



Doctoral Thesis

## Neuromorphic implementation of a saliency-based visual selective attention system

**Author(s):**

Sonnleithner, Daniel Eduard

**Publication Date:**

2013

**Permanent Link:**

<https://doi.org/10.3929/ethz-a-010105959> →

**Rights / License:**

[In Copyright - Non-Commercial Use Permitted](#) →

This page was generated automatically upon download from the [ETH Zurich Research Collection](#). For more information please consult the [Terms of use](#).

DISS. ETH No. 21347

# **Neuromorphic Implementation of a Saliency-based Visual Selective Attention System**

A dissertation submitted to

**ETH ZURICH**

for the degree of

**Doctor of Sciences**

presented by

**DANIEL EDUARD SONNLEITHNER**

Dipl.-Ing., Universität Karlsruhe (TH)

23. December 1982

citizen of

Vully-les-Lacs, Switzerland

accepted on the recommendation of

Prof. Dr. Rodney Douglas, examiner

Prof. Dr. Giacomo Indiveri, co-examiner

Prof. Dr. Ernst Niebur, co-examiner

2013

# Abstract

Visual perception is one of our most important senses. To be able to see our eyes transform the light signal into spiking data streams. Our brain extracts from this data relevant information. To provide vision to a mobile robotic system both computational steps have to be implemented. With today's technology it is possible to build visual sensors with high resolution and high recording frame rates that dissipate little power. The sensors provide a clear, detailed view of the robot's environment. Therefore the first aspect of vision is implemented. The second aspect is to extract relevant information from the visual data. Information are relevant for the robot if they allow its interaction with the environment, e.g. determining its location or recognizing obstacles. Therefore, the extraction has to happen in real time.

The problem for the robot is to extract relevant information from the huge amount of visual data provided by its sensors in real time with limited computational resources.

In biology an analogy of this phenomenon can be observed: Our eyes provide far more data than the human brain can process. Nevertheless we are able to interact with our environment in real time. The mechanism that allows us to extract the relevant information from the data provided by our eyes in real time is called *selective attention* [Treisman and Gelade, 1980]: Only a subset of the visual data is processed in detail, the rest is discarded. A preprocessing system identifies regions in the visual space that are salient. The visual data from these regions is processed further by our brain in a serial fashion. An alternative to master the described problem is to adapt this bio-inspired strategy to robotics. In this thesis I present a neuromorphic multi-chip system that is derived from a saliency-based selective attention model [Koch and Ullman, 1985]. My proposed solution uses building blocks derived from the brain: emulates of neurons and synapses. Therefore, it achieves very efficient computations. To estimate the most relevant regions of the input scene the model uses a "saliency map": this map assigns to each pixel of the input image a value for its saliency. In the model, saliency is computed by using center-surround operations [Itti et al., 1998]. In this thesis I implement this operation by making use of a 2D-array of silicon neurons with excitatory and inhibitory synapses. The synaptic weights are realized with the help of a probabilistic mapping device. The selective attention model scans sequentially the regions of high saliency. I implement this operation by using a neuromorphic chip implementing a 2D-Winner-Take-All (WTA) network with Inhibition of Return (IOR) functionality. In order to control the operational parameters of the neuromorphic chips used in this thesis as well as for the communication of the individual chips with a workstation, I developed a custom hardware/software infrastructure. Furthermore, I present results of experiments conducted with the visual selective attention system to show its functionality.

By implementing the bio-inspired method of selective attention a mobile robot can better assign its computational resources to certain regions in the robot's visual input space. It is the first time that such an implementation of a selective attention system based on a neuromorphic

multi-chip system is presented.

# Zusammenfassung

Sehen ist eines unserer wichtigsten Sinne. Um zu sehen, wandelt unser Auge das Licht in elektrische Impulse um. Das Gehirn extrahiert daraus relevante Informationen. Um einen mobilen Roboter mit der Fähigkeit des Sehens auszustatten, müssen beide Verarbeitungsschritte implementiert werden. Mittels heutiger Fertigungstechniken ist es möglich, sparsame Bildsensoren mit hoher Auflösung und hoher Aufnahme­frequenz zu bauen. Sie liefern scharfe, detailreiche Bilder der Umgebung des Roboters. Damit wird der erste Aspekt des Sehens ermöglicht. Der zweite ist, die relevanten Informationen aus den Bilddaten zu extrahieren. Informationen sind für den Roboter relevant, wenn sie ihm ermöglichen mit seiner Umgebung zu interagieren, z.B. seine Position zu bestimmen oder Hindernissen auszuweichen. Um dies zu gewährleisten, ist es nötig, dass die Bilddaten in Echtzeit verarbeitet werden.

Problematisch dabei ist, mit begrenzten Rechenkapazitäten die für den Roboter relevanten Informationen aus der riesigen Menge an visuellen Daten, die der Sensor zur Verfügung stellt, in Echtzeit zu extrahieren.

In der Biologie lässt sich ein analoges Phänomen beobachten: Unsere Augen liefern ein Vielfaches der Daten, die vom menschlichen Gehirn überhaupt verarbeitet werden können. Dennoch können wir mit unserer Umgebung in Echtzeit interagieren. Den Mechanismus, der die relevanten Informationen aus den Daten, die unsere Augen liefern, extrahiert, wird *selektive Aufmerksamkeit* genannt [Treisman and Gelade, 1980]: Nur ein Teil der Bilddaten wird im Detail verarbeitet, der Rest wird verworfen. Ein vorverarbeitendes System erkennt Bereiche im Sehfeld, die salient sind. Die Bilddaten dieser Bereiche werden von unserem Gehirn in sequentieller Weise weiterverarbeitet. Eine Möglichkeit das Datenproblem zu lösen, ist, diese Strategie aus der Biologie auf die Robotik zu übertragen. In dieser Arbeit beschreibe ich ein neuromorphisches Multi-Chip-System, das von einem selektiven Aufmerksamkeitsmodell abgeleitet ist [Koch and Ullman, 1985]. Das neuromorphische Multi-Chip-System nutzt dem Gehirn nachempfundene Bausteine: Emulationen von Neuronen und Synapsen. Dadurch kann das System Berechnungen sehr effizient ausführen. Um die relevanten Bereiche im Eingangsbild zu erkennen, nutzt das Modell eine Salienzkarte. Diese Karte ordnet jedem Pixel des Eingangsbildes einen Wert für dessen Salienz zu. Im Modell wird die Salienz mittels Zentrum-Umfeld-Funktion bestimmt [Itti et al., 1998]. In dieser Arbeit implementiere ich diese Funktion mit Hilfe eines neuromorphen Chips bei dem künstliche Neuronen und erregende und hemmende Synapsen in zwei Dimensionen angeordnet sind. Die Gewichte der Synapsen werden mittels eines wahrscheinlichkeitsgesteuerten Mappers realisiert. Das Aufmerksamkeitsmodell scannt sequentiell alle Bereiche mit hoher Salienz. Ich implementiere diese Funktion mit Hilfe eines zweidimensionalen Winner-Takes-All Netzwerkes mit Inhibition of Return Funktionalität. Um zum Einen die Betriebsparameter der neuromorphen Chips einzustellen und zum Anderen die Kommunikation zwischen den Chips und einem Computer zu ermöglichen, habe ich eine Hardware/Software Infrastruktur entwickelt. Ausserdem beschreibe ich die Resultate

---

von Experimenten, die mit dem selektiven Aufmerksamkeitssystem durchgeführt wurden, um seine Funktionsweise zu zeigen.

Durch die Anwendung der selektiven Aufmerksamkeitsmethode, die von der Biologie abgeleitet ist, können mobile Roboter ihre Rechenkapazitäten besser den relevanten Bereichen im Sichtbereich zuordnen. Das hier dargestellte System ist das erste selektive Aufmerksamkeitssystem, das auf einem neuromorphen Multi-Chip-System aufbaut.