

Titel der Dissertation: A Neuro-Dynamic Architecture for Autonomous Visual Scene Representation

Name des Antragstellers: Stephan Klaus Ulrich Zibner

## Kurzfassung

Unsere Fähigkeit mit Gegenständen in unserer Umgebung zu interagieren ist einzigartig. Eine Grundlage dafür ist die visuelle Perzeption von Szenen, welche zur Generierung interner Repräsentationen führt, auf denen nachfolgende Verhalten aufbauen. Greifbewegungen, sowie Sprachgenerierung und -verständnis sind Beispiele für solche Verhalten. In der Robotik stellt die visuelle Verarbeitung von Szenen eine große Hürde dar, insbesondere für a priori unbekannte oder dynamische Szenen. In dieser Arbeit stelle ich eine neurodynamische Architektur zur visuellen Verarbeitung von Szenen vor, welche ein Arbeitsgedächtnis aufbaut, dieses bei Veränderungen der Szene aktualisiert und das akkumulierte Wissen über die Szene nutzt, um die Suche nach einem Zielobjekt effizient durchzuführen. Kern der Repräsentation sind dreidimensionale neuronale Felder, welche die Position von Objekten mit deren visuellen Eigenschaften wie ihrer Farbe oder Größe assoziieren. Ich lege dabei meinen Fokus auf die Organisation der involvierten Verhalten und zeige, wie die internen Prozesse autonom ablaufen. In Experimenten auf einer Roboterplattform evaluiere ich die von der Architektur generierten Verhalten und Prozesse und ziehe Parallelen zu Erkenntnissen über die menschliche Szenenrepräsentation. Ich stelle zwei Erweiterungen der Szenenrepräsentation vor, in denen ich die zugrunde liegenden Prinzipien nutze, um Objekterkennung in die Szenenrepräsentation zu integrieren und die Szenenrepräsentation nutze, um auf Objekte gerichtete Armbewegungen zu steuern. Die Verbindung mit Objekterkennung erlaubt es, Zielobjekte anhand abstrakter Label zu suchen. Da die Objekterkennung rechenaufwändig ist, ist die Erzeugung der Label aus der Szene nicht als paralleler Prozess der visuellen Wahrnehmung gestaltbar, sondern nutzt die entwickelte sequentielle Bearbeitung von visuellen Objekten. Für die Bewegungsgenerierung leistet die Architektur, dass Bewegungsverhalten sich jederzeit an Positionsveränderungen des Ziels anpassen können, gewährleistet durch die kontinuierliche Anbindung der Szenenrepräsentation an die visuelle Wahrnehmung und durch die autonome Organisation der Prozesse der kognitiven visuellen Wahrnehmung. Solches "online updating" ist auch beim Menschen bekannt. Ich schließe mit einem Ausblick auf weiterführende Integration im Kontext von Greifbewegungen.