

Erkennen von anforderungsrelevantem Text aus Bildschirmaufnahmen mittels NLP-Methoden

Hintergrund

Mit Eye Tracking bezeichnet man das Aufzeichnen von Blickbewegungen einer Person. In den letzten Jahren wurde diese Methode auch zunehmend im Bereich des Software Engineering genutzt. Die Anwendungen sind vielfältig, wie beispielsweise der Ansatz, Eye Tracking Daten zu nutzen, um automatisch Trace-Links zwischen angeschauten Artefakten zu erkennen. Ein Problem dabei ist aber nach wie vor der große manuelle Aufwand, der die Analyse mit sich bringt. Vor allem dynamische Umgebungen, in denen die Nutzer zwischen Fenstern wechseln, scrollen und Inhalte editieren, sind eine Herausforderung für eine automatische Verarbeitung der Blickdaten.

In einer vorangegangenen Arbeit wurde daher mittels einer OCR (Optical Character Recognition) der Text aus Bildschirmaufnahmen der Eye Tracking Daten automatisch erkannt, um festzustellen, welche Inhalte sich an den x- und y-Koordinaten der Blickpunkte befanden.

Aufgabe

Im Rahmen dieser Arbeit sollen die zuvor erkannten Textbereiche semantisch interpretierbar gemacht werden. Bislang ist damit nur bekannt, in welchen Bildschirmbereichen, zu welchen Zeitpunkten, welcher Text stand. Nicht jeder Text ist aber für die weitere Auswertung wirklich relevant. Die URL im Browser, die Beschriftung eines Buttons oder ein eingeblendeter Hilfetext beispielsweise sind in der Regel nicht von Bedeutung für die weitere Analyse – natürlich abhängig vom konkreten Anwendungsfall.

Für diese Arbeit soll die Texterkennung für die Zielsetzung Trace-Links bzw. Verknüpfungen zwischen in Beziehung stehenden Artefakten in Softwareprojekten mittels Eye Tracking zu erkennen, optimiert werden. Das heißt, es müssen irrelevante Textbereiche herausgefiltert werden und die Granularität erkannter Textbereiche so angepasst werden, dass die Textlänge sinnvoll für die weitere Auswertung ist.

Die konkreten Aufgaben sind:

- Sichtung einschlägiger Literatur über verwandte Ansätze
- Konzeption geeigneter Kriterien und Heuristiken zur Filterung und Verfeinerung der erkannten Textbereiche – dazu Nutzung von Natural Language Processing (NLP)
- Evaluation des entwickelten Konzepts anhand einer kleinen Beispielstudie, in der Eye Tracking Aufnahmen bei der Arbeit mit Anforderungsartefakten aufgezeichnet werden
- Auswertung der Daten
- Kritische Bewertung der Ergebnisse und Diskussion möglicher Verbesserungen

Die Ausarbeitung kann auf Deutsch oder Englisch angefertigt werden. Bei der Bewertung wird besonderen Wert auf die methodische Umsetzung, sowie die gewissenhafte und wissenschaftlich nachvollziehbare Bearbeitung und Dokumentierung gelegt.

Organisatorisches

Betreuer: M.Sc. Maike Ahrens, maike.ahrens@inf.uni-hannover.de, Raum G304, M.Sc. Martin Obaidi, martin.obaidi@inf.uni-hannover.de, Raum G302

Prüfer: Prof. Dr. Kurt Schneider

Beginn: ab sofort möglich