From Neuroscience to Mechatronics

Fabian Diewald

19th April 2006

Contents

Over millions of years nature has developed great techniques in solving a large variety of problems. A typical example for such a technique, which was adopted by science, is the lotus effect as an ingenious method for the self-cleaning of surfaces. But in this paper the focus is set on a much more complicated phenomenon: the human brain. The ability to solve a large variety of problems much better than a classic computer makes it obvious that we can learn a lot from the human brain and use these advantages for technical purposes. Here, one example is presented in particular – the vestibulo-ocular reflex whose technical application is the key to affordable driver-assistance systems. Moreover, the problem of face recognition is presented.

1 Introduction

The major advantage the human brain has in contrast to classic computer systems is the ability to learn an to adapt to changes by itself. Of course, this can be very useful to react to changes in the environment of the system, but also inexact production of sensors can be compensated, which allows using cheap sensors and components.

2 Architecture of the human brain

The elementary components of our central nerve system are the neurons. The inputs of these cells are the so-called dendrites, the output is the bouton which contacts other neurons or muscles. Consequently a network structure is possible – the neural network. Transmission of information from neuron to neuron works electrically – the information is encoded in changes of electrical potential. The form of the signals is quite similar to pulse rate modulation known from technical applications. Data transfer in the human brain works by using short impulses with fixed intensity, and the information is contained in the frequency of these signals, not in the intensity. If we have a closer look on the synapses, the links between the neurons, we will find the key to the human ability of adaption and learning. Normally, synapses are gaps against electrical transmission – transmission is controlled by chemicals, the so-called neurotransmitters. Consequently, the intensity of the information float over this gap can be influenced permanently according to the amount of neurotransmitters.

3 The cerebellum – responsible for motorical issues

The cerebellum – which means ”little brain” – is the part of the human brain which is responsible for accurate movements. This fact is evident as people with a damage of the cerebellum show similar difficulties in accurate movements like after the consumption of alcohol. Consequently the instructions by the forebrain are insufficient – they have to be ”translated” by the cerebellum in a very special way. All in all, the cerebellum is an essential part in learning motor skills.

4 The cerebellar cortex – the surface layer of the cerebellum

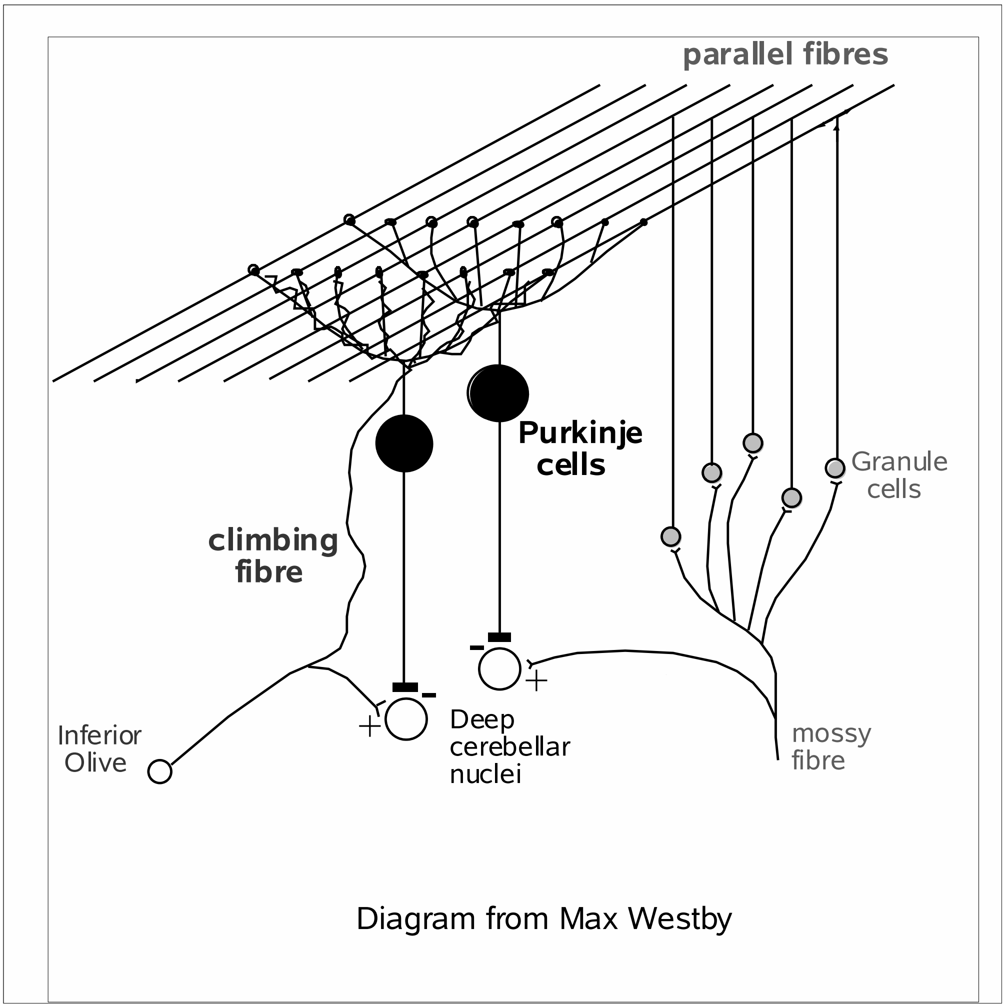

Diagram from Max Westby

Figure 1: Structure of the cerebellar cortex

A very important aspect in copying neuroscience is the need to understand the information input of the cerebellum over the cerebellar cortex. The information input is transferred over the so-called mossy fibres leading to the granule cells. One single mossy fibre leads to about 20 granule cells. The axes of the granule cells are called parallel fibres. These axes synapse on the Purkinje cells – the ”central” cells of the cerebellar cortex. One Purkinje cell has about 200 000 synapses. These synapses – a dendritic tree – are wrapped around by the so-called climbing fibre, only one for each Purkinje cell. The climbing fibres are the axons of the inferior olive and contain a ”teaching signal” influencing the synapses of the Purkinje cells. Consequently they are responsible for controlling the intensity of certain inputs to the Purkinje cells and thus for the ability of learning. This principle was twice independently proposed by David Marr in 1969 and James S. Albus in 1971, and is therefore called the "Influential Theory of Marr-Albus". But in 1982 Marr himself expressed his doubt about the possibility to gain technical benefits from his own theory:

> In my own case, the cerebellar study (...) disappointed me, because even if the theory was correct, it did not enlighten one about the motor system — it did not, for example, tell one how to go about programming a mechanical arm.

Obviously, finding an algorithm simulating cerebellar functions is a rather difficult field.

5 The vestibulo-ocular reflex

If we stare at an object in our environment, we maintain a stable picture of this object although our head is moved. Essential for this function is the vestibular organ, situated in our inner ear and detecting the motion of our head. This motion information is more or less directly transmitted to the extra-ocular eye muscles, which results in extremely brief processing delays of about 5-10ms. Picture stabilization only by optical means would lead to processing delays of about 50-100ms, which is far too slow. Nevertheless also visual information is needed for the vestibulo-ocular reflex as a learning signal as only the remaining optical flow on the retina is a criterion for errors of eye muscle movements. The excellent learning abilities of the vestibulo-ocular reflex (VOR) has been proved in experimentations with monkeys. After magnifying or miniaturizing glasses had been put on monkeys' eyes, the animals showed problems concerning the motion of their head as the glasses caused abnormal image motion speeds. After some days, these problems did not exist any more. Obviously, the monkeys adapted to the different correlation between head angular rates and optical flow. After having taken the glasses away, the monkeys showed similar problems again until their movements got normal after some days. As the VOR is used unconsciously, a motoric task and calibration is needed, its function seems to be situated in the cerebellum.

The VOR seems to work with two different types of paths of information: a direct, very quick path from the vestibular organ to the extra-ocular eye muscles and an adaptive path from the vestibular organ and the eye as well over the cerebellum to the eye muscles. Although the direct path is very quick, it is not sensible to changes within the system, i.e. to changes in muscular strength concerning the eye or to magnifying or miniaturizing glasses, as the experimentation with monkeys has shown. Therefore, there is a second path over the cerebellum: the head motion information from the vestibular organ and the optical flow from the eye are transferred to the synapses of the Purkinje cells via parallel fibres. As already mentioned, these synapses can be influenced by teaching signals from the so-called climbing fibre. Regarding the VOR, the climbing fibre carries the optical flow on the retina as a teaching signal. The interaction of the direct path and the "cleaning" effect of the adaptive path over the cerebellum guarantee a stable picture of our environment.

6 Technical applications of the vestibulo-ocular reflex

The need for the VOR concerning our eye is obvious. But there is also a wide range of fields where it can be useful or necessary in technical applications. Besides the stabilization of camera systems in general there is also a field which will get more and more important in the future: driver-assistance systems. These systems often need optical input (e.g. from road signs, other cars, number plates etc.), therefore a telephoto lens is needed which is quite sensitive to motion caused by road bumps and other shocks. Consequently the image from the camera has to be stabilized. Stabilization by optical means leads to processing delays of about 50-100ms, which is too slow. Therefore, the angular velocity of the camera has to be measured in order to rotate the camera by servo motors to obtain a stable picture. On the other hand, cars and driver-assistance systems have to be affordable, consequently also cheap, inaccurate sensors must be allowed to be used. This is only possible if inaccuracies are compensated automatically, and therefore nature has an excellent method: the VOR. Anomg others, the compensationing qualities of the VOR have been seen during the experimentation with monkeys. But how can we "copy" the VOR for our technical intentions?

7 Hardware implementation of the vestibulo-ocular reflex

For the implementation of the VOR, adequate hardware has to be found or newly developed as a biological analogon. The function of the cerebellum can be implemented software-based on a digital signal processor (DSP), e.g. in Matlab, so there is no need to implement new hardware. The vestibular organ detecting motion has to be implemented as a six-degrees-of-freedom inertial measurement unit (6-DOF IMU), where the most significant problem is size. For driver-assistance systems, the sensor should more or less vanish behind the rear-view mirror, and the smallest commercially available intelligent IMU measures 50x38x25mm^3. Consequently, a new IMU was developed at TU Munich, Institute of Applied Mechanics, the μCube with an edge length of only 15mm. Inside this cube, three gyroscopes are situated to detect rotation in each axis, and an accelerometer with three axes integrated on one chip. The direct path from the vestibular organ to the eye muscles can be implemented as a CAN bus system, and the analogon to the eye muscles are servo motors. The camera itself – a CMOS sensor with an effective resolution of 750x400 pixels – is installed in a serial gimballed configuration with two perpendicular axes. The pan actuator is driven by a 4.5 Watt motor, the inner frame by a 11 Watt Maxon motor with a backlash free HarmonicDrive gearbox. Each of the two axes is controlled by a 512 steps differential encoder. In the future, only a mirror projecting the picture onto a camera is moved by the configuration, which leads to further miniaturization.

8 An example for a VOR algorithm

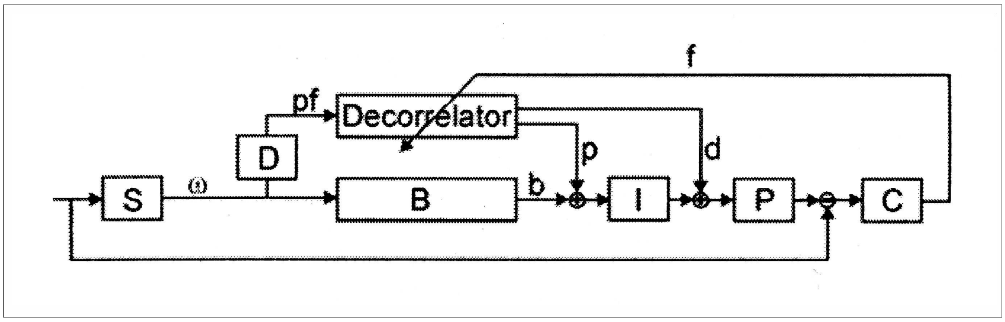

Figure 2: Block diagram of biomimetic control for gaze stabilization, from [1]

Primary, the angular velocity of the IMU is measured and according to these data, the adequate motion of the camera is calculated ("direct path"). If the used sensors were perfect, this would be sufficient. But as already mentioned, sensors in driver-assistance systems have to be affordable, consequently there will always be relative movement of the environment ("optical flow") because of inaccuracies is production and assembling. To compensate these inaccuracies, the relation between angular rates and the optical flow is measured and these data influence a matrix w of a so-called decorrelator:

$$\dot{w} = -\vec{f} \cdot \vec{pf}^T \cdot \beta = - \begin{bmatrix} f_x \\ f_y \\ f_z \end{bmatrix} \begin{bmatrix} \omega_x \; \omega_y \; \omega_z \; \frac{d\omega_x}{dt} \; \frac{d\omega_y}{dt} \; \frac{d\omega_z}{dt} \end{bmatrix} \begin{bmatrix} \beta_p \cdot E^{3x3} & 0^{3x3} \\ 0^{3x3} & \beta_d \cdot E^{3x3} \end{bmatrix}$$

where $\vec{f}$ is the optical flow and $\vec{pf}$ the teaching signal ("parallel fibre signal"), i.e. the angular rates and their derivatives:

$$\vec{pf} = \begin{bmatrix} \omega \\ \frac{d\omega}{dt} \end{bmatrix}$$

The matrix w is needed to calculate a certain output out of the angular velocity in order to "clean" the information via the direct path additively. After some time, the matrix w is nearly perfect and always adds the right correction amount to the direct path data, i.e. optical flow and angular velocity are decorrelated.

The decorrelator produces two outputs: a proportional output and a derivative output. The proportional output is added to the direct path and can be expressed as

$$\vec{p} = J \, [w_p \; w_d] \begin{bmatrix} E^{3x3} & 0^{3x3} \\ 0^{3x3} & 0^{3x3} \end{bmatrix} \vec{pf}$$

$$w = [w_p \; w_d] \in \mathbb{R}^{3\times 6}$$

The output of the derivative path is added after the integrator:

$$\vec{d} = J\,[w_p\; w_d] \begin{bmatrix} 0^{3x3} & 0^{3x3} \\ E^{3x3} & 0^{3x3} \end{bmatrix} \vec{pf}$$

$$w = [w_p\; w_d] \in \mathbb{R}^{3\times 6}$$

8.1 Possible variants of the basic algorithm

8.1.1 Eligibility trace

Processing delays of the optical flow, i.e. the climbing fibre signal, are about 100ms, which is far longer than the rather short delays of the motion data about 10ms. The decorrelator compares these signals in order to calculate the matrix w. As the amounts of delay are different, learning can get instable for high frequencies. Therefore, the parallel fibre signals can be delayed by about 100ms, and high frequencies can be removed by a smoothing filter.

8.1.2 Simplification of the teaching signal

Another possibility of modifying the algorithm is the simplification of the teaching signal. Under certain circumstances it may be adequate to use only the sign of the optical flow, not the exact values in regard to saving capacities of the climbing fibre pathway. Experiments have shown that the learning process is not much worse, although the speed of convergence may be worse.

8.1.3 Modification of the learning rate

The closer the values of the matrix w are to the optimum, the more often the sign of the optical flow changes. If the matrix w is quite perfect, big changes in its values due to changing signs of the optical flow are unwanted. Consequently the algorithm can be improved by lowering the learning rate if w has reached the environment of its optimum, i.e. when the sign of the optical flow changes often.

9 Other problems in regard to driver-assistance systems

Besides the described reduction of optical flow, some more features are needed to keep the line of sight on the road. E.g., so-called saccades (rapid changes in viewing direction) and visual tracking for the closer examination of objects like road signs have to be implemented.

10 Testing the camera stabilization system

10.1 Testing under laboratory conditions

Before testing the camera stabilizing system mounted in a car, it was tested under laboratory conditions on a hexapod. The platform of a hexapod can be moved in all three degrees of freedom of translation and in all three degrees of freedom of rotation, creating sensed angular rates of about $100\frac{\circ}{s}$. For our purposes, the tool center point of the platform stayed fixed, and only the rotaional degrees of freedom corresponding to the degrees of freedom of our motion device were used. In the beginning of the experimentation, the optical flow was about 6 pix/frame, and after 2-3 minutes, it was reduced to less than 1 pix/frame. The improvement was not only measured, but also subjectively viewable.

10.2 Testing under real conditions on the road

In the next step, the system was tested under real conditions on the road. Therefore, the stabilizing system was mounted near the rear-view mirror, and additionally, a wide-angle camera was installed to get an overwiev of the environment and to detect points of interest for the stabilized camera with a telephoto lens. The described algorithm was tested in association with saccades and visual tracking, and in order to improve the performance of the system in regard to tracking space fixed objects, data like vehicle velocity and yaw rate were added via CAN bus.

The performance of the system was tested in three modes. In the first mode, a lane marker was focused on in a distance of 60m, followed up to a distance of 20m, then the next one was focused on. In the second mode, the lane separation was focused on in a constant distance of 40m. In the third mode, nothing was focused on, but the camera was stabilized around a constant line of sight. In all three modes road bumps could be compensated.

11 Another field of neuroscience and technical application: face recognition

There is another field where learning from the human brain is necessary: face recognition. Humans have excellent face recognition abilities under all kinds of different circumstances: different kinds of illumination, different viewing angles, different pose, different facial expressions and so on, whereas classical computer systems do not. The need to make researches in this field is obvious: the fear of the global threat of terrorism increases, and there is a trend towards the surveillance of citizens.

Here, only a very brief look onto a possible algorithm of face recognition according to Valentin, Abdi and Edelman (1997) is described. The neural network we use consists of the so-called input units, the hidden units and the output units. As an input for the n input units, an n-dimensional vector as a projection of the face properties on a certain eigenspace for face description is used. The hidden units are a kind of inner representations of faces. There can be one hidden unit for each face from a certain angle of view, two hidden units for a profile view and a front view or in the best case more hidden units for a lot of angles of view.

The more similar the input is to a certain hidden unit, the more active this unit gets. During the learning step, the output weights of the hidden units can be changed so that the output is 1 for the viewed person. During the testing step, a face in an angle of view not seen yet by the system is shown to the system, but nevertheless the highest output (close to 1) is produced for the viewed person.

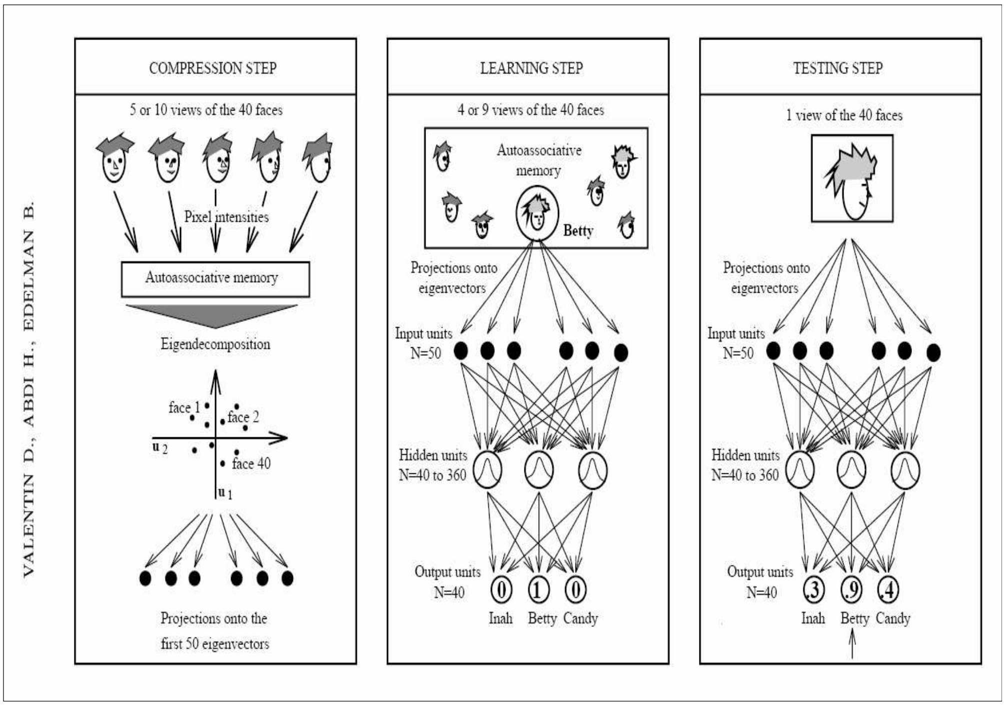

Figure 3: Principle of a face recognition algorithm, from [2]

12 Conclusion

All in all, biologically inspired algorithms for mechatronic systems are the key to the solution of many different problems. First, the technical use of neuroscience can give machines typical human abilities, like face recognition. But furthermore, also economic aspects can be improved, e.g. service intervals can be optimized, the complexity of the installation of systems can be minimized by self-learning abilities – this all leads to cost reduction as shown by the example of the camera stabilizing system by Günthner et al. (2005). Consequently the field of neuroscience is interesting to a large variety of technical fields.

References

[1] Günthner, W.; Glasauer, S.; Wagner, P.; Ulbrich, H.: Biologically inspired multi-sensor fusion for adaptive camera stabilization in driver-assistance systems, Advanced Microsystems for Automotive Applications AMAA, Berlin, April 25-27, 2006 (in press)

[2] Valentin, D.; Abdi, H..; Edelman, B.: What represents a face: A Computational Approach for the Integration of Physiological and Psychiological Data, 1997