

# 1

## Introduction

Our world is a visual world. Visual perception is by far the most important sensory process by which we gather and extract information from our environment. Light reflected from objects in our world is a very rich source of information. Its short wavelength and high transmission speed allow us a spatially accurate and fast localization of reflecting surfaces. The spectral variations in wavelength and intensity in the reflected light resemble the physical properties of object surfaces, and provide means to recognize them. The sources that light our world are usually inhomogeneous. The sun, our natural light source, for example, is in good approximation a point source. Inhomogeneous light sources cause shadows and reflectances that are highly correlated with the shape of objects. Thus, knowledge of the spatial position and extent of the light source enables further extraction of information about our environment.

Our world is also a world of motion. We and most other animals are moving creatures. We navigate successfully through a dynamic environment, and we use predominantly visual information to do so. A sense of motion is crucial for the perception of our own motion in relation to other moving and static objects in the environment. We must predict accurately the relative dynamics of objects in the environment in order to plan appropriate actions. Take for example the following situation that illustrates the nature of such a perceptual task: the goal-keeper of a football team is facing a direct free-kick toward his goal.<sup>1</sup> In order to prevent the opposing team from scoring, he needs an accurate estimate of the real motion trajectory of the ball such that he can precisely plan and orchestrate his body movements to catch or deflect the ball appropriately. There is little more than just visual information available to him in order to solve the task. And once he is in motion the situation becomes much more complicated because visual motion information now represents the relative motion between himself and the ball while the important coordinate frame remains

<sup>1</sup>There are two remarks to make. First, “football” is referred to as the European-style football, also called “soccer” elsewhere. Second, there is no gender-specific implication here; a male goal-keeper was simply chosen so-as to represent the sheer majority of goal-keepers on earth. In fact, I particularly would like to include non-human, artificial goal-keepers as in robotic football (RoboCup [Kitano et al. 1997]).

static (the goal). Yet, despite its difficulty, with appropriate training some of us become astonishingly good at performing this task.

High performance is important because we live in a highly competitive world. The survival of the fittest applies to us as to any other living organism, and although the fields of competition might have slightly shifted and diverted during recent evolutionary history, we had better catch that free-kick if we want to win the game! This competitive pressure not only promotes a visual motion perception system that can determine quickly what is moving where, in which direction, and at what speed; but it also forces this system to be efficient. Efficiency is crucial in biological systems. It encourages solutions that consume the smallest amount of resources of time, substrate, and energy. The requirement for efficiency is advantageous because it drives the system to be quicker, to go further, to last longer, and to have more resources left to solve and perform other tasks at the same time. Our goal-keeper does not have much time to compute the trajectory of the ball. Often only a split second determines a win or a defeat. At the same time he must control his body movements, watch his team-mates, and possibly shout instructions to the defenders. Thus, being the complex sensory-motor system he is, he cannot dedicate all of the resources available to solve a single task.

Compared to human perceptual abilities, nature provides us with even more astonishing examples of efficient visual motion perception. Consider the various flying insects that navigate by visual perception. They weigh only fractions of grams, yet they are able to navigate successfully at high speeds through a complicated environments in which they must resolve visual motions up to 2000 deg/s. [O'Carroll et al. 1996] – and this using only a few drops of nectar a day.

## 1.1 Artificial Autonomous Systems

What applies to biological systems applies also to a large extent to any artificial autonomous system that behaves freely in a real-world<sup>2</sup> environment. When humankind started to build artificial autonomous systems, it was commonly accepted that such systems would become part of our everyday life by the year 2001. Numberless science-fiction stories and movies have encouraged visions of how such agents should behave and interfere with human society. Although many of these scenarios seem realistic and desirable, they are far from becoming reality in the near future. Briefly, we have a rather good sense of what these agents should be capable of, but we are not able to construct them yet. The (semi-)autonomous rover of NASA's recent Mars missions,<sup>3</sup> or demonstrations of artificial pets,<sup>4</sup> confirm that these fragile and slow state-of-the-art systems are not keeping up with our imagination.

Remarkably, our progress in creating artificial autonomous systems is substantially slower than the general technological advances in recent history. For example, digital microprocessors, our dominant computational technology, have exhibited an incredible development. The integration density literally exploded over the last few decades, and so did

---

<sup>2</sup>The term *real-world* is coined to follow an equivalent logic as the term *real-time*: a real-world environment does not really have to be the "real" world but has to capture its principal characteristics.

<sup>3</sup>*Pathfinder 1997, Mars Exploration Rovers 2004*: <http://marsprogram.jpl.nasa.gov>

<sup>4</sup>e.g. *AIBO* from SONY: <http://www.sony.net/Products/aibo/>

the density of computational power [Moore 1965]. By contrast, the vast majority of the predicted scenarios for robots have turned out to be hopelessly unrealistic and over-optimistic. Why?

In order to answer this question and to understand the limitations of traditional approaches, we should recall the basic problems faced by an autonomously behaving, cognitive system. By definition, such a system perceives, takes decisions, and plans actions on a cognitive level. In doing so, it expresses some degree of intelligence. Our goal-keeper knows exactly what he has to do in order to defend the free-kick: he has to concentrate on the ball in order to estimate its trajectory, and then move his body so that he can catch or deflect the ball. Although his reasoning and perception are cognitive, the immanent interaction between him and his environment is of a different, much more physical kind. Here, photons are hitting the retina, and muscle-force is being applied to the environment. Fortunately, the goalie is not directly aware of all the individual photons, nor is he in explicit control of all the individual muscles involved in performing a movement such as catching a ball. The goal-keeper has a nervous system, and one of its many functions is to instantiate a *transformation layer* between the environment and his cognitive mind. The brain reduces and preprocesses the huge amount of noisy sensory data, categorizes and extracts the relevant information, and translates it into a form that is accessible to cognitive reasoning (see Figure 1.1). This is the process of perception. In the process of action, a similar yet inverse transformation must take place. The rather global and unspecific cognitive decisions need to be resolved into a finely orchestrated ensemble of motor commands for the individual muscles that then interact with the environment. However, the process of action will not be addressed further in this book.

At an initial step perception requires sensory transduction. A sensory stage measures the physical properties of the environment and represents these measurements in a signal the

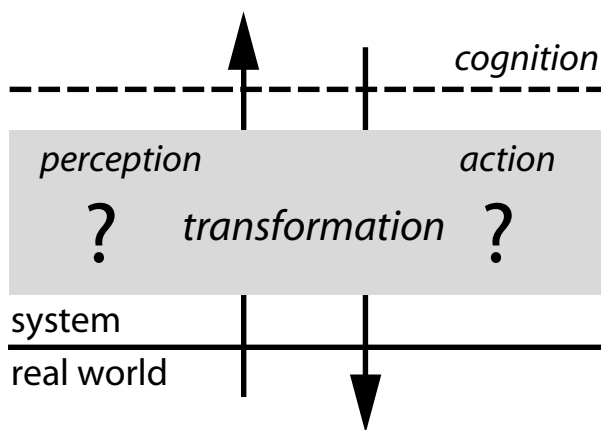


Figure 1.1 *Perception and action.*

Any cognitive autonomous system needs to transform the physical world through perception into a cognitive syntax and – vice versa – to transform cognitive language into action. The computational processes and their implementation involved in this transformation are little understood but are the key factor for the creation of efficient, artificial, autonomous agents.

rest of the system can process. It is, however, clear that sensory transduction is not the only transformation process of perception. Because if it were, the cognitive abilities would be completely overwhelmed with detailed information. As pointed out, an important purpose of perception is to reduce the raw sensory data and extract only the relevant information. This includes tasks such as object recognition, coordinate transformation, motion estimation, and so forth. Perception is the *interpretation* of sensory information with respect to the perceptual goal. The sensory stage is typically limited, and sensory information may be ambiguous and is usually corrupted by noise. Perception, however, must be robust to noise and resolve ambiguities when they occur. Sometimes, this includes the necessity to fill in missing information according to expectations, which can sometimes lead to wrong interpretations: most of us have experienced certainly one or more of the many examples of perceptual illusions.

Although not described in more detail at this point, perceptual processes often represent large computational problems that need to be solved in a small amount of time. It is clear that the efficient implementation of solutions to these tasks crucially determines the performance of the whole autonomous system. Traditional solutions to these computational problems almost exclusively rely on the digital computational architecture as outlined by von Neumann [1945].<sup>5</sup> Although solutions to all computable problems can be implemented in the von Neumann framework [Turing 1950], it is questionable that these implementations are equally efficient. For example, consider the simple operation of adding two analog variables: a digital implementation of addition requires the digitization of the two values, the subsequent storage of the two binary strings, and a register that finally performs the binary addition. Depending on the resolution, the electronic implementation can use up to several hundred transistors and require multiple processing cycles [Reyneri 2003]. In contrast, assuming that the two variables are represented by two electrical currents flowing in two wires, the same addition can be performed by simply connecting the two wires and relying on Kirchhoff's current law.

The von Neumann framework also favors a particular philosophy of computation. Due to its completely discrete nature, it forces solutions to be dissected into a large number of very small and sequential processing steps. While the framework is very successful in implementing clearly structured, exact mathematical problems, it is unclear if it is well suited to implement solutions for perceptual problems in autonomous systems. The computational framework and the computational problems simply do not seem to match: on the one hand the digital, sequential machinery only accepts defined states, and on the other hand the often ambiguous, perceptual problems require parallel processing of continuous measures.

It may be that digital, sequential computation is a valid concept for building autonomous artificial systems that are as powerful and intelligent as we imagine. It may be that we can make up for its inefficiency with the still rapidly growing advances in digital processor technology. However, I doubt it. But how amazing would the possibilities be if we could find and develop a more efficient implementation framework? There must be a different, more efficient way of solving such problems – and that's what this book is about. It aims to demonstrate another way of thinking of solutions to these problems and implementing

---

<sup>5</sup>In retrospect, it is remarkable that from the very beginning, John von Neumann referred to his idea of a computational device as an explanation and even a model of how biological neural networks process information.

them. And, in fact, the burden to prove that there are indeed other and much more efficient ways of computation has been carried by someone else – nature.

## 1.2 Neural Computation and Analog Integrated Circuits

Biological neural networks are examples of wonderfully engineered and efficient computational systems. When researchers first began to develop mathematical models for how nervous systems actually compute and process information, they very soon realized that one of the main reasons for the impressive computational power and efficiency of neural networks is the collective computation that takes place among their highly connected neurons. In one of the most influential and ground-breaking papers, which arguably initiated the field of computational neuroscience, McCulloch and Pitts [1943] proved that any finite logical expression can be realized by networks of very simple, binary computational units. This was, and still is, an impressive result because it demonstrated that computationally very limited processing units can perform very complex computations when connected together. Unfortunately, many researchers concluded therefore that the brain is nothing more than a big logical device – a digital computer. This is of course not the case because McCulloch and Pitts' model is not a good approximation of our brain, which they were well aware of at the time their work was published.

Another key feature of neuronal structures – which was neglected in McCulloch and Pitts' model – is that they make computational use of their intrinsic physical properties. Neural computation is physical computation. Neural systems do not have a centralized structure in which memory and hardware, algorithm and computational machinery, are physically separated. In neurons, the function is the architecture – and vice versa. While the bare-bone simple McCulloch and Pitts model approximates neurons to be binary and without any dynamics, real neurons follow the continuous dynamics of their physical properties and underlying chemical processes and are analog in many respects. Real neurons have a cell membrane with a capacitance that acts as a low-pass filter to the incoming signal through its dendrites, they have dendritic trees that non-linearly add signals from other neurons, and so forth. John Hopfield showed in his classical papers [Hopfield 1982, Hopfield 1984] that the dynamics of the model neurons in his networks are a crucial prerequisite to compute near-optimal solutions for hard optimization problems with recurrent neural networks [Hopfield and Tank 1985]. More importantly, these networks are very efficient, establishing the solution within a few characteristic time constants of an individual neuron. And they typically scale very favorably. Network structure and analog processing seem to be two key properties of nervous systems providing them with efficiency and computational power, but nonetheless two properties that digital computers typically do not share or exploit. Presumably, nervous systems are very well optimized to solve the kinds of computational problems that they have to solve to guarantee survival of their whole organism. So it seems very promising to reveal these optimal computational strategies, develop a methodology, and transfer it to technology in order to create efficient solutions for particular classes of computational problems.

It was Carver Mead who, inspired by the course “The Physics of Computation” he jointly taught with John Hopfield and Richard Feynman at Caltech in 1982, first proposed the idea of embodying neural computation in silicon *analog very large-scale integrated (aVLSI)* circuits, a technology which he initially advanced for the development of integrated digital

circuits.<sup>6</sup> Mead's book *Analog VLSI and Neural Systems* [Mead 1989] was a sparkling source of inspiration for this new emerging field, often called *neuromorphic* [Mead 1990] or *neuro-inspired* [Vittoz 1989] circuit design. And nothing illustrates better the motivation for the new field than Carver Mead writing in his book: "Our struggles with digital computers have taught us much about how neural computation *is not* done; unfortunately, they have taught us relatively little about how it *is* done."

In the meantime, many of these systems have been developed, particularly for perceptual tasks, of which the *silicon retina* [Mahowald and Mead 1991] was certainly one of the most popular examples. The field is still young. Inevitable technological problems have led now to a more realistic assessment of how quickly the development will continue than in the euphoric excitement of its beginning. But the potential of these neuromorphic systems is obvious and the growing scientific interest is documented by an ever-increasing number of dedicated conferences and publications. The importance of these neuromorphic circuits in the development of autonomous artificial systems cannot be over-estimated.

This book is a contribution to further promote this approach. Nevertheless, it is as much about network computation as about hardware implementation. In that sense it is perhaps closer to the original ideas of Hopfield and Mead than current research. The perception of visual motion thereby only serves as the example task to address the fundamental problems in artificial perception, and to illustrate efficient solutions by means of analog VLSI network implementations. In many senses, the proposed solutions use the same computational approach and strategy as we believe neural systems do to solve perceptual problems. However, the presented networks are not designed to reflect the biological reference as thoroughly as possible. The book carefully avoids using the term *neuron* in any other than its biological meaning. Despite many similarities, silicon aVLSI circuits are bound to their own physical constraints that in many ways diverge from the constraints nervous systems are facing. It does not seem sensible to copy biological circuits as exactly as possible. Rather, this book aims to show how to use basic computational principles that we believe make nervous systems so efficient and apply them to the new substrate and the task to solve.

---

<sup>6</sup>There were earlier attempts to build analog electrical models of neural systems. Fukushima et al. [1970] built an electronic retina from discrete(!) electronic parts. However, only when integrated technology became available were such circuits of practical interest.