



Most traffic accidents in cities occur at intersections. It's therefore no surprise that researchers at DaimlerChrysler are working on supporting drivers at these high-risk locations. To do so, they are employing various methods aimed at achieving their long-term goal of creating an urban assistant.

Objekte mit der Kamera erkennen gehört in der Automation inzwischen zum Standard. Soll die Bildverarbeitung jedoch Autofahrer unterstützen, wird sie zur Königsdisziplin. In Echtzeit muss sie aus dem Straßenbild die richtigen Objekte herausfiltern. Lesen Sie, wie das geht.

Intersections are the most critical locations with regard to urban traffic, accounting for 60 percent of all accidents in inner cities. Almost half of the accidents at intersections with traffic lights are caused by motorists driving through red lights or ignoring the right of way. Moreover, incorrect interpretations of who has the right of way account for 95 percent of all accidents at intersections with traffic signs. In view of these figures, Stefan Hahn, head of DaimlerChrysler's "Machine Perception" research department in Ulm, will not rest until something has been done about the problem. His team is therefore specifically concerned with city traffic and is looking into assistance systems that can quickly warn drivers of imminent danger, thus helping to prevent accidents from occurring.

Hahn points out that the focus of his work is on computer vision. The reason for this is that despite the relatively low speeds in city traffic compared to highway driving, urban situations tend to be far more complex. "Radar sensors can only be used to a limited degree for what we're trying to accomplish." Hahn's team is therefore working with video sensors – cameras – that provide color or black-and-white pictures and are operated individually or in pairs as stereo systems. However, a sensor image is of no use by itself. Not until the information contained in the image is analyzed and interpreted can imminent danger be recognized. The vehicle's computer has very little time to accomplish this task, however – real-time process-

ing is required here. Luckily, big advances have been made in this sector in the last few years. 20 years ago engineers had for the first time managed to develop image processing software that could recognize movement with the help of cameras. However, it took the computers an entire weekend to analyze a single second of footage. Today's image recognition systems, on the other hand, require only 80 milliseconds to recognize the head of a child that suddenly runs into the street from between two parked cars. This

<i>accomplish, to</i>	<i>erreichen, erfüllen</i>
<i>account for, to</i>	<i>verantwortlich sein</i>
<i>achieve, to</i>	<i>erreichen</i>
<i>advance</i>	<i>Fortschritt</i>
<i>aim, to</i>	<i>abzielen</i>
<i>due to</i>	<i>aufgrund</i>
<i>effort</i>	<i>Anstrengung</i>
<i>employ, to</i>	<i>anwenden</i>
<i>extent</i>	<i>Ausmaß, Umfang</i>
<i>footage</i>	<i>Film(material)</i>
<i>head, to</i>	<i>zusteuern</i>
<i>imminent</i>	<i>unmittelbar bevorstehend</i>
<i>intersection</i>	<i>Kreuzung</i>
<i>machine perception</i>	<i>Maschinelle Wahrnehmung</i>
<i>occur, to</i>	<i>passieren</i>
<i>pattern</i>	<i>Muster</i>
<i>spatial</i>	<i>räumlich</i>
<i>urban</i>	<i>städtisch</i>
<i>utilization</i>	<i>Gebrauch, Verwendung</i>



Stereo cameras are used to measure a vehicle's distance to objects. Close objects are marked red, distant objects are colored green. Only one camera is used to calculate movement from frame to frame. The movement vectors are shown as colored lines.

(Bilder: DaimlerChrysler)

short processing time is only partially *due to* the enormous increase in computing power – it is just as much a result of new types of image processing algorithms, which DaimlerChrysler researchers played a very important role in developing. And this is exactly the area in which Hahn's team is focussing its *efforts*.

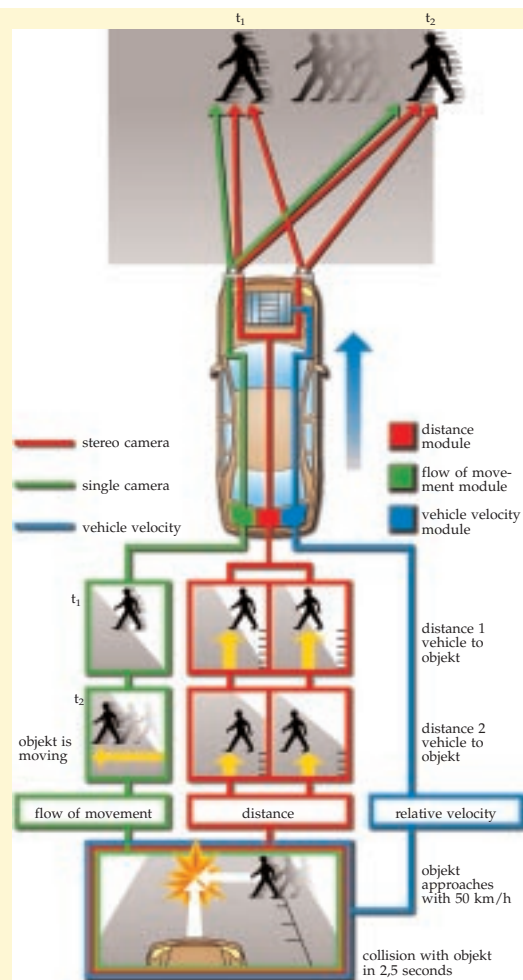
Machine perception is a three-part process. The first step is object detection: All of the approximately ten million pixels that sensor images provide to the computer each second potentially belong to some object. If the computer and software were to try to completely process all of these pixels, it would take virtually an eternity to analyze the

Something's moving

The research team at DaimlerChrysler combined two computing methods to make it possible to track moving objects quickly and accurately.

The first method, which involves measuring the flow of movement, requires the *utilization* of only a single camera. In a sequence of images, the light reflected by any moving objects always appears at different pixel locations. Through a comparison of the *patterns* in the images, the program calculates the object's movement vector and then informs the computer that something is moving and also tells it in which direction the object is *headed*.

Using the *spatial* image created by a stereo camera, it is possible to precisely calculate the distance between the object and the vehicle as well as the *extent* to which this distance changes from image to image. Comparing the distance between two successive images allows the relative velocity between vehicle and object to be directly measured.



In city traffic, huge amounts of information need to be processed and potential dangers assessed at lightning speeds.



images. However, not all objects are equally important for traffic situations. That's why the system does the same thing that humans do: It *perceives* its surroundings in a selective manner and filters out unimportant data. This is exactly what the next two steps of the image processing procedure are designed to do.

When monitoring traffic, it is particularly important to notice movement by people or objects. This requires that the objects in question be followed over time or *tracked*. When tracking an object, the researchers take advantage of the fact that the position of a car or pedestrian does not suddenly change from one video image to the next. The objects have a certain "*inertia*", which means the direction and speed of their movement rarely changes abruptly. Hahn's team takes advantage of this "law of inertia" to accelerate image processing by letting the software calculate the position of the pixels at which a moving object should appear if it moved at the same relative *velocity* and in the same direction as in the immediately *preceding* image. In addition, the system expands the computing process to cover the areas surrounding the pixel position at which the object was last detected. To put it simply, the software greatly accelerates the detection process by focusing on a relatively small area of the image that follows the one in which the object was last registered.

In monitoring movements, the algorithms operate according to the "condensation tracking" principle. "The software doesn't make instant decisions, but instead *assigns* the same plausibility to various hypotheses concerning changes in the object's position

from one image to the next," Hahn explains. "However, although they're initially treated in the same way, the brightness of most of the potential object locations begins to *conform* less and less to that of the tracked object as the film progresses." The software therefore serves to *separate the wheat from the chaff*, since the less *likely* a position hypothesis becomes, the more likely it is to be rejected. The computer will then begin tracking the pixel area in which the brightness values are most likely to *match* those of the object.

But it is not enough to simply track the object, as this does not *provide* any information about what kind of object it is. Here, an immediate reaction by human beings – recognizing an object as a traffic light or car – must be learned by the electronic image processing system with great effort. Not until it has *thoroughly* mastered this process can it accomplish the third step – classification. To *enable* the computer to recognize stop signs, for example, it is shown up to 10,000 digitized images of such a sign. These are images made under various lighting conditions, by day and by night, and from all *imaginable* perspectives.

The problem with this method is that in order to ensure that the computer knows where to find the stop sign and can also recognize it, the signs must be manually highlighted with the help of the PC mouse – a very *tedious* business. This is why Hahn's team is cooperating with a research group headed by Prof. Tomaso Poggio from the Massachusetts Institute of Technology (MIT) in Boston. The two teams are working on getting computers to learn on their own. In a sort of *preliminary course*, the researchers load

a set of manually highlighted images of the object to be classified – the basic set – into the computer’s memory. Although this process is similar to the old method, only about 300 images are needed now. Thanks to special algorithms that the transatlantic team developed and plans to test in coming years, the computer should be able to classify most of the other, unmarked images of the objects. If the system is unsure about how to classify certain images, it asks the human instructor. In the end, the same rate of recognition is achieved as with the older method, but with only a *fraction* of the human interaction.

Anyone visiting Hahn in Ulm quickly realizes that the *long-term goal* of creating an urban assistant can only be achieved if a combination of new hardware and software systems are used. For Hahn, however, it will be well worth the effort: “After all, there’s more *at stake* here than just *fender-benders* and *property damage*.” ■

<i>assign, to</i>	<i>zuteilen, zuschreiben</i>
<i>at stake</i>	<i>auf dem Spiel</i>
<i>conform, to</i>	<i>entsprechen</i>
<i>enable, to</i>	<i>ermöglichen</i>
<i>fender-bender</i>	<i>verbeulte Stoßstange</i>
<i>fraction</i>	<i>Bruchteil</i>
<i>imaginable</i>	<i>erdenklich</i>
<i>inertia</i>	<i>Trägheit</i>
<i>likely</i>	<i>wahrscheinlich</i>
<i>long-term goal</i>	<i>Fernziel</i>
<i>manner</i>	<i>Art</i>
<i>match, to</i>	<i>entsprechen</i>
<i>perceive, to</i>	<i>wahrnehmen</i>
<i>preceding</i>	<i>vorhergehend</i>
<i>preliminary course</i>	<i>Anfangs-, Vorkurs</i>
<i>property damage</i>	<i>Sachschaden</i>
<i>provide, to</i>	<i>zur Verfügung stellen</i>
<i>separate the wheat from the chaff</i>	<i>die Spreu vom Weizen trennen</i>
<i>surroundings</i>	<i>Umgebung</i>
<i>tedious</i>	<i>langwierig</i>
<i>thoroughly</i>	<i>gründlich</i>
<i>track, to</i>	<i>verfolgen</i>
<i>velocity</i>	<i>Geschwindigkeit</i>

Dieser Artikel erschien ursprünglich im Hightech Report, Ausgabe 2/2002, und wird hier mit freundlicher Genehmigung der DaimlerChrysler AG wiedergegeben.