

Kurzfassung

Die visuelle Aufmerksamkeit ist zentraler Bestandteil der menschlichen visuellen Informationsverarbeitung. Sie findet zunehmend auch in der Modellierung von Systemen des Computer-Sehens Verwendung. Aufmerksamkeit ist immer da entscheidend, wo es um die Verteilung von Ressourcen, die Auswahl von relevanten Informationen und die Priorisierung von Aufgaben geht. Die positiven Effekte von Aufmerksamkeit liegen in der effizienteren Verarbeitung visueller Informationen sowie in der Unterdrückung von ablenkenden Informationen. Schließlich spielt Aufmerksamkeit eine wichtige Rolle, wenn es um die Verbindung von Wahrnehmung und Handlung und somit um die Lösung des Bindungsproblems geht.

Konventionelle Modellierungen visueller Aufmerksamkeit zeichnen sich jedoch entweder durch eine Fixierung auf statische zweidimensionale Bilder aus oder zeigen eine stark vereinfachte Modellierung der Selektion. Dabei sind es gerade dreidimensionale dynamische Umgebungen, in denen der Einsatz visueller Aufmerksamkeit den größten Nutzen verspricht.

Als Konsequenz sollten einerseits fortgeschrittene Aspekte der Selektivitätsmodellierung wie objektbasierte Aufmerksamkeit, Integration datengetriebener und modellgetriebener Aufmerksamkeit und dynamische Selektion und Inhibition in einem Modell visueller Aufmerksamkeit berücksichtigt werden. Andererseits sollen die Selektionsmechanismen dahingehend modifiziert werden, dass sie einer dynamischen dreidimensionalen Umgebung gerecht werden. Eine zentrale Herausforderung ist es dabei, in einer dynamischen Umgebung mit einer seriellen attentiven Verarbeitungsstufe ein internes Modell mit den wichtigsten Objekten aktuell zu halten.

Die vorliegende Arbeit stellt ein Modell visueller Aufmerksamkeit für Systeme des Aktiven Sehens vor, das sich durch einen neuartigen Selektionsmechanismus auszeichnet. Seine Relevanz wird sowohl aus Sicht der Effektivität im Computer-Sehen als auch hinsichtlich der Modellierung menschlicher visueller Aufmerksamkeit belegt. Dieser Selektionsmechanismus ist durch die Anwendung einer Aufmerksamkeitssteuerung auf dynamische räumliche Szenen motiviert, die sich durch konventionelle Modellierungen nicht ohne weiteres erreichen lässt. Das Modell ist in der Lage, seltener beachtete experimentelle Daten zum *multiple object tracking* oder der objektbasierten *inhibition of return* zu erklären. Wichtiger jedoch ist die Leistungsfähigkeit als Bestandteil eines Computer Vision-Systems. Sie zeigt sich daran, dass mit wenigen Ressourcen ein Weltmodell der wichtigsten Objekte in dynamischen Szenen bestimmt und aufrecht erhalten werden kann.

Abstract

Visual attention is a substantial aspect of the way humans perceive and process visual information. Increasing use of it is made in models of computer vision. Attention is relevant whenever resources are to be distributed, relevant information has to be selected, and tasks have to be prioritized. The positive effects of attention are due to the more effective processing of visual information as well as the suppression of distracting elements. At last, attention plays an important role in connecting perception and action and thus in solving the binding problem.

Conventional models are mostly either focussed on static two-dimensional images, or are equipped with a strongly simplified selection mechanism. But especially in three-dimensional dynamic environments the use of attention seems most profitable.

As one consequence advanced aspects in modelling visual attention like object-based attention, the integration of data-driven and model-driven attention, and the dynamic selection and inhibition should be integrated. On the other side, selection mechanisms need to be modified in order to cope with dynamic three-dimensional environments. A central challenge lies in using a serial attentive computations stage in a dynamic environment providing an up-to-date world model of the most relevant objects.

This work introduces a model of visual attention for active vision systems using a novel selection mechanisms. Its quality will be established regarding the effectiveness as a computer vision process as well as regarding the modelling of natural visual attention. The selection mechanism is motivated by applying attentional mechanisms to dynamic spatial scenes that cannot be accomplished by conventional models. The model serves to explain some of the more disregarded experimental data on multiple object tracking and object-based inhibition of return. Even more important is its usefulness as a module of computer vision systems. This usefulness is evident in its ability to compute and update a world model of the most relevant objects in a dynamic scene with most efficient use of resources.