

# ERmed

26. November 2012

Cebit Innovation  
Award

Beitrag:

ERmed -  
Erweiterte Realität in  
der Medizin



## Erweiterte Realität in der Medizin ERmed

**Autor: Dr. Daniel Sonntag**

Was kommt als nächstes? Was als übernächstes? Die Antwort auf beide Fragen ist einfach: **Erweiterte Realität**. Diese bedeutet für uns neue visuelle Möglichkeiten mit Projektionen auf Datenbrillen, kombiniert mit der Dateneingabe durch Augenbewegungen und digitalem Stift. Und ganz wichtig: natürlicher Sprach-Dialog mit digitalen Assistenten. Wir glauben, dass diese Techniken die Zukunft prägen werden. Auch Google glaubt daran und stellt Konzepte vor: "Die Daten-Brille ist das nächste große Ding." Wir sind aber schon weiter: Mit ERmed entsteht ein multimodaler Dialog für die mobile Medizinversorgung. Völlig

neue Aspekte mobiler, intelligenter und ambienter Umgebungen für den Arzt werden bereitgestellt, damit dieser sich optimal auf den Patienten einstellen kann. Als Ergebnis steht ein multimodales Interaktionssystem zur Verfügung.

Der Arzt wird aktiv in den Design-Prozess eingebunden. Die Echtzeit-Schnittstelle für den medizinischen Benutzer in Form der visuellen Unterstützung durch die Brillenprojektion wird kontextsensitiv und patientenspezifisch gestaltet.





## Einleitung

Wir verfolgen einen neuen Ansatz des human-centred Designs, in dem das deskriptive Wissen über die anvisierten Benutzer des Interaktionssystems ein zentrales Element bildet: Neben dem Wissen über die Bedürfnisse des Arztes und des Patienten, spezielle medizinische Abläufe und die physische Umgebung der Benutzer und beteiligten Personen, die eine aktive Rolle im Designprozess einnehmen, verfolgen wir ein neues Konzept der Multimodalität, um sprachliche, visuelle und haptische Modalitäten optimal als Ein- und Ausgabekanäle zu nutzen. Neue visuelle Darstellungsmethoden finden sich in den zur Verfügung stehenden Echtzeitmethoden der erweiterten Realität, um Objekte und Patientengesichter in Echtzeit zu erkennen. Entsprechende Zusatzdaten werden direkt in eine elektronische Brille projiziert. Zu den teils physischen, teils digitalen Artefakten unserer erweiterten Realität zählt auch das digitale Papier als neue Eingabemöglichkeit, ein sprachbasiertes Dialogsystem, und ein Head-Mounted Display (HMD) in einer Arztbrille, die auch Augenbewegungen lesen kann.

Als Ergebnis steht ein multimodales Interaktionssystem bei Anwesenheit der Patienten zur Verfügung. Hauptbestandteil ist die visuelle Unterstützung durch die Brillenprojektion, mit deren Hilfe die Patientenbefundung kontextsensitiv und patientenspezifisch gestaltet wird.

Die wissenschaftlichen Forschungspotentiale ergeben sich an der Schnittstelle zwischen Mensch-



**Erweiterte Realität schließt erweiterte visuelle Möglichkeiten mit Projektionen auf Brillen, die Dateneingabe mit digitalem Stift und natürlichen Sprachdialog mit digitalen Assistenten mit ein. Es entsteht ein natürlicher multimodaler Dialog für mobile Medizinversorgung. Wir stellen intelligente ambiente Umgebungen für den Arzt bereit, damit dieser sich optimal auf den Patienten einstellen kann.**

Maschine-Interaktion und der kognitiven Psychologie als „ambientes patientengerechtes Arbeiten“, sowie der technischen Schnittstelle der Echtzeit-Detektoren und Synchronisation einzelner Eingabe- und Ausgabekanäle: „synchronisierte Echtzeit-Erweiterung“.

Einen Zusatzaspekt bildet die Anwendung mobiler Technik-Komponenten: man erzeugt eine erweiterte Realität als intelligente Umgebung, ohne dass aufwändige oder ungewünschte Sensor-Festinstallationen nötig werden: „Mobile Intelligenz als Arbeitsumgebung“.

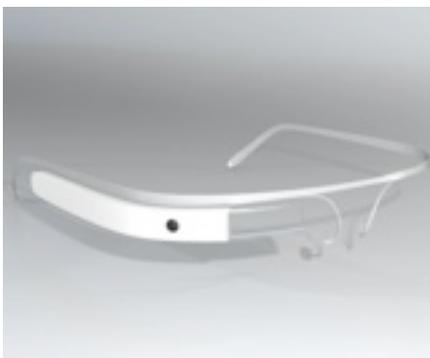
### Das Team:

Daniel Sonntag (Interaktionsdesign und Usability), Christian Husodo Schulz (Doktorand, Dialogsystem), Markus Weber (Doktorand, Gesten-Interaktion und Stifteingabe), Takumi Toyama (Doktorand, Head-Mounted Display), Mikhail Blinov (Graphik-Design), Tobias Zimmermann (Diplomand, digitale Annotierung), Matthias Hammon (Radiologe), Alexander Cavallaro (Radiologe)

### Projekt-Koordination:

Dr. Daniel Sonntag

Deutsches Forschungszentrum für Künstliche Intelligenz



### Design-Studie einer Projektions-Brille

Geeignete Brillen für die Projektion digitaler Inhalte sind in Entwicklung. Die Technologie für das Designkonzept (Quelle Google) wird in wenigen Jahren zur Verfügung stehen.

Das Vorgängerprojekt RadSpeech wurde aus 130 THESEUS-Demonstrationen ausgesucht, um Bundeswirtschaftsminister Dr. Philipp Rösler bei einem Rundgang auf der Computermesse CeBIT 2012 in Hannover präsentiert zu werden und hat außerdem den German HighTech Award 2011 gewonnen. Ein YouTube-Video erläutert die Bedienbarkeit von RadSpeech und damit die Grundlage von ERmed, per Sprachdialog Patientinformationen anzufordern.

<http://www.youtube.com/watch?v=6TANtNdRR6U>

Mit RadSpeech als Basissystem geht es in ERmed, um die Künstliche Intelligenz (KI)-gestützte Datenakquisition im Patientenumfeld. In ERmed wird der Sprachdialog erweitert, um dem Arzt weitere Interaktionsmöglichkeiten und Kontrollmöglichkeiten über neue Ein- und Ausgabekanäle wie die Stifteingabe oder Datenvisualisierung in der Arztbrille zu geben. Das erweiterte Interaktionsdesign basiert also auf physischen anthropozentrischen Interaktions-Objekten, die digital veredelt werden: Arztbrille, Stift, Schreibunterlage (Pad) mit Papier (und die digitale Armbanduhr, Secret!).



## „ERmed – Künstliche Intelligenz, Mobiler Sprachdialog und Erweiterte Realität mit digitaler Vernetzung im Dienste der Medizin“

Wir zeigen an einem medizinischen Beispiel auf, wie HMDs zu einer Unterstützung der Arbeitsabläufe führen können. Datenbrillen wurden bereits in verwandten Forschungsprojekten in der computergestützten Chirurgie, der Gynäkologie und der Rehabilitation prototypisch eingesetzt. Bei ERmed geht es erstmals um die computergestützte klinische Entscheidungsfindung und Diagnose. Durch die intuitive Benutzbarkeit unterstützt ERmed den Arbeitsablauf des Arztes und kann in Situationen genutzt werden, in denen keine feste Arbeitsstation zur Verfügung steht, z.B. bei der Visite oder in Besprechungen. Vor allem aber während der Patientenuntersuchung und während des anschließenden Patientengesprächs, wo feste Arbeitsstationen Störfaktoren sind. Mit herkömmlichen Benutzeroberflächen auf Arbeitsstationen können behandelnde Ärzte zwar Patientendaten

### ERmed - Ambientes patientengerechtes Arbeiten (Human Aspects):

- Ästhetik
- Perzeption
- Mentales Modell
- Soziale Umgebung
- Entscheidungsunterstützung
- Multimodale Dokumentation
- Entscheidungs-Feedback und Patienteninfo (patient grounding)

### Synchronisierte Echtzeit-Erweiterung:

- Menschliche Informationsaufnahme
- Menschliche Reaktionen
- Entscheidungsfindung
- Multimodalität
- Interaktivität

### Mobile Intelligenz als Arbeitsumgebung:

- Arbeitsgedächtnis und Arbeitsumgebung
- Erweiterte Kognition
- Physikalische Freiheit
- Gewonnene mentale Parallelverarbeitung



abrufen oder neue Daten einpflegen. Bei der Interpretation der Bilddaten allerdings bieten diese Oberflächen keine weitergehende Assistenz an. Außerdem kommt der Patient oft zu kurz, da der Arzt sich zu sehr auf das Bedienen des Computers konzentrieren muss. Die kritisierte mangelnde Ergonomie und Integrationsschwierigkeit bestimmen den Einsatz aktueller Computer-systeme im medizinischen Alltag. In ERmed wird daher versucht, die nächste Generation intelligenter, skalierbarer und intuitiver Benutzerschnittstellen in medizinischen Bereichen, und vor allem bei Patientenuntersuchungen, zu entwickeln. Wir setzen darauf, erweiterte Realität für die Patientenroutine anwendbar zu machen. Dabei gehen wir noch einen Schritt weiter, als Google es im neusten Konzeptvideo im Freizeitbereich andeutet, <http://www.technik-aktuell.com/augmented-reality-google-brille>:

# Intelligente KI-gestützte Datenakquisition mit digitalem Stift und erweiterter Realität

Denn in ERmed wird multimodale erweiterte Realität für sensible Patientenuntersuchungen und Patientengespräche sowie Arztkonferenzen anwendungs- und gesellschaftstauglich gemacht!

**Es geht uns in ERmed vor allem um die vier Punkte, die bei RadSpeech schon wichtig und entscheidend für den Erfolg waren:**

## Digitaler Stift

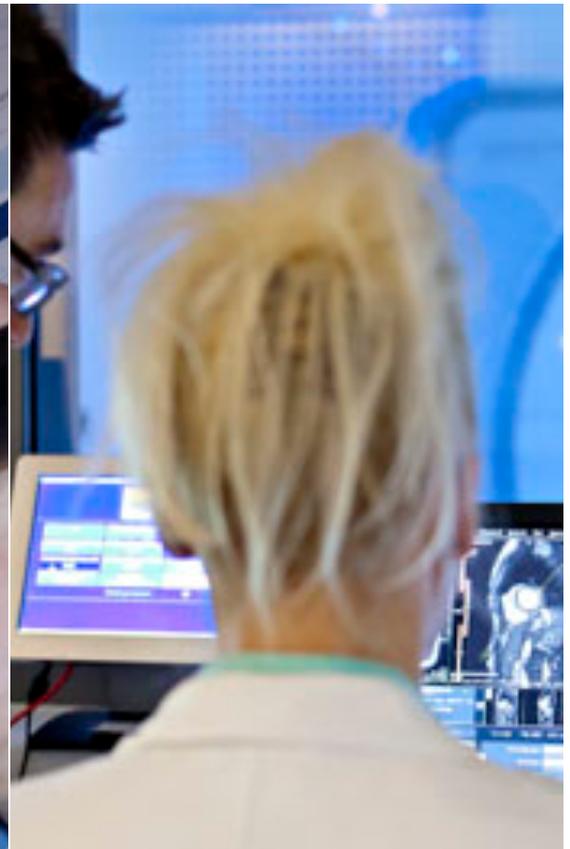
Der digitale Stift ist ein stiftähnliches Eingabegerät, mit dem man auf normalem Papier schreiben kann. Dieses Papier kann direkt in medizinischen Abteilungen eingesetzt werden. Dabei setzen wir auf die Technologie der Firma Anoto, die den Stift mit einer Bluetooth Funktion vertreibt. Mithilfe dieser Zulieferer-Technologie lassen sich bestimmte Bögen für medizinische Anwendungen kreieren, bei denen Patientenuntersuchungen festgehalten werden. Einen solchen komplexen Bogen konnten wir bereits prototypisch für die Mammographie entwickeln. Dabei stellt das Design des Bogens und die softwaretechnische Erweiterung, Daten in Echtzeit zu übertragen, einen Wettbewerbsvorteil gegen-

- Aufbauend auf den MEDICO\*-Zielen medizinische Bildinformationen automatisch zu erkennen und zu strukturieren, und die strukturierten semantischen Daten auch für intelligente Benutzerinteraktionen zu nutzen.
- Sprachbasierte multimodale Entscheidungshilfe bei diagnostischen und therapeutischen Entscheidungen zu leisten.
- Schneller und effektiver auf medizinisches Fachwissen und Patientendaten der Vorsorgeuntersuchungen zuzugreifen.
- Die Interaktion mit dem Benutzer insgesamt intuitiver und intelligenter zu gestalten.

*\*MEDICO ist ein Anwendungsszenario des Großprojekts THESEUS (2007-2012), in dem rund 30 öffentliche und industrielle Forschungspartner an Basistechnologien und Standards für die nächste Generation des Internets arbeiten ([www.theseus-programm.de](http://www.theseus-programm.de)). Das Bundeswirtschaftsministerium fördert THESEUS mit rund 100 Millionen Euro. Nochmals den gleichen Betrag müssen die beteiligten Industriepartner aufbringen.*

### Kommen wir nun zur Vision!

Wir verknüpfen RadSpeech (ohne iPad!) mit dem digitalen Stift und einer AR (Augmented Reality) Projektionsbrille. In ERmed wird der Sprachdialog in Richtung Meta-Dialog erweitert werden, um dem Arzt mehr Kontrolle über die Stifteingabe und Datenvisualisierung in der Brille zu geben. Somit kann man per Sprache verschiedene Patientenakten und Mammographiebögen aus Vorsorgeuntersuchungen in verschiedenen Modalitäten (Röntgen, Sonographie, MRT) öffnen, Korrekturen der Stifteingabe machen und Behandlungsempfehlungen aussprechen, sowie den leitenden Arzt zur Zweitdiagnose miteinbeziehen.



über Konkurrenten dar (die Technologie ist momentan nur am DFKI verfügbar). Der Bogen wurde bereits im Universitätsklinikum Erlangen getestet, und als Ergebnis steht eine Evaluationsstudie zur Verfügung, anhand derer nun weitere Optimierungen für den erweiterten funktionsfähigen Cebit 2013 Demonstrator ERmed, Erweiterte Realität für mobile Medizin, angegangen wurden.

Dabei dient der HMD als real-time Ausgabemedium. In diesem Zusammenhang ist das entscheidende Merkmal, dass wir auf den Desktop-Computer zur Kontrolle der Eingabe verzichten können. Die gesamte Sprach- und Stift-basierte Interaktion mündet in ein neues Interaktions-Konzept mittels erweiterter Realität und Augmented Vision (AV).

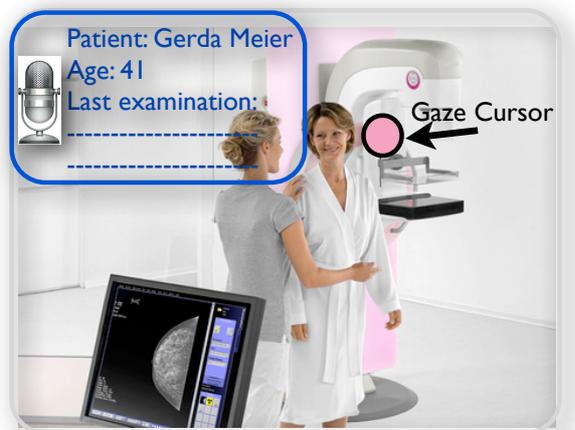
Sensor-Technologien zum Motion-Tracking und Gestenerkennung stehen ebenfalls zur Verfügung. Hinzu kommt ein digitaler Stift, der natürliche Handschrift und Schreibgesten interpretiert und in Echtzeit in der AV-Umgebung Text- und Bild Darstellungen erlaubt, sowie ein Sprachdialogsystem, mit dem man über Artefakte in der AV-Umgebung einen Dialog führen kann.

Anhand des mit den Klinikern erarbeiteten Interaktionsdesigns setzen wir zusätzlich auf mobile Eyetracker-Technologie, um Augenbewegungen als Anwendungsspezifisches Eingabemedium zu verwenden; die Blickposition (gaze cursor) bestimmt die kontextabhängige Anzeige im mobilen HMD.

## Erweiterte Realität

Erweiterte Realität kann mehrere Formen haben. Wir gehen davon aus, dass dabei virtuelle Realität und die Realität miteinander kombiniert und teilweise überlagert sind, dass Interaktivität in Echtzeit möglich ist, und dass reale und virtuelle Objekte zueinander in Bezug stehen. Augmented Vision und Sensor-Fusion stehen dabei im Mittelpunkt und bieten Technologien, die ins medizinische Umfeld übertragen werden:

<http://av.dfki.de/images/stories/Video/ARHandbook-InstallingRAM.flv>



Mobile Eye Tracker



Mobile HMD



Ein entferntes Ziel erweiterter Realität ist es, ohne Projektionsbrille ein holographisches digitales Bild in der Umgebung erzeugen zu können. Schon heute stehen technische Möglichkeiten bereit, erweiterte Realität konzeptionell umzusetzen, z.B. die Projektion digitaler Artefakte in den Blickbereich des Benutzers in Kombination mit Aktivitätserkennung und Gestenkontrolle. Das Gleiche gilt für Sprachdialogsysteme mit robuster Sprach-Erkennung und täuschend echter Sprachsynthese.



# Intelligente multimodale Interaktion in der Medizin: die klinische Sicht

Heutzutage sind medizinische Bilder in der Krankheits-Diagnose und Differenzierung, Eingriffsplanung oder zur Überwachung der Be-

Wir stellen nun unser spezifisches Test-Anwendungsfeld, die Mammographie, vor:

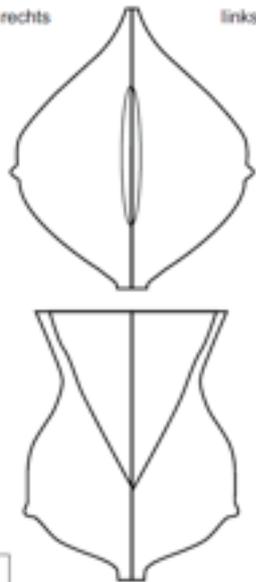
Mammographie ist eine Röntgenmethode zur Früherkennung von Brustkrebs (Mammakarzinom), die in den meisten Ländern häufigste Krebserkrankung der Frau. Die Mammographie ist in erster Linie ein Verfahren der Radiologie zur Diagnostik der weiblichen, gegebenenfalls aber auch der männlichen Brust, und wird als Synonym für die Röntgenmammographie verwendet, die durch eine klinische Untersuchung, die Sonographie und ggf. durch MRT ergänzt werden (ERmed bildet alle Szenarien ab).

**Der Faktor Mensch ist einer von vielen Faktoren, industrielle Assistenz-Systeme zur kognitiven Verstärkung und physischen Entlastung erfolgreich in Medizinanwendungen zu integrieren. Multimodale industrielle Assistenzsysteme rücken somit in den Mittelpunkt unserer Forschung in der Mensch-Maschine Kommunikation für die Medizin. Sie bieten eine Möglichkeit, der individualisierten Medizin Vorschub zu leisten. Denn gerade im Medizinbereich ist es wichtig, bei steigender Autonomisierung durch KI-Methoden die Ärzte bestmöglich einzubeziehen.**

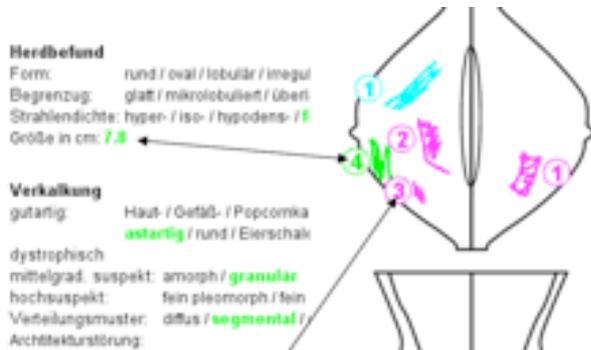
Dass neuerdings in vielen Krankenhäusern das strukturierte Berichten eingeführt wird, hat zur Folge, dass vordefinierte (elektronische) Formulare verwendet werden müssen. Dies empfinden beispielsweise Radiologen als Behinderung und fürchten, dass sie von den Bildern und dem Patienten zu sehr abgelenkt werden. Das strukturierte Berichten besteht im Wesentlichen aus dem Ausfüllen eines Papierbogens, der dann abgetippt wird und zur Kontrolle und Unterschrift nach 24-48 stündiger Verzögerung zu Arzt zurückkommt.

handlung jedes einzelnen Patienten unverzichtbar geworden. Um die KI-gestützte Diagnose realisieren zu können, muss die nötige Bildsemantik auch im Bild annotiert werden. Wie jedoch wie lässt sich dies in die radiologische Arbeit integrieren ohne den Arbeitsablauf des Arztes zu sehr zu stören? Zur Beantwortung dieser Frage ist der Faktor Mensch in einem bestimmten Anwendungsfeld entscheidend. Er soll die Schreibkultur behalten dürfen, und die funktionale Brille (und die Uhr, Surprise) sollen ihre Symbolik nicht verlieren, sondern funktional erweitert werden.

Mit ERmed soll dem Radiologen zukünftig ermöglicht werden, den digitalen Stift zu benutzen, um direkt Eingaben in digitaler Form zu tätigen und MRT Bildannotationen mit natürlichem Sprachdialog zu ergänzen bzw. zu korrigieren. Diese intelligente Interaktion wird kombiniert mit der Anzeige des letzten Voruntersuchungsbogens in der AR Brille. Zusätzlich wird das Resultat der momentanen Befundung (siehe Beispielbogen unten) in Echtzeit in die Brille projiziert.

<p><b>Dichtetyp:</b> I. lipomatös II. fibroglandulär III. inhomogen dicht IV. extrem dicht</p> <p><b>Herdbefund</b> Form: rund / oval / lobulär / irregulär Begrenzung: glatt / mikrolobuliert / überlagert / unscharf / spikuliert Strahlendichte: hyper- / iso- / hypodens / fettäquivalent Größe in cm: <input style="width: 80px;" type="text"/></p> <p><b>Verkalkung</b> gutartig: Haut- / Gefäß- / Popcornkalk                   astartig / rund / Eierschalen- / Kalkmilchkalk dystrophisch mittelgrad. suspekt: amorph / granulär hochsuspekt: fein pleomorph / fein linear / -verästelt Verteilungsmuster: diffus / segmental / regional / linear / gruppiert</p> <p>Architekturstörung: <input style="width: 100px;" type="text"/></p> <p>Spezialfälle: <input style="width: 100px;" type="text"/></p> <p>Lokalisation (cm v. Mamille / h): <input style="width: 100px;" type="text"/></p>	<p><b>Mammographie</b></p> <p>rechts                      links</p> 	<p><b>Dichtetyp:</b> I. lipomatös II. fibroglandulär III. inhomogen IV. extrem d</p> <p><b>Herdbefund</b> Form: rund / oval / lobulär / Begrenzung: glatt / mikrolot Strahlendichte: hyper- / iso- Größe in cm: <input style="width: 80px;" type="text"/></p> <p><b>Verkalkung</b> gutartig: Haut- / Gefäß- / Po                   astartig / rund dystrophisch mittelgrad. suspekt: amorph hochsuspekt: fein pleomorph Verteilungsmuster: diffus / s</p> <p>Architekturstörung: <input style="width: 100px;" type="text"/></p> <p>Spezialfälle: <input style="width: 100px;" type="text"/></p> <p>Lokalisation (cm v. Mamille /</p>
1		1

# Die Hauptsache ist, was der Arzt in der Brille sieht: die Voruntersuchung, das reale Bild und die Erkennung des neuen digitalen Befundes in Echtzeit:



## Aspekte: Perzeption, Mentales Modell, Erweiterte Kognition

Die Umsetzung erfolgt durch die technischen Methoden der AV Umgebung. Menschen sind sehr gut darin, schnelle visuelle Abläufe zu verfolgen ohne die Konzentration zu verlieren. Die beste Unterstützung durch die technischen Hilfsmittel soll darin bestehen, ablenkende Annotierungsaufgaben am strukturierten Patienten-Bogen zu reduzieren und dabei die Komplexität zu nehmen, sowie die kognitive Belastung bei Korrekturen zu verringern. Durch das Real-Time Tagging mit dem Stift, dessen Ergebnis in der Brille visualisiert wird, wird eine deutliche Effizienzsteigerung erreicht, obwohl das manuelle Annotierungsverfahren insgesamt nicht skalierbar ist (es sind keine Weiterverarbeitungsprozesse möglich, bevor das Ergebnis nicht (direkt) vom Arzt überprüft wird). Zusätzlich kann ein zusammenhängendes Dialog-Verhalten implementiert werden, ohne die Überprüfung der Angaben zu stören.

Durch Aktivitätserkennung (des Echtzeit-Videobildes) wird der spezielle Bereich im Annotierungsbogen und der Arzttroutine in der HMD-Visualisierung in den Vordergrund gebracht, der der aktuellen Aktivität des Arztes entspricht: klinische Untersuchung am Patienten, Aufnahme Röntgen, Kontrolle Röntgenbilder, Kontrolle MRT Bilder, oder Sonographie. Diese Bereiche können mehrmals bei einer Patientenbegutachtung gewechselt werden.

### Gesten-basierte Erweiterungen und Embodied Interaction

Räumliche Interaktionssysteme, die mit 3D Zeige-Gesten und verkörperter Interaktion meistens zur 3D Navigation eingesetzt werden, erleben nicht umsonst einen Siegeszug als neue Benutzerkonzepte und Interaktionsformen: Unser Leben ist geprägt von räumlichen Eindrücken, und der evolutionär gewachsene kognitive Verstehens-Apparat kommt gut zurecht mit räumlichen



Konfigurationen und Interaktions-Metaphern, die räumlich ausgedrückt werden können. Bei Suchaufgaben, die einen analytischen Charakter haben, lassen sich potentielle Anwendungsfelder leicht finden, z.B. die Suche in mehreren Bildstapeln.

Andererseits ist die momentane Interaktions-Forschung in diesem Gebiet, und die à la TED Talk, Minority Report und damit verbundener „Mission Impossible“ gearteter Showeffekt-hascherischer Präsentation neuer Gesteninteraktionen kritisch zu beurteilen: In den meisten Fällen wird kein realistisches (kollaboratives) Einsatzszenario aufgezeigt, das den Einsatz der neuen Technologie überhaupt rechtfertigen könnte. Mit anderen Worten, viele der gezeigten Anwendungen haben wenig oder gar nichts mit human-centred Design zu tun, da man keine vom speziellen Anwender ausgehende Motivation für ein Task-basiertes Interaktionsverhalten mit Gesten feststellen kann. Aus diesem Grund ist die Anwendung im medizinischen Bereich ganz besonders auf Usability-Aspekte hin zu prüfen: Mit dem neuen Leap Motion Sensor ließe sich eine Gesten-basierte Interaktion mit den augmentierten 3D Modellen relativ einfach umsetzen. Zwar gibt es Bedienungsstationen im OP, die man bislang aufgrund der fehlenden Genauigkeit anderer Sensoren nicht umsetzen konnte, doch einziges Argument für eine Gestensteuerung im OP ist die Sterilität beim operativen Eingriff.

In unserem Anwendungsszenario Mammographie lassen sich MRT Daten potentiell in die AV Brille und/oder einem großen Befundungsbildschirm (RadComet) in den Arzt-Patienten-Szenarien (mit AV Unterstützung) einspielen. Doch würden dadurch wichtige Funktionen des beschriebenen patientengerechten Arbeitens verletzt. Die soziale Interaktion am und mit dem Patienten bei der Untersuchung erlaubt keine Gestenkontrolle, die der Situation angemessen ist. Auch muss beispielsweise das Sonographie-Gerät gesteuert werden, das der Arzt in einer Hand hält. Ein gleichzeitiges Gesten-basiertes Annotieren der Befunde ist dadurch nicht möglich. Demzufolge ist das momentan herrschende Szenario mit Papierbogen und Kugelschreiber ein viel besseres (asynchrones) Eingabemittel für klinische Befunde. Die Erweiterung durch den digitalen Stift und direkter Digitalisierung in der AV Brille knüpft nahtlos daran an.

Die verkörperte Interaktion im Patientengespräch als Kombination aus Bewegungsgesten und Sprache findet allerdings ein anderes Anwendungsszenario im Sinne einer kollaborativen Multimodalität (collaborative multimodality):

Schwierige und schwer zu deutende Fälle werden immer vom Leiter der radiologischen Abteilung oder Frauenklinik mit dem behandelnden Arzt in der Routine besprochen. Die Gesteninteraktion komplettiert hier das Arbeiten mit digitalem Stift, Sprachdialog und der AV Brille in sinnvoller Weise: alle MRT-, Sonographie- und Röntgenaufnahmen der Vorsorgeuntersuchungen werden zusammen mit den ausgefüllten digitalen Bögen geöffnet und gemeinsam inspiziert, wobei eine zusätzliche Suchaufgabe mit analytischem Charakter entsteht. Dafür bietet sich eine Gesten-basierte Interaktion am Großbildschirm an, wobei der Patient (sofern er dabei ist) eine passive Betrachter-Rolle einnimmt. Neben dem Aspekt der sozialen Umgebung ist auch Entscheidungsfeedback und gleichzeitige Patienteninformation (patient grounding) gewährleistet.

Dieses erweiterte Szenario wurde ebenfalls mit neuester DFKI-Technologie in ERmed umgesetzt; momentan wird der WiiMote Controller für Handgesten (Pose und bewegte Gesten) verwendet, der Anfang 2013 durch den viel präziseren Leap-Motion Sensor ausgetauscht werden kann. Mit Leap-Motion können auch weitere Integrationsmöglichkeiten von Zeigegesten prototypisch aufzeigt werden, die in dieser Form bislang nicht als Bedienkonzept zur Datenakquisition zur Verfügung standen. Ein Beispiel ist die Gesten-basierte Erkennung von handschriftlichen Bildannotationen als Behandlungs-Empfehlung mit Unterschrift, ohne einen (neuen) Papierbogen ausfüllen zu müssen.



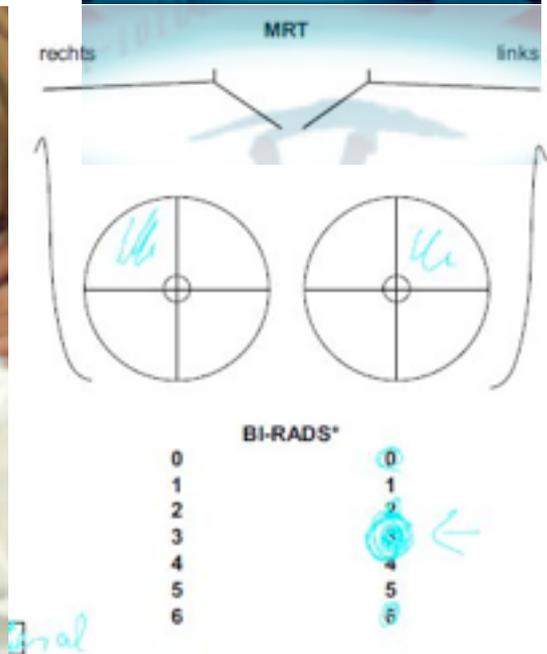
Gestik im OP wie hier bei Syngo.via (Siemens). Gibt's nicht noch andere professionelle Anwendungsfelder?



**Röntgenaufnahme**

Die Röntgenaufnahme der Mammographie erfolgt mit speziellen Geräten. Bereits heute können computerbasierte klinische Assistenzsysteme (CAD) bei der Auswertung der Aufnahmen unterstützen. Zur weiteren Abklärung können MRT Aufnahmen erstellt werden, wo mit dem digitalen Stift tumorverdächtige Befunde annotiert und digitalisiert werden können, bevor das Verfahren zur weiteren Abklärung festgesetzt wird.





# Diskussion der Evaluationskriterien

## Aspekte: Praxistauglichkeit und Innovations- höhe

In ERmed wird erweiterte Realität für sensible Patientenuntersuchungen und Patientengespräche sowie Arztkonferenzen anwendungs- und gesellschaftstauglich gemacht. Das Mittel ist die anthropozentrische digitale Veredelung.

Unser Ziel kann allerdings nur erreicht werden, wenn weitere multimodale Interaktionsformen miteingebunden werden und der HMD als "Onbody-Interaktion" als Schnittstelle und Vermittler mehrerer Zusatz-Interfaces dient. Zum Beispiel erhält die digitale Stützeinteraktion durch die HMD-Darstellung den Echtzeitcharakter zur direkten Nachkontrolle der Akquise medizinischer Falldaten. Das Dialogsystem erhält ein "On-body" Ausgabemedium zur Anzeige von Texten, Bildern und Videos. Extrem spannend wird es, den Markt dieser total-immersiven neuen Interaktionsform auch im Home-Entertainment-Bereich und in Kombination mit (autonomen) Industrie-Anwendungen (4.0) zu beobachten.

Ergonomische Einschränkungen durch Größe, Gewicht und ungewohnte Bedienung der Geräte haben gerade in der Medizin dazu geführt, dass die Akzeptanz eher gering ausfiel. Die Integrationsschwierigkeiten der HMD Technologie in den klinischen Arbeitsablauf als hochgradig organisierte Umgebungen waren bislang der Hauptgrund, warum HMDs im klinischen Alltag keine Anwendung fanden.

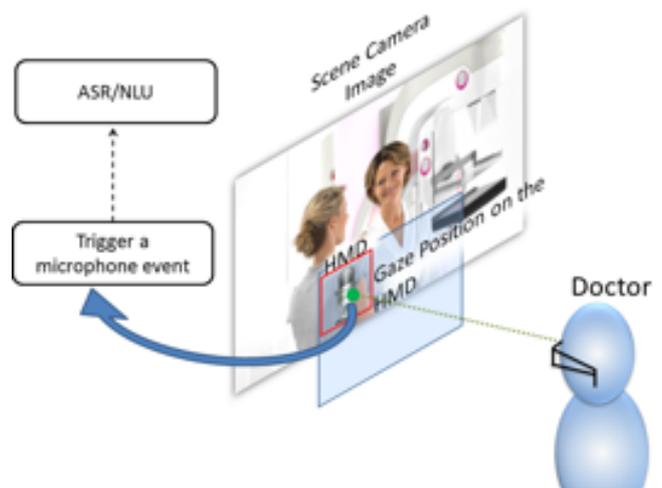
In ERmed wird der klinische Indikator für den Einsatz der Geräte um die Aspekte der ambienten Intelligenz für den Arzt und den Patienten erweitert. Außerdem setzen wir auf neuste HMD Technologien im Bereich der "see-through" HMDs, bei denen ein großes peripheres Sehfeld auf die reale Welt besteht, kombiniert mit den Vorteilen zusätzlich eingblendeter Informationen. Dabei ist die Registrierung von virtuellen und realen Objekten genau und synchronisiert. Der lichtdurchlässige HMD liefert reale Bildqualität im Vergleich zu lichtundurchlässigen Video-basierten HMDs. Der besondere Clou der ERmed Technologie ist die Kombination mit dem sprachbasierten kollaborativen RadSpeech-System, das als gemeinsame Projektionsfläche dient. So können HMD Träger untereinander kommunizieren und anderes Personal ohne HMD miteinbeziehen.

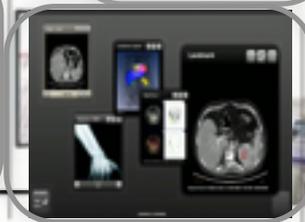
Die nächste Stufe der Hardware-Entwicklungen geht einher mit Retina-Displays, bei denen das Bild mit einem schwachen Laser direkt auf die Netzhaut des Arztes projiziert werden kann.

Durch andere aktuelle Entwicklungsarbeiten wird es bald möglich sein, den realen Hintergrund pixelgenau auszublenden. Dadurch können radiologische CT und MRT Bilder mit vollständig schwarzem Hintergrund eingblendet werden. Wir planen heute schon Evaluationen zur offiziellen Diagnosefreigabe, wie es die FDA in den USA letztes Jahr mit dem iPad und iPhone durchgeführt und die Freigabe erteilt hat (siehe mobile DICOM Viewer). Durch den zusätzlich angebrachten Eyetracker liefern wir völlig neue Möglichkeiten, am Körper angebrachte Ein- und Ausgabegeräte zu kombinieren. Eine besondere Anwendung ist unser "Gaze-Trigger" im medizinischen Bereich. Der Arzt blickt auf das Mikrophon in der Datenbrille während eines Patiententermins und löst auf diese Weise die Spracherkennung (ASR) und das Sprachverstehen (NLU) aus.

**Abgerundet wird das neue Interaktionskonzept kombinierter Ein- und Ausgabemethoden "anthropozentrische digitale Veredelung" mit der Möglichkeit, Patienten automatisch zu erkennen, neue Patienten hinzuzufügen (Online Learning), und über den Sprachdialog die Technik der erweiterten Realität als externes Gedächtnis zu verwenden: ein Patient wird (wieder)erkannt, sobald der Arzt den Blick auf ihn richtet!**

Mit der Kombination von HMD und Eyetracker bieten wir auch den Technologieanbietern der HMDs und der mobilen Eyetrackertechnologien neue Anwendungsfelder und Marktpotentiale. Ganz neue Forschungspotentiale im KI-Anwendungsbereich ergeben sich aus den integrierten maschinellen Lernverfahren.





Aufbau des ERmed - Demonstrators als Cebit - Messestand.

Die AR Brille (von Brother) und der Eyetracker (von SMI) liegen in mehreren Ausfertigungen vor, das Videosignal kann gesplittet werden. Ein zusätzliches Mikrophon erlaubt den Anschluss ans Sprachdialog-System.

Schematisch aufgebaut wird ein radiologisches Befundungszimmer, das mit dem Untersuchungszimmer dank erweiterter Realität kombiniert wird.

## Storyboard Mammographie-Demo: Aspekte: Darstellbarkeit auf der Cebit Messe und gestalterische Leistung

Im Medizinbereich ist es wichtig, bei steigender Autonomisierung durch KI-Methoden die Ärzte bestmöglich einzubeziehen. „Allerdings dürfen wir die Industrialisierung der Medizin nicht so verstehen, dass das Schicksal der Patienten der Apparatedizin überlassen wird, sondern vielmehr im Sinne einer Produktivitäts- und Effizienzsteigerung bei qualitativ hochwertiger medizinischer Versorgung.“, sagt Herrmann Requardt (Siemens). Hieraus lassen sich zwei zentrale Aspektfelder für die Verflechtung von technischen und sozialen Komponenten ableiten.

**Aspekte: Ästhetik, Soziale Umgebung, Entscheidungsunterstützung, Physikalische Freiheit**

Ästhetik hat hier die Bedeutung eines harmonisierten Ablaufes der Begutachtung für den Arzt und den Patienten. Dabei soll eine interaktive Benutzerschnittstelle mit ansprechendem Kommunikations- und Interaktionsdesign für den Arzt entworfen werden, die der Patient ebenfalls als angenehm wahrnimmt.

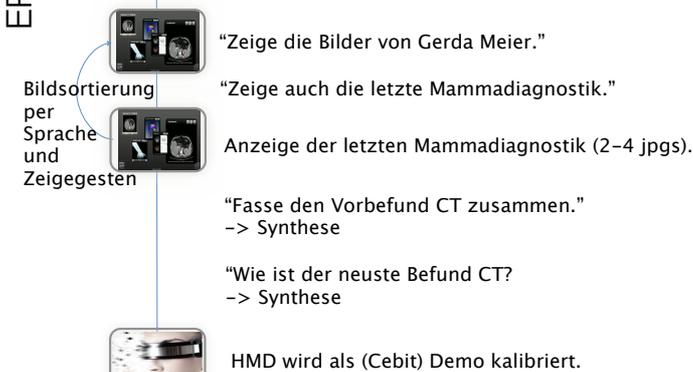
Die soziale Umgebung setzt Restriktionen, welche Interaktionsformen (Ein- und Ausgabemodalitäten) angemessen sind. Die physikalische Freiheit ist ein aus der Verwendung der AV Umgebung resultierender Effekt, der extrem wichtig ist: die robuste (ambiente) visuelle Wahrnehmung (Perzeption) des Menschen erlaubt das unbewusste, aber präzise Navigieren in (Untersuchungs-) Räumen zusammen mit einer bewussten Patientenuntersuchung bzw. bewusster Stift-Dokumentation und Kontrolle.

Der „Performanz-Task“ der Annotation geht über in eine ambiente Schreibhandlung mit ambienter Visualisierung und automatischem Sprachdialog.

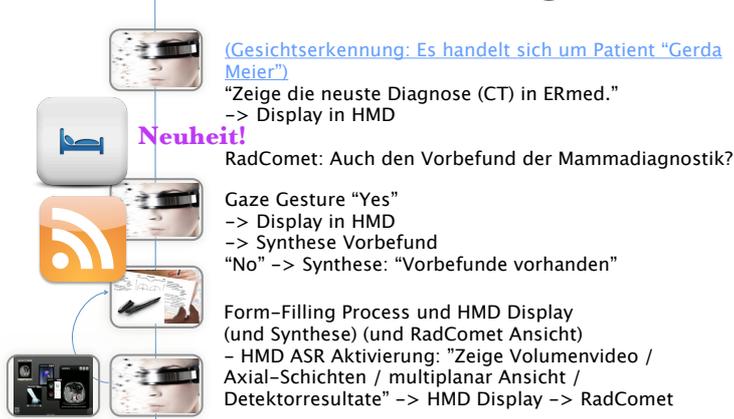
# Evaluation in Berlin, Storyboard



## (1) Vorbereitung



## (2) Untersuchung



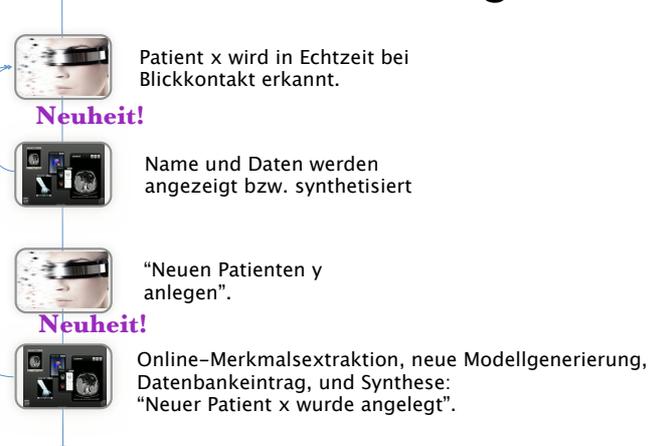
## (3) Interview



## (4) Kollaboration



## (5) Online Learning



Die technische Integration von ERmed in den domänen-spezifischen Ablauf (1-5) birgt vier echte **Weltneuheiten** in sich:

**HMD bei Patientenbegutachtung in Kombination mit RadSpeech!**

**Digitaler Stift mit Schreibgestenerkennung und Echtzeitkontrolle im mobilen HMD!**

**Sprachaktivierung und multimodale Entscheidungsfragen mit Durchsicht-HMD und Augengesten!**

**Lernen und Wiedererkennen von Patientengesichtern durch Blickinterpretation und Sprachinteraktion: integriertes Arztgedächtnis!**

## Fazit

Erweiterte Realität, das bedeutet für uns erweiterte visuelle Möglichkeiten mit Projektionen auf Brillen, die Dateneingabe mit digitalem Stift und natürlichen Sprachdialog mit digitalen Assistenten. Es entsteht ein technischer multimodaler Dialog für mobile Medizinversorgung. Völlig neue Aspekte mobiler, intelligenter und ambienter Umgebungen werden für den Arzt bereitgestellt, damit dieser sich optimal auf den Patienten einstellen kann.

Der adäquate Umgang mit neuen Interaktionsformen schließt neben den technischen Aspekten des synchronisierten Echtzeitverhaltens und der mobilen Intelligenz als Arbeitsumgebung das ambiente, patientengerechte Arbeiten mit ein. Dieser Faktor Mensch ist bei Mammographie-Untersuchungen sogar das zentrale Element. Menschliche Emotionen über die Sprache, Mimik und Körperhaltung zu erkennen ist ein aktives Forschungsfeld der KI. Der erste wichtige Schritt für deren Integration als Bedienkonzept ist allerdings, dass, wie in unserem sensiblen medizinischen Anwendungsfall, mit den Emotionen der Patienten nicht gespielt wird. Das bedeutet, dass neue (unsichere) Interaktionstechnologie nicht zum Selbstzweck verwendet werden darf, die den Patienten verunsichert, sondern ganz spezifisch und „human-centred“ eingesetzt werden sollte.

Und doch erstaunt die Datenbrille und der Eyetracker selbst. Insbesondere kann die anthropozentrische Angepasstheit der verschiedenen angesprochenen Wahrnehmungskanäle evaluiert werden. Echtzeit-ERmed-Systeme mit zusätzlicher passiver Sensorik in Smartphones zur Aktivitätserkennung werden im EIT Programm in der Aktivität „Medial Cyber-Physical Systems“ unter Leitung des Autors und mit Hilfe der TUM weiterentwickelt, wo ein Sensorraum gebaut wird. Evaluationsfragestellungen für weitere Forschung sollten sich vor allem um das Verständnis physiologischer Fakten der Mensch-Maschine Interaktion (und kontext-sensitiver Verhaltensmuster drehen. Individuelle Besonderheiten der Nachrichtenaufnahme im AR Bereich bieten weitere Forschungsfelder.

Der subtile Mechanismus, der bestimmte Emotionen beim Patienten vermeidet, ist eine weitere Stärke des Interaktionskonzepts, das mit digitalem Stift, Brille, Sprache (und Armbanduhr, Secret) eine ruhige und beruhigende Grundstimmung bei der Untersuchung erzeugt, die vom Arzt ausgeht, weil er die technische Interaktion stets souverän und ästhetisch abwickeln kann. Beispielsweise wird der Stift mit Leichtigkeit bedient und der Arzt muss keine Verrenkungen machen.

Potentielle Erkennungsfehler können von ihm korrigiert werden, da er nur das visuelle, lautlose Feedback in der AV Brille wahrnimmt, und beispielsweise störendes sprachliches Feedback vermieden wird. Diese Verflechtung, die in neusten Forschungsansätzen als „Coactive Design“ bezeichnet wird, geht nicht mehr von steigender Autonomie der Technologie aus, sondern der Interpendenz zwischen den Faktoren Mensch und Maschine, wenn die kontinuierliche Interaktion im Vordergrund steht. Unser Interaktions-Konzept ERmed, bietet diese Grundlage als modernen Anschauungsfall und verknüpft die neusten multimodalen Eingabe- und Ausgabetechnologien zu einem bis zur Cebit 2013 realisierbarem Gesamtkonzept mit vier Weltneuheiten.

**Intelligente Umgebungen werden in erweiterte Realität münden. Solange wir Sprache, Augen und Ohren besitzen und Feinmotorik nutzen, wird erweiterte Realität die Beziehung des Menschen zum Computer bedeuten.**

## Referenzen

Debbie Stone, Caroline Jarrett, Mark Woodroffe, Shailey Minocha: User interface design and evaluation. Morgan Kaufmann/Elsevier, Los Altos (2005)

Daniel Sonntag: Collaborative Multimodality. KI 26(2): 161-168 (2012)

Matthew Johnson, Jeffrey M. Bradshaw, Paul J. Feltovich, Robert R. Hoffman, Catholijn M. Jonker, Birna van Riemsdijk, Maarten Sierhuis: Beyond Cooperative Robotics: The Central Role of Interdependence in Coactive Design. IEEE Intelligent Systems 26(3): 81-88 (2011)

Daniel Sonntag, Colette Weihrauch, Oliver Jacobs and Daniel Porta: Usability Guidelines for Use Case Applications, DFKI Technical Report, THESEUS CTC-WP4, Vol. 1, 4 (2010)

## Sponsoren

**SIEMENS**

**brother**

**SMI** **TUM**  
 SensoMotoric Instruments TECHNISCHE  
 UNIVERSITÄT  
 MÜNCHEN



**Universitätsklinikum  
 Erlangen**





# Notizen

