

Kognitive Objekterkennung in dynamischer Feldtheorie
(Cognitive object recognition in dynamic field theory)
Kurzfassung in deutscher Sprache

Oliver Lomp
17. Oktober 2016

Unser Gehirn ist überragend gut darin, die überwältigende Menge visueller Informationen zu verarbeiten, die von unseren Augen aufgenommen wird. Dieses Niveau in künstlichen Sehsystemen nachzubilden stellt eine große Herausforderung dar. Aktuell vorherrschende Ansätze bilden dabei die schnellen Vorwärtspfade des Kategorisierungssystems im visuellen Cortex nach. Andere wichtige Fähigkeiten des Sehsystems werden dabei oft nicht modelliert, wie zum Beispiel die Fähigkeit, neue Objekte schnell zu lernen und in neuen Posen wiederzuerkennen. Dies ist zum Teil auf Lücken in unserem Verständnis der funktionellen Architektur biologischer Sehsysteme zurückzuführen. Ein Beispiel für ein Modell dessen Ziel es ist, einige dieser Lücken zu schließen nutzt lokale Histogramme um neue Objekte schnell zu erlernen und wiederzuerkennen während es gleichzeitig ihre Pose schätzt. Dieses System basiert auf dynamischer Feldtheorie, einer mathematischen Methode zur Modellierung kognitiver Prozesse. In meiner Arbeit erweitere ich dieses Objekterkennungsmodell um die Fähigkeit, aufmerksamkeitsbasiert Objekte in einer Szene sequenziell zu betrachten und zu erkennen sowie ein Gedächtnis für die erkannten Identitäten. Für diese Erweiterung integriere ich das Objekterkennungssystem in eine Szenenrepräsentationsarchitektur, die ebenfalls in der Sprache der dynamischen Feldtheorie modelliert ist. Ein Teil dieser Integration besteht daraus, die Prozesse des Objekterkennungssystems so zu erweitern dass rein auf Basis neuronaler Mechanismen mehrere Objekte in Folge betrachtet werden können. Ergebnis dieser Integration ist ein großes dynamisches System das auf den Prinzipien der dynamischen Feldtheorie beruht. Ich evaluiere dieses System mit Hilfe einer eigens erstellten Datenbank und zeige, dass auch vom kombinierten System weiterhin akzeptable Erkennungsraten erreicht werden.

Ich beschreibe außerdem eine Erweiterung, die es dem Objekterkennungssystem erlaubt eine neue Repräsentation zu nutzen, die die räumliche Anordnung der verschiedenen Merkmale der Objekte besser erhält als die zuvor genutzten lokalen Histogramme. Diese neue Repräsentation erdet das System stärker in neuronalen Prinzipien und ersetzt Aspekte des Modells, die zuvor nicht innerhalb der Sprache der dynamischen Feldtheorie gelöst wurden. Außerdem erlaubt sie dem System, die Skalierung des Inputs, sowie die Orientierung über den gesamten Winkelbereich zu schätzen. Weiterhin deckt die Verwendung dieser Repräsentation einige Grenzfälle auf, die bisher unbehandelt waren. Um diese Fälle zu behandeln führe ich Änderungen an den neuronalen Dynamiken des Objekterkennungssystems ein. Ich demonstriere, dass das System mit der neuen Repräsentation bessere Erkennungsleistung liefert als der histogrammbasierte Ansatz sofern vergleichbare Abstraten für die Objektbeschreibungen verwendet werden. Weiterhin analysiere ich die Eigenschaften des Systems um einen Vergleich mit dem menschlichen Sehsystem zu ziehen. Hier zeigt das ansichtsbasierte Objekterkennungssystem poseninvariantes Verhalten, was der in der Literatur gängigen Meinung widerspricht, dass solches Verhalten nur in Systemen auftritt, die nicht ansichtsbasiert sind.