. Two very big super-classes of objects have to be distinguished if scene understanding on the semantic level is the goal: 1. Objects that are unable to initiate motion on their own (called here 'objects proper'), and 2. those objects that are able to sense information about the environment and that then activate some control output affecting their physical state; the latter ones will be dubbed 'subjects' here. All animals and robots fall into this category.

For understanding scenes including subjects it is mandatory to have knowledge available about how these subjects transform their measurement data into own behavior. If this triggering of behavior is not a fix program, like in humans, the closed-loop sequence of: sensing, behavior decision and acting is of importance. Since direct access to mental processes of subjects is not possible, the best approach is to try to grasp a subject's intention by observing onsets of maneuvers known in principle. This is possible only if typical maneuvers of members of the class of subjects observed are represented in the knowledge base of the observer; in the context of the situation given, the likely candidates for a maneuver just started have to be recognized from measurement data on their onset. In road vehicle guidance, typical such maneuvers are lane changes or turn-offs in the vicinity of a cross-road. In aircraft or helicopter guidance such maneuvers are *e.g.* change of flight direction by banking or phases of a landing approach near an airport.

Considerations with respect to system architecture for observing human actions have been performed for some time, but only in rather general terms ([8] and the first 8 references given there); the goal here is to discuss "maneuvers" as specific knowledge components for robotic systems based on well-defined mathematical models that can help realizing an efficient merger of CS/AI and 'Control Engineering' methods towards general cognitive real-time systems. In this generality this is a new contribution surpassing single examples as given before. Memory requirements are reduced to a minimum by paying attention mainly to the control variables in the process.

*Research supported by BMVg Germany.

Author is with the University of the Bundeswehr (UniBw) Munich, Germany (phone: +49 8104 1361; e-mail: edd@dyna-vision.de).

II. EFFICIENT REPRESENTATION OF SUBJECTS

To a large extent, knowledge about the world is linked to classes of subjects and to their individuals. Beside geometrical shape and body articulation these classes and the individuals are characterized by their capabilities of a) sensing, b) data processing and perception on higher mental levels, c) decision making in a situational context, and d) control actuation for achieving some goal.

As has been shown recently [9], more reliable visual perceptions and higher discrimination rates in complex scenes can be achieved by using bottom-up models (from features to objects) and top down models (scenes with objects) in parallel. The approach described there seems well suited to start tracking of individual members in the scene. In order to understand the semantics of what they are doing, however, it is necessary to have knowledge about the maneuvers performed and about the context these maneuvers are applied in, usually. This means that three levels should be represented and used in parallel:

1. **The visual feature level** with links to objects / subjects,
2. **The object / subject level** with:
   - 2a) Relative distribution of features on the surface of 3-D bodies,
   - 2b) 3-D shape and articulation of the body and its elements,
   - 2c) typical movements of limbs, head/neck and the trunk as part of maneuver elements for locomotion or some other goal; this means the use of dynamical models with state variables and control actuation schemes.
   - 2d) Typical goals of subjects in certain situations: hypothesize the most likely decisions and behaviors to be expected.
3. **The task domain on the situation level** containing typical environmental conditions (like geometry, lighting and weather) and the types of object- / subject-classes to be encountered (probably ranked according to the likelihood of their appearance).

The basic task of subjects is to come up with well-suited decisions for their own behavior given the conditions perceived. Thus, since deeper understanding of movements of subjects depends on the task domain and the situation, on the one side, and since visual recognition of subjects depends on sets of features and typical movements, on the other side, the whole range from features of objects / subjects to situations should be considered in parallel.

A. Visual features

Fig. 1 visualizes the ranges needed in parallel both in 3-D space (vertical) and in 1-D time (horizontal). All measurements are done at the point 'here and now' in the upper left corner of the matrix. For an image this means that primary feature detection can be done purely bottom-up without reference to previous images; only local neighborhoods in the image plane are taken into account. This yields features like: 1. local regions with non-planar intensity distributions (white regions in the synthetic image in the lower left corner of Fig. 1; the figure is built only from features extracted; no original pixel values are shown), 2. edge elements (red and green line segments there), 3. corner features (blue crosses), and 4. larger regions with homogeneous gray shading).

B. Objects / subjects in motion

A human observer looking at the synthetic image in Fig. 1 cannot but immediately recognize a three-lane road with heavy traffic. The gray regions in the lower part ('nearby') with typical 'lane markings' (both the white local regions as locations of non-planar gray value distribution and the red-colored edge elements within them, forming almost-straight longer line segments) enforce this interpretation. For an experienced human driver seven 3-D-objects on the road are readily detected: Three in the right neighboring lane (truck nearby, car, bus in front), two in the own lane (car behind a truck ahead), and two cars in the left neighboring lane. Eleven objects (1 road surface, 2 lane markings, 7 vehicles, and the sky region) can thus readily be hypothesized. The remaining features (essentially non-homogeneous small regions to the side of the road) have to be understood from feature flow over time. It usually takes three to seven video cycles to achieve a stable interpretation with small sums of squared prediction errors in these types of road scenes.

The additional degrees of freedom of subjects relative to 'objects proper' require that for scene understanding these objects and 'subjects' have to be treated differently. While for 'objects proper' knowledge about laws of motion is sufficient, for subjects the self-decided variation of movements is an additional degree of complexity for adequate perception and understanding of motion processes.

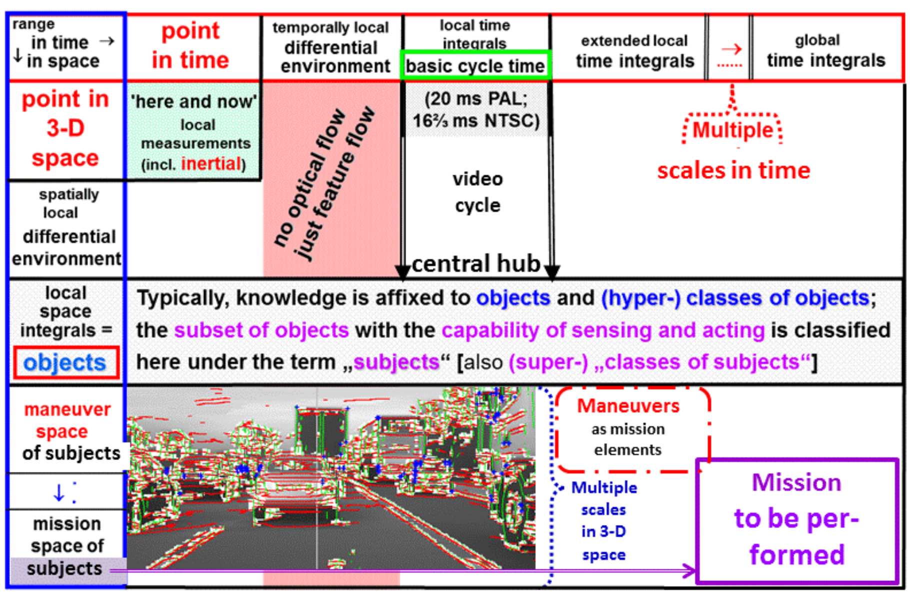

Figure 1. Various scales of 3-D space (vertical) and of 1-D time (horizontal) span the matrix for understanding complex scenes (here: 'Dense traffic on a highway', the situation shown by sets of edge, blob, and corner features in the lower left corner). All measurements are taken 'here and now' in the upper left corner. The central hub is the intersection of the 'object'-row with the column 'basic cycle time' (from video-imaging).

The distinction for dynamical systems between state variables and control variables, introduced by Caratheodory in the first half of the last century for treating optimal control problems, may be the key for better ‘understanding’ of movements of subjects. The following definition holds: State variables in a dynamical system are all those variables, the value of which cannot be changed at one point in time; they evolve over time with differential equations describing the constraints holding. If formulated properly, state variables contain all effects of the past. Only the actual state and the control variables − to be chosen freely (within limits) at each moment − determine its future development. It is interesting to note that the presence of control variables in dynamical systems is a precondition for developing a free will. If there is no such variable available in a system, its future development cannot be controlled (‘free will’ is thus meaningless).

On the other side, if there are no measurement data available (encompassing audio-, visual-, tactile, inertial- or other conventional sensor signals), there is no base for proper decision making with respect to behaviors in the (then unknown) environment given. Thus, the ‘sensing – acting loop closures’ are the driving factors for ‘mental’ data processing of subjects. Beside body shape and articulation as well as kinds of locomotion, it is the capability of sensing and data processing that determines the classes of subjects among animals / robots.

A rather direct link from sensed data to control actuation is dubbed a reflex in biological systems; it characterizes early ‘animals of lower order’ but continues to exist in some data paths of highly developed species. It is the idea of evolution that during its process more and more senses developed providing various data in parallel. Those animals that happened to use them in a fashion leading to superior results in the environment met, had better chances to survive, to generate more descendants and to spread. Combined use of data from separate paths must have been one important ingredient. This process continued (with interruptions) over hundreds of millions of years generating a multitude of classes of living beings. Most of them developed specialized organs for data combination: nervous systems and the brain. The sciences of biology and archaeology have established a well-founded theory of the tree of living beings by now.

C. *Situations in task domains*

A ‘situation’ is defined as the complete collection of all conditions relevant for decision making for a subject. It encompasses all relevant environmental conditions in the task domain, indoors or outdoors. In the latter case: Weather conditions, lighting- and visibility conditions, surface conditions for ground vehicles, local geometrical structure and buildings / objects / subjects in the vicinity. In all cases, the mission to be performed, timing conditions and own health state are of importance. It also makes a difference whether the mission is to be performed by oneself as the only acting subject towards this goal or whether it has to be done together with cooperative partners. With direct opponents in the field, the situation again is a different one.

All potential situations constitute such a tremendous volume that subdivision into specific task domains is mandatory. In human society, this is the reason for the many professions existing. The basic organizational structure for handling different task domains may be the same to a large extent. However, environments, objects / subjects likely to be met, as well as typical behaviors of subjects may vary widely.

Within each task domain there are characteristic missions to be performed; usually, each mission can be subdivided into a sequence of mission elements that can be treated with the same set of behavioral components. For example, in a transport mission on a road network the following mission elements will have to be performed: A) Get started from the actual parking position and move into the traffic flow on the local road; B) perform vehicle guidance on local roads with b1) lane following, b2) taking proper turn-offs, b3) get onto entrance of freeway. C) Cruise on multi-lane freeways: c1) Move into freeway traffic, c2) select lane according to both speed and turn-offs desired, c3) perform transition onto the freeway exit. D) Move into traffic flow on local road (similar to B), but with d3) move into final destination and park vehicle according to the facilities found.

For all these mission elements the capabilities for: 1) perception and scene understanding, 2) proper decision making, and 3) control actuation have to be available. There is also a need for evaluating the performance levels achieved and for keeping track of their changes over time under different environmental conditions (improvements / deteriorations). These values form the basis for adapting maneuver parameters and for selecting maneuvers in the future in accordance with the situation encountered. This constitutes learning of (dynamical) behavioral components.

Note that all these elements for mission performance are described as processes in temporal terms as usually done in control engineering. The nominal part of state transitions is specified by parameterized feed-forward control time histories; learning which one of these parameter sets should be used in which situations is what constitutes “experience in the field”. This experience allows recognizing snapshots as part of a process; on this basis, expectations can be derived that allow: a) focusing attention in feature extraction on special events (occlusion / uncovering of features in certain regions of next image); b) increased resolution in regions of the real world by gaze control for a multifocal system.

Crucial situation-dependent decisions have to be made for transitions between mission phases; here, switches between behavioral capabilities for maneuvers are required. That is why representation of specific knowledge for “maneuvers” is important (second block on the diagonal from the lower right corner in Fig. 1 and Fig. 2). Nominal maneuvers describe the transition process from the present state to the desired one disregarding perturbation effects. The latter ones are dealt with using feedback control superimposed, based on the situation actually encountered.

III. MANEUVERS AS 4-D-KNOWLEDGE ELEMENTS

It is important to realize that representing maneuvers as mental objects extended along the time axis with their own scale, introduces a new dimension for dynamic scene understanding. The state in a maneuver as element of a mission links the physical state at the moment ‘now’ first with the overall task context to achieve the desired goal (on a global time scale) and second with the momentary behavior (on a locally extended time scale). Perturbations actually encountered are counteracted steadily as best as possible taking actual systems dynamics into account. This may well be a step towards developing an initial kind of consciousness; temporal aspects on several scales are considered in parallel for the actual situation in the mission context.

Lane changes of cars may 'feel' quite different depending on the parameters of their suspension system and on the time scale of their performance. When slow changes are made with relatively long maneuver time, the wheels may be assumed as rigid (like for railways); however, when fast lane changes with short maneuver times are realized (amplitude 'A' large at high speed) the tires act like springs, most likely with nonlinear characteristics. In this case, a more refined dynamical model for the behavior of the vehicle has to be chosen taking the additional degrees of freedom into account. In this case, the order of the simple dynamical bicycle model increases from three to five; a more detailed discussion with figures demonstrating the changes in dynamic behavior may be found in section 3.4 of [7]. An even more realistic model with two parallel wheel tracks becomes much more complicated and can be handled numerically only.

Knowledge about maneuvers that allow realizing state transitions over extended periods of time are important both for planning own actions in mission performance and for understanding movements of other subjects of any kind. Maneuvers are specific to classes of subjects and even to individual subjects since they depend on body articulation and degrees of freedom actually available. Such a component for the transition from state $S_i(t_i)$ to $S_f(t_f$ + delta-t-dyn) contains for each maneuver (S_i to S_f) in a task domain the following information on a normalized time scale

$$\tau_i = (t - t_i) / (t_f - t_i). \quad (1)$$

1. The nominal control time histories $\underline{u}(\tau)$, for $0 \le \tau \le 1$ in parameterized form for the transition; note that τ may become > 1 depending on delta-t-dyn. (2a)
2. The dynamical model for generating the nominal paths of the state variables corresponding to $u(\tau)$. (2b)
3. The code for generating the coefficients of the feedback control laws suited for counteracting perturbations corresponding to 1. and 2. (above). (2c)
4. Conditions under which the maneuver may be used with which proper set of parameters. (2d)
5. Codes for evaluating one or more pay-off functions that allow judging the quality of the maneuver actually achieved. (2e)

This process-oriented approach geared to the control variables of dynamical systems is more efficient than centering on state variables. Fig. 2 shows a visualization of the central role maneuvers play for subjects as the link between the actual state (designated as 'central hub'), both of the own body and that of all relevant other subjects observed, and the overall mission; they all, together with the environmental conditions actually encountered, form the situation in the context of the mission to be performed in a certain task domain. It has to be kept in mind that the sensor signals all originate at the point 'here and now' (upper left corner in the matrix); however, the knowledge-based interpretation of evaluated image sequences becomes available only several video cycles later (up to several tenths of a second). When inertial data with almost no time delay and results from vision are interpreted in conjunction, this effect has to be taken into account; it can be easily handled if temporal representations of motion processes are used. [Note that in humans perturbed processing of these signals may lead to sea-sickness 'nausea'; for example, this is the case when experiencing unusual rotations temporally in parallel around several axes with different orientation, like in devices of fun parks.]

In the lower right corner of Fig. 2 the connection − diagonally downward − between maneuvers and the mission signifies the sequencing of the overall mission over extended periods of time into specific mission elements that can be handled with well-defined sets of capabilities for perception, situation assessment, for behavior decision, and application of feed-forward as well as feedback control.

The availability of payoff functions for maneuvers, for mission elements, and for the entire mission allows judging the performance achieved with the capabilities and the parameters used. Comparing these with results from previous applications with different parameters and different environmental conditions allows learning optimal parameter sets for certain conditions. This also provides a rich base for communication with other subjects dealing with the same field of application. Teaching newcomers is much enhanced when using maneuvers as knowledge elements.

Intent recognition of other subjects is possibly the most useful application of knowledge about maneuvers in traffic scenes since it allows more time for reaction. For example, looking with a high-resolution camera to the front wheel of the vehicle in the neighboring lane next to oneself provides the most possible lead time for recognizing the intention of making a lane change since it has to be initiated by turning the front wheel in the direction desired. This can be noticed before the body of the vehicle starts moving sideways. Tracking the lateral distance between the front wheel and the line perceived by fitting a curve through broken lane markings over some distance yields the best indication for a lane change maneuver intended. Note that in this simple case extended interpretations are made both over time (the

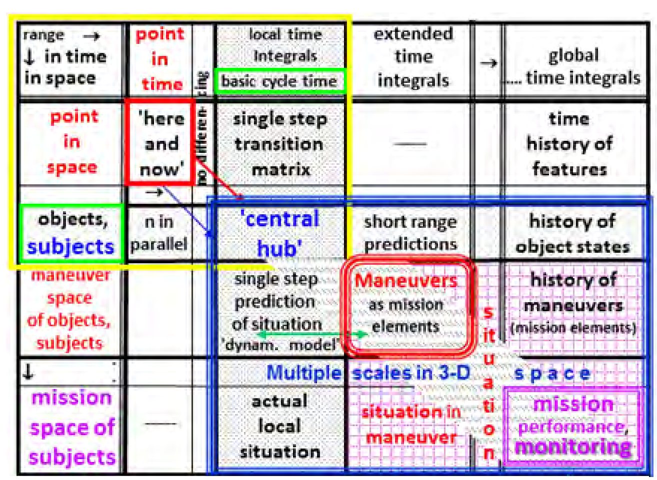

Figure 2. Temporally extended maneuvers, possibly consisting of several maneuver elements, together with cruise phases form the link between the actual state and the overall mission.

maneuver) and over space (the marked lane). This is typical for cognitive processes that they refer to abstract mental objects (here 'lane markings' in space and 'lane change' over time).

A. Maneuver elements for more complex maneuvers

In order to increase efficiency in representing maneuvers, standard elements for building more complex maneuvers are represented as separate units. For example, in road vehicle guidance a passing maneuver consists of several elements that are used in conjunction in reverse forms. The single pulse in steer rate shown in Fig. 3a) results in a change of curvature of the trajectory driven. Driving straight initially, a circular arc will be driven for $t > T_{SP}$, the radius of which depends on both the amplitude A and the duration T_{SP} of the maneuver element. If two opposite point-symmetric pulses are applied (Fig. 3b), the resulting change in the trajectory is a different constant heading angle since the steer angle lambda goes back to zero (dashed line). In Fig. 3c) two symmetric double pulses (with intermediate zero-input phases) are used to generate a lateral offset as needed for obstacle avoidance or for lane changes; the parameters a, T_{SP}, τ and T_D are used to find agreeable maneuvers depending on speed driven, time available and limits on trajectory characteristics like heading changes or maximum lateral acceleration (depending on the square of speed driven and on the maximum curvature of the trajectory as a function of lambda-max.

Passing a slower vehicle ahead requires two such lane changes, the second one in opposite direction after a safe lead margin has been achieved at higher speed. A difference in speed requires, of course, adapted maneuver parameters.

B. Experimental result with a complex maneuver

Unfortunately, these complex maneuvers with idealized elements almost never occur in real life with all its dirt effects like laterally hanging road surface, side-winds blowing, different friction on the two wheel tracks, or asymmetric deviations from a smooth surface (small potholes or small objects on the road etc.). In order to be able to counteract these perturbations immediately and effectively, the ideal trajectory corresponding to the parameters chosen is computed in parallel to the application of the control output. With the computing power available nowadays this is no problem at all. Essential state variables measured, like the heading rate and -direction as well as the lateral position relative to the ideal trajectory, are then used to compute compensatory feedback control components that force the actual trajectory towards the ideal one; these additive outputs are superimposed on the feed-forward components of the idealized maneuver. The advantage of this approach is that the actual state variables in connection with the dynamical model of the system allow computing feedback parameters that are comfortable and lead to desired eigen-behavior of the controlled system.

Fig. 4 shows experimental results achieved with the test vehicle 'VaMP', a large sedan; here, lane width is 3.5 meter, and the internal sub-phases with zero-control output are eliminated. Properly scaled, in the nominal case without perturbations this entire maneuver needs four numbers for its specification, the sequence of solid straight (green) rectangular lines in the upper left plot of Fig. 4. The steer angle as time integral of the green curve on the left is shown top right as dashed black straight line segments ('commanded value'). During this dynamic maneuver both the front and the rear hemisphere have to be observed whether the lane change maneuver continues to be safe. This means that up to half a dozen other subjects in the own and the neighboring lanes have to be tracked. For example, if a lane change is performed for passing a vehicle in front in the own lane, but then this vehicle also starts doing a lane change, momentarily the situation changes and the own passing maneuver has to be cut off. Handling a situation like this has been demonstrated by the UniBwM-group in 1994 [10]. The subject makes a decision to switch into a 'convoy-driving'-mode behind the vehicle now in front again in the lane changed to (with the goal state: same speed as the vehicle ahead, speed-dependent safe distance behind it).

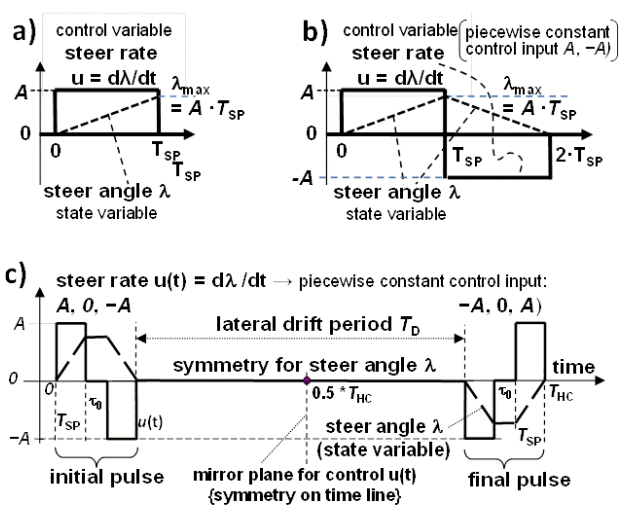

Figure 3. Maneuver elements for lateral guidance of road vehicles: a) Pulse in steer rate (as realistic control variable) generates a change in curvature of the trajectory. b) Opposite double pulse from zero results in a heading change. c) Two symmetric opposite double pulses with proper parameters result in a lane change (lateral offset).

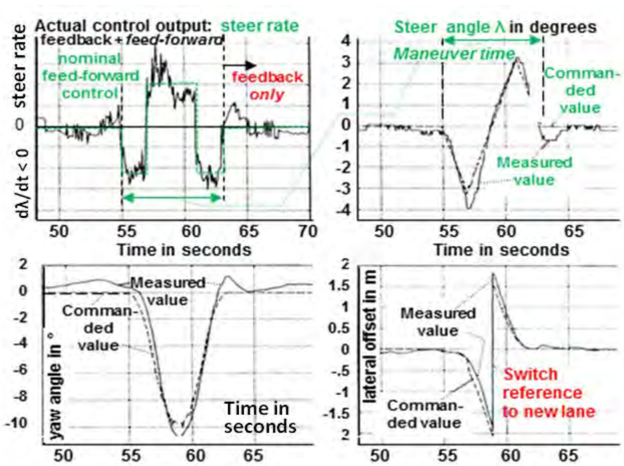

Figure 4. Experimental results with a large sedan of a lane change maneuver realized by feed-forward and feedback control of steer rate (steer angle λ = time integral).Figure 1. Various scales of 3-D space

Now the typical sequence of subtasks for initiating a lane change may be started again.

The noisy solid black line in the upper left of Fig. 4 shows the actually applied steer angle rate; the resulting steer angle as the time integral of this input is the solid black line top right. The yaw angle time history of the vehicle is shown in the lower left part, while the lateral offset of the center of gravity from the center of the relevant lane is shown in the lower right plot; the jump in reference value at the center of the maneuver is from -1.75 to + 1.75 m.

C. *Typical maneuvers*

For classes of vehicles there exist typical maneuvers as elements for transitioning between cruise phases. These (probably extended) cruise phases are performed with proper feedback control around pre-specified driving (flying) states.

Road vehicles (beside lane changes):

- Turnoffs onto a visually perceived crossroad [11] of unknown width and orientation, to the left and right.
- Starting and stopping at the side of the road.
- Obstacle avoidance by proper control of speed and lateral position [12], including emergency stops and negative obstacles [13]
- Merging into traffic flow on highways.
- Moving into and out of a parking lot at an oblique or a right angle ….etc.

Air vehicles: Fixed wing aircraft with separate propulsion systems and ailerons as well as rudder as control devices require quite different control activities from single- or multiple-rotor helicopters that use the rotor(s) also for pitch and bank control. Therefore, maneuvers for these types of aircraft are quite different.

Typical maneuvers for fixed wing aircraft are: a) Taxying to the start of the runway; b) assume proper position, orientation and parameter settings for the start. c) Accelerate to lift-off speed (brakes off and throttle position), d) turn in pitch angle (elevator control time history), adjust parameter settings (landing gear, flaps, throttle); e) make transition into steady climb state. f) Assume cruising conditions by proper acquisition of altitude, speed and flight direction; this may include banking maneuvers (control of ailerons and rudder). g) At waypoints adjust flight direction and flight altitude. h) Avoid unexpected obstacles by horizontal or vertical maneuvering, i) approach inclined trajectory for landing, j) horizontal and vertical transition into steady descent with proper parameter adjustments. k) Begin and perform flare, l) touch-down and parameter adjustments in all control variables, m) decelerate to taxying speed, and n) use taxiways to parking position. [13; 15]

Vision and 4-D-maneuvering for helicopters has been investigated since the end of last century at several places around the globe. [16] is one of the early approaches using 4-D models and recursive estimation for maneuvers like a) lift-off and transition to cruising. b) landmark navigation: following and flying relative to landmarks like road junctions, taxiways, runways and their special markings, e.g. b1) the large white stripe at their ends consisting of a sequence of white rectangles, or b2) the helicopter landing spot marked by the capital letter H. Space does not allow going into more details of this currently very active field of development. The concept of maneuvers as 4-D-knowledge elements centered around control time histories and proper feedback control laws with dynamical models for the motion process has proven very practical and effective in all cases (see [10], sections A.5 and A.6).

IV. Conclusion

Using parameterized finite control time histories for maneuvers in connection with realistic dynamical models as knowledge elements for representing actions of subjects yields a v ery efficient approach to control and to deeper understanding of motion: see (1) and (2a) to (2e). Species of subjects may be distinguished − beside 3-D shape and articulation − by their capabilities of perceiving and of reacting to activities observed.

References

[1] R. D. Kalman, “A new approach to linear filtering and prediction problems. Trans. ASME, Series D, Journal of Basic Engineering, 1960, pp. 35–45

[2] D. G. Luenberger, “Observing the state of a linear system. IEEE Trans. on Military Electronics 8, 1964, pp. 74–80; pp. 290–293

[3] H.G. Meissner and E.D. Dickmanns, “Control of an Unstable Plant by Computer Vision”. In T.S. Huang (ed): Image Sequence Processing and Dynamic Scene Analysis. Springer, Berlin, 1983, pp. 532-548

[4] E. D. Dickmanns and A. Zapp, “A Curvature-based Scheme for Improving Road Vehicle Guidance by Computer Vision”. In: 'Mobile Robots', SPIE Proc. Vol. 727, Cambridge, Mass., 1986, pp. 161-168

[5] H.-J. Wuensche, “Detection and Control of Mobile Robot Motion by Real-Time Computer Vision. In N. Marquino (ed): Advances in Intelligent Robotics Systems. Proc. of the SPIE, Vol. 727, 1986, pp. 100-109.

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

[7] E. D. Dickmanns. “Dynamic Vision for Perception and Control of Motion”. Springer-Verlag, 2007 (474 pp.)

[8] D. Vernon, G. Metta and G. Sandini, “A survey of Artificial Cognitive Systems . Implications for the Autonomous Development of Mental Capabilities in Computational Agents”. IEEE Trans. On Evolutionary Computation, Vol. 11, No.2, 2007, pp. 151-180

[9] J.C. Niebles, B. Han, Li Fei-Fei, “Efficient Extraction of Human Motion Volumes by Tracking”, 2012, Stanford Vision group, Comp.Sci; on their website

[10] www.dyna-vision.de: Section 3.1.3.04 (Lane change) and 3.1.3.06 (Obstacle avoidance); especially: Video: ‘Twofold lane change’, demo 1994 on Autoroute 1 near Paris.

[11] N. Müller, S. Baten. “Image Processing based Navigation with an Autonomous Car”. Int. Conf. on Intelligent Autonomous Systems, IAS-4, Karlsruhe, 1995, pp. 591-98

[12] F. Thomanek. „Visuelle Erkennung und Zustandsschätzung von mehreren Straßenfahrzeugen zur autonomen Fahrzeugführung“. Diss., UniBwM / LRT, 1996

[13] M. Pellkofer, M. Luetzeler, E.D Dickmanns. „Vertebrate-type Interaction of perception and gaze control for autonomous road vehicles.. In: Jarvis R.A. and Zelinski A.: Robotics Research, The Tenth International Symposium. Springer Verlag, 2003, pp. 271-288

[14] F.R Schell. Bordautonomer automatischer Landeanflug aufgrund bildhafter und inertialer Meßdatenauswertung, Diss., UniBwM, 1992

[15] S. Werner, S. Fürst et al. „A vision-based multi-sensor machine perception system for autonomous aircraft landing approach”. Enhanced and Synthetic Vision AeroSense '96, SPIE, Vol. 2736, Orlando, FL, April 1996, pp. 54-63

[16] S. Werner. „Maschinelle Wahrnehmung für den bordautonomen automatischen Hubschrauberflug“. Diss., UniBwM / LRT, 1997