

SPIRIT - Ereignisgesteuerte Informationsvermittlung, Inspiration und Unterhaltung im urbanen Umfeld auf Basis mobiler Augmented Reality Technologien

Antonia Kampa¹, antonia.kampa@hs-rm.de
Habiburrahman Dastageeri², habiburrahman.dastageeri@hft-stuttgart.de
Martin Storz², martin.storz@hft-stuttgart.de
Volker Coors², volker.coors@hft-stuttgart.de
Ulrike Spierling¹, ulrike.spierling@hs-rm.de

¹Hochschule RheinMain, Unter den Eichen 5, 65195 Wiesbaden

²Hochschule für Technik Stuttgart, Schellingstr. 24, 70174 Stuttgart

Abstract: Im Rahmen des Projektes SPIRIT wird eine innovative, prototypische Anwendung entwickelt, die die Sensorik mobiler Smartphone und Tablets verwendet, um ortsbasiert und kontextabhängig deren Realkamerabild mit eigens vorproduzierten Videoaufnahmen zu überlagern. In der mobilen Augmented Reality Anwendung sollen so ortsbezogene Informationen im Außenbereich mittels GPS, aber auch innerhalb von Gebäuden über eigene Indoor Navigationslösungen getriggert werden. Dargebotene Inhalte werden durch die narrative Metapher einer Geistersuche mit dem technischen Konzept integriert. Somit wird Nutzern der Anwendung durch Interactive Digital Storytelling ein situationsbezogenes unterhaltsames Erlebnis ermöglicht. In diesem Beitrag wird an Hand eines Szenarios ein erster lauffähiger Demonstrator beschrieben, sowie die Motivation für den videobasierten Ansatz erläutert.

Keywords: Augmented Reality, Interactive Storytelling, Indoor Navigation, Location-based Services, Story Engine

1 Einleitung

Im Projekt SPIRIT wird eine Anwendung entwickelt, die Besuchern historischer Kulturstätten einen persönlichen Bezug zu Informationen über die Vergangenheit besuchter Orte erlaubt. Eine mobile App verwandelt ein Smartphone oder Tablet in ein „magisches“ Medium und erweckt mittels Augmented Reality (AR) darstellbare „Geister“ in interaktiven Storys zum Leben. Technisch wird im Projekt eine mobile, location-based AR Software für dynamische Video-Inhalte entwickelt. Sie fußt thematisch auf dem Projekt GEIST [KCS+01], dessen Mobile-Game-Design-Konzept des „magischen Equipments“ übernommen wird, mit dem Besucher Kontakt zum „Geist der Vergangenheit“ aufnehmen. Unter Verwendung von Richtungssensoren, Kamera-, GPS- und Indoor Navigations-Daten werden ortsbezogene Informationen multimedial in das Umfeld integriert. Es sollen dazu passende interaktive Storys auf mobilen Geräten dynamisch von einer Engine präsentiert werden. Schauplatz für den ersten Prototyp ist das Römerkastell Saalburg, dessen Museumseinrichtung als Projektpartner beteiligt ist.

Ziel dieses Beitrags ist die Vorstellung des SPIRIT Demonstrators und eine Diskussion des Einsatzes von technischen und erzählerischen Elementen bezogen auf Augmented Reality (AR). Nach einer Einführung der wichtigsten Begriffe des Interactive Digital Storytelling (IDS) folgt eine Beschreibung der SPIRIT Demonstrator AR App, sowie verwandter Projekte und Herausforderungen bei der Erstellung von Konzept und Inhalten. Die mediale Inhaltsstruktur wird vorgestellt und in Hinblick auf die Realisierung als AR App diskutiert. Im Anschluss werden die Anforderungen an die Entwicklungsplattform genannt und bereits implementierte technische Grundlagen des Trackings der SPIRIT Demonstrator App beschrieben.

2 Verwandte Anwendungen und Stand der Technik

2.1 Szenario und Story-Metapher im SPIRIT Demonstrator

Die konzipierte SPIRIT App „erzählt“ Spaziergängern abhängig von ihrer Position Teile einer Geschichte mit historischem Bezug, in dem sie ad-hoc aneinander gereimte Video-„Schnipsel“ abspielt. Abb. 1 symbolisiert den ersten implementierten Demonstrator – mit Hilfe eines Tablets gelingt einer Spielerin der Kontakt zum Geist eines römischen Soldaten, der sie auffordert, mit ihm das Römerkastell Saalburg zu besuchen.



Abbildung 1: Mockup-Szenariodarstellung des SPIRIT Demonstrator

Mit Hilfe des magischen Equipments interagieren Besucher demnach mit videobasierten virtuellen Figuren. Das Neue an der Konzeption ist weniger die erstmalige Entwicklung einzelner vorhandener Techniken (mobile, ortsbezogene, videobasierte AR App, Entertainment und Wissensvermittlung) als deren Integration. Vorproduziertes Videomaterial wird auf das Kamerabild mobiler Geräte gelegt, um für Besucher eine erweiterte Realität (AR) zu schaffen. Technische Herausforderungen liegen dabei in der glaubwürdigen Darstellung und der Interaktion mit virtuellen Figuren eingebettet in den AR-Kontext.

Für die Anpassung des gezeigten Inhalts auf das Verhalten des Benutzers wird eine Plot-Engine entwickelt. Die Plot-Engine setzt für einen Nutzer innerhalb einer oder mehrerer „Begegnungen“ mit Geistern aus atomaren Story-Elementen ad hoc eine Abfolge zu einem Story-Erlebnis zusammen. Die Plot-Engine reagiert auf User-Input, wie Bewegung, Standortveränderung und Interaktion mit graphischen Elementen (GUI), und berücksichtigt Regeln der Storyworld, die vorab von Autoren definiert wurden. Basierend auf derselben Storyworld erzeugt eine Plot-Engine für verschiedene Benutzerverhalten also verschiedene Story-Abfolgen. Die Repräsentation der Story-Elemente erfolgt im Projekt SPIRIT mit vorproduzierten Videos, Ton und Text, die mit dem Kamerabild des Tablet überlagert werden (AR). Alle Story-Elemente und ihre Repräsentation müssen vorab erstellt werden. Das Ziel im Projekt SPIRIT ist es auch, ein Authoring-Framework zur Verfügung zu stellen, mit dem Inhalte von Nicht-Programmierern, wie z.B. Mediengestaltern, eingepflegt werden können. Zu diesem Zweck wird im späteren Verlauf des Projekts ein Autorentool erstellt.

2.2 Verwandte Anwendungen

Eine Übersicht über den Trend zu AR Spielen gibt [BLH+08]. „Haunted Planet“ betreibt mobile AR für Schlossbesuche mit „Geister-Kontakt“. Im Gegensatz zu SPIRIT beschränkt sich diese Anwendung auf das einmalige Geister-„Aufspüren“ ohne Story-Zusammenhang. Mobile Urban Drama [HKG12] ist eine dänische Gruppe aus Theaterdarstellern und Entwicklern, die verschiedene mobile ortsbasierte Anwendungen im historischen und kommerziellen Kontext erstellt haben. Ohne Einsatz von AR unterstützen dabei reale Schauspieler die Anwendung für Shops und Betriebe vor Ort.

Eine typische Herausforderung im Interactive Storytelling ist das ad-hoc Anpassen eines Geschichtenverlaufs (Plot) an Nutzereingaben. Zwei Beispiele hierfür sind Façade [MS05] und Office Brawl [MSS13]. Beide Anwendungen sind hochinteraktiv, d.h. Nutzer greifen mit einer hohen Frequenz in das Geschehen ein. Beide werden durch Texteingaben gespielt, auf die virtuelle Figuren (Charaktere) sofort verbale Antworten geben. Bei Office Brawl soll vom Spieler ein Konflikt zwischen zwei virtuellen Mitarbeitern gelöst werden. In informellen Tests sind Nutzern dabei menschliche Stimmen glaubwürdiger erschienen als synthetische [MSS13]. In Façade [MS05] soll analog ein Ehekrach geschlichtet werden. Zusätzlich zu einer Desktop-Version wurde Façade auch in einer AR-Version entwickelt [DML+06]. Diese ist ein interaktives Drama mit einer angestrebten zwanglosen, natürlichen Interaktion. Letzteres wird durch reale Requisiten in einem physisch realen, dem Desktop Façade Setting nachempfundenen Raum realisiert.

Android verwendet Java als Programmiersprache und bietet einen direkten Zugriff auf die notwendigen Hardwarekomponenten, wie beispielsweise Kamera, Magnetsensoren und OpenGL. Anpassungen waren hierbei im Bereich der Videodarstellung notwendig, da der Android MediaPlayer keine Unterstützung für beispielsweise den Alphakanal bietet. Die Transparenz musste manuell implementiert werden siehe. Dadurch sind flexiblere Effekte und Skalierungen zur Laufzeit sowie generelle Anpassungen an unsere Anforderungen möglich.

Andere verbreitete AR-Frameworks (Wikitude, Metaio, AndAR) bieten dagegen nur bedingt nutzbare Lösungsansätze für unsere Anforderungen.

Das Video wird vom „Android MediaPlayer“ abgespielt und auf eine Textur gerendert. Diese Textur wird daraufhin mit OpenGL ES 2.0 Shadern bearbeitet und dargestellt. Beim Erreichen bestimmter Positionen wird dabei ein Referenzfoto des aktuellen Kamerabildes erstellt, das zur Positionierung des Videos auf dem Display verwendet wird. Bei Veränderungen des Blickwinkels oder der Entfernung wird das Video entsprechend skaliert und positioniert.

3 Interactive Digital Storytelling

Interactive Digital Storytelling folgt in seiner Bedeutung im Projekt SPIRIT einer Arbeitshypothese, bei der einzelne Nutzer durch die Interaktion mit einer Applikation eine Story „erleben“ [Spi12]. Mit Story bzw. Geschichte meinen wir das Ergebnis eines Ablaufs von Ereignissen, die nicht nur Objekte, sondern „menschliche“ Umstände, Beziehungen und Gefühle betreffen und erfunden oder historisch recherchiert („wahr“) sein können. *Interaktiv* bezieht sich auf die Erstellung des Handlungsablaufs der Story, den der Benutzer aktiv mitbestimmen kann.

3.1 Interactive Storytelling allgemein

Beim Storytelling legt ein Autor eine strikte Abfolge von Szenen in einem linearen Szenario der Story fest. Beim Interactive Storytelling im Projekt SPIRIT brechen wir diese lineare Struktur auf und erstellen dagegen eine Story-World mit sinnvoll annotierten, atomaren Elementen. Auf Basis dieses Story-Modells können wir dynamisch geeignete Szenen auswählen, die einen neuen Handlungsablauf (Plot) der Story ergeben. Im Projekt GEIST [SGB+02], [KCS+01] definierte eine „Story-Engine“ nicht konkret die nächste abzuspielende Szene, sondern entschied lediglich auf Basis des Story-Modells, welche narrative (erzählerische) Funktion nach Propp [Pro58] die nächste Szene erfüllen soll. Die Entscheidung wurde an eine Scene-Engine weiter gegeben, die ihrerseits eine passende Szene auswählte. Wurde keine passende Szene gefunden, musste die Story-Engine das Story-Modell erneut konsultieren und eine alternative narrative Funktion herausuchen.

Eine Story-Engine generiert ad-hoc unter Berücksichtigung von Nutzereingaben Antworten, die eine interaktive Story ergeben. Sie plant eine Handlungsabfolge aus mehreren erzählerischen Elementen, die nicht der Reihenfolge der vom Autor vorgegebenen Story entsprechen muss [KCS+01]. Eine Story Engine koordiniert dynamisch den Ablauf einer Story auf ihrem strukturell höchsten Level [SGB+02]. Hierbei ergibt sich ein Problem für Autoren, die meist Nicht-Programmierer sind, eine Story in geeignete, maschinell lesbare kleinere Einheiten aufzubrechen und zu annotieren. Bei [HKG12] geben Autoren alle Verzweigungen zwischen den Einheiten der Story vor. Solch ein vollständiges so genanntes „Branching“ funktioniert nur, wenn alle Entscheidungsmöglichkeiten von Benutzern im Ablauf bekannt sind. Bei umfangreichen Storys müssten sehr viele Verzweigungen konstruiert werden, um alle möglichen Nutzereingaben zu berücksichtigen. Somit müssten auch Verzweigungen für unwahrscheinliche Kombinationen von Nutzereingaben erstellt werden, die möglicherweise

niemals oder selten auftreten. Daher wird in der Fachgemeinschaft des Interactive Storytelling manuelles „Branching“ nach Möglichkeit vermieden und Verzweigungen dynamisch erzeugt.

3.2 Interactive Storytelling in SPIRIT

Im Projekt SPIRIT muss die geplante Plot-Engine verschiedene Auftritte von Geistern mit Nutzereingaben unter Berücksichtigung von Vorbedingungen und Effekten für die Story koordinieren. Außerdem sollen Auftritte von Geistern – visualisiert durch Videoschnipsel – auch proaktiv initiiert werden können. Im Demonstrator werden vorerst direkte Verzweigungen (Branching) benutzt. Im Verlauf des Projekts SPIRIT wird dieser Ansatz erweitert, um dynamisch Handlungsabfolgen aus möglichst kurz erzählbaren Teilen der Story zusammenstellen zu können. Dabei wird nicht komplett auf Verzweigungen verzichtet, da im Projekt SPIRIT nicht so oft Antworten generiert werden müssen wie zum Beispiel in Façade [MS05].

Um eine Brücke zwischen den Autoren, die Nicht-Programmierer sind, und der Erzeugung von maschinenlesbaren, annotierten, atomaren Handlungen zu schlagen, werden wir im weiteren Verlauf des Projekts SPIRIT ein Autorentool erstellen.

4 Herausforderungen und Lösungsansätze beim Einsatz von Animation und Video in AR Anwendungen

4.1 Glaubwürdigkeit in der erweiterten Realität

Die Theorie des Uncanny Valleys [Mor05] deutet an, dass der Versuch, animierte Charaktere und Roboter visuell nahezu menschengleich und realistisch zu gestalten, einen negativen Effekt erzeugt, wenn die Animation nicht realistisch gelingt. Das Resultat ist eine geringere Glaubwürdigkeit sowie Akzeptanz. In informellen Untersuchungen sind Nutzern von Office Brawl [MSS13] aufgenommene, menschliche Stimmen glaubwürdiger erschienen als synthetische Stimmen. In AR Façade [DML+06] wird nach eigenen Aussagen versucht, Glaubwürdigkeit durch einen visuellen Abstraktionsgrad (Cartoon) zu erzeugen. Im Projekt SPIRIT wollen wir eine glaubwürdige Erfahrung durch die ad-hoc Kombination von kurzen, live gedrehten Videoschnipseln für die Repräsentation von aneinander gereihten Story-Elementen erreichen. Da der Effekt alberner Spukgeschichten vermieden werden soll, wurden Cartoon-Repräsentationen im Vorfeld von den Partnern abgelehnt. Die resultierende Glaubwürdigkeit der Charaktere soll im Projekt durch Entwicklung von darauf abgestimmten non-linearen Video-Produktionsverfahren verbessert und anschließend evaluiert werden.

4.2 Darstellung von virtuellen Charakteren in der erweiterten Realität

In diesem Abschnitt werden verschiedene Aspekte der Darstellung der Geistmetapher in unserem SPIRIT Demonstrator beschrieben. Damit die AR Repräsentation der nicht realen Geister möglichst konsistent sind, wollen wir Logikbrüche für den Besucher vermeiden. Für den ersten

Entwurf eines Geister-„Stils“ wurden zunächst Eigenschaften identifiziert, die Geistern zugeschrieben werden, um sie in unsere Geistmetapher mit aufzunehmen: Unschärfe, Transparenz, Schweben und Wabern, sowie eine leicht violette Färbung.

Alle Auftritte der virtuellen Charaktere müssen als Video vorproduziert werden. Dabei wird der Nachteil gegenüber Animationen in Kauf genommen, dass Videos weniger Freiheitsgrade wie z.B. Transformation, Skalierung und Translation erlauben, und deswegen nur schwer an Änderungen wie Blickwinkel dynamisch angepasst werden können. Weiterhin müssen alle Möglichkeiten dafür vorab eingeplant werden, um in einer Produktion ausreichend Videomaterial zu erstellen. Unser „Geist“ im ersten Demonstrator schaut immer in eine vorgegebene Richtung, beim Dreh meist in die Kamera und in der AR App wird er auf einem Billboard abgebildet, um immer den Benutzer anzusehen. Versucht der Benutzer um den Geist zu gehen, wird dieser immer in seine Richtung ausgerichtet.

Der Benutzer könnte für seinen Input mehr Zeit benötigen, als das Videomaterial lang ist. Um Stillstände der Videos zu vermeiden wurden nach dem Vorbild von Office Brawl [MSS13] neutrale Videos produziert, in denen der Geist merklich auf den Benutzer wartet. Unterschiedliche Sequenzen aus Video und Ton in unserem Projekt können nicht ad hoc ohne Brüche in Bild und Ton aneinander gereiht werden wie vergleichsweise reiner Text, da Körperhaltung und Stimmlage von Aufnahme zu Aufnahme variieren. Unsere Vorüberlegungen für eine glaubwürdige Geistmetapher hierzu waren den Darsteller des Geists in der immer gleichen Körperhaltung jedes Video beenden und beginnen zu lassen, was aber beim Video-Dreh nicht realisiert werden konnte. Auf dieser Grundlage musste das Anschluss-Problem in der Postproduktion der Video Schnipsel gelöst werden. Wir brauchten Zwischensequenzen im Video und Ton, die zu jedem Zeitpunkt im Videomaterial auftreten können. Als glaubwürdige Erklärung für diese Zwischensequenzen benutzen wir neben der Geistmetapher die Metapher des „magischen Equipment“ als Übertragungsmedium mit variablem Empfang. Ähnlich dem gestörten Empfang von Fernseher und Radio, bei dem Störeffekte wie Verzerrungen und Aussetzer in Bild und Ton auftreten, erzeugten wir mittels Filter der Postproduktionssoftware Störungen im Videomaterial (siehe Abb.2). Diese Störungen fügen wir ad hoc zu Beginn jedes Video Schnipsels ein, um konsistente Übergänge zu schaffen.



Abbildung 2: Effekte für den „gestörten Empfang bei einer Geistübertragung“

4.3 Interpretation von Bewegungen im realen Raum und Nutzereingaben in AR

In diesem Abschnitt werden verschiedene Möglichkeiten der Interpretation von Bewegungen der Benutzer im Raum und Eingaben der Nutzer diskutiert. Die aktuelle Umsetzung im SPIRIT Demonstrator sowie die geplante Umsetzung im Projekt SPIRIT wird erläutert.

Zu Beginn des SPIRIT Demonstrators haben wir uns ähnlich der App „Haunted Planet“ [BLH+08] für ein „Aufspüren“ unseren Geists entschieden. Dem römischen Geist wird eine zufällig bestimmte, dem Benutzer aktuell nahe GPS-Koordinate zugewiesen. Nähert sich der Benutzer dieser Koordinate, beginnt das „magische Equipment“ die „Verbindung zum römischen Geist“ aufzubauen und das „Signal zum Geist“ wird immer stärker. In diesem Schritt wird das halbtransparente Video schrittweise in das Kamerabild eingeblendet und an einen bestimmten Punkt im Kamerabild verankert (markerloses Tracking), um dadurch eine erweiterte Realität (AR) zu schaffen. Damit Geister zukünftig an der richtigen Stelle und Größe im Setting dargestellt werden, wollen wir im Projekt SPIRIT das Einblenden der vorproduzierten Videos an bestimmten historischen Schauplätzen durch markerloses Tracking auslösen.

Wie auch in AR Façade [DML+06] umgesetzt, wollen wir Entscheidungen auf dem Story-Level mit einer Story-Engine realisieren; somit sind die Verzweigungspunkte für den Benutzer nicht offensichtlich. Im Projekt SPIRIT wollen wir diese Nebenläufigkeit durch bereits implementiertes markerloses Tracking unterstützen, sowie durch weitere Eingabemöglichkeiten wie z.B. Sprache.

Bei AR Façade [DML+06] wurde festgestellt, dass die meisten der 80 Teilnehmer in einer Demonstration sprachen, ohne zu den virtuellen Charakteren zu sprechen. Sie überlegten laut oder sprachen zu den Entwicklern oder den Zuschauern. Wenn man nun die Sprache und Bewegungen der Benutzer in eine AR Erfahrung einbeziehen will, steht man somit vor zwei Problemen. Zum einen muss man bestimmen, ob eine Handlung vom Benutzer für die Story gedacht war, und zum anderen, ob man die außerhalb der Story gemachten Eingaben interpretieren will. Kommentare oder Hilfen könnten parallel zur Story angezeigt werden. Im Projekt SPIRIT wollen wir das magische Equipment auf diese Art und Weise nicht nur Metapher für ein Übertragungsgerät nutzen, sondern auch als Begleiter, der für Fragen und Anmerkungen offen steht und Hilfestellung geben kann. Auf einen Modus für ein Tutorial außerhalb des eigentlichen Erlebnisses könnten wir dadurch verzichten.

4.4 Interaktion mit realen Gegenständen in AR Räumen

Eine andere Frage betrifft, wie man mit Interaktion mit realen Gegenständen im Raum umgeht, da reale Gegenstände nicht von virtuellen Charakteren, jedoch vom Benutzer manipuliert werden können. In den ersten aus AR Façade gezogenen Lehren [DML+06] wird der Vorteil beschrieben, dass die benötigte Infrastruktur des AR Systems durch allgemeines Referenzieren von realen Objekten im Raum vereinfacht wurde. Das System interpretiert somit keine detaillierten Interaktionsformen des Benutzers mit Objekten, wie ansehen, anfassen, nehmen, benutzen, weglegen.

Das Problem, dass AR Charaktere keine realen Gegenstände manipulieren können, bleibt noch ungelöst. In AR Façade nehmen die AR Charaktere Weingläser auf, indem ihre Handhaltung so verändert wird, als ob sie ein Weinglas halten: Das reale Weinglas bleibt stehen und die Hand des AR Charakters bleibt leer. Es wäre auch denkbar gewesen, dass ein animiertes, virtuelles Weinglas aus dem Nichts entsteht, sobald Getränke zubereitet wurden.

Der Benutzer kann reale Objekte in der erweiterten Realität (AR) manipulieren. In AR Façade [DML+06] werden ferngesteuerte Computer als Telefon und Anrufbeantworter verwendet, sodass beim Abnehmen des Hörers der Benutzer die Stimme aus dem realen Telefon hören kann, bzw. bei Abrufen einer Nachricht eine Audiodatei im realen Anrufbeantworter ausgegeben und nicht über den üblichen Audio-Output abgespielt wird. Die Macher von AR Façade [DML+06] erhoffen sich davon ein realeres Erlebnis für den Benutzer. Im Projekt SPIRIT wollen wir in Zukunft ähnlich diesem Ansatz reale Objekte in die AR Erfahrung unserer App einbeziehen.

Bei Interaktionen zwischen Benutzern und virtuellen Objekten oder Charakteren fehlt haptisches Feedback, deswegen halten die Macher von AR Façade [DML+06] eine solche Interaktion für unnatürlich für den Benutzer. Als Lösung für dieses Problem im Projekt SPIRIT schlagen wir hier einen Reality Mix vor, in dem ein virtuelles Objekt ein reales Objekt durch markerloses Tracking überlagert. Am Beispiel des Weinglases beschrieben könnten Benutzer ein reales leeres Weinglas nehmen, das mit einer Aufnahme eines gefüllten, virtuellen Weinglases überlagert wird. Als Interaktion leert man das virtuelle Weinglas nach jedem zum Mund führen nach und nach.

5 AR-Konzept

5.1 Anforderungen an die Entwicklungsplattform

Die Voraussetzungen für die Darstellung von videobasierter Augmented Reality auf mobilen Geräten ist die plattformabhängige Unterstützung von Videos auf Texturen. Die hierfür benötigten Methoden werden ab Android 4.0.3 (Min. Api Level 15) zur Verfügung gestellt. Weitere Voraussetzungen sind die Unterstützung von OpenGL ES 2.0, Sensorik (Kompass, Lagesensoren und GPS), Kamera sowie die notwendigen Codecs für die Wiedergabe der Video- und Audioformate. Zwar kann die hohe Gerätefragmentierung den Entwicklungsprozess verlängern, allerdings traten Probleme nur im hardwarenahen Bereich, z.B. OpenGL und vereinzelt auf. Laut Android 4.4 Compatibility Definition Document (Stand: 27. November 2013) ist die Unterstützung diverser Codecs, wie z.B. VP8 verpflichtend und kann daher als vorausgesetzt betrachtet werden.

5.2 Eingesetzte Verfahren des Image- und GPS-Trackings

Ziel des Projekts SPIRIT ist es, videobasierte Augmented Reality berührungsfrei zu realisieren. Bei der Implementierung der Trigger erwies sich eine Kombination aus GPS- und Kompassdaten

als zu ungenau. Alternative Vorgehensweisen mussten entwickelt und getestet werden. Daher wurde ein Image-Tracking-Verfahren entwickelt, das die Kamerabilder vor Ort mit Vergleichsbildern der Ziele auf Gemeinsamkeiten hin überprüft. Markante Merkmale, sogenannte Keypoints, werden dabei aus dem Bild extrahiert und abgeglichen. Die Fläche, die dabei für jeden Keypoint betrachtet wird, nennt man Patch. Für jeden dieser Keypoints des Kamerabilds wird ein Gegenpunkt im Vergleichsbild gesucht. Je größer die Übereinstimmung, desto kleiner die Differenz (distance). Hierzu wurde der Algorithmus ORB (Oriented FAST and Rotated BRIEF [RRK+11]) der Open-Source Bibliothek OpenCV (Computer Vision) verwendet. Ausschlaggebend für das Ergebnis sind somit insgesamt vier Werte; Anzahl der zugewiesenen Vergleichspunkte, minimale Differenzen (Min_dist), maximale Differenzen (Max_dist) und durchschnittliche Differenzen (Avg_dist). Allerdings erwies sich ein ungefilterter Abgleich der Vergleichspunkte als sehr unzuverlässig, da sehr viele falsche Zuordnungen stattfanden und die Ergebnisse somit nicht aussagekräftig waren. Daraufhin wurden drei Filtermethoden implementiert. Im ersten Schritt erfolgt ein „symmetrischer Abgleich“. Es sollen nur Vergleichspunkte, die beim Vergleich von Bild 1 mit Bild 2, aber auch in entgegengesetzter Reihenfolge übereinstimmen, betrachtet werden. Vergleichspunkte ohne diese Eigenschaft wurden entfernt. Folgend wurden im zweiten Schritt Vergleichspunkte mit schlechten Bewertungen eliminiert. Dazu zählen diejenigen, deren Max_dist über dem doppelten Wert des Min_dist liegt. Jedoch kann es auch hierbei zu Zufallstreffern kommen, wie z.B. auf dem Bild befindliche ähnliche aber nicht angepeilte Elemente, wie im Hintergrund befindliche Fensterrahmen. Um auch diese Fehlerquelle zu entfernen, wurde der RANSAC Filter [FB81] verwendet. Entfernte Punkte können somit als Ausreißer erkannt und gelöscht werden.

Maßgebliche Parameter für das Ergebnis und Performance des Image-Trackings ist die Anzahl der Keypoints und die Auflösung der Vergleichsbilder. (Tab.1) zeigt einen Ausschnitt der Testtabelle für Messwerte mit variierten Keypoints. Die Bilder hatten dabei eine Auflösung von 400x300 Pixel für das Kamerabild und 175x395 Pixel für das Vergleichsbild.

| Messungen | Keypoints | Minimale Zeit | Maximale Zeit | Durchschnitt |
|-----------|-----------|---------------|---------------|--------------|
| 100 | 500 | 439 ms | 1246 ms | 468 ms |
| 100 | 750 | 527 ms | 701 ms | 571 ms |
| 100 | 2500 | 1120 ms | 1437 ms | 1194 ms |
| 100 | 7500 | 1604 ms | 1941 ms | 1699 ms |

Tabelle 1: Auszug aus den Testergebnissen der zeitlichen Messung bei ansteigenden Keypoints

Ab 2500 Keypoints konnten zuverlässig erfolgreiche Zuweisungen erfolgen. Zudem ist es notwendig, die übereinstimmenden durchschnittlichen Differenzen (Avg_dist) als Parameter festzulegen. Experimentell erwiesen sich im Innenbereich Zuweisungen ab einem Wert von 30 Avg_dist, im Außenbereich ab 35 als erfolgreich.



Abbildung 3: Ausschnitte aus Original mit hervorgehobenem Patch:
links (400x30px) mit 8x8 Pixel Patch, Mitte (1600x120px) mit 64x64 Pixel Patch, rechts
(1600x120px) mit 8x8 Pixel Patch

Die gewählte geringe Auflösung des Vergleichsbildes hat zwei Vorteile, eine deutlich bessere Performance beim Ermitteln der Keypoints sowie eine geringere Fehleranfälligkeit bezüglich der Eindeutigkeit der Keypoints. Dies liegt daran, dass bei kleinen „Patchgrößen“ bei der Berechnung der Keypoints bei geringer Auflösung viele Informationen einbezogen werden können siehe (Abb.3, links). Bei größeren Auflösungen müssten für vergleichbare Ergebnisse sehr große Patchgrößen verwendet werden siehe (Abb.3, Mitte), da ansonsten nur sehr wenige Informationen berücksichtigt werden können siehe (Abb.3, rechts). Da bei hoher Auflösung exponentiell mehr Pixel beachtet werden müssen und die Qualität der Ergebnisse auch bei reduzierter Auflösung gegeben ist, wird die Auflösung generell gegebenenfalls automatisch verringert. Entsprechend werden ausschließlich Bilder in geringer Auflösung verwendet.

(Abb.4) stellt das eingesetzte Verfahren visuell beim Vergleich des Haupteingangs des Römerkastells Saalburg dar. Hierbei wurden 8 Vergleichspunkte (Row_Count) erkannt. Das beschriebene Verfahren erwies sich als zeitaufwändig, da lediglich 1-4 Bilder pro Sekunde verglichen werden konnten. Zur Steigerung der Performance kann per GPS eine Filterung der vorher definierten Bilder stattfinden, und somit ein Vergleich mit lediglich in nahem Umkreis liegenden Zielen erfolgen.



Abbildung 4: Darstellung der Zuweisung von Vergleichspunkten

Bei bereits im Voraus festgelegten Vergleichsbildern sind dynamische Umwelteinflüsse zu berücksichtigen, wie tageszeitabhängige Lichtverhältnisse, Witterungsbedingungen oder auch jahreszeitliche Faktoren beim Pflanzenwachstum. Dies sollte bei der Auswahl der

Vergleichsobjekte berücksichtigt werden. Daher werden für den Außenbereich mehrere angepasste Vergleichsbilder verwendet. Im Innenbereich kann dies vernachlässigt werden.

6 Fazit und Ausblick

In diesem Beitrag haben wir verschiedene Design-Ideen und Konzepte für videobasierte AR diskutiert und Lösungen genannt, die im Projekt SPIRIT in einem ersten Demonstrator umgesetzt wurden oder im weiteren Verlauf noch umgesetzt werden sollen.

Durch eine glaubwürdige Metapher der Interaktion mit „Geistern“ sowie mit einem „magischen Equipment“ sollen möglichst durch Design-Entscheidungen eventuelle bleibende technische Begrenzungen abgepuffert werden. Darüber hinaus werden Interactive Storytelling Lösungen entwickelt, die eine dynamische Anpassung von vorgefertigtem Video-Material an verschiedene Situationen von Nutzern ermöglichen.

Das beschriebene AR-Verfahren eignet sich, um ohne Platzierung eigens erstellte Marker (QR-Codes) erfolgreich Videos zu triggern. Videodateien können so an Georeferenzen gebunden werden, ohne den historischen Standort Saalburg selbst zu verändern. Damit konnte den Anforderungen des Römerkastells Saalburg nachgegangen werden, da aufgrund des Status des Kastells als UNESCO-Weltkulturerbes keine eingreifenden Änderungen möglich sind.

Die Performance der Videodarstellung auf den von uns benutzten mobilen Geräten reicht aus, um auch HD-Videos mit üblichen 24 Frames pro Sekunde darzustellen. Empfehlenswert ist die Beobachtung der weiteren Entwicklung des Android-Mediaplayers bzgl. der Unterstützung des VP8 Alpha-Kanals (<http://wiki.webmproject.org/alpha-channel>), da wir dadurch einen direkten Zugriff auf den Alpha-Kanal erhalten und somit die Dateigröße der Videos verringert wird. Speziell berücksichtigt wurde ein möglichst hoher Grad an Unabhängigkeit der Ausgabehardware. Die evolutionäre Entwicklung über PC, Laptop, Netbook und schließlich zum Tablet, Smartphone und „See-through glasses“ wird noch nicht als abgeschlossen angesehen [LFM+14]. Derzeit werden Tablets als Ausgabemedium verwendet, zukünftig könnten aber auch Smartphones, „See-through glasses“ und weitere „Wearables“ berücksichtigt werden.

Danksagung

Das Projekt SPIRIT wird unter dem Förderkennzeichen 03FH035PA3/B3 vom BMBF gefördert. Dank gilt K. Stöbener und dem Saalburgmuseum für die Unterstützung bei der Inhaltsgestaltung.

Literatur

- [BLH⁺08] Wolfgang Broll, Irma Lindt, Iris Herbst, Jan Ohlenburg, Anne-Kathrin Braun, and Richard Wetzel. Toward Next-Gen Mobile AR Games. *IEEE Comput. Graph. Appl.* 28, 4 (July 2008), 40-48, 2008.
- [DML⁺06] Dow, S., Mehta, M., Lausier, A., MacIntyre, B., & Mateas, M. Initial lessons from AR Façade, an interactive augmented reality drama. In *Proceedings of the 2006*

ACM SIGCHI international conference on Advances in Computer Entertainment Technology, p. 28, 2006.

- [FB81] Martin A. Fischler and Robert C. Bolles: Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. In *Proceedings of Communications of the ACM* 24 (6), pp. 381–395, 1981.
- [HKG12] Frank Allan Hansen, Karen Johanne Kortbek, and Kaj Grønbaek. Mobile Urban Drama: interactive storytelling in real world environments. *New Review of Hypermedia and Multimedia*, Taylor & Francis, 18(1-2):63-89, 2012.
- [KCS⁺01] Ursula Kretschmer, Volker Coors, Ulrike Spierling, Dieter Grasbon, Kerstin Schneider, Isabel Rojas, and Rainer Malaka. Meeting the Spirit of History. In *Proceedings of the International Symposium on Virtual Reality, Archaeology and Cultural Heritage, VAST 2001*, Glyfada, Greece, pp. 161-172, 2001.
- [LFM⁺14] D. Lanman, H. Fuchs, M. Mine , I. McDowall, M. Abrash. Put on your 3D glasses now: the past, present, and future of virtual and augmented reality. In *Proceedings SIGGRAPH '14*, 2014.
- [Mor05] Mori, M., On the Uncanny Valley. *Proceedings of the Humanoids-2005 workshop: Views of the Uncanny Valley*. Tsukuba, Japan, 2005.
- [MS05] Michael Mateas and Andrew Stern. Structuring Content in the Façade Interactive Drama Architecture. In *Proceedings of the First Artificial Intelligence and Interactive Digital Entertainment Conference, AIIDE 2005*, pp. 93-98, 2005.
- [MSS13] Wolfgang Müller, Ulrike Spierling, and Claudia Stockhausen. Production and Delivery of Interactive Narratives Based on Video Snippets. In *Interactive Storytelling, Proceedings of ICIDS 2013*, Springer LNCS 8230:71-82, 2013.
- [Pro68] Vladimir Propp. *Morphology of the Folktale*. The American Folklore Society and Indiana University (Translation 1968, orig. 1928), Second Edition, University of Texas Press, 1968.
- [RRK11] Ethan Rublee, Vincent Rabaud, Kurt Konolige, and Gary R. Bradski. ORB: An efficient alternative to SIFT or SURF. In *Proceedings of IEEE Computer Vision (ICCV 2011)*, pp. 2564-2571, 2011.
- [SGB⁺02] Ulrike Spierling, Dieter Grasbon, Norbert Braun, and Ido Iurgel. Setting the scene: playing digital director in interactive storytelling and creation. *Computers & Graphics*, Vol. 26(1):31-44, 2002.
- [Spi12] Spierling,U. ‘*Implicit Creation*’ – *Non-Programmer Conceptual Models for Authoring in Interactive Digital Storytelling*. Dissertation. University of Plymouth, Plymouth, UK, 2012.