Computer Vision in Road Vehicles – Chances and Problems

E.D. Dickmanns*

Abstract: A brief review of the development of vehicle and road systems is given discussing the functions of vision in this context. It is analyzed what services may be provided by computer vision systems of the future. These functions are grouped into three areas: 1. "Internalization" of actual traffic regulation and navigation status, 2. monitoring functions and 3. partly or fully autonomous control. For each group several exemplary proposals are discussed, leaving the technical problems of image sequence processing aside and focusing on man-machine interface problems. The flexibility of the future knowledge based monitoring and automatic control system using active vision will require special emphasis on cognitive ergonomy-engineering to achieve acceptance and to make the full potential of the system accessible to the driver for selecting the parameters according to his preferred style of driving.

ICTS-Symposium on
"Human Factors Technology for Next-Generation Transportation Vehicles"
Amalfi, Italy, June 16 - 20, 1986

---

*Dr.-Ing., Prof. for Control Engineering, Universität der Bundeswehr München;
W.Heisenberg-Weg 39, 8014 Neubiberg, FRG

1. Introduction

Today's road traffic systems are based almost exclusively on the human visual system for guidance and control. Before the advent of the automobile, the animal power source also provided some guidance function by its eye/brain-system. Horses, for example, as relatively intelligent animals, learned to move a certain distance from the (right) road boundary and even memorized the road system around their home. So they were able to find home even if the driver had fallen asleep. At road junctions they could take oral commands and proceed in the right direction without further control action by the driver. These, of course limited, autonomous guidance and control capabilities have been traded for higher performance and endurance when switching to a technical propulsion system onboard the vehicle.

The development of microminiaturized electronic devices like CCD-TV cameras and digital computers will allow the introduction of technical vision systems onboard the vehicle, which can take over monitoring, guidance and control functions in order to increase safety and/or to reduce workload from the driver. An early attempt is described in [1]. However, the cooperation between the human driver, who always will have to carry the burden of ultimate responsibility for actions taken in a mixed mode, and the automatic system has to be carefully designed and tuned. That is why this topic should be discussed and investigated in the near future even though the introduction of autonomous guidance systems is one or two decades away. Since vision requires intelligence for understanding what is being seen, this problem may be imbedded in the more general paradigm of cognitive ergonomy dealing with human and artificial intelligence interfaces.

Up to now, microprocessors in automatic subsystems of cars either work in isolated, well defined tasks like engine fuel flow control, anti-blocking systems (ABS, where they act as feedback devices to ensure the best possible braking action while the driver requests it) and speed regulation, or they may be part of modern gadgetry like voice output in non-vital areas. In an autonomous guidance system the computer has to take the initiative depending on his interpretation of what the state of its environment is. It will have to show competence, reliability and good manners in order to be accepted by the public.

To give you an idea of where we stand now with respect to computer vision for vehicle guidance a video film is presented showing real-time road boundary analysis on a German Autobahn and a closed loop guidance simulation with real image sequence processing hardware in the loop [2, 3].

The paper will discuss typical environments of different road systems and the roles which vision and scene understanding do play here. Then it is asked, what services may be provided by computer vision systems using knowledge bases on traffic situations. A wide range of options may become available in the long run. A sequence of steps will be discussed for gradual development and deployment, emphasizing the man machine interface as a crucial point for success and acceptance.

2. Road net standards

The different road system standards will be discussed in the sequence of their historical development with increasing structuredness. For introducing autonomous guidance functions by computer vision it is proposed to do this in the reverse order, starting with the most structured limited access highway (motorway, Autobahn) since it puts least demand on image processing and understanding.

2.1 Rural roads

They may range from unprepared but regularly used tracks to paved roads, usually without markings. Depending on weather conditions, they may well be unpassable. The surface has to be checked continuously for its state. Any type of vehicle, animal or obstacle may be expected on the road as well as almost any type of human activity. No standard can be expected.

2.2 Local and city roads

They usually do have a hard surface (e.g. paved, bitumen, concrete) and a centerline marking, if they are wide enough. Surface state checking may be widely reduced; however, almost any type of living being or object may be expected on the road. In cities or villages there is a close intermingling of pedestrian, bicycle and motorized traffic; therefore collision courses have to be checked even for erratic movements. Due to the wide range of environmental factors there may be many different traffic signs and even unusual (improvised) regulations and indications. Right of way is either coded in general rules or is posted by traffic signs including road markings like zebra stripes for pedestrians and by time varying traffic lights.

2.3 State roads (long distance), simple highways

Built for fast long distance travel, they usually have more than two lanes, limited curvature and early warning indicators for nonstandard situations like same level road crossings or sudden curvature changes. Traffic lights are usually sparse, and slow traffic often is prohibited. Road boundaries in general are well marked and guidance information for passing is coded in the type of lane marking. To alleviate road recognition in the night or under snow condition, reflecting poles are posted at defined distances.

In areas with relatively dense motorized traffic these roads are avoided by pedestrians and bicyclists, in remote cross country environment the situation may resemble that under 2.2.

2.4 High standard highways (e.g. Autobahn)

These roads for high speed (transcontinental) traffic are well structured. They
have to have the following properties by definition

a) at least two lanes of minimum width (e.g. 3,6 m) on separate tracks for each direction
b) limited access with ac-/deceleration lanes at entries/exits; only motorized vehicles able to drive at a minimum speed (e.g. 40 km/h) are admitted; therefore no rural traffic.
c) no same level road crossings, no traffic lights (except for non-standard situations (like construction), which have to be announced well in advance by traffic signs)
d) road curvature and slope are smaller than a defined maximum
e) well marked lanes, standardized traffic signs and early warning indicators, boundary markings by reflecting poles.

In regions with wild life activity the tracks often are separated by fences from the environment to reduce the probability of obstacles on the road.

3. The functions of vision in car driving

For the human driver, a distinction has to be made between the visual activities which support the guidance function and those which serve as man/ machine interface for information transfer about the state of the machine. The latter ones (like reading the speedometer, fuel flow, rpm-meter, engine and electrical system state indicators (oil pressure, temperature, electrical generator, battery) or warning devices for brakes, choke, fuel reserves etc.) may be discarded from a list of vision functions for an automated vehicle, since they would be input into the computer directly from the sensors.

So, a vision system for guidance support should provide the following functions:

3.1 Recognition of road and lanes and of the position of the camera relative to the road (lanes). This includes information on the road state a certain look-ahead-distance into the driving direction yielding road curvature, lane width and road surface state.

3.2 Indication of the velocity vector and the turn rate of the vehicle relative to the road; the driving direction (path azimuth angle or transversal speed component) as a signal containing lead information should be determined rather precisely,

3.3 Detection of obstacles, including motion extrapolation:
3.3.1 Obstacles on the lane (road): determine their size and both absolute and relative speed, predict relative trajectory in order to control the avoidance of collision.
3.3.2 Detect objects around the predicted egomotion-trajectory (lane) and check for collision avoidance, also for objects close to but off the road.

3.4 Continuous checking of the nearest other traffic participants for transmission of information: stopping lights, blinking for lane change, warning signals.

3.5 Pickup of local traffic regulation information and warnings:

3.5.1 Lane markings: interrupted or solid lines; both in parallel (to regulate lane changes); exits, change in width or number of lanes, parking lane.

3.5.2 Understanding of traffic signs:

- speed limits (general, lane - or vehicle specific)
- passing allowed/prohibited (passenger cars, trucks)
- lane information: merging, deviation, different standard (e.g. construction site), narrowing, approaching exit
- cautioning signals: flash lights, detour, curve, steep slope, slippery when wet, danger of ice/snow, road surface bumps, accident indication by triangle etc.
- traffic lights or regulations by humans at construction sites or "moving maintenance" (like grass cutting)

3.6 Reading of navigation information

- distance to cities and exits of highway, milestones
- directional signs for junctions (forks) or interconnections to different directions (road labels or target cities or general directions (e.g. west))
- early indications of junctions, exits, interconnections

Aside from these specific functions for vehicle guidance on well-structured roads, vision is able to provide information on the general state of the environment. Whether it is rainy or snowing can be detected directly; wind or storm may be inferred from watching certain objects (bending of twigs on trees, wind bags on bridges etc.).

The special importance of vision for car driving rests in the fact that it is literally possible to look into the future to points in space where interactions between the vehicle and the environment are going to happen soon. This provides some lead time for proper reaction; it increases safety margins and allows higher speeds.

4. What services can be provided by "intelligent" computer vision systems?

Until the functions listed in section 3 above can be autonomously and reliably carried out by onboard systems for general traffic situations, it certainly will need further development steps in computer systems and corresponding software. The very rapid current progress in VLSI-technology, however, shows promise, that already now research in this field may yield significant results for opening up flexible automation chances for the future. This may develop into systems for driver's assistance or even for partly or fully autonomous control. Many aspects will have to be checked by different authorities before such systems may eventually pass the examination for public introduction. The topics discussed below and their sequence are not intended to propose a specific mode of introduction. It certainly is too early for this.

They are meant to stimulate discussion and considerations, which of those (or others) might be useful functions and how their realization should be tackled. They are grouped according to three categories which will be discussed below: 1. "internalization" of the actual traffic regulation- and navigation-status, i.e. useful support functions for the driver, 2. monitoring functions about how the car is being driven and probably giving warnings depending on the environmental status. 3. partly or fully automatic control.

In any one of these categories the man/machine interface is of crucial importance and deserves special attention [4].

4.1 "Internalization" of actual traffic regulation- and navigation-status

In order to acquire a drivers license, a person - among other items - has to learn and store general rules and regulations, some of which are basic and always valid, while others are activated by traffic signs along the road for controlling driving behavior. These signs are posted at certain distances and in between their actual visibility the driver has to memorize them. One relatively simple function of a vision system could be to map the actual traffic signs on certain areas of a display within the car, so that the driver can always refer to them, e.g. whether passing is allowed or not. More involved internal representations may be convenient as will be discussed by some examples below.

4.1.1 Speedometer with actual speed limit display

In order to reduce instrument scanning by the driver it seems favorable to integrate the speed limit display into the speedometer, so that the difference between the actual speed, shown e.g. by the tip of the needle, and the actual limit can be grasped at one glance. This may be realized by a swiveling red region, the angular position of which is controlled by a computer according to the limit actually valid. The vision system in addition to reading numbers from the proper signs, has to understand implicit regulations as for example speed limits in cities or after crossing state borders. A change in the actual speed limit may be indicated by a short auditory signal.

4.1.2 Distance to go till exit or to a certain waypoint

As a navigational aid the computer vision system may read posted navigational information, screen it for relevance to the mission and use the relevant part for updating the navigation display. In this display the odometer output may be integrated to always show actual values.

4.1.3 Display of rearward view

In certain situations like backing up or preparing for lane change an unobstructed rearward view is desirable and the mirrors in use today are but a good price/performance compromise for partly achieving this goal. An additional rearward looking camera feeding a computer which, like an expert associate, continuously analyses the rearward scene and on request displays the results in symbolic form on a screen in the normal viewing range of the driver, would be of great help, especially if it is capable of estimating and indicating the relative velocity of other objects precisely.

4.2 Monitoring functions

Powerful monitoring functions become available by computer vision if the corresponding knowledge base checks the driving behavior of the driver against its own standards. There is of course a danger that the automatic system is designed using very conservative standards and that warnings may be given too often thereby becoming a. nuisance to the driver. It is proposed to introduce several threshold values which may be set by the driver according to his actual mood, however, always in a safe range.

Since computer vision also allows the detection of rain and snow, threshold values for warnings may be adapted automatically to the general state of the environment. Some exemplary functions are discussed in the sequel:

4.2.1 Violation of the actual speed limit: depending on the magnitude of the violation, the type or some parameters of the warning may be adapted: e.g. the region in the speedometer (see 4.1.1) indicating the actual speed limit may change color or blinking frequency in an electronic display. If voice output is available benevolent warnings may be uttered or the fine regulated by law may be quoted as a possible consequence.

4.2.2 Distance keeping to the vehicle in front: The distance to an object touching the ground can be estimated both by the apparent size, when the actual size is roughly known, and by the elevation angle to the point where it touches the ground, provided the camera altitude above the planar ground is known. This distance is checked against the relative braking distance for the actual speed and road surface state (friction coefficient), including reaction time delay. A warning sign may be given depending on a threshold value to be set by the driver.

4.2.3 Safety distance to vehicle behind: The rearward looking system mentioned in section 4.1.3 may do a similar evaluation as above for the vehicle following. If it closes in too much, either, the driver of the own vehicle may be warned to take some action, or the system may switch on blinking warning lights itself, (which in some countries, however, is not allowed).

4.2.4 Checking for lane changes: If the driver indicates that he intends to perform a lane change, the operations under 4.2.2 and 4.2.3 may be performed for the intended lane with the purpose of indicating to the driver the relative state of his vehicle and possibly drawing his attention to areas of conflict for a safe performance of the maneuver. If passing is not allowed, the system may remind the driver of this actual regulatory status by a proper output.

4.2.5 Absolute speed against visibility and road surface state: On high standard highways the mismatch between these opposing factors sometimes leads to accidents involving dozens or even hundreds of vehicles causing considerable death tolls, injuries and damage. An automatic system may be more neutral in assessing the danger when sharp curves or fog patches show up or when the surface state changes from dry to wet; automatic reading of outside temperatures may indicate the danger of ice on the road. Vision may detect the beginning of snow fall. All this information may be combined to compute a safe speed which may be utilized in connection with 4.2.1 above to warn the driver, overruling the general speed regulation.

4.2.6 Checking for collision courses: Using vision, a collision course can be detected very easily from image sequences by determining the displacement vectors of feature points on the circumference of an object. If the figure connecting the base points of the displacement vectors lies completely inside the figure connecting the tip points, the corresponding object is on collision course; future changes in speed and direction may, however, change the situation. This is also true in the opposite sense: Objects not on collision course may change their velocity vector, thereby suddenly being on collision course, e.g. human beings, animals; therefore, also objects in the vicinity of the driving lane (or road) have to be watched. If an object motion comes close to a collision course the driver's attention has to be directed towards this object by proper means (e.g. synthetic voice output, with loudness depending on the magnitude or immediateness of the danger of collision).

4.2.7 Quality of driving behavior: An automatic system able of driving a vehicle completely autonomously has all the means installed that allow it to check the driving behavior of a driver in a monitoring mode. Here, the computed control is not output to the actuators but is taken as a reference against which the driver's activities are checked; simultaneously the path performance relative to the lane centerline may be monitored. If there are erratic control inputs or the path performance deteriorates beyond threshold values, the system may issue warnings in several output channels (e.g. auditory, blinking warning lights (control wheel shaking (?)) etc.). Thus the driver may be kept from falling asleep or from continuing driving in his present dangerous mental state.

In order to keep the legal situation as clear as possible it is argued that no control action should be taken over by the automatic system in the monitoring mode. Since the driver had chosen to stay in command by selecting the monitoring and not the fully automatic driving mode, he has to carry full responsibility for his action. Similar to the flight recorder in modern transport aircraft, a "drive recorder" (encapsulated memory) should be installed to record the last seconds (minutes?) of a selected group of signals that allows experts to reconstruct what has happened shortly before the end of recording.

4.3 Partly or fully autonomous control

Systems for automatic speed control (speedomat, tempomat...) are on the market. They activate the engine throttle depending on the difference between commanded and measured speed; there is no checking for obstacles, other vehicles in front or for road curvature. The driver still has to take care of these situations.

A vision system can expand this primitive automatic mode in several directions; some possible steps are described below.

4.3.1 Constant speed lane keeping: Here, the vision system detects the lane markings and an added automatic steering control keeps the vehicle near the center of the lane. The driver has his hands free for other activities. The technology for this step may be considered well in hand, although developmental work remains to be done.

The driver still has to check for obstacles or other vehicles and has to intervene for speed adjustments necessary due to curves or slower traffic ahead. It is likely that this function is but an intermediate step to the following system.

4.3.2 Automatic lane driving: Here the vision system takes over the checking for obstacles or other vehicles in the lane and adjusts the engine throttle or activates braking, both in correspondence to the traffic situation (safe distance to vehicle ahead) and to road curvature. The driver for this part of the mission has become a passenger; however, cruising speed will settle down to the slowest vehicle in the lane if the driver does not take over control and guides the vehicle past the slow traffic. If situations occur which the automatic system cannot interpret, it may alert the driver and slow down to increase reaction time. If there is a static obstacle in the lane the system should bring the vehicle to a safe stop unless the driver takes over control,

4.3.3 Automatic driving with passing (lane change): In this mode the automatic control system handles both longitudinal and lateral control in agreement with its mission goals. If the traffic in the lane does not allow the commanded travel speed, in addition to 4.3.2 the system must be able to check the traffic status in a neighboring lane in order to safely initiate and perform an overtaking maneuver.

The normal traveling speed may be an input of the driver or the system may have enough knowledge about its performance capabilities so that it is able to translate global inputs like "best economic cruise" or "as fast as safely possible", into proper control action.

4.3.4 Autonomous navigation

In addition to 4.3.3 this system has a route planning and -checking capability that enables it to take the correct interconnections or directions at junctions in order to arrive at the destination point. It will need a data base on geographical relations and road networks. The actual navigation input at an interconnection has to be taken from traffic signs, since local routing may be opposite to the general long range direction. Proper lane changing is the essential control activity.

In a very advanced system, route selection for minimum travel time may be done using recent broadcasting inputs on traffic congestions or accidents; i.e. the data base has to be updated continuously while driving, using the local broadcasting services in an automatic mode.

In the long run, when enough experience on limited access highways has been gathered and when computer capabilities have grown further, autonomous systems may become able to even handle car driving on state or country roads. Infrared or Radar imaging sensors will possibly contribute to improve night- and all-weather-driving capabilities.

5. Problems to be solved in the field of man-machine interface

Aside from the technical problems with respect to image sequence understanding, which have to be solved in order to provide the capability of performing the vision functions listed in section 3, there are a number of man-machine interface problems to be solved in connection with the tasks described in 4.1.1 to 4.3.4 (see also [4]).

The "internalization"-function of section 4.1 requires display capabilities in the car which may be realized using color-CRT technology similar to the one nowadays available in modern aircraft electronic display systems for piloting and guidance. Important traffic regulation signs (like passing prohibited or allowed, speed limits etc.) have to be displayed in an area of good visibility with little scanning activity, while navigation information may be removed more to the periphery.

A legal question may arise if, because of some malfunction, the internal display does not correspond to the actual outside situation and the driver relies on the internal display causing some damage. He will hold the manufacturer responsible while police and judges may stick to the responsibility of the driver. Procedures for calibration and checking at the beginning of each mission may contribute to improving dependability and confidence.

While these functions require almost no human operator input, this is different for the monitoring function of section 4.2: Threshold values have to be set and levels of warning activities should be adaptable to the driver and his actual mood. On request, the system should be able of giving reasons for some general type of warning (like self-inferred absolute speed limit 4.2.5). In addition, the results of complex evaluations like checking for lane change (4.2.4) or collision courses (4.2.6) should be conveyed to the driver in an efficient way directing his attention immediately to critical areas and indicating the type of danger eventually discovered.

In a learning mode for the driver, the system should be able to introduce its user into the services offered and the control options available; this may be done using a menu technique in conjunction with the on board multi-function display. Since the spectrum of possible drivers and their intellectual capabilities will be rather broad, the problems of cognitive ergonomy may be hard. Possibly several optional standards providing different degrees of flexibility have to be offered. Most of the monitoring functions should be selectable individually with the choice of auditory or visual output adaptable by parameter adjustment in a certain range. The acceptance by the driver will depend on the degree of influence he or she can bring to bear in order to adapt the system to the personally preferred style of interaction.

In the autonomous control mode (section 4.3) the engage/disengage procedure has to be carefully developed; it may be necessary to provide extra training for the driver and to possibly introduce additional type ratings like for aircraft autopilots.

Functions like 4.3.1 and 4.3.2 are activated when the vehicle is in a nominal traveling state, while automatic driving according to 4.3.3 and .4 may be initiated either when driving or starting from a parking state. The initialization procedure should guide the driver by presenting menues for the selection of options depending on the state of the vehicle. The automatic system has to adapt its parameters in order to achieve a smooth transition; only after a complete self-check with positive outcome a "system ready"-sign should allow the driver to transfer control to the autonomous system. The automatic checking procedure must be safe against spurious commands but should be able to recognize meaningful ones even when stated in a slightly incorrect way; misunderstandings may be excluded by repeating the interpretation chosen and waiting for confirmation.

Disengagement of the autopilot should be possible either by activating the proper control switch or by just taking over control with a force beyond some threshold level (to eliminate the effects of spurious touching).

In the development of such a system special attention has to be paid to the wide variety of possible errors with which the system may be confronted by unskilled users. In view of the large computer power and the knowledge base necessary to realize the vision functions it can be expected that such an "intelligent" automation system will also be able to politely and helpfully react to repeated faulty usage. In this field of cognitive ergonomy the contribution of psychology to systems design becomes of increasing importance. Easy to handle help-functions may be crucial to long range acceptance.-

Besides simple performance parameters like desired cruising speed these systems may be able to accept complex statements to determine the actual control: e.g. "go as fast as safely possible" or "go at best economic cruise speed" or a mixture of both. In the first case the system will select speed always at the upper limit of the allowed or safe range, while in the second it exploits its knowledge about engine fuel consumption and adjusts both acceleration rate and cruise speed to the vehicle characteristics. In the mixed mode a specification on a percentage base may be convenient. Travelling comfort may be adjusted by some "sportivity"-parameters like acceleration levels, both longitudinally and laterally. The former ones may vary the style of ac- and deceleration (say max V = 0.1 to 0.5 g) and the latter one determines the speed on curvy roads by predicting centrifugal acceleration as a function of road curvature and speed (say a = 0.05 to 0.2 g).

6. Conclusions

Highly parallel VLSI computer systems of the future in conjunction with digital real time image sequence analysis and knowledge based control methods do offer a chance for very flexible automation systems for guiding vehicles along well structures limited access standard highways. They even may show the potential for autonomous mobility on less well defined road systems.

Automatic lane keeping and speed control on curved roads at speeds up to 60 km/h with a real CCD-TV-camera as sensor and rather simple image processing microcomputer hardware has been demonstrated in a simulation facility; real time lane detection and tracking of lane markings on an "Autobahn" was shown to work for speeds up to 100 km/h (speed limit of vehicle used)$_0$ It seems that speed itself is not the limiting factor for automatic visual guidance but rather scene complexity, provided the world model for data interpretation and control is chosen properly.

The functions of vision in vehicle guidance have been discussed and grouped into three areas: 1. "Internalization" of the actual traffic regulation- and navigation status, 2. monitoring functions (warnings) and 3. active control determination. In each area some specific proposals have been made, certainly not exhausting the potential. The man-machine interface has been emphasized as a crucial point for acceptance.
Even though mass production systems may be more than one vehicle generation away it seems timely to start research into this application on a broader basis now. These systems do have a tremendous growth potential and the great advantage, that for their introduction they do not need any additional installations along the roads but take full advantage of the installations provided for the human driver, thereby allowing gradual deployment and mixed traffic modes without any interference with the present system.

7. Literature

[1]-Tsugawa, S.; et al: "An Automobile with Artificial Intelligence". 6th IJCAI, pp.893-895, Tokyo, Japan, 1979.

[2] Dickmanns, E.D.; Zapp, A.; Otto, K.-D.: "Ein Simulationskreis zur Entwicklung einer automatischen Fahrzeugführung mit bildhaften und inertialen Signalen". ASIM 1984, Wien (Sept. 84). Informatik Fachberichte Nr. 85, Springer-Verlag, pp.554-558.

[3] Dickmanns, E.D.; Zapp, A.: "Guiding Land Vehicles Along Roadways by Computer Vision". AFCET-Conf.: "Automatique 85 - Tools for Tomorrow", Toulouse, Oct. 1985, Proceedings pp.233-244.

[4] Madni, A.M.; Freedy, A.: "Intelligent Interfaces for Human Control of Advanced Automation and Smart Systems". IEEE-EASCON 84 Proceedings, 17"th Annual Electronics and Aerospace Conference, Washington D.C., Sept. 1984, pp.227-232.