

Analyse von Digitalkameras im Infrarotbereich für die 3D-Rekonstruktion von Personen

Marina Leontopoulos

Technische Universität Berlin

Abstract

Einer der wichtigsten Aspekte bei Virtual Reality-Inhalten ist die Immersion. Gerade bei Darstellungen von Personen ist der menschliche VR-Betrachter in der Lage kleinste Ungereimtheiten ausfindig zu machen. Im Gegensatz zum *Computer Graphics*-Ansatz wird beim volumetrischen Video keine Figur per Computer generiert, die einen echten Menschen nachbilden soll. Stattdessen wird eine reale Person in *Volumetric Video Studios* mit 32 Kameras gleichzeitig aufgenommen und anschließend zu mehreren 3D-Modellen rekonstruiert. Diese 3D-Modelle lassen sich danach in eine VR-Umgebung einfügen und als begehbarer Film betrachten.

Um die Qualität der Rekonstruktionen weiter zu verbessern, wurde untersucht, inwiefern sich infrarote und rote Strahlung nutzen lassen, um auf glatter Haut Muster zu erzeugen. Zudem wurde ein Messkörper entwickelt und aufgebaut, der einem menschlichen Kopf in seiner Form, Struktur sowie seinen spektralen Reflexionseigenschaften nachempfunden ist.

1 Einleitung

Es existieren zwei *Volumetric Video Studios*, die aus der Forschung des Fraunhofer Heinrich-Hertz-Instituts entstanden sind. Bei ihnen handelt es sich um Rotunden aus diffus hinterleuchtetem Material von fünf bis sechs Metern Durchmesser, auf deren Zentren 32 Videokameras gerichtet sind. Aus den aufgenommenen 25 Bildern pro Sekunde in 5K werden anschließend für jeden Zeitabschnitt aus den 32 Kameraperspektiven automatisch Pixelkorrespondenzen ermittelt, aus denen schließlich 3D-Modelle einer Person erzeugt werden. Auf diesem Weg entsteht eine 3D-Modellsequenz, die in einer beliebigen virtuellen Umgebung eingebettet als volumetrisches Video abgespielt werden kann.

Problemstellung

Dadurch, dass die Anzahl der verwendeten Kameras im Vergleich zu 3D-Scan-Studios deutlich geringer sein muss, um die anfallenden Datenströme der Videos verarbeiten zu können (ca. 1 TB/min), ist der erreichbare Detailreichtum eingeschränkt. Für den Zuschauer macht sich dies bei genauem Hinsehen u.a. durch fehlende scharfe Kanten am Übergang zwischen Mund und Nase bemerkbar. Auch



die Rekonstruktion von glatter Haut stellt eine Herausforderung dar, weil es bei strukturlosen Oberflächen kaum möglich ist, Pixelkorrespondenzen zu erkennen.

Methodik

In dieser Arbeit wurden daher Verfahren untersucht, die die Rekonstruktionsqualität von Gesichtern verbessern. Dabei soll anders als in zahlreichen Untersuchungen, die sich mit der Optimierung der Algorithmen von 3D-Rekonstruktionsverfahren beschäftigen, ein neuer Weg gezeigt werden, der sich ausschließlich mit Anpassungen des Quellmaterials befasst, bevor es überhaupt für die Rekonstruktion eingesetzt wird. Die Eigenschaften von Kamerasensoren in Bezug auf ihre Interpretation von Strahlung wurden hierfür als Grundlage gewählt, wobei besonderes Augenmerk auf den Infrarotbereich gelegt wurde. Die Informationen aus den Bereichen roter und infraroter Strahlung wurden genutzt, um auf glatter Haut tiefere Strukturen hervorzuheben.

Konkret wurden dafür

- 1) die Sensoren von drei Kameras unterschiedlicher Kategorien spektral vermessen und darauf aufbauend die Farbörter bestimmt, in die infrarote Strahlung im sRGB-Farbraum abgebildet wird. Diese wurden in infrarotbehaftetem Bildmaterial detektiert und ihre Helligkeit reduziert, wodurch Muster auf der Haut sichtbar wurden.
- 2) ein Messkörper für die Validierung verschiedener Beleuchtungssituationen im Studio hergestellt, der anstelle eines Menschen aufgenommen und rekonstruiert werden kann. Für die Gestaltung der Oberfläche wurden die spektralen Reflexionseigenschaften von menschlicher Haut betrachtet und mehrere Lederproben auf ihre Ähnlichkeit hin untersucht.

2 Vermessung der Kamerasensoren

Bei einer Digitalkamera dienen Fotodioden als Empfänger von Licht. Dabei wird pro

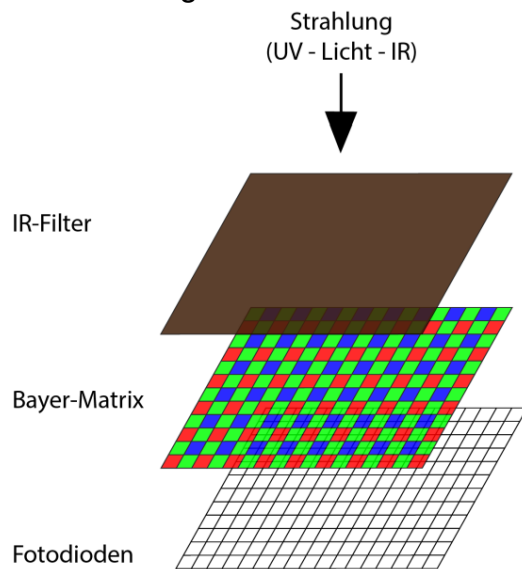


Abbildung 1: Schematischer Aufbau eines Bayer-Sensors

Diode ein Pixel im späteren Bild gewonnen. Da diese Empfänger von sich aus lediglich Schwarzweißbilder erzeugen können, werden Kamerasensoren mit einer sog. Bayer-Matrix versehen – einer Matrix auf roten, grünen und blauen Filtern (siehe Abbildung 1). Folglich entstehen dabei drei Bilder, die jeweils in ihrer Auflösung deutlich unterhalb der gesamten Sensor-Auflösung liegen. Auch unterscheiden sich die einzelnen Pixel leicht von ihrer Position. Um ein Farbbild zu erzeugen, muss daher aus den drei Teilbildern ein Gesamtbild geschätzt werden. Diesen Vorgang nennt man

auch *debayering*. Die Bayerfilter sind allerdings durchlässig für infrarote Strahlung, weshalb ein zusätzlicher Infrarotsperfilter

eingesetzt werden muss, um die Farbinterpretation der Kamera nicht zu verfälschen.

Am empfindlichsten sind Digitalkameras zwischen ca. 400 nm und 1000 nm (ohne Infrarotsperfilter) bzw. 400 nm und 700 nm (mit Infrarotsperfilter). Die Gewichtung der Kanalhelligkeiten hängt dabei von der spektralen Empfindlichkeit des Sensors sowie der Umsetzung des Sensormaterials in ein Bild in einem konkreten Farbraum ab. Als Farbraum wurde der im Workflow des Volumetric Video Studios eingesetzte sRGB-Farbraum gewählt.

Während sich gängige Messverfahren für die Qualität von Kamerasensoren stets am menschlichen Auge als Referenz orientieren, interessiert hier lediglich, wie Wellenlängen in Farbörter umgesetzt werden.

Die Sensoren wurden hierzu schmalbandig gefilterter Strahlung ausgesetzt und dabei Bilder aufgenommen.

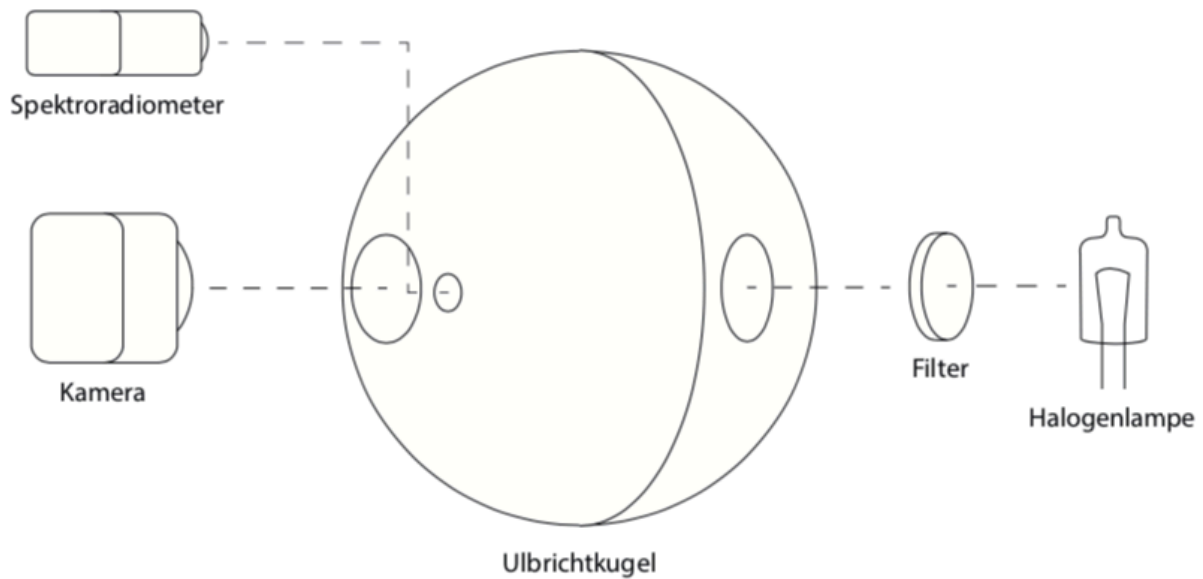


Abbildung 2: Blockdiagramm des Messaufbaus

Der Messaufbau ist in Abbildung 2 dargestellt. Dabei wurde ein Temperaturstrahler als Lichtquelle eingesetzt, dessen Spektrum durch wechselbare Interferenzfilter im Bereich von 350 nm bis 1100 nm angepasst wurde, was dem messbaren Bereich des verwendeten Spektroradiometers entspricht.

Bei den drei vermessenen Kameras handelt es sich um folgende Modelle:

- Ximea CB200CG-CM (Industriekamera ohne Infrarotsperfilter)
- Panasonic AU-EVA1 (Cinecam mit entfernbarem Infrarotsperfilter)
- Canon EOS 5D Mk III (DSLR)

Bei Ersterer handelt es sich um das Modell, das in den *Volumetric Video Studios* zum Einsatz kommt.

Die aufgenommenen Bilder von jeder der drei Kameras wurden entsprechend ihrer Eigenheiten vorverarbeitet und in 8-Bit TIFF-Einzelbilder im sRGB-Farbraum konvertiert sowie anschließend in Normfarbwerte umgerechnet. Dadurch konnte der Zusammenhang zwischen den Wellenlängen einer Lichtquelle und ihren korrespondierenden Farbörtern im Bildmaterial hergestellt werden.

Folgende Messwerte (Abbildung 3 bis Abbildung 5) konnten dabei ermittelt und in die CIE-Normfarbtafel eingefügt werden. Dabei unterscheiden sich die drei vermessenen Kameras deutlich in Bezug auf ihre Interpretation von monochromatischem Licht im sRGB-Farbraum.

Auswertung

Damit der sRGB-Farbraum vollständig genutzt werden kann, muss monochromatisches Licht (schwarze Kreise) stets auf dem Rand des aufgespannten Dreiecks abgebildet werden. Ist dies nicht der Fall, wie bei Abbildung 3 für die im *Volumetric Video Studio* verwendete Ximea CB200CG-CM zu erkennen, weist der

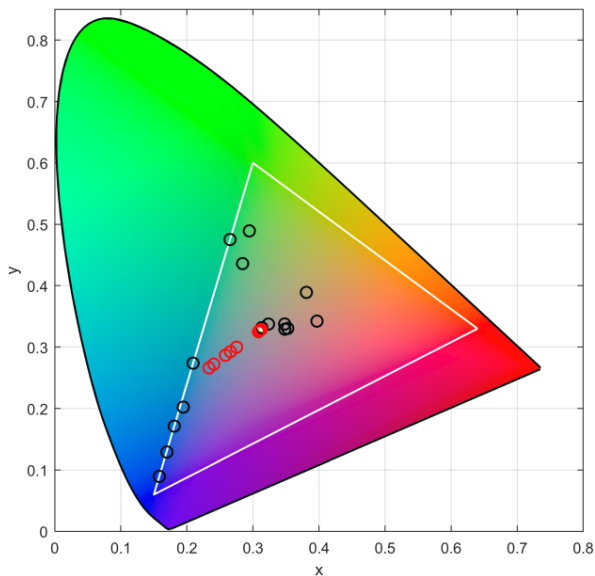


Abbildung 3: Ximea CB200CG-CM
Messpunkte innerhalb der CIE-
Normfarbtafel

Sensor keine ausreichende Kanaltrennung auf. Infrarote Strahlung hingegen ist kein Teil des sichtbaren Lichts und wird von Kamerasensoren i.d.R. als Farbmischung interpretiert, da alle Filter der Bayer-Matrix infrarotdurchlässig sind. Ein Infrarotsperrfilter sperrt allerdings bereits im sichtbaren Bereich, weshalb Kameras ohne Filter generell eine schlechtere Kanaltrennung aufweisen, wie auch beim Vergleich von Abbildung 4 links und rechts zu sehen.

Bei der Panasonic AU-EVA1 lässt sich der Einfluss des Infrarotsperrfilters näher betrachten.

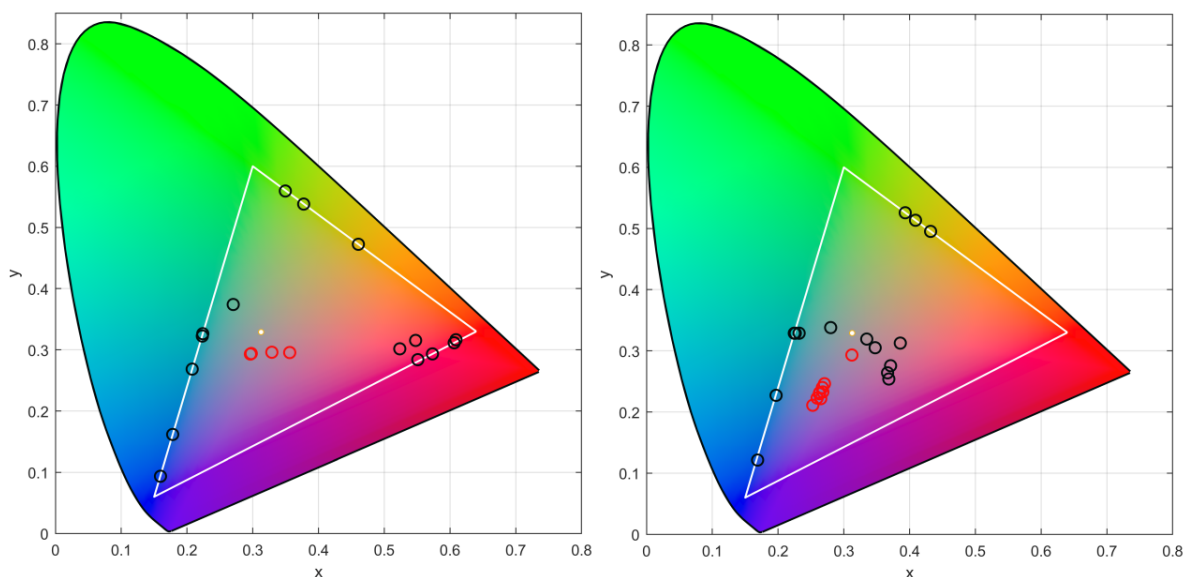
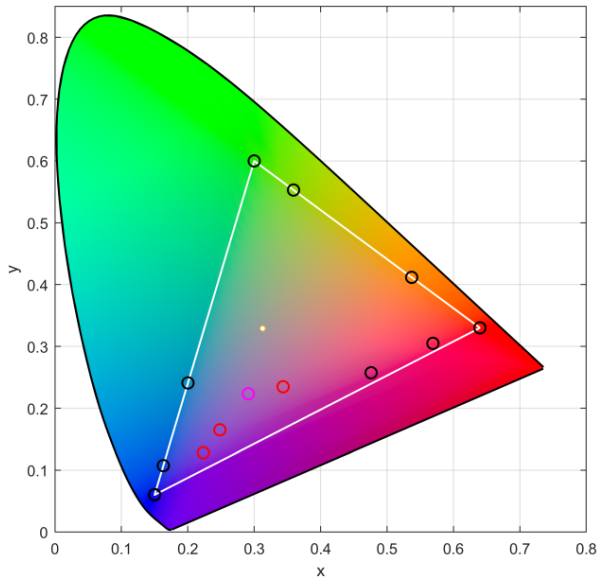


Abbildung 4: Panasonic AU-EVA1 Messpunkte innerhalb der CIE-Normfarbtafel
(links mit Infrarotsperfilter, rechts ohne Infrarotsperfilter)

Während sich bei aktivem Sperrfilter die Kanaltrennung bis auf geringe Abweichungen wie gewünscht verhält (Abbildung 4 links), sind ohne den Filter keine Rottöne mehr darstellbar (Abbildung 4 rechts).



Die beste Kanaltrennung unter den vermessenen Kameras wies die Canon 5D Mk III auf. Die in Abbildung 5 eingezeichneten roten Kreise zeigen die Farbörter von Infrarotstrahlung. Der rosa Kreis beschreibt die Interpretation der Kamera von einer Wellenlänge von ca. 700 nm, die sich zwar im sichtbaren Bereich befindet, vom eingebauten Infrarotsperfilter allerdings herausgefiltert wird, da er ab ca. 675 nm greift (vgl. hierzu Abbildung 6).

Abbildung 5: Canon 5D Mk III
Messpunkte innerhalb der CIE-Normfarbtafel

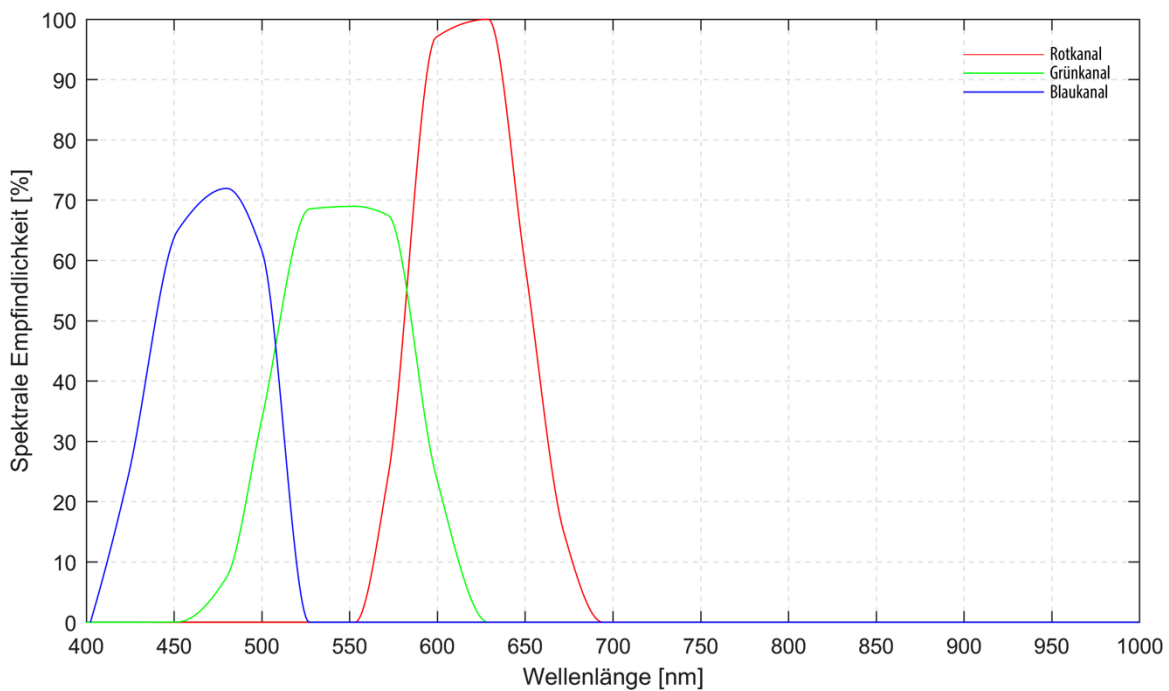
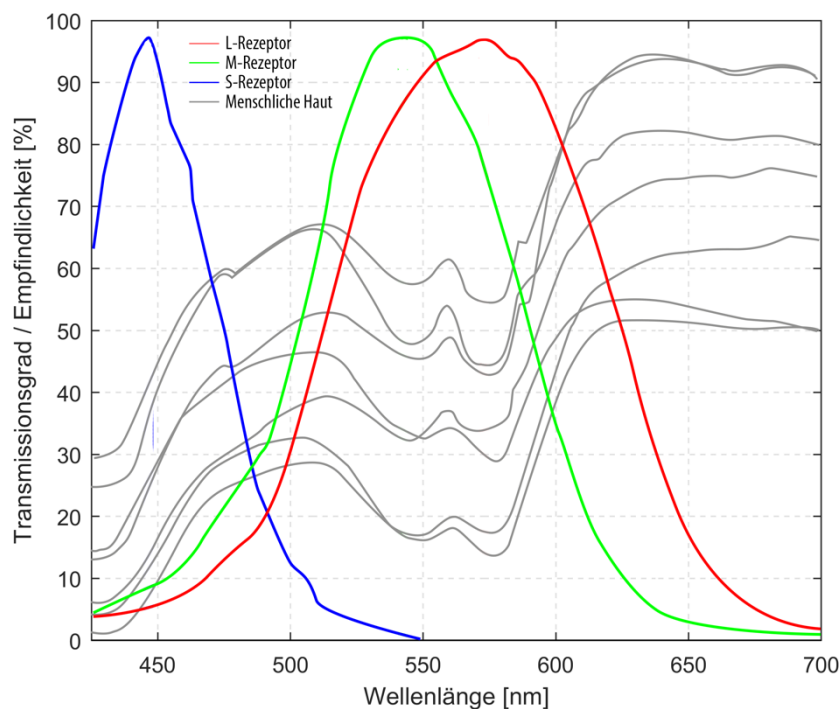


Abbildung 6: Spektrale Empfindlichkeit der Canon 5D Mk III

3 Konstruktion und Einsatz des Messkörpers

Eine Person ist nicht in der Lage vollständig still zu bleiben. Für den Vergleich zweier Aufnahmen zur 3D-Rekonstruktion ist eine vollkommen unbewegte Person allerdings unabdingbar. Daher wurde ein Messkörper angefertigt, der einem menschlichen Kopf nachempfunden ist. Dabei war neben der Erscheinungsform vor allem wichtig, dass sich die Oberfläche des Messkörpers bzgl. ihrer Struktur und ihrer spektralen



Reflexionseigenschaften ähnlich wie menschliche Haut verhält.

Die gewünschten spektralen Reflexionseigenschaften sind durch die Messungen von [1] bekannt und wurden nach [2] mit den spektralen Empfindlichkeitskurven der Zapfen des menschlichen Auges (nach [3]) in Verbindung gebracht (siehe Abbildung 7).

Abbildung 7: Spektrale Reflexionseigenschaften von menschlicher Haut und Zapfenempfindlichkeit des Auges (zusammengefasst und skaliert) nach [2]

Das Material, das am geeignetsten erschien, um menschliche Haut nachzuahmen, ist Leder. Aus diesem Grund wurden verschiedene Lederproben mit lebendiger menschlicher Haut verglichen und dabei festgestellt, dass mit Make-up geschminktes Schafsleder sowohl ähnliche spektrale Reflexionseigenschaften (siehe Abbildung 8), als auch Strukturen aufweist.

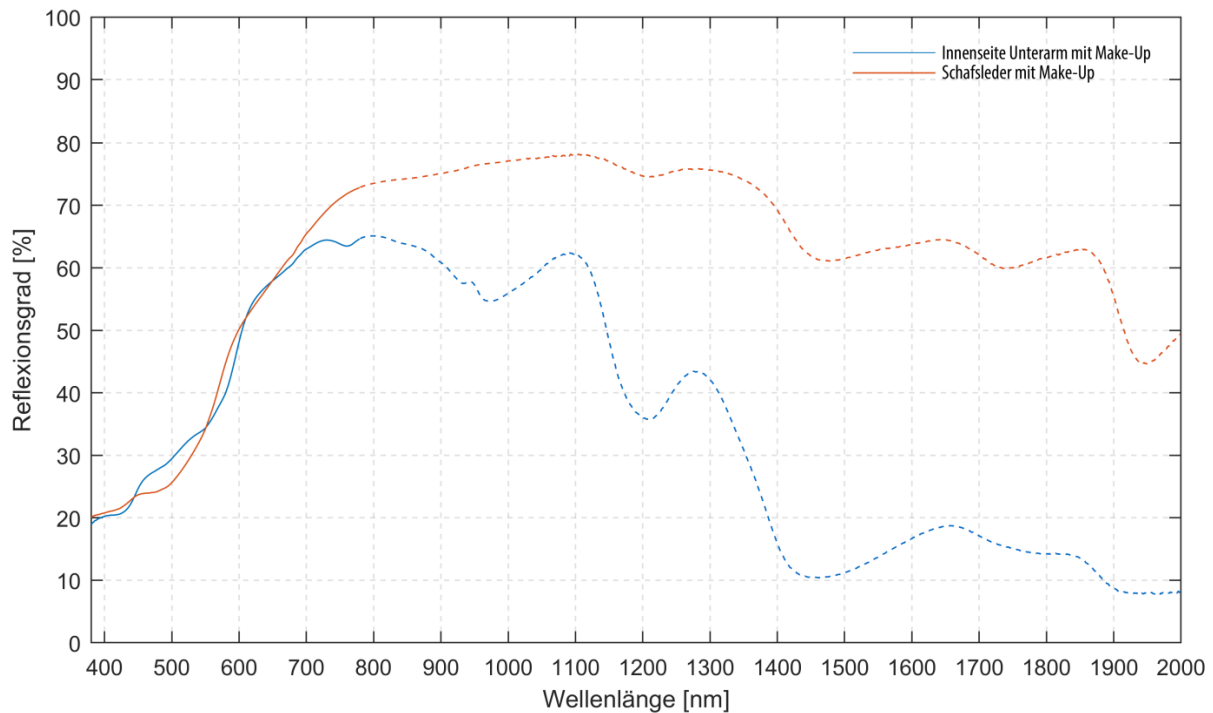


Abbildung 8: Vergleich der spektralen Reflexionseigenschaften von menschlicher Haut

Herstellung

Der Messkörper selbst wurde mithilfe von 3D-Modellierungs-Software gestaltet und mit einem 3D-Drucker in Lebensgröße aus PLA gefertigt (siehe Abbildung 9).

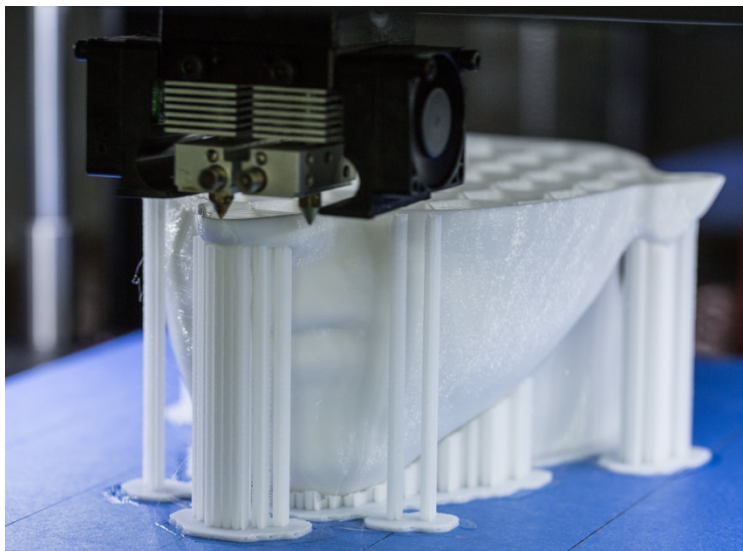


Abbildung 9: 3D-Druck des Messkörpers

Anschließend wurde der Plastikträger mit Schafsfleder überzogen und geschminkt. Der fertige Messkörper ist in Abbildung 10 dargestellt.



Abbildung 10: Fertiger Messkörper mit Make-up und Behaarung

Testaufnahmen und Nachbearbeitung



Abbildung 11: Abschwächung des Infraroteinflusses (links), Abschwächung des Rotkanals (rechts)

Mit dem fertigen Messkörper wurden bei zwei verschiedenen Beleuchtungssituationen (Halogen und LED) Aufnahmen gemacht. In Abbildung 11 sind zwei bearbeitete Aufnahmen von der Ximea CB200CG gegenübergestellt.

- 1) Im Bild auf der linken Seite wurde der Rekonstruktionskörper mittels Halogenscheinwerfer beleuchtet. Anschließend wurden diejenigen Farben in ihrer Helligkeit reduziert, die mit infraroten Wellenlängen korrespondieren.
- 2) Das rechte Bild zeigt den Messkörper unter LED-Beleuchtung ohne Infrarotanteil, nachdem der Rotkanal aus der Aufnahme entfernt wurde. Diese Möglichkeit wurde insbesondere untersucht, da bei den *Volumetric Video Studios* LED-Leuchten eingesetzt werden, die frei von Infrarotstrahlung sind. Wenn auch deutlich schwächer, lassen sich hier ebenfalls Muster auf der Haut erkennen.

Demnach sollte idealerweise bei vorhandenem Infrarotanteil bei der Beleuchtung Variante 1) bevorzugt werden, wenn die Qualität des Gittermodells selbst am wichtigsten ist. Durch die Einschränkung der darstellbaren Farben eignen sich derart aufgenommene Bilder allerdings kaum zur Gewinnung von Texturen. Demgegenüber stellt Variante 2) einen Kompromiss dar, wenn das Bildmaterial ebenfalls Texturen liefern soll bzw. keine infrarote Strahlung ausgenommen werden kann.

4 Anwendung im *Volumetric Video Studio*

Die in den *Volumetric Video Studios* eingesetzte Beleuchtung ist frei von Infrarotstrahlung. Aus diesem Grund wurde die Entfernung des Rotkanals von Bildern für die 3D-Rekonstruktion mit vorhandenem Material durchgeführt.



Abbildung 12: Maria RGB

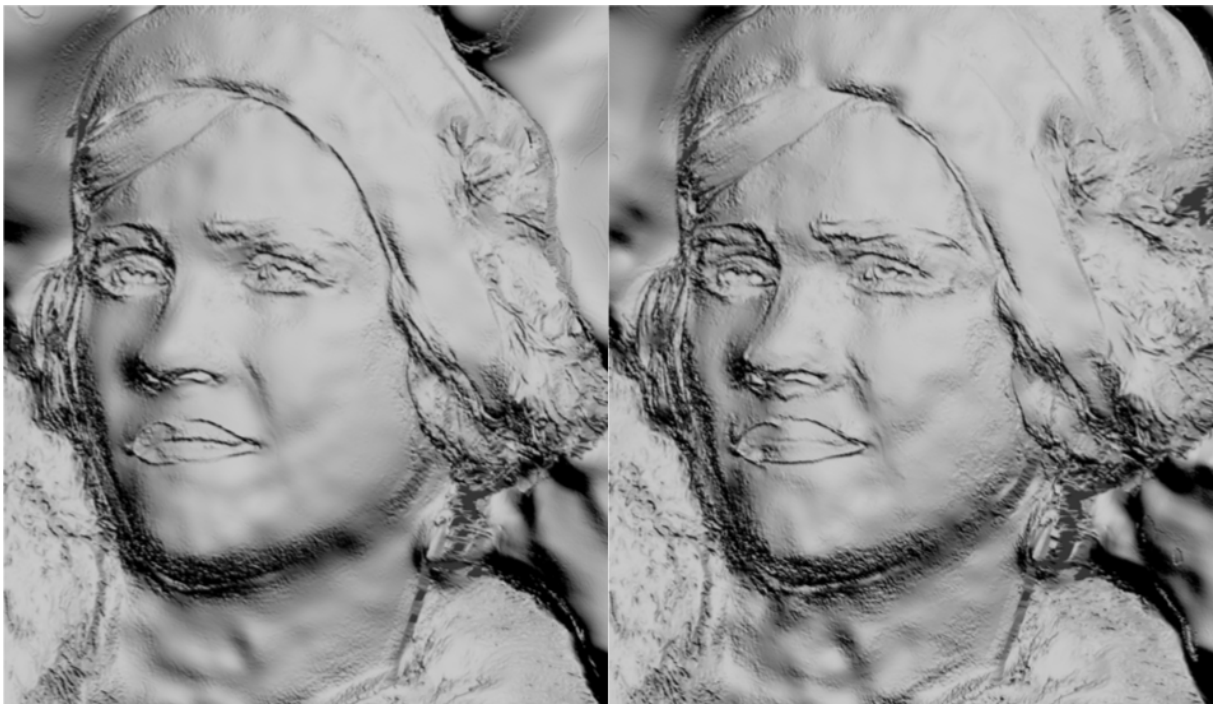
Abbildung 12 und Abbildung 13 zeigen Franziska Brandmeier als Maria aus „Ein Ganzes Leben / 100 Jahre UFA“.

Während bei Abbildung 12 Rot-, Grün- und Blaukanal in gleichen Anteilen gewichtet wurden, um das Schwarzweißbild zu erzeugen, besteht Abbildung 13 lediglich aus den Informationen des Blau- und Grünkanals. Auch hier ist ähnlich wie bei Abbildung 11 (rechts) mehr Musterung auf der glatten Haut zu erkennen.



*Abbildung 13: Maria ohne
Rotkanal*

Beide Datensätze wurden zur 3D-Rekonstruktion herangezogen und bezüglich ihres Detailgrads miteinander verglichen.



*Abbildung 14: 3D-Rekonstruktion von Maria – links gleichgewichtete RGB-Kanäle,
rechts entfernter Rotkanal*

In Abbildung 14 (links) ist der Übergang zwischen Mund und Nase weniger scharf als in Abbildung 14 (rechts). Auch die Nasenspitze konnte durch die Entfernung des Rotkanals etwas detaillierter rekonstruiert werden. Die glatte Haut der Wangen ist

durch diese Maßnahme zwar etwas rauschbehafteter, aber dafür konnte die Falte zwischen Mund und Wange definierter rekonstruiert werden.

5 Fazit

Es konnte gezeigt werden, dass infrarote Strahlung von Digitalkameras im sRGB-Farbraum abgebildet wird. Mithilfe der Kenntnis der genauen Farbörter lässt sich Bildmaterial für die 3D-Rekonstruktion von Personen gezielt nachbearbeiten, um Muster auf der Haut zu erzeugen, die sonst nicht sichtbar sind.

Auch ohne den Einfluss von Infrarotstrahlung lässt sich der Detailreichtum einer 3D-Rekonstruktion von Personen erhöhen, indem die zu rekonstruierenden Schwarzweißbilder ohne Rotkanal erzeugt werden.

6 References

- [1] Elli Angelopoulou, "The Reflectance Spectrum of Human Skin," Technical Report, University of Pennsylvania Department of Computer und Information Science Technical, 1999.
- [2] Alexis van Hurkman, Color Correction Handbook: Professional Techniques for Video and Cinema. Digital Video & Audio Editing Courses. Pearson Education, 2013.
- [3] Andrew Stockman, Donald I. A. MacLeod und Nancy E. Johnson, "Spectral sensitivities of the human cones," J. Opt. Soc. Am. A 10.12 (Dez. 1993), S. 2491–2521.