

# Reinforcement Learning for Behaviour from Similarity of Visual Output

Projektgruppe „Praxis der Forschung“  
Sommersemester 2018

## 1 Hintergrund

Im Bereich der Medizinrobotik sind lernende Systeme, die physisch mit ihrer Umgebung interagieren, mit untragbar großen Risiken verbunden. Ein System, das sich noch nicht optimal und sicher verhält, darf nicht in Kontakt mit Patienten kommen. Lernen in der Simulation ist daher von großer Bedeutung, da dort aus Fehlern gelernt werden kann, die in der Realität zu schwere Folgen hätten.

## 2 Projektbeschreibung

Fokus dieser Arbeit soll ein Kameraroboter sein, der eine Szene überwachen soll und seine Position so verändert, dass relevante Teile der Szene immer sichtbar sind. Oft ist es der Fall, dass diese Aufgabe durch Telemanipulation des Roboters durch einen Menschen gut lösbar, aber schwer formalisierbar und damit schwer automatisierbar ist. Moderne Methoden des Deep Reinforcement Learning stellen hierbei einen vielversprechenden Lösungsansatz dar.

Nachdem zunächst in einer Simulation ein Szenario einer Szenenüberwachung implementiert wird, soll das System so designt werden, dass ein Mensch per Telemanipulation das Szenario durchspielen kann und sowohl seine Steuersignale als auch der visuelle Output aufgezeichnet wird.

In einem zweiten Schritt soll ein Agent das Szenario durchspielen und anhand eines Rewards sein Verhalten so ändern, dass es dem Verhalten des Menschen angenähert wird. Hierdurch soll das Verhalten eines Menschen gelernt werden, ohne es explizit formalisieren zu müssen. Der Reward soll sich daran orientieren wie ähnlich die visuellen Outputs des Agenten den visuellen Outputs des Menschen sind (hierbei kann die Bewertung des Bildes sowohl auf Pixel- als auch auf symbolischer Ebene durch eine Vorverarbeitung geschehen, z. B. durch semantische Segmentierung). Die Steuersignale können nachfolgend zur Evaluation genutzt werden.

Als konkreter Anwendungsfall soll ein Szenario im Operationssaal umgesetzt werden, bei dem eine Kamera den Kopf des Patienten dauerhaft beobachten soll. Die Bewegung der Kamera soll hierdurch Verdeckungen verhindern, wenn sich z. B. der Chirurg zwischen Kamera und Kopf stellt.

## 3 Kontakt / Betreuer

Paul Maria Scheikl

[paul.scheikl@kit.edu](mailto:paul.scheikl@kit.edu), Raum 002.2 (Geb. 40.28)