

HIMBA KUNSTHANDWERK UND AUGMENTED REALITY

Robert Beer^a, Matti Mäki-Kihniä^b, Crescience Kouba^c, und Jürgen Sieck^d

^aHochschule für Technik und Wirtschaft Berlin, Deutschland, s0545347@htw-berlin.de; ^bMetropolia University of Applied Sciences, Finnland, matti.makikihnia@gmail.com; ^cNamibian University Of Science and Technology, Namibia, crescience01@gmail.com; ^dHTW Berlin, Deutschland und NUST Windhoek J.Sieck@htw-berlin.de

KURZDARSTELLUNG: Dieser Beitrag beschreibt die Umsetzung einer Anwendung für mobile Geräte, bei der einem Nutzer anhand visueller Objekterkennung und Augmented Reality Inhalten Informationen zu Himba-Armringen auf interessante und anschauliche Weise zur Verfügung gestellt werden.

1. EINFÜHRUNG

Die Himba sind eine, vor allem im nördlichen Namibia und zum Teil im südlichen Angola lebende, Volksgruppe mit einer stark ausgeprägten und ausgelebten eigenen Kultur und Geschichte. Diese zeigen viele von ihnen auch äußerlich durch besondere Körperbemalung, Kleidung und Schmuck. Besonders letzteres bieten sie auf vielen Märkten auch Besuchern zum Kauf an, wobei vor allem aufwändig verzierte Armringe ein beliebtes Andenken sind. Oftmals ist es jedoch schwierig, zusätzliche Informationen wie Bedeutung, Besonderheiten der Fertigung, verwendete Materialien oder die Geschichte zu den einzelnen Schmuckstücken zu erhalten.

Daher entstand die Idee, zu untersuchen, ob die Muster dieser Armringe geeignet sind, um sie eindeutig mithilfe der Kamera eines Smartphones visuell zu erkennen. Wenn dies gelänge, könnten eben diese zusätzlichen Informationen sehr benutzerfreundlich und unabhängig von der äußeren Umgebung spezifisch für die unterschiedlichen Armringe zur Verfügung gestellt werden.

Außerdem sollten Möglichkeiten gesucht werden, wie verschiedene Augmentierungen in das Kamerabild eingeblendet oder beispielsweise als Ton abgespielt werden können.

Das Ergebnis des Projektes ist eine prototypische Anwendung, welche die genannten Anwendungsfälle umsetzt und demonstriert.

2. VORBETRACHTUNG

Zentrale Anforderung der Anwendung ist die visuelle Erkennung der Armringe mit der Kamera eines Smartphones. Hierfür eingesetzte Verfahren arbeiten in der Regel mit Key-Points, also Punkten, welche über bestimmte lokale Merkmalseigenschaften verfügen. Dies können beispielsweise in der näheren Umgebung eines Punktes vorhandene Farbunterschiede sein, welche auf einen Eckpunkt hindeuten. Dabei ist nicht unbedingt entscheidend, dass dieser Punkt absolut eindeutig ist, sondern lediglich, dass er mit sehr hoher Wahrscheinlichkeit wiedererkannt werden kann [1]. Aus einer Menge von Key-Points und deren Relation zueinander ergibt sich dann für ein Muster eine Art „Fingerabdruck“, welcher eine Erkennung ermöglicht.

Besonders gut funktioniert dies für zweidimensionale Objekte, wie sie in Abbildung 1 dargestellt sind. Auf der linken Seite der Abbildung sind die auf einer Spielkarte erkannten Key-Points beispielhaft dargestellt. Um nun ein Objekt wiederzuerkennen, werden Matching-Algorithmen eingesetzt, welche die gefundenen Key-Points eines Bildes mit denen eines Referenzbildes vergleichen. Dabei können je nach Algorithmus zwei Bilder selbst bei Rotationen oder einfachen Verzerrungen relativ gut und effizient einander zugeordnet werden, wie auf der rechten Seite von Abbildung 1 zu sehen ist [2].

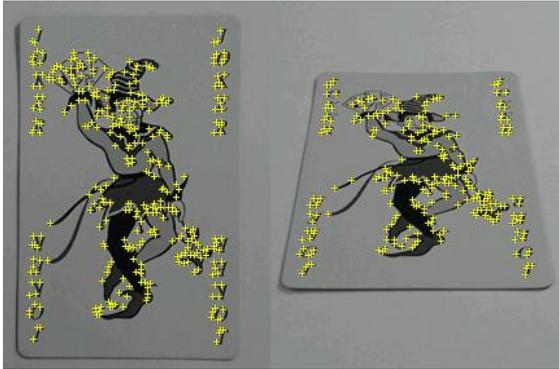


Abb. 1: Von Vuforia bestimmte Key-Points auf einem zweidimensionalen Marker aus unterschiedlichen Blickwinkeln

Deutlich schwieriger ist jedoch die Erkennung von komplexeren dreidimensionalen Objekten, da hier sehr unregelmäßige Verzerrungen und Verdeckungen möglich sind. Für die in der Anwendung verwendeten Armringe ist dies der Fall. Erschwerend kommt hinzu, dass diese relativ klein sind, sodass sie selbst bei aufwändigen Verzerrungen nur über eine sehr begrenzte Anzahl von für die Erkennung notwendigen Merkmalen verfügen, was durch Abbildung 2 verdeutlicht wird.

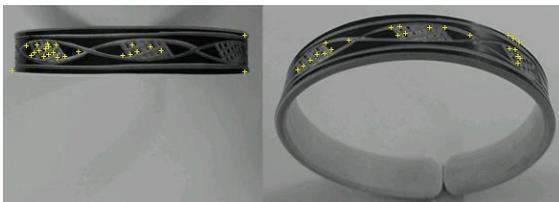


Abb. 2: Von Vuforia bestimmte Key-Points auf einem dreidimensionalen Objekt (Himba-Armring) aus unterschiedlichen Blickwinkeln

3. ANFORDERUNGEN

Das Ziel des Projektes war es, dem Nutzer durch eine interessante Präsentation Informationen zu Himba-Armringen zur Verfügung zu stellen. Neben der bereits beschriebenen Integration einer robusten Erkennung dreidimensionaler Objekte sollte der Nutzer die Möglichkeit haben, die Armringe auch ohne deren physische Verfügbarkeit betrachten zu können. Hierzu war es notwendig, texturierte Modelle von diesen erzeugen und in der Anwendung darstellen zu können.

Außerdem sollte durch eine entsprechende Gestaltung der grafischen Nutzeroberfläche Bezug auf die Thematik genommen werden, welche gleichzeitig ein hohes Maß an Intuitivität und Erwartungskonformität aufweist.

4. GENUTZTE WERKZEUGE

Für die Realisierung der Anwendung fiel die Wahl der Werkzeuge auf Unity [3] als Haupt-Framework für die Darstellung, Datenhaltung und Steuerung, sowie Vuforia [4] für die Objekterkennung und Augmentierung. Unity bietet neben der einfachen Implementierung von 3D Inhalten vor allem die Möglichkeit, die Anwendung mit minimalen Anpassungen für verschiedene Zielplattformen wie iOS oder Android zu entwickeln. Vuforia ermöglicht es dem Entwickler, visuelle Objekterkennung zu nutzen, ohne selbst die sehr aufwändige und anspruchsvolle Implementierung von Algorithmen zur Objekterkennung sowie deren Handhabung vornehmen zu müssen. Ein weiterer Vorteil ist die sehr gute Integration in Unity.

Die Erstellung der 3D-Modelle der Armringe erfolgte mit Autodesk ReMake, welches Objekte räumlich mittels Photogrammetrie aus Fotos ohne Tiefeninformation rekonstruieren kann. Dazu wird von dem zu modellierenden Objekt eine große Menge von Fotos aus verschiedenen Blickwinkeln benötigt. Durch die bereits beschriebenen Verfahren zur Erkennung und Verknüpfung von Bildmerkmalen werden die Bilder analysiert und so iterativ abgeschätzt, aus welcher Position die Fotos gemacht wurden und welche Form das Objekt hat.

5. UMSETZUNG UND TEST

Die aus den Anforderungen erarbeiteten Anwendungsfälle für die Umsetzung in der Anwendung sind in Abbildung 3 dargestellt.

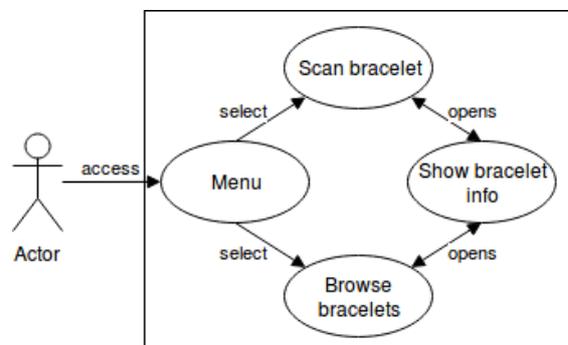


Abb. 3: Anwendungsfälle für die Anwendung

Dabei erfolgte für die prototypische Implementierung der Anwendung ein rein lokaler Ansatz, ohne Online-Backend zur Verwaltung und Wartung der Inhalte. Dies hat auch den Vorteil, dass die Anwendung weniger

Schnittstellen nach außen hat, welche potentielle Schwachstellen für die Sicherheit oder bei Aktualisierungen der Betriebssysteme darstellen. Außerdem wird während der Nutzung keine Internetverbindung benötigt, was wiederum für Touristen von Vorteil ist. So war es möglich, mehr Ressourcen für die Erkennung der Armringe einzusetzen.

Dies stellte sich als große Herausforderung dar, denn anders als zunächst angenommen, boten die Verzierungen der Armringe oft nicht genügend Merkmale für eine robuste Erkennung. Hier trat auch einer der größten Nachteile von Vuforia zum Vorschein, da es dem Entwickler zwar Schnittstellen auf einer sehr hohen Abstraktionsebene liefert, was die Einbindung in die Anwendung erheblich erleichtert, gleichzeitig im Falle von Problemen jedoch nur wenig Möglichkeiten der Optimierung bietet. Da Vuforia nicht quelloffen ist, bleiben auch Details wie die verwendeten Algorithmen zur Objekterkennung im Verborgenen.

Verfügt ein Marker über eine große Anzahl von aussagekräftigen Merkmalen, so kann dieser oft trotz vorhandener Störungen sicher erkannt werden. Häufig auftretende Störgrößen sind beispielsweise eine schlechte Qualität der Kamera mit einem hohen Bildrauschen, niedriger Auflösung, schlechter Fokussierung oder wenig Kontrasten. Äußere Störgrößen können beispielsweise schnelle Bewegungen, schlechte Lichtverhältnisse, ungünstige Schattenwürfe oder eine teilweise Verdeckung des Objektes sein. Bei einer entsprechend hohen Anzahl von Merkmalen kann oft, trotz dieser Störgrößen, eine genügend hohe Anzahl von Merkmalen gefunden und verknüpft werden. Wie aber bereits in Abbildung 2 dargestellt, verfügen die Armringe nur über sehr wenige Merkmale, sodass selbst unter sehr guten Bedingungen eine Erkennung schwierig ist. Treten hierzu noch Störgrößen auf, was in einem realen Umfeld immer der Fall ist, erschwert dies die Erkennung zusätzlich. Bereits bei anderen Projekten, bei denen Vuforia mit zweidimensionalen Markern verwendet wurde, sank die Erkennungsrate bei sehr kleinen Markern. Als grober Richtwert für eine gute Erkennungsrate wurde eine Mindestgröße von etwa 8 x 8 cm für die Marker ermittelt. Selbst wenn diese über sehr viele Merkmale verfügen, sinkt dort die Nutzerfreundlichkeit sehr stark, da das Gerät sehr nah an den Marker gehalten werden

muss und so nur wenig Bewegungsspielraum besteht. Außerdem treten hier auch bei geringen Abständen zum Objekt oft Fokussierungsprobleme auf.

Da die Armringe in der Frontalansicht eine Projektionsfläche von etwa 1-2 x 6 cm aufweisen, waren die Voraussetzungen für eine zuverlässige Erkennung also denkbar schlecht. Dennoch ist es gelungen, für einige Armringe eine zuverlässige Erkennung und Augmentierung zusätzlicher Inhalte zu realisieren, wie in Abbildung 4 auf der linken Seite zu sehen ist.



Abb. 4: Screenshots der Anwendung: links - Augmentierung einer Himba anhand eines erkannten Armrings; rechts - Informationsansicht zu einem Armring.

Neben der Erkennung der Armringe und den Augmentierungen sollten auch die Informationen in einer ansprechenden Form dargestellt werden. Dazu wurde auch hier versucht, das Thema der Himba in die optische Gestaltung der Anwendung einfließen zu lassen. Dies zeigt sich beispielsweise in den verwendeten Farben und Darstellungen, welche sich sehr an natürlichen und mit den Himba assoziierten Rohstoffen wie Leder, Sand und Ton orientieren.

Da der Nutzer auch die Möglichkeit haben soll, die Informationen zu allen Armringen anzuzeigen ohne dass diese physisch vorliegen, wurde ein Menü erstellt, in dem Abbildungen der Ringe so angeordnet sind, als ob sie auf einem Arm aufgereiht sind (Abbildung 5). Bei der Informationsseite selbst war es wichtig, diese möglichst übersichtlich zu gestalten. Hierfür wurden Texte und die photogrammetrisch erzeugten 3D-Modelle jeweils mit einem halbtransparenten Hintergrund unterlegt, sodass diese gut

voneinander getrennt sind, optisch jedoch dennoch harmonieren (Abbildung 4, rechts). Mittels Wisch-Gesten können diese rotiert und so genauer untersucht werden.



Abb. 5: Das Menü zur Auswahl der Armringe

Durch Interviews mit den Himba in ihrer Muttersprache Otjherero und einem Übersetzer konnten die zusätzlichen Informationen recherchiert werden. Da diese Sprache nur noch von sehr wenigen Menschen gesprochen wird, wurden als ein weiteres Feature zum Teil die Audiodateien mit den Texten verknüpft. So ist es für den Nutzer möglich, die Texte in Englisch zu lesen und diese auch in Otjherero zu hören und den Klang dieser Sprache zu erleben.

Bei der Entwicklung wurden sehr gute Erfahrungen mit Unity als Entwicklungswerkzeug gemacht. Für sehr viele Komponenten existieren bereits Muster, welche mit anderen Komponenten verknüpft, den eigenen Wünschen angepasst oder erweitert werden können.

7. ZUSAMMENFASSUNG UND AUSBLICK

Im Rahmen des Projektes konnte gezeigt werden, dass mit Augmented Reality auch für sehr spezielle Anwendungsfälle interessante Ansätze möglich sind. Die Leistungsfähigkeit der vorhandenen Werkzeuge ermöglicht es dabei, auch für kleinere Projekte nutzerfreundliche und innovative Anwendungen zu gestalten.

So ist es gelungen, in dem Prototypen die Erkennung für einige Armringe zu erreichen, wie es auf der linken Seite von Abbildung 4 zu sehen ist.

Leider ist es jedoch nicht gelungen, die gewünschte Robustheit der Erkennung zu erreichen, welche es zulässt, die jeweiligen Armringe auch bei ungünstigen Lichtverhältnissen oder bei nur wenigen verfügbaren Merkmalen sicher zu erkennen. Aus diesem Grund konnten einige Armringe nicht für die Anwendung verwendet werden. Die Gründe hierfür wurden bereits im vorherigen Abschnitt diskutiert. Für eine zukünftige Weiterentwicklung wäre somit eine robustere Erkennung der Armringe der erste Ansatzpunkt. Durch ausführliche Versuche mit größeren und wirklich kreisrunden Ringen als Marker könnten eventuelle Fehlerquellen eingegrenzt und festgestellt werden. Außerdem könnten Versuche mit anderen Werkzeugen wie OpenCV oder ARToolkit zeigen, ob sie eine robustere Erkennung der Armringe ermöglichen. Besonders das quelloffene OpenCV stellt sowohl im akademischen als auch industriellen Rahmen ein oft genutztes und leistungsfähiges Werkzeug dar. Jedoch sind Vuforia und OpenCV nur bedingt vergleichbar. Während Vuforia bereits viele Funktionen wie Objekterkennung und -verfolgung auf einer hohen Abstraktionsebene zur Verfügung stellt, ist OpenCV eine Programmierschnittstelle für Bildverarbeitung und maschinelles Sehen, die vor allem eine Vielzahl von effizienten und schnellen Algorithmen in diesen Gebieten anbietet. Es müssen also viele Funktionalitäten die Vuforia mitbringt, wie Datenhaltung, Marker-Erstellung, Objektverfolgung, etc., selbst implementiert werden. Dadurch wäre es möglich, einen speziell für den Anwendungsfall optimierten Code zu schreiben, was jedoch auch aufgrund der Komplexität des Themas sehr viel Zeit in Anspruch nimmt. Ein weiterer Vorteil von OpenCV ist die Auswahl aus einer Vielzahl verschiedener Algorithmen für die einzelnen Schritte der Objekterkennung. Diese sind bei Vuforia unbekannt, da es nicht quelloffen ist und der Hersteller diese Information nicht veröffentlicht. Für eine relativ schnelle Implementierung ist also Vuforia durchaus sehr gut geeignet. Wenn die Erkennung wie im Falle dieses Projektes an ihre Grenzen stößt, ist eine Nutzung oder zumindest Untersuchung anderer Werkzeuge sinnvoll.

Außer Frage steht jedoch, dass bereits heute verschiedene Algorithmen in der Lage sind, auch sehr schwierige Objekte zu erkennen. Die hierfür benötigte Rechenleistung ist jedoch

enorm und vergrößert sich mit steigender Auflösung der Bilder sowie der gewünschten Wiederholrate. Da für das Nutzererlebnis in AR-Anwendungen möglichst keine Verzögerungen bei der Erkennung und Verfolgung von Objekten auftreten sollen, sind entsprechende Vereinfachungen notwendig, welche zwar die benötigte Rechenleistung, unter Umständen aber auch die Qualität der Erkennung verringern [6]. So ist es beispielsweise möglich, den Bereich, in dem sich das Objekt wahrscheinlich im Kamerabild befindet, anhand dessen Position in den vorherigen Bildern vorherzusagen. Ist das Objekt in diesem geschätzten Bereich nicht zu finden, kann es sein, dass die Erkennung, verloren geht. Um es nun wieder zu finden muss es erneut im Bild gefunden werden, was sehr viel Rechenzeit benötigt. Dieses Problem tritt insbesondere bei schnellen Richtungswechseln auf. Eine Möglichkeit, die Qualität dieser Vorhersage zu verbessern, ist, weitere Sensoren wie Accelerometer, Gyroskope oder der Kompass zu verwenden. Deren Messdaten können ohne großen Ressourcenaufwand abgefragt und verarbeitet werden und sie sind in der Regel in Smartphones vorhanden. Ein weiterer Ansatzpunkt für eine Optimierung ist das Wissen über die Gemeinsamkeit aller Arminge bezüglich ihrer Form und Farbe. Diese haben in der Regel eine bestimmte Kontur, welche sich mathematisch beschreiben lässt sowie einen bestimmten Anteil von weißen Stellen. Somit lässt sich ebenfalls schnell über die Farbwerte und eine grobe Suche der äußeren Form die wahrscheinliche Position im Bild suchen. Die genaue Position und Orientierung kann dann mithilfe aufwändiger Algorithmen in einem kleinen Bildbereich erfolgen. Eben diese speziellen Fälle ermöglichen Optimierungen, die bei eigenen Entwicklungen mit Vuforia oft nicht möglich sind, sehr wohl aber mit OpenCV.

Unabhängig davon haben die letzten Jahre eine rasante Entwicklung bei der Leistungsfähigkeit der Kameras und Prozessoren von mobilen Geräten sowie im Bereich der AR- und Bildverarbeitungs-Software gezeigt. Hinzu kommt, dass Objekterkennung und -verfolgung in immer mehr Bereichen wie beispielsweise der Automobiltechnik oder der Unterhaltungselektronik eine wichtige Rolle spielt. Daher ist anzunehmen, dass in naher Zukunft einige der hier angesprochenen Probleme gelöst werden können.

8. LITERATURHINWEIS

- [1] Lowe, David G.: Distinctive Image Features from Scale-Invariant Keypoints, *International Journal of Computer Vision*. Band 60, Nr. 2, S. 91-110, 2004
- [2] Bay, Herbert et. al.: SURF: Speeded Up Robust Features, *9th European Conference on Computer Vision*, Graz, 7.-13.05.2006
- [3] Unity, Unity Technologies, www.unity.com, Abgerufen am: 19.10.2017
- [4] Vuforia, PTC Inc., www.vuforia.com, Abgerufen am: 19.10.2017
- [5] OpenCV (Open Source Computer Vision Library), Willow Garage, www.opencv.org, Abgerufen am: 19.10.2017
- [6] Iñigo Barandiaran, Céline Paloc, Manuel Graña, *Real-time optical markerless tracking for augmented reality applications*. *Journal of Real-Time Image Processing*, Volume 5, Number 2, S. 129–138, 2010