

Diss. ETH No. 28704

# **Towards Advanced User Guidance and Context Awareness in Augmented Reality-guided Procedures**

A thesis submitted to attain the degree of

Doctor of Sciences of ETH Zurich  
(Dr. sc. ETH Zurich)

presented by

Julian Christian Wolf

M.Sc. Karlsruhe Institute of Technology (KIT)  
born on 10.05.1990  
citizen of Germany

Prof. Dr.-Ing. Mirko Meboldt, examiner  
Prof. Dr. Philipp Fürnstahl, co-examiner

2022

# Abstract

Procedural tasks are common in many professions, such as maintenance, assembly, or surgery, and are characterized by an operator performing a predefined sequence of steps to achieve a specific goal. Because these tasks often involve elaborated machines, devices, or even patients, they place the highest requirements on correct task execution.

Augmented reality (AR) head-mounted displays (HMDs) have been shown to provide effective support during procedural tasks. Compared to conventional information mediums, where information is often spread among multiple documents (e.g., maintenance) or external screens (e.g., surgery), AR HMDs display contextual information directly into the field of view of the operator without occupying the operators' hands. While with AR, displayed information is only changed based on manual user input, context-aware AR promises to further improve the support provided by automatically adapting displayed information to best address the operator's current needs and by providing feedback. Understanding the strengths and weaknesses of these two technologies is key to developing support systems that can improve the quality of task execution, making procedural tasks safer and improving outcomes. Previous studies on context-aware systems have focused primarily on manual execution without consideration of an important part of human interaction, the perception. Eye tracking allows to measure perception and provides deep insights into cognitive processes, and might therefore bring benefits to context-aware systems that are important to be investigated.

This work investigates different concepts of how AR and context-aware AR support systems can be designed, how they work, and how they affect operators' task performance. It further aims to advance context-aware AR support by integrating eye tracking and by deriving a suitable system model to describe the relationships between human behavior, AR, and context-aware AR. Three studies are presented in this work.

Study I investigates the benefits of contextual information in AR over traditional information mediums to provide training instructions. A study

## *Abstract*

was conducted with 21 medical students performing an extracorporeal membrane oxygenation (ECMO) cannulation on a physical simulator setup. The evaluation comprised of a detailed error protocol with both a categorization into knowledge- and handling-related errors and an error severity ranking. The results showed clear benefits of AR over conventional instructions while pointing out certain limitations that might be improved by context-aware AR.

Study II investigates effective visualization strategies when real-time feedback is provided continuously. A study was conducted with 4 expert surgeons and 10 surgical residents performing surgical drilling on a physical simulator setup. The results show that continuous performance feedback generally levels task performance between novice and expert operators, reveal clear advantages and preferences of certain AR visualizations, and give insights into how AR visualizations guide visual attention. In particular, the peripheral field around the area of execution proved to be promising for displaying information as the operator can simultaneously perceive feedback and coordinate hand movement.

Study III investigates the suitability of eye and hand tracking features for predicting and preventing an operator's erroneous actions. A study was conducted on a memory card game to explore the potential and limitations of this approach. The first experiment, which involved 10 participants, recorded participants' eye and hand movement to derive a method for target prediction. The second experiment with 12 participants examined the timeliness and accuracy of the implemented method end-to-end and showed the method to be highly effective in preventing a user's erroneous hand actions.

One of the key conclusions of this work is that context-aware AR support can significantly improve procedural outcomes and even raise the task performance of less experienced operators to the level of experts. In addition, analyzing hand-eye coordination patterns in real-time allows for predictive AR support and error prevention, which might eventually provide a safety net for operators performing their first independent task executions. For future work, important research directions include integrating and advancing predictive AR support for more complex

procedures, investigating effective visualization strategies in environments with multiple dynamic visual stimuli, as well as effective feedback and support strategies while operators transition from their first training to independent execution and eventually become experts.

# Zusammenfassung

Prozedurale Aufgaben sind in vielen Berufen, wie z. B. in der Wartung, Montage oder Chirurgie, üblich und zeichnen sich dadurch aus, dass ein Bediener eine vordefinierte Abfolge von Schritten ausführt, um ein bestimmtes Ziel zu erreichen. Da diese Aufgaben oft in Wechselwirkung mit komplexen Maschinen, Geräten oder sogar am Patienten durchgeführt werden, werden allerhöchste Anforderungen an die korrekte Ausführung gestellt.

Augmented Reality (AR) Head-Mounted Displays (HMDs) können eine wirksame Unterstützung bei prozeduralen Aufgaben bieten. Im Vergleich zu herkömmlichen Informationsmedien, bei denen die Informationen oft auf mehrere Dokumente (z.B. bei der Wartung) oder externe Bildschirme (z.B. in der Chirurgie) verteilt sind, zeigen AR HMDs kontextbezogene Informationen direkt im Sichtfeld des Bedieners an, ohne dabei die Hände zu blockieren. Während bei AR die angezeigten Informationen nur auf Grundlage manueller Benutzereingaben geändert werden, versprechen kontextbewusste AR-Systeme eine weitere Verbesserung der Unterstützung, indem sie die angezeigten Informationen automatisch an die aktuellen Bedürfnisse des Bedieners anpassen und den Bediener durch Feedback unterstützen. Ein Verständnis der Stärken und Schwächen dieser beiden Technologien ist der Schlüssel zur Entwicklung von Unterstützungssystemen, die die Qualität der Aufgabenausführung verbessern können, um prozedurale Aufgaben sicherer zu machen und die Endresultate zu verbessern. Frühere Studien über kontextbewusste AR-Systeme haben sich in erster Linie auf die manuelle Ausführung konzentriert und dabei einen wichtigen Teil der menschlichen Interaktion, nämlich die Wahrnehmung, außer Acht gelassen. Eye Tracking ermöglicht die Messung der Wahrnehmung und bietet tiefe Einblicke in kognitive Prozesse und könnte daher Vorteile für kontextbewusste Systeme bringen, die es zu untersuchen gilt.

Das Ziel dieser Dissertation ist es, zu untersuchen, wie sich AR und kontextbewusste AR Systeme gestalten lassen, wie sie funktionieren, und

wie sie die Ausführungsqualität des Bedieners beeinflussen. Dabei soll die kontextbewusste AR-Unterstützung insbesondere auch durch die Integration von Eye-Tracking und die Herleitung eines geeigneten Systemmodells vorangetrieben werden. In dieser Arbeit werden drei Studien vorgestellt.

Studie I untersucht die Vorteile kontextbezogener Informationen in AR zum Vermitteln von Ausführungsinstruktionen gegenüber herkömmlichen Informationsmedien. Eine Studie wurde mit 21 Medizinstudenten durchgeführt, die eine extrakorporale Membranoxygenierung (ECMO) an einem physikalischen Simulator durchführten. Die Auswertung umfasste ein detailliertes Fehlerprotokoll mit einer Kategorisierung in wissens- und handhabungsbezogene Fehler sowie eine Einstufung des Fehlerschweregrads. Die Ergebnisse zeigen deutliche Vorteile von AR gegenüber konventionellen Anleitungen und weisen gleichzeitig auf bestimmte Limitationen hin, die durch kontextbewusstes AR verbessert werden könnten.

Studie II untersucht effektive Visualisierungsstrategien für den Fall, dass Echtzeit-Feedback kontinuierlich bereitgestellt wird. Es wurde eine Studie mit vier erfahrenen Chirurgen und zehn Assistenzärzten durchgeführt, wobei Bohrungen an einem physikalischen Wirbelsäulenmodell gesetzt werden mussten. Die Ergebnisse zeigen, dass kontinuierliches Feedback im Allgemeinen die Aufgabenleistung zwischen Anfängern und Experten angleicht, zeigen klare Vorteile und Präferenzen bestimmter AR-Visualisierungen und geben Einblicke in die Art und Weise, wie AR-Visualisierungen die visuelle Aufmerksamkeit lenken. Insbesondere das periphere Feld um den Ausführungsbereich erwies sich als vielversprechend für die Anzeige von Informationen, da der Bediener gleichzeitig Feedback wahrnehmen und die Handbewegung koordinieren kann.

Studie III untersucht die Eignung von Blick- und Handbewegungsmessungen zur Vorhersage und Vermeidung von Fehlhandlungen des Bedieners bei prozeduralen Aufgaben. Um das Potenzial und die Grenzen dieses Ansatzes zu untersuchen, wurde eine Studie mit einem Memory-Kartenspiel durchgeführt. Im ersten

## *Zusammenfassung*

Experiment mit 10 Teilnehmern wurden nur die Augen- und Handbewegungen der Teilnehmer aufgezeichnet, um eine Vorhersagemethode herzuleiten. Das zweite Experiment mit 12 Teilnehmern untersuchte die Geschwindigkeit und Genauigkeit der implementierten Methode mit Nutzern und zeigte, dass die Methode sehr effektiv darin ist fehlerhafte Handaktionen des Benutzers rechtzeitig zu verhindern.

Eine der wichtigsten Schlussfolgerungen dieser Arbeit ist, dass eine kontextbewusste AR-Unterstützung die Ergebnisse erheblich verbessern und sogar die Aufgabenleistung von weniger erfahrenen Bedienern auf das Niveau von Experten anheben kann. Darüber hinaus ermöglicht die Echtzeitanalyse von Hand-Augen-Koordinationsmustern eine vorausschauende AR-Unterstützung und Fehlervermeidung, die zukünftigen Bedienern ein Sicherheitsnetz bieten könnte, wenn diese zum ersten Mal eigenständig eine neue Aufgabe ausführen müssen. Wichtige Forschungsrichtungen für zukünftige Arbeiten sind die Integration und Weiterentwicklung von präventiver AR-Unterstützung für komplexere Prozeduren, die Untersuchung effektiver Visualisierungsstrategien in Umgebungen mit mehreren dynamischen visuellen Stimuli, sowie effektive Feedback- und Unterstützungsstrategien, die den Bediener vom ersten Training bis zur eigenständigen Ausführung bestmöglich unterstützen.