



FAKULTÄT FÜR **INFORMATIK**

# Evaluierung von Fehler- metriken für Image Matting

DIPLOMARBEIT

zur Erlangung des akademischen Grades

**Diplom-Ingenieur/in**

im Rahmen des Studiums

**Medieninformatik**

eingereicht von

**BSc Pamela Rott**

Matrikelnummer 0025316

an der

Fakultät für Informatik der Technischen Universität Wien

Betreuung:

Betreuerin: Prof. Mag. DI Dr. Margrit Gelautz

Mitwirkung: Univ.-Ass. Mag. DI Christoph Rhemann

Wien, 11.05.2009

\_\_\_\_\_  
(Unterschrift Verfasser/in)

\_\_\_\_\_  
(Unterschrift Betreuer/in)

BSc Pamela Rott  
Winzerstraße 6-8/1/2  
1130 Wien

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

---

Ort, Datum, Unterschrift

## **Kurzfassung**

Image Matting zielt darauf ab, ein Vordergrundobjekt aus einem beliebigen Bild herauszulösen, welches als eine Kombination aus Vorder- und Hintergrund beschrieben wird. Das Resultat eines Mattingverfahrens ist eine Alpha Matte, welche den Einfluss des Vorder- sowie Hintergrundes an jedem Pixel im Bild definiert. Die Vielzahl der Mattingverfahren bringt Resultate von unterschiedlicher Qualität hervor. Um die visuelle Qualität einer Alpha Matte festzustellen, können sowohl subjektive als auch objektive Evaluierungsverfahren eingesetzt werden. Der Mensch spielt hier eine wesentliche Rolle, da das menschliche Auge Kriterien zur Beurteilung heranzieht, die entscheidender sein können als die pixelweise Differenz zum Referenzbild. Subjektive Evaluierungsverfahren in Form von Benutzerstudien sind jedoch sehr aufwändig, weshalb automatisierte Berechnungen notwendig sind. Diese Berechnungen werden aber nur dann erfolgreich sein, wenn die Bewertung des Menschen damit korreliert. Dies motiviert die Verwendung von HVS (Human Visual System)-basierten Methoden. In dieser Studie werden Faktoren wie die räumliche Verbundenheit (Connectivity), der Gradientenverlauf und die auftretenden Artefakte im Vorder- und Hintergrund herangezogen und durch eine Auswahl an Probanden sowie automatisierte Berechnungen bewertet. Anschließend werden die Resultate der Probanden und der automatisierten Berechnungen verglichen. Die Auswertung der Studie hat gezeigt, dass Objekte mit starker Verbundenheit und niedrigem Gradientenverlauf von den Probanden tendenziell als visuell hochwertig eingestuft werden. Traditionell verwendete pixelbasierte Fehlermaße korrelieren in diesem Fall nicht mit den Probanden und dementsprechend nicht mit der visuellen Wahrnehmung. Bei der im Vorfeld getroffenen Annahme hingegen, dass Artefakte im Hintergrund störender wirken als im Vordergrund, konnte im Rahmen dieser Studie keine Einigkeit unter den Probanden festgestellt werden.

## **Abstract**

Image Matting aims to separate a foreground object from an arbitrary image, which is explained as a combination of foreground and background. The result of a matting procedure is an alpha matte, which defines the influence of foreground and background on every pixel. The variety of matting approaches creates results of different quality. To measure the visual quality of an alpha matte, one can use subjective and objective evaluation methods. The human viewer plays a decisive role, as the human eye uses criteria for evaluation which may be even more important than the pixel-wise difference to the reference image. Subjective evaluations in the course of user studies are very time-consuming, thus automated calculations are necessary. These calculations will succeed, if they correlate with human perception. This motivates the use of methods that take the human visual system into account. In this work, a user study and automated calculations are conducted to infer the visual quality of matting results in the presence of three different classes of artefacts. These classes are spatial connectivity, gradient and possibly arising artefacts in foreground and background. Afterwards the results of study subjects and automated calculations are compared. The evaluation of the study has shown that objects with high connectivity and low gradient are classified as of high visual quality by experimentees. Traditionally used pixel-wise error measures do not correlate well with the visual quality as perceived by the study subjects. The assumption that artefacts in the background have a worse effect than those in the foreground, could not be confirmed by study subjects of the current study.

# Inhaltsverzeichnis

<b>1. Einleitung und Motivation</b> .....	7
<b>2. Was ist Image Matting?</b> .....	9
2.1. Alpha Matte.....	9
2.2. Das Mattingproblem.....	10
2.3. Trimap.....	10
<b>3. Methoden für Image Matting</b> .....	12
3.1. Blue Screen Matting.....	13
3.2. Triangulation Matting.....	13
3.3. Bayesian Matting.....	13
3.4. Closed Form Solution.....	15
3.5. Easy Matting.....	16
3.6. Random Walk Matting.....	18
3.7. Robust Matting.....	19
<b>4. Menschliche Wahrnehmung</b> .....	21
4.1. Anatomie des menschlichen visuellen Systems.....	21
4.2. Wahrnehmung eines Bildes.....	23
4.3. Wahrnehmungspsychologie.....	24
4.3.1. Gestaltpsychologie.....	24
4.3.2. Objektwahrnehmung.....	26
<b>5. Bildqualitätsbeurteilung</b> .....	28
5.1. Objektive Qualitätsmetriken.....	29
5.2. HVS basierte Systeme.....	30
5.3. Gradientmetriken.....	31
5.4. Verbundenheitsmetriken.....	32
5.5. Evaluierung von Segmentierungen.....	32
5.5.1. Evaluierung mittels Klassifikators.....	33
5.5.2. Evaluierung mittels Normalized Probabilistic Rand (NPR) Index.....	34
5.6. Evaluierung bei dieser Studie.....	35
<b>6. Benutzerstudie</b> .....	36
6.1. Erstellung der Bilder.....	37
6.2. Vorstudie.....	38

6.2.1.	Einteilung der Bilder in Kategorien .....	38
6.2.2.	Auswahl der Bildausschnitte für die Vorstudie.....	39
6.2.3.	Planung.....	41
6.2.4.	Durchführung .....	42
6.2.5.	Ergebnisse der Vorstudie .....	43
6.3.	Automatisierte Berechnungen .....	45
6.4.	Studie.....	46
6.4.1.	Kategorien .....	47
6.4.2.	Auswahl der Bildausschnitte.....	48
6.4.3.	Planung.....	51
6.4.4.	Durchführung .....	53
<b>7.</b>	<b>Auswertung der Ergebnisse.....</b>	<b>54</b>
7.1.	Analyse der Einzelergebnisse.....	55
7.1.1.	Kategorie: Verbundenheit .....	55
7.1.2.	Kategorie: Gradient .....	62
7.1.3.	Kategorie: Vordergrund-Hintergrund-Problem.....	70
7.2.	Schlussfolgerungen der Studie .....	77
<b>8.</b>	<b>Zusammenfassung.....</b>	<b>79</b>
<b>9.</b>	<b>Literaturverzeichnis.....</b>	<b>81</b>
<b>10.</b>	<b>Abbildungsverzeichnis.....</b>	<b>84</b>

# 1. Einleitung und Motivation

Image Matting stellt eine Schlüsseltechnik für die Bildbearbeitung sowie Filmproduktionen dar. Ziel ist es, ein Vordergrundobjekt aus einem beliebigen Bild herauszulösen und auf einem neuen Hintergrund wieder einzusetzen. Zur Lösung dieser Aufgabe wurden bereits unterschiedlichste Ansätze sowie Strategien entwickelt, deren Resultate ebenfalls von unterschiedlicher visueller Qualität sind.

Bei heutigen Multimedia-Technologien steht die Bildqualität im Vordergrund. Deshalb wurden und werden viele Evaluierungsstrategien entwickelt, um die visuelle Qualität automatisiert zu messen und laufend zu verbessern. Bei der Bewertung der Bildqualität ist jedoch der Mensch als Endnutzer auch der wesentlichste Entscheidungsträger. Deshalb ist es von essenzieller Bedeutung, dass die Evaluierungsverfahren mit der Bewertung des Menschen übereinstimmen.

Um das Problem zu verdeutlichen, soll dies anhand eines Beispiels näher erklärt werden. Die zugehörigen Darstellungen werden in Abbildung 1 gezeigt. Dabei sind ein Originalbild (linke Seite) sowie zwei Segmentierungsergebnisse des Originalbildes zu sehen. Die Resultate wurden erzeugt, indem das Vordergrundobjekt vom originalen Hintergrund herausgelöst und auf einem neuen Hintergrund wieder eingesetzt wurde. Beide Resultate weisen jedoch deutliche Fehler auf.



Abbildung 1: Bildausschnitt mit pixelbasierter Qualitätsmessung (links: Originalbild)

Zusätzlich wurde eine pixelbasierte Qualitätsberechnung durchgeführt. Bei den Kalkulationen ergeben sich ein Fehlerwert von 1215 beim ersten Bild und ein Fehlerwert von 806 beim zweiten Bild. Die Höhe des Fehlerwertes bestimmt die Qualität, wodurch die Wertung des ersten Bildes schlechter ausfällt. Der Mensch als Entscheidungsträger hingegen würde

vermutlich das erste Bild besser bewerten. Deshalb müssen Kriterien gefunden werden, die auf die menschliche Wahrnehmung eingehen, um tatsächlich die visuelle Qualität eines Bildes feststellen zu können.

Deshalb soll im Zuge dieser Arbeit die visuelle Qualität von Image Matting Resultaten mithilfe einer Benutzerstudie sowie bestehender Fehlermaße - welche bisher zum Vergleich von Matting Algorithmen verwendet wurden - gemessen werden. Die Problematik der Evaluierung entsteht durch die menschliche Wahrnehmung, daher soll anfänglich auf das menschliche visuelle System näher eingegangen werden. Denn nur durch dieses Verständnis können die visuellen Elemente eines Bildes entsprechend bewertet und verarbeitet werden. Anschließend werden eine Vielzahl von objektiven Evaluierungsmöglichkeiten vorgestellt. Der Grund dafür ist, dass am Ende dieser Arbeit herausgefunden werden soll, inwieweit Artefakte bestimmter Fehlerkategorien, wie beispielsweise die räumliche Verbundenheit, für die Qualität von Matting Resultaten wichtig sind.

Dazu werden Bildausschnitte – Resultate von Matting Verfahren - von unterschiedlicher visueller Qualität sowohl vom Menschen als auch von automatisierten Evaluierungsberechnungen bewertet. Dafür werden zunächst Fehlerklassen kategorisiert und entsprechend dieser Bildausschnitte ausgewählt. Die Bildausschnitte weisen dabei nur Fehler einer bestimmten Kategorie auf. Bei der Auswertung werden die subjektiven und objektiven Ergebnisse verglichen. Dabei soll einerseits festgestellt werden, ob bestimmte Kriterien tatsächlich vom menschlichen visuellen System bevorzugt wahrgenommen werden und andererseits wird überprüft werden, ob die objektiven Evaluierungsmaßnahmen mit der subjektiven Wertung übereinstimmen.

## 2. Was ist Image Matting?

Formally, image matting methods take as input an image  $I$ , which is assumed to be a composite of a foreground image  $F$  and a background image  $B$  [LLW06]. Aufgrund dieser Erklärung beschäftigt sich Image Matting mit der Vordergrundschtzung von Bildern. Dabei wird ein Vordergrundobjekt aus einem beliebigen Bild wie in Abbildung 2a herausgelöst und auf einen neuen Hintergrund wieder eingesetzt (Abbildung 2c). Vorder- und Hintergrund werden durch den Opazitätswert  $\alpha$  gemischt und somit durch eine Alpha Matte definiert (Abbildung 2b). [WC07]

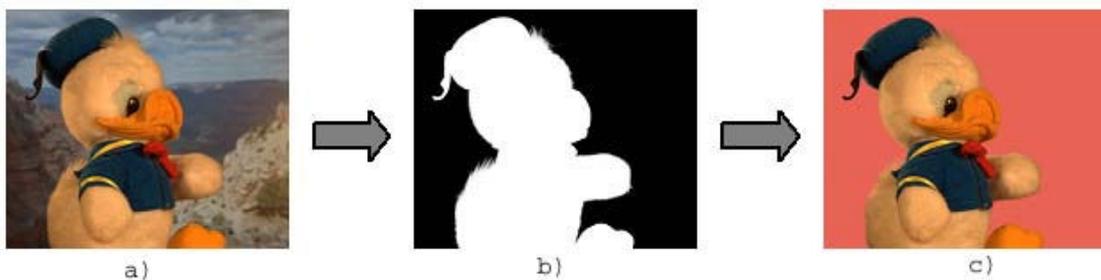


Abbildung 2: Image Matting: a) Original, b) Alpha Matte und c) Photomontage auf neuem Hintergrund

### 2.1. Alpha Matte

A matte originally meant a separate strip of monochrome film that is transparent at places, on a corresponding strip of color film, that one wishes to preserve and opaque elsewhere [SB96]. Der Grund für die Entstehung einer Matte war der Analogfilm und bei der Projektion eines Films sollte Licht durch den farbigen Filmstreifen gelangen. Uninteressante Stellen wurden stattdessen blockiert und nicht mit dem Licht bestrahlt. Die Alpha Matte hingegen entspricht dem Gegenteil der Matte des Analogfilms. Diese ist undurchsichtig an Stellen, die erhalten bleiben sollen und ansonsten transparent. Die transparenten Stellen einer Alpha Matte sind nicht von Relevanz und betreffen nur den Hintergrund, der beim reproduzierten Bild nicht übernommen werden soll. [SB96]

Mathematically, the observed image  $I_z$  ( $z = (x,y)$ ) is modelled as a convex combination of foreground image  $F_z$  and background image  $B_z$  by using the alpha matte  $\alpha_z$ :

$$I_z = \alpha_z F_z + (1 - \alpha_z) B_z \quad (1)$$

where  $\alpha_z$  can be any value in  $[0,1]$  [WC07]. Das resultierende Bild  $I_z$  wird meist in der Literatur als Composite bezeichnet, weshalb dieses mit C beschrieben wird. Alpha nimmt die Werte 0 oder 1 an, wodurch ein Pixel z als Vorder- bzw. als Hintergrund gekennzeichnet werden kann. Ein Alpha-Bild wird mit den Farben Weiß für Vordergrund und mit Schwarz für Hintergrund visualisiert. Jedoch weisen die meisten Bilder Regionen auf, die sich zwischen dem Vorder- und Hintergrund befinden, wodurch das Pixel z einen gemischten Wert annimmt. [WC07]

## 2.2. Das Mattingproblem

Ein beliebiges Eingabebild verfügt über drei unbekannte Werte: F für Vordergrund, B für Hintergrund und  $\alpha$ . Ziel ist es nun, diese aufgrund des Eingabebildes I zu bestimmen. Bei Farbbildern sind F und B dreidimensionale Vektoren, welche die Farbwerte (bzw. RGB-Werte) beinhalten. Somit ergibt sich ein Gleichungssystem mit sieben Unbekannten in nur drei Gleichungen. Sieben unbekannte Werte machen Image Matting zu einem unterbestimmten Problem. Deshalb arbeiten viele Ansätze mit Benutzerinteraktion, das bedeutet, dass die Benutzer dem Algorithmus Informationen über Vorder- und Hintergrund liefern. Durch diese zusätzlichen Informationen kann das Problem eingegrenzt werden und eine gute Schätzung der Vordergrundregion kann erwartet werden. [WC07]

## 2.3. Trimap

Eine häufig verwendete Form der Benutzerinteraktion ist die Trimap. Eine Trimap teilt ein Bild in drei Regionen:

- Vordergrund
- Hintergrund und
- unbekannte Region.

Diese Regionen werden vorab vom Benutzer angegeben und als weiterer Input zusätzlich zum Originalbild eingesetzt. *The matting problem is thus reduced to estimating F, B and  $\alpha$  for pixels in the unknown region based on known foreground and background pixels* [WC07]. Idealerweise sollte die unbekannte Region nur Pixel enthalten, deren Werte zwischen 0 und 1 liegen. In anderen Worten, diese Region sollte rund um die Vordergrundobjektsgrenze laufen und dabei so dünn als möglich sein. Der Grund dafür ist offensichtlich, denn wenn diese

Region möglichst klein ist, kann der Algorithmus durch die vorbestimmten Regionen auch nur wenige Pixel mit der falschen Region kennzeichnen. [WC07]

Da die Trimap manuell vom Benutzer spezifiziert wird, werden üblicherweise zwei Pinsel für Vorder- bzw. Hintergrund eingesetzt und der Benutzer zeichnet interaktiv die unterschiedlichen Regionen ein. Eine zweite Variante wurde von Juan und Keriven [JK05] entwickelt, bei der eine exakt spezifizierte Trimap durch grobe Angaben vom Benutzer berechnet wird. Eine dritte Möglichkeit wäre, dass der Benutzer entlang der Kante des Vordergrundobjektes eine Linie zieht und dementsprechend wird eine Trimap angepasst, wobei die unbekannte Region in der Größe variiert werden kann [WAC07]. Die Abbildung 3 zeigt zwei Beispiele der Trimapvariationen. Eine weitere Möglichkeit wäre der Einsatz von Segmentierungsansätzen wie beispielsweise dem Blue Screen Matting [SB96], bei dem der Benutzerinput nicht erforderlich ist. [WC07]



Abbildung 3: Trimap links: Exakt spezifizierte Trimap; Trimap rechts: Grob spezifizierte Trimap

### 3. Methoden für Image Matting

*Matting is identical to an optimization problem: given the mixed color  $C$  of each pixel, it aims to reconstruct optimal values of the foreground color  $F$ , the background color  $B$ , and the alpha value  $\alpha$  [GCLDP06].* Um dieses Optimierungsproblem zu lösen, wurde bereits eine Vielzahl an unterschiedlichen Ansätzen entwickelt. Deshalb sollen an dieser Stelle sechs ausgewählte Matting-Verfahren näher erläutert werden. Für weiterführende Informationen wird der interessierte Leser auf Wang und Cohen mit "*Image and Video Matting: A Survey [WC07]*" verwiesen.

Eine in der Filmproduktion verwendete Variante des Image Matting ist das Blue Screen Matting [SB96], dass die Vordergrundschatzung auf einem Bild durchgeführt, dessen Hintergrund durch seine Farbe bekannt ist. Eine Erweiterung wäre das Differenzmatting [CCSS01], bei dem ein beliebiges Objekt mit einem nicht konstanten, aber bekannten Hintergrund abgelichtet wird. Aufgrund des Wissens über den Hintergrund kann die Differenz zwischen dem bekannten Hintergrund und dem Vordergrund leicht errechnet werden. Ebenso kann die Statistik für bekannte Regionen eingesetzt werden, um so die Grenzen zwischen Vorder- und Hintergrund zu schätzen. *This scheme works well when the contrast between the foreground and background is adequately large to be captured. However, for images with complicated color distribution or low contrast, it will probably lead to erroneous solutions due to the mis-classification of color samples [GCLDP06].*

Statistisch gesehen haben Pixel, deren Nachbarschaftspixel die gleiche Farbe aufweisen, auch den gleichen Alpha-Wert. Diese lokale Korrelation [WC07, GCLDP06] stellt sicher, dass diese Pixel zusammen gehören und in sich abgeschlossen sind. Dieser Ansatz beschreibt somit die Abgeschlossenheit der Farbwerte. Andere Ansätze hingegen beziehen sich auf den Übergang zwischen Vorder- und Hintergrund, der möglichst glatt sein sollte und damit auf eine abgeschlossene Kante hinweist. [GCLDP06]

Welcher Ansatz auch gewählt wird, vorweg müssen Fragen bezüglich der Nachbarschaft, der Farbmodelle und die Vorgehensweise der Vorder- und Hintergrundschatzung beantwortet werden. Die Nachbarschaft muss genau definiert werden, um sagen zu können, welche Pixel zur Nachbarschaft zählen und welche nicht. Die Vorgehensweise bzw. der Ablauf der

Vordergrundschtzung kann unterschiedlichste Variationen annehmen und die existierenden Image Matting Ansätze zeigen bereits diese Vielfalt. [WC07]

### 3.1. Blue Screen Matting

Eine Lösung des Mattingproblems wurde von Smith und Blinn mit dem Blue Screen Matting [SB96] vorgeschlagen. Da dieses Problem eine schwierige Aufgabe darstellt, wurde ein Verfahren entwickelt, das mit einer besonderen Form der Separation arbeitet und dadurch sehr gute Resultate hervorbringt. Damit ist gemeint, dass ein beliebiger Vorder- und ein konstanter Hintergrund eingesetzt werden. Dieser konstante Hintergrund hat meist eine bestimmte Farbe wie beispielsweise Blau, wodurch der Algorithmus auch seinen Namen erhalten hat. Das bedeutet, dass bei der Gleichung (1)  $B_z$  immer konstant ist und  $F_z$  sowie  $\alpha$  unbekannt.

Wie bereits erwähnt, weist bei diesem Verfahren der Hintergrund eine vordefinierte Farbe auf. Durch diese Vorgabe ist die Abgrenzung des Vordergrundobjektes leichter möglich. Im Gegenzug darf der Vordergrund diese Farbe nicht beinhalten. Wenn der Vordergrund die gleiche Farbe aufweisen würde wie der Hintergrund, würde an dieser Stelle der Hintergrund durchscheinen. [SB96]

### 3.2. Triangulation Matting

Eine Erweiterung des Verfahrens aus Kapitel 3.1. führt Aufnahmen des Vordergrundobjekts gegen mindestens zwei verschiedene Hintergrundfarben durch. In diesem Fall erhöht sich die Anzahl der Gleichungen mit bekanntem  $B$ , wodurch ein überbestimmtes Problem entsteht. Durch Lösen des Gleichungssystems kann so der Vordergrund leicht ermittelt und der Hintergrund eliminiert werden. Anders ausgedrückt, erhalten die Regionen der Bilder, die sich unterscheiden - somit den Hintergrund betreffen – den Alpha-Wert null. [SB96]

### 3.3. Bayesian Matting

Dieses Verfahren wurde von Y. Chuang, B. Curless, D. Salesin und R. Szeliski [CCSS01] entwickelt und basiert auf der Methode von Ruzon und Tomasi [RT00]. Dieser Algorithmus erstellt Alpha aufgrund von parametrischen Verteilungen der Vorder- und Hintergrundfarben entlang der Objektgrenzen. [WC07]

*The goal of our algorithm, then, is to solve for the foreground color  $F$ , background color  $B$ , and opacity  $\alpha$  given the observed color  $C$  for each pixel within the unknown region of the*

image [CCSS01]. Um die unbekannt Variablen des Mattingproblems zu berechnen, wird zuerst eine Trimap vom Benutzer spezifiziert. Anschließend werden Vorder- und Hintergrundwahrscheinlichkeitsverteilungen aufgrund der Farbverteilung in der Trimap erstellt, was von der Ruzon und Tomasi [RT00] Methode übernommen wurde. Dieser Ansatz lässt ein kontinuierliches Fenster einwärts über die Bildregionen gleiten, ersichtlich ist dieser Vorgang in Abbildung 4. Bei der Verschiebung des Fensters werden für Vorder- und Hintergrund Gauß'sche Verteilungen berechnet. Bei den Berechnungen fließt das Wissen über die bekannten Regionen des Vorder- und Hintergrundes mit ein.

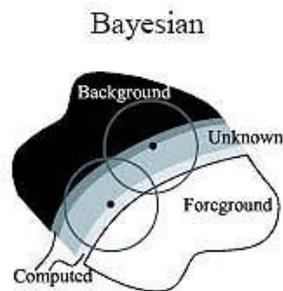


Abbildung 4: Bayesche Verfahren: Abbildung aus [CCSS01]

Die Alpha Matte wird nun in einem Bayesischen Framework mithilfe der Maximum-APosteriori-Methode (MAP) erstellt. Dabei soll die optimale Schätzung für  $\alpha$  gefunden werden.

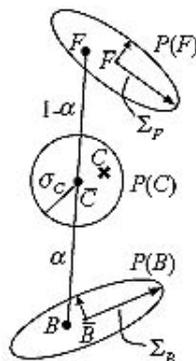


Abbildung 5: Bayesche Verfahren: Verteilungen der optimalen F, B und  $\alpha$  Parameter: Abbildung aus [CCSS01]

Um die Gauß'schen Verteilungen zu berechnen, werden zu Beginn die Farben der Vordergrundpixel in der räumlichen Nachbarschaft zu Gruppen zusammengefasst. Für jede dieser Gruppen werden Gauß'sche Verteilungen erstellt, indem der Mittelwert sowie die Kovarianz berechnet werden. Diese Vorgehensweise ist auch für den Hintergrund gültig. Eine Vorder- und eine Hintergrundgruppe bilden jeweils ein Paar. Weiters werden Maximum-

Likelihood-Kriterien eingesetzt um optimale Alpha-Werte für jedes Vorder- und Hintergrundpaar sowie deren Pixel zu schätzen. Das Composite C ergibt sich durch den Differenzvektor von F und B, die eine Linie im Farbraum bilden. Eine Verdeutlichung dieses Absatzes ist in Abbildung 5 zu finden. [CCSS01]

### 3.4. Closed Form Solution

Das Closed Form Solution-Verfahren wurde von A. Levin, D. Lischinski und Y. Weiss [LLW06] entwickelt. *We derive a cost function from local smoothness assumptions on foreground and background colors, and show that in the resulting expression it is possible to analytically eliminate the foreground and background colors to obtain a quadratic cost function in alpha. This allows us to find the globally optimal alpha matte by solving a sparse linear system of equations [LLW06].* Genauer gesagt, basiert dieser Ansatz auf der Annahme, dass die Vorder- und Hintergrundfarben in einem lokalen Fenster auf einer Linie im Farbraum liegen. Auch dieser Ansatz arbeitet mit einer grob vordefinierten Trimap, wobei wenige Linien für Vorder- und Hintergrund vom Benutzer hinreichend sind.

Für die Berechnungen der unbekanntenen Regionen wird ein kleines Fenster rund um jedes Pixel benötigt. Der Wert eines Pixels für Vorder-  $F_i$  und Hintergrund  $B_i$  innerhalb eines kleinen Fensters liegt auf einer Linie im RGB-Farbraum, wodurch von einer Mischung aus zwei Farben oder dem Farblinienmodell gesprochen werden kann. Der Alpha-Wert innerhalb eines kleinen Fensters kann nun wie folgt ausgedrückt werden:

$$\alpha_i \approx \sum_c a^c I_i^c + b, \forall i \in w \quad (2)$$

$c$  ist dabei der Farbkanal.  $a^c$  und  $b$  sind konstante Werte innerhalb eines bestimmten Fensters. Die Variable  $w$  definiert die Pixel in einem Bildfenster. Das Ziel ist nun durch Minimierung der folgenden Kostenfunktion  $\alpha$  zu finden:

$$J(\alpha, a, b) = \sum_j \left( \sum_{i \in w_j} \left( \alpha_i - \sum_c a_j^c I_i^c - b_j \right)^2 + \varepsilon \sum_c a_j^{c^2} \right) \quad (3)$$

$a^c$  und  $b$  können von der Kostenfunktion entfernt werden und so ergibt sich:

$$J(\alpha) = \alpha^T L \alpha \quad (4)$$

L ist eine  $N \times N$ -Matrix und wird als matting Laplacian bezeichnet und wie folgt berechnet:

$$\sum_k \sum_{(i,j) \in w_k} \left( \delta_{ij} - \frac{1}{|w_k|} (1 + (I_{i-\mu_k}) \left( \sum_k + \frac{\varepsilon}{|w_k|} I_3 \right)^{-1} (I_j - \mu_k)) \right) \quad (5)$$

$\Sigma_k$  ist eine  $3 \times 3$  Kovarianzmatrix,  $\mu_k$  ist ein  $3 \times 1$  Mittelwertvektor mit den Farben des jeweiligen Fensters  $w_k$ ,  $I_3$  ist eine  $3 \times 3$  Einheitsmatrix und  $\varepsilon$  ist ein konstanter Vektor. Die optimalen Alpha-Werte werden nun mit

$$\alpha = \operatorname{argmin} \alpha^T L \alpha, \quad \text{s. t. } \alpha_i = s_i, \quad \forall i \in S \quad (6)$$

berechnet.  $S$  ist dabei die Gruppe von Pixeln aus  $s_i$ , die vom Benutzer in einer Trimap gekennzeichnet wurden. [LLW06]

### Rekonstruktion von Vorder- bzw. Hintergrund

Wenn die Alpha Matte berechnet ist und ein gewünschtes Resultat geliefert hat, wird weiters der Vorder- oder eventuell auch der Hintergrund des Originalbildes rekonstruiert. Bei der Rekonstruktion von komplexen Vorder- bzw. Hintergrundmustern treten häufig unerwünschte Artefakte in den Vorder- bzw. Hintergrundfarben auf. Der Vorder- und Hintergrund erhält in einem Optimierungsverfahren unter lokalen Glattheitskriterien seine Farben. Die Abbildung 6 zeigt eine erstellte Alpha Matte sowie einige Rekonstruktionen. [LLW06]

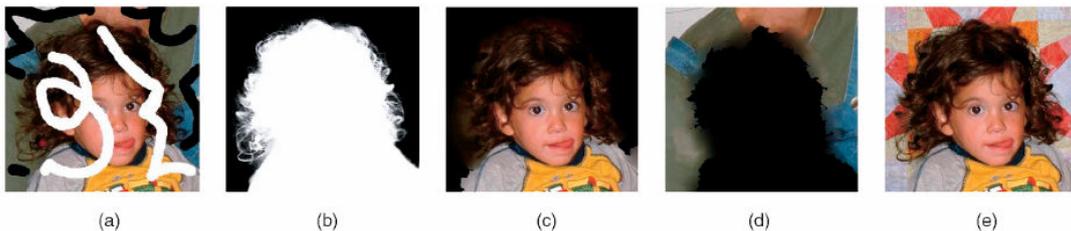


Abbildung 6: Vorder- und Hintergrundrekonstruktion: a) Eingabebild mit Vorder- und Hintergrundmarkierung, b) Alpha Matte, c) Vordergrundrekonstruktion, d) Hintergrundrekonstruktion und e) Vordergrund mit einem neuen Hintergrund: Abbildung aus [LLW06]

### 3.5. Easy Matting

Der Ansatz des Easy Matting Algorithmus wurde von Y. Guan, W. Chen, X. Liang, Z. Ding und Q. Peng [GCLDP06] entwickelt. *Our approach is easy in the sense that it is fast and*

*requires only few user-specified strokes for marking the foreground and background [GCLDP06].* Bei diesem Ansatz kann auch von einem iterativen Optimierungsframework gesprochen werden.

Bevor der Algorithmus starten kann, sind einige Markierungen des Vorder- und Hintergrunds vom Benutzer in Form einer sehr ungenauen Trimap notwendig. Für die unbekannte Region wird bei jeder Iteration ein Markov Random Field (MRF) erzeugt. Das MRF wird als Graph repräsentiert, wobei jedes Pixel als ein Knoten bezeichnet wird und Verbindungen zu den Nachbarn hat. Nach jeder Iteration wird der Graph neu berechnet. Anschließend wird folgende Kostenfunktion erstellt, die als MRF dargestellt wird:

$$E = \sum_{p \in \Omega} (E_d(p) + \lambda E_s(p)) \quad (7)$$

$E_d$  ist ein Datenausdruck und schätzt  $\alpha$  aufgrund der Differenz zwischen dem gegenwärtigen Farbwert  $C_p$  und den geschätzten  $F_p$ ,  $B_p$  und  $\alpha_p$ . Um  $p$  zu berechnen, wird die lokale Nachbarschaft hinzugezogen. Dafür werden die räumlich nächsten Pixel in der Vorder- und Hintergrundregion ausgewählt. Die Distanz zwischen den Farben wird mittels der Euklidischen Distanz im RGB-Farbraum berechnet.

$E_s$  beschäftigt sich mit der Glattheit. Dabei wird der Grad der Differenzen der Nachbarpixel näher betrachtet. Eine quadratische Funktion und der Einsatz des Farbgradienten sollen die Ebenheit garantieren. Diese zwei Ausdrücke erzwingen zusammen eine Glattheit in der Alpha Matte, auch wenn der Kontrast niedrig ist.

Der Parameter  $\lambda$  gewichtet die beiden Terme der Kostenfunktion. Das Gewicht wird nach jeder Iteration neu bestimmt und wechselt schrittweise von einem hohen bis hin zu einem niedrigen Wert. Diese dynamische Gewichtung ist für eine reibungslose Ausbreitung der Regionen während der Iterationen notwendig. Der Grund dafür ist, dass viele Benutzer ihre Markierungen innerhalb der Vorder- bzw. Hintergrundregion ansetzen und nicht an den Grenzen. Somit wird am Beginn dem Glattheitsausdruck ein größeres Gewicht zugeschrieben, um möglichst schnell eine Lösung zu entwickeln und die Objektgrenzen zu finden. Wenn jedoch die Grenzregionen geschlossen werden, wird der Datenausdruck dominanter. [GCLDP06]

### 3.6. Random Walk Matting

Diese Methode wurde von L. Grady, T. Schiwietz, S. Aharon und R. Westermann [GSAW05] entwickelt. Diese Technik hat laut den Entwicklern folgende Vorteile: *First, by leveraging a recent technique from manifold learning theory, we effectively use RGB values to set boundaries for the random walker, even in fuzzy or low-contrast images. Second, the algorithm is straightforward to implement, requires specification of only a single free parameter (set the same for all images), and performs the segmentation and alpha-matting in a single step. Third, the user may locally fine tune the results by interactively manipulating the foreground/background maps [GSAW05].*

Die Idee des Random Walk Algorithmus ist, dass die finalen Alpha-Werte auf der Basis der Ähnlichkeit berechnet werden. Dafür unterteilt der Benutzer die Bilder in drei Regionen (Vorder-, Hintergrund und unbekannte Region). Die Pixel der unbekannt Region erhalten ihren Alpha-Wert als Wahrscheinlichkeit. Der Random Walker startet nun von einer Pixelposition in der unbekannt Region. Die Wahrscheinlichkeit ergibt sich durch das Auftreffen des Random Walkers auf die Vorder- bzw. Hintergrundregion und welche Region davon zuerst erreicht worden ist.

Diese Wahrscheinlichkeiten können durch das Lösen eines linearen Gleichungssystems berechnet werden. Um das Gleichungssystem zu erstellen, wird eine diskrete Struktur wie beispielsweise ein Graph benötigt. Die Pixel sind dabei die Knoten und die Kanten befinden sich jeweils dazwischen. Die vier direkten Nachbarn nehmen dabei die Hauptrichtungen an. Anschließend wird der Graph bzw. seine Kanten gewichtet:

$$w_{ij} = \exp\left(-\frac{\|z_i - z_j\|^2}{\sigma^2}\right) \quad (8)$$

$z_i$  ist ein Vektor, der die RGB-Farben eines Pixels  $i$  repräsentiert.  $\sigma$  ist ein freier Parameter und kann für alle Resultate beispielsweise auf  $\sigma = 1/30$  gesetzt werden. Um die Objektgrenzen so gut als möglich zu erkennen, wird die Local Preserving Projections (LPP) Technik [GSAW05] von He und Niyogi eingesetzt. Die Vorteile dieser Technik sind Linearität, Generalisierung außerhalb der Trainingspunkte und Robustheit bei Ausreißern. Für die LPP Technik ist die Lösung des Eigenvektorproblems erforderlich:

$$ZLZ^T x = \lambda ZDZ^T x \quad (9)$$

Z ist eine 3 x N Matrix mit  $z_i$  in jeder Spalte. D ist eine Diagonalmatrix und wird mittels  $D_{ii} = d_i$  definiert.  $d_i$  gibt den Grad eines Knotens an und ist durch  $d_i = \sum w_{ij}$  definiert. L ist die Laplacian Matrix und ist gegeben durch:

$$L_{w_{ij}} = \begin{cases} d_{vi} & \text{if } i = j, \\ -w_{ij} & \text{if } v_i \text{ and } v_j \text{ are adjacent nodes,} \\ 0 & \text{otherwise.} \end{cases} \quad (10)$$

Die Lösung des Eigenvektorproblems wird mit Q angegeben. Die Gewichte oder die Verwandtschaft zwischen den Pixeln werden letztendlich mit

$$w_{ij} = \exp\left(\frac{(z_i - z_j)^T Q^T Q (z_i - z_j)}{\sigma^2}\right) \quad (11)$$

berechnet. Die neuen Alpha-Werte basieren auf diesen Verwandtschaftsberechnungen und bilden ein Set von Wahrscheinlichkeiten. Die Wahrscheinlichkeiten werden mittels eines Systems von linearen Gleichungen gelöst. Somit ist jedes neu berechnete  $\alpha$  ein gewichteter Durchschnitt von den Alpha-Werten der Nachbarpixel. [GSAW05]

### 3.7. Robust Matting

Dieses Verfahren wurde von J. Wang und M. F. Cohen [WC07\_R] entwickelt. *In our approach we also sample foreground and background colors for unknown pixels, but more importantly, analyze the confidence of these samples [WC07\_R].*

Das Verfahren basiert auf einer optimierten, robusten Farbabtastungsmethode, welche nicht nur die Vorder- und Hintergrundfarbwerte für die unbekannt Pixel schätzt, sondern insbesondere das Vertrauen der Schätzung bewertet. Das bedeutet, dass zur Berechnung der Alpha Matte nur verlässliche Pixel herangezogen werden.

Als Eingabe dienen ein Eingabebild sowie eine vordefinierte Trimap. Zu Beginn sammelt der Algorithmus eine Vielzahl an Vorder- und Hintergrundsamples entlang der Vorder- und Hintergrundregion der benutzerspezifizierten Trimap. Diese Samples dienen der Schätzung der wahren Vorder- und Hintergrundfarben an einer bestimmten Position. Wichtig dabei ist,

dass gute Samples aus der Vielzahl ausgewählt werden. Nun wird das Vertrauen für jedes Paar von Vorder- und Hintergrundsamples berechnet. Paare mit hoher Verlässlichkeit sollten drei Kriterien erfüllen:

1. Gute Samplepaare beschreiben die Farbe des Pixels in der unbekanntem Region als lineare Kombination.
2. Vorder- und Hintergrundfarben sollten im Farbraum weit auseinander liegen.
3. Die Farbe sollte ähnlich der des Pixels in der unbekanntem Region sein.

Zur Evaluierung wird das Distanzverhältnis zwischen der gegenwärtigen Pixelfarbe sowie der Farbe, die das Pixel aufweisen sollte, berechnet. Anschließend werden die Vertrauenswerte hinzugezogen. Der durchschnittliche Alpha-Wert sowie das Vertrauen der Samplepaare werden als Werte übernommen. All diese Angaben werden in einem gewichteten Graphen eingetragen. Die Gewichte der Kanten ergeben sich durch die Angabe von Vorder- und Hintergrund und den dadurch ergebenden relativen Wahrscheinlichkeiten. Mittels dieser Schätzung wird Alpha für jedes Pixel berechnet und im Framework von [LLW06] optimiert. [WC07\_R]

## 4. Menschliche Wahrnehmung

Die spezielle Empfindlichkeit der Sinnessysteme ist darauf abgestimmt, die für den natürlichen Lebensraum des Menschen signifikanten Informationen aufzunehmen und zu verarbeiten [G02]. Nachdem bei dieser Diplomarbeit die Bildverarbeitung im Vordergrund steht, wird in den folgenden Abschnitten auf das visuelle System des Menschen näher eingegangen.

### 4.1. Anatomie des menschlichen visuellen Systems

Das visuelle System umfasst vier Hauptkomponenten, diese werden in Abbildung 7 schematisch dargestellt:

- das Auge
- das Corpus geniculatum laterale (CGL; laterale geniculate nucleus LGN)
- den primären visuellen Cortex
- den extrastriären Cortex

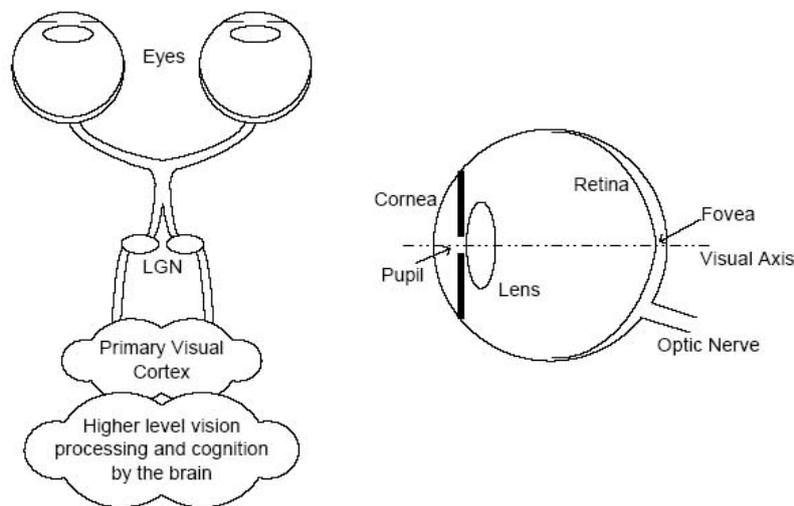


Abbildung 7: Schematisches Diagramm des menschlichen visuellen Systems: Abbildung aus [WSB03]

Das optische System des Auges besteht aus der Hornhaut, der Pupille, der Linse und der Flüssigkeit, die das Auge füllt. Die schematische Darstellung des Auges ist in Abbildung 7 ersichtlich. Die Hornhaut ist die transparente Oberfläche des Auges. Die Pupille befindet sich zwischen der Hornhaut und der Linse. Je nach Lichtmenge bzw. Beleuchtung wird die Pupille

größer oder kleiner. Diese Helligkeitsanpassung bezeichnet man als Adaption. Die entscheidende Rolle bei der Adaption spielt jedoch nicht die Pupille, sondern die Netzhaut, denn auf diese gelangt das Licht. Die optischen Komponenten des Auges befinden sich schließlich unter der Retina bzw. der Netzhaut im hinteren Bereich des Auges, die helligkeits- und farbempfindlich sind. Die Linse befindet sich hinter der Iris. Das Linsenband wird mittels Muskeln gespannt und dient dem Fokussieren des Auges auf ein Objekt. [R02, A06, G02]

Der Sehvorgang beginnt damit, dass Licht vom Auge aufgenommen und in ein Netzhautbild umgewandelt wird. Dafür passiert das Licht die Hornhaut, die Pupille und die Linse. Die Hornhaut sowie die Linse lenken das Bild auf die Netzhaut ab. In der Netzhaut befindet sich eine Schicht von Neuronen. Die Neuronen teilen sich in Stäbchen und Zäpfchen, die ihre Namen aufgrund ihrer Form erhalten haben. Diese Neuronen sind Photorezeptoren, deren Hauptaufgabe es ist, Licht in elektrische Signale umzuwandeln. Diese Aufgabe wird mittels lichtempfindlicher chemischer Substanzen gelöst. Ein Teil der Zäpfchen befindet sich in der Fovea, welcher ein kleiner Fleck in der Netzhaut ist. Rund um die Fovea befinden sich weitere Zäpfchen und auch Stäbchen. [G02, A06]

Nun erfolgt der Übertrag des Signals vom Auge ins Gehirn durch das Neuronennetzwerk. Dabei werden zuerst die Ganglienzellen im Auge erreicht. Alle Ganglienzellenfasern zusammen bilden den Sehnerv bzw. den Nervus opticus, über den die Signale das Auge verlassen. Diese Stelle wird auch als blinder Fleck bezeichnet. Wie der Name dieses Flecks schon besagt, ist der Mensch an dieser Stelle blind. Der blinde Fleck müsste uns somit hindern einen Gegenstand wahrzunehmen. Allerdings fällt meistens das Bild nicht in beiden Augen auf den blinden Fleck und ausgefallene Stellen werden mit Bildteilen von Nachbarschaftsstellen aufgefüllt. [G02, A06]

Die Stäbchen und Zäpfchen wandeln die Signale in Nervenimpulse um und ermöglichen eine Kommunikation der Sinnessysteme. Die Impulse erreichen den CGL, dieser stellt eine wichtige Verarbeitungsstation für die Regelung des Informationsflusses zwischen der Netzhaut und dem Cortex dar. [G02]

Nach mehreren Verarbeitungsstufen gelangen die Impulse über die Sehbahn zum primären visuellen und extrastriären Cortex. Wenn die Signale im Gehirn angekommen sind, werden

diese weiter verarbeitet. Die neu aufgenommene Information wird mit gespeicherten Daten verarbeitet, verglichen und verbunden. [G02, A06]

*Das Sehen ist ein komplizierter psychologischer Vorgang, und die optischen Eindrücke werden mit den von den anderen Sinnesorganen vermittelten Signalen in Einklang gebracht [A06].* Nur das Zusammenspiel eines visuellen Eindrucks mit den Meldungen anderer Sinnesorgane sowie den Erfahrungsberichten aus früherer Zeit ergeben einen endgültigen optischen Eindruck. [A06]

## 4.2. Wahrnehmung eines Bildes

Der Großteil der Menschen nimmt die optischen Eindrücke der Außenwelt unterschiedlich und zum Teil falsch wahr. Deshalb spricht man von einer Objektivität des Sehens und zeigt damit, dass Denken das Sehen vervollständigt. Die unterschiedlichen Ausprägungen des Sehens weisen auf eine Subjektivität sowie Relativität der optischen Wahrnehmung hin. Trotz dieser Probleme erhalten wir Menschen verhältnismäßig genaue Angaben über die Geschehnisse der Welt. [A06]

Licht gibt uns Menschen die Möglichkeit sehen zu können, doch das für den Menschen sichtbare Licht befindet sich lediglich innerhalb eines kleinen Bereichs der elektromagnetischen Wellen. Dieser kleine Abschnitt reicht von einer Wellenlänge von 400 bis 700 Nanometer. [G02]

Wenn das Licht eines reflektierten Gegenstandes auf das Auge trifft, wird dieser von der Augenlinse auf die Netzhaut vertikal gespiegelt übertragen, das bedeutet, dass das Bild auf dem Kopf steht. Trotzdem wird der Gegenstand in aufrechter Haltung wahrgenommen, weil bei der Verarbeitung im Gehirn die Drehung des Bildes korrigiert wird. Zusätzlich ist zu erwähnen, dass das Auge kugelförmig ist und auch diese Verzerrung wird bei der Wahrnehmung des Bildes nicht bemerkt. Beim Sehen wird nicht nur ein Gegenstand wahrgenommen, sondern immer auch der Hintergrund. Die lichtempfindlichen Elemente der Netzhaut werden separat für ein Objekt und für den Hintergrund gereizt. Somit kann das Auge das Objekt am Kontrast zum Hintergrund erkennen. Dies wird als Kontrastempfindlichkeit des Auges bezeichnet. Auch die Unterscheidung mittels Helligkeit ist möglich, weshalb hier von Helligkeitskontrast gesprochen wird. Zu der Helligkeitsunterscheidung ist zu bemerken,

dass eine Figur auf dunklem Hintergrund heller und umgekehrt auf hellem Hintergrund etwas dunkler erscheint. [A06]

Die Rezeptoren der Netzhaut ermöglichen das Farbsehen. Durch Variation und Sättigung gibt es eine riesige Vielfalt an Farben, die für uns Menschen unterscheidbar ist. *Farben erleichtern die Gliederung der Szenen, unterstützen die Unterscheidung und das Erkennen von Objekten. Ferner haben sie wichtige Signalfunktionen [G02].* Insbesondere liefert Farbe Kontrast, sodass Gegenstände vom Hintergrund leichter unterscheidbar sind. Zusätzlich werden Hinweise auf Zusammengehörigkeit geliefert, auch wenn Teile eines Objektes verdeckt sind. Somit helfen Farben bei der Täuschung durch Schatten, Spiegelungen und ungleichmäßiger Beleuchtung. [A06]

### 4.3. Wahrnehmungspsychologie

Die Wahrnehmungspsychologie ist ein umfassendes Gebiet, für die Bildbearbeitung und speziell für Image Matting sind insbesondere die Gestaltpsychologie sowie die Objektwahrnehmung von Bedeutung. Diese zwei Punkte werden infolgedessen näher erläutert werden.

#### 4.3.1. Gestaltpsychologie

Der Hauptvertreter der Gestaltpsychologie ist Max Wertheimer (1880 – 1943). Seine Experimente und theoretischen Beiträge führten zur Gründung der Gestaltpsychologie. Die Gestaltpsychologie beschäftigt sich mit der Wahrnehmungsorganisation. Eine Grundregel ist: *"Das Ganze ist mehr als die Summe seiner Teile [G02]."* In der Gestaltpsychologie werden mehrere Faktoren beschrieben, die als Regeln verstanden werden. Diese beschreiben die Wahrnehmung, die bei bestimmten Reizbedingungen hervorgerufen werden. Goldstein [G02] erläutert in seinem Buch "Wahrnehmungspsychologie" diese Faktoren, welche im Folgenden angeführt werden. Davor sollte jedoch erwähnt werden, dass es heute eine Neuinterpretation der Gestaltpsychologie gibt, wodurch von Heuristiken gesprochen wird und nicht von Gesetzen. [G02]

#### ***Faktor der Prägnanz, guten Gestalt oder der Einfachheit:***

Dabei soll jede Struktur oder jedes Muster als so einfach wie möglich angesehen werden. Beispielsweise bei der Abbildung eines Quadrats und eines Dreiecks, die ineinander

verwoben sind, wird trotzdem ein Quadrat und Dreieck wahrgenommen und nicht ein komplexes Gebilde. [G02]

***Faktor der Ähnlichkeit:***

Ähnliche Formen scheinen zusammen zu gehören und bilden eine Gruppe. Ein Beispiel dafür wäre, wenn auf einer Darstellung mehrere Kreise und Quadrate abgebildet sind, dann bilden sowohl die Kreise als auch die Quadrate eine Gruppe. [G02]

***Faktor der gestaltgerechten Linienfortsetzung:***

Einzelne Punkte, die als gerade oder geschwungene Linie gesehen werden können, gehören zusammen. Grundsätzlich folgen Linien dem einfachsten Weg. Anders gesagt, die Verfolgung einer Linie ist trotz mehrmaliger Verdeckungen und Unterbrechungen möglich. [G02]

***Faktor der Nähe:***

Formen, die sich nahe beieinander befinden, scheinen zusammen zu gehören. [G02]

***Faktor des gemeinsamen Schicksals:***

Objekte, die sich in die gleiche Richtung bewegen, gehören zusammen. Ein Beispiel dafür wäre die Tanzbewegung von zwei Personen in die gleiche Richtung. [G02]

***Faktor der Bedeutung oder Vertrautheit:***

Objekte bzw. Subjekte bilden Gruppen, wenn diese vertraut erscheinen bzw. etwas bedeuten. Ein Beispiel dafür wären die einzelnen Teile des Gesichts. [G02]

***Figur-Grund-Trennung:***

Die Figur-Grund-Trennung ist eine entscheidende Grundlage um Objekte zu erkennen. Die Eigenschaften der Figur-Grund-Trennung sind, dass eine Figur dinghaft ist und so leichter im Gedächtnis bleibt als der Hintergrund. Weiters steht eine Figur immer vor dem Hintergrund. Die Konturen, die sich zwischen der Figur und dem Hintergrund befinden, scheinen zur Figur zu gehören. Der Hintergrund hingegen stellt meist ein ungeformtes Material dar. Symmetrisch geformte Bereiche, konvexe Formen und kleine Flächen werden eher als Figur angesehen. [G02]

In den 1990er Jahren wurden die Faktoren in der Wahrnehmungsorganisation von Stephen Palmer und Irvin Rock mit drei weiteren Punkten ergänzt. [G02]

***Faktor der gemeinsamen Region:***

Elemente oder Formen, die sich innerhalb einer gemeinsamen Region befinden, gehören zu einer Gruppe. Das bedeutet, dass, wenn beispielsweise zwei Punkte von einem größeren Kreis umgeben sind, diese zusammen gehören. [G02]

***Faktor der Verbundenheit der Elemente:***

Bei diesem Faktor ist entscheidend, dass Elemente miteinander verbunden werden. Denn nur verbundene Objekte können eine Einheit bilden. [G02]

***Faktor der zeitlichen Synchronizität***

Einzelne Punkte, die in der gleichen Farbe und Zeit aufleuchten, werden als Gruppe betrachtet. Die Zusammengehörigkeit wird durch eine räumliche Trennung nicht gestört. [G02]

#### 4.3.2. Objektwahrnehmung

*Für die Interpretation eines Bildes muss ein breites Spektrum des Wahrnehmens, Wiedererkennens und Kategorisierens eingesetzt werden [G02].* Um ein Objekt als solches erkennen zu können, sind mehrere Arbeitsschritte des Gehirns notwendig. Zu Beginn wird ein Gegenstand erfasst und damit wahrgenommen. Um das Objekt auch wieder zu erkennen, muss man sich an das schon früher Gesehene bzw. Gehörte erinnern. Dafür wird das Objekt entsprechend kategorisiert und in eine bereits bestehende Ordnung einsortiert. [G02]

All diese Schritte sind erforderlich um ein Objekt zu erkennen, doch die Vielzahl an Mustern, Formen oder Lichteffekten erschwert diese Aufgabe. Deshalb müssen die vielen unterschiedlichen Erscheinungen zu größeren Einheiten zusammengefasst werden. Anders ausgedrückt: um eine Szene richtig zu deuten, müssen sowohl Details als auch globale Informationen verarbeitet werden. Hinzu kommt, dass bei einer wahrnehmungsbasierten Verarbeitung viele fehlerhafte optische Eindrücke hervorgerufen werden, da Objekte oder Figuren immer in Beziehung zu anderen Objekten oder einem Hintergrund stehen und dementsprechend wahrgenommen werden. [G02]

Um ein bestimmtes Objekt zu erkennen, müssen die Oberfläche und der Gegenstand als solcher erkannt werden. Dafür muss ein Bezug zu bestehenden Gedächtnisinhalten sowie Handlungsmöglichkeiten hergestellt werden. Für die Erkennung werden alle Einzelteile, wie Konturen, Ecken, Farben, Texturen und Geometrien des Objekts analysiert und zu einer Einheit zusammengefasst. Grundlegende Kriterien können von der Gestaltpsychologie herangezogen werden, wie beispielsweise der Faktor der Ähnlichkeit. Dies ist notwendig, da ein Objekt auch nach einer Transformation wie einer Drehung oder einer Verdeckung erkannt werden soll. [G02]

## 5. Bildqualitätsbeurteilung

Die Bildqualitätsbeurteilung spielt eine bedeutsame Rolle bei einer Vielzahl von Bildbearbeitungsapplikationen. Ob ein Vordergrundobjekt herausgelöst oder ob ein Bild komprimiert wird, das Resultat soll immer von möglichst hoher visueller Qualität sein.

Bei der Evaluierung der Bildqualität wird insbesondere die Bildähnlichkeit zwischen einem Originalbild und einem reproduzierten Bild analysiert. Um die visuelle Qualität eines Bildes zu messen, kann die subjektive oder die objektive Methode angewendet werden. *The HVS is the terminal of image processing systems, thus the most correct method of quantifying image quality may be subjective evaluation [CYX06]*. HVS steht für das menschliche visuelle System (human visual system). Doch der Mensch als Entscheidungsträger ist bei Qualitätsmessungen sehr zeitintensiv, teuer und die Resultate sind abhängig von Faktoren wie dem Vorwissen des Probanden und dessen Motivation. [ASS02]

Hingegen ist die objektive Messung wiederholbar, einfacher und stellt eine Automation der Bildqualitätsbeurteilung dar. *A good objective quality measure should reflect the distortion on the image well due to, for example, blurring, noise, compression, and sensor inadequacy [ASS02]*. Allerdings ist der Mensch der Endnutzer der Multimediaapplikation und deshalb sollten die Bildqualitätsmessungen auf das menschliche visuelle System eingehen, um eine Akzeptanz vom Endnutzer zu erhalten. [ASS02, LKS04, SPA07]

Zur objektiven Evaluierung gibt es eine Vielzahl an unterschiedlichen Ansätzen, einige davon werden in den nächsten Kapiteln vorgestellt werden. Diese Ansätze können das gesamte Bild mit einbeziehen oder nur auf bestimmte Teile näher eingehen. Die Idee der Evaluierung kann ebenfalls unterschiedliche Formen annehmen. Zum einen kann das menschliche visuelle System (HVS) als Grundlage dienen, bzw. Teile davon können in den Evaluierungsprozess miteingebunden werden. Als Beispiel wird hier der Structural Similarity Index [SB07] vorgestellt, da dieser entsprechend dem HVS die strukturellen Bildinformationen eines Bildes entnimmt. Zum anderen können einfache Übereinstimmungsmaße die visuelle Qualität messen, indem pixelweise die Differenz zwischen zwei Bildern berechnet wird. Weitere Methoden beziehen sich auf bestimmte Bereiche eines Bildes, wie der Gradient, der die Grenzregion eines Objektes anzeigt. Auch die Verbundenheit eines Objektes kann als

Bewertungsmaß miteinander verglichen werden. Bei der letzten vorgestellten Evaluierungsmöglichkeit wird die Segmentierung des Bildes näher betrachtet. Dabei soll untersucht werden, ob die Segmentierung eines Bildes erfolgreich war oder nicht. All diese Beurteilungsmöglichkeiten können auch für Image Matting von Relevanz sein, denn die unter Kapitel 3 vorgestellten Algorithmen produzieren eine Vielzahl an unterschiedlichen Bildfehlern.

### 5.1. Objektive Qualitätsmetriken

Die objektiven Qualitätsmetriken lassen sich in sechs Gruppen kategorisieren. Die Einteilung bezieht sich auf die spezielle Nutzung der Bildinformation:

1. Maßeinheit der Pixeldifferenz: Das Ergebnis ergibt sich durch die pixelweise Differenz zweier Bilder. Beispiele für diese Kategorie sind Standard-Distanz-Metriken wie der mittlere quadratische Fehler (MSE – Mean Square Error), mittlere absolute Fehler (MAD – Mean Absolute Distance) oder Peak Signal to Noise Ratio (PSNR).
2. Korrelationsbasierte Maßeinheit: Die Ähnlichkeit zwischen zwei Bildern kann mithilfe einer Korrelationsfunktion von Pixels gemessen werden. Eine Möglichkeit wäre, dass die Winkel durch die Korrelation berechneter Vektoren überprüft werden. Bei übereinstimmender Farbe würde auch die Richtung des Vektors identisch sein.
3. Kantenbasierte Maßeinheit: Die Kanten innerhalb einer Szene stellen den informativsten Teil eines Bildes dar. Das gilt sowohl für die menschliche Wahrnehmung als auch für die unterschiedlichsten Algorithmen. Somit werden bei dieser Maßeinheit die Kantenpositionen auf Ähnlichkeit hin überprüft.
4. Maßeinheit für die Spektraldistanz: Bei dieser Maßeinheit wird das Fourierspektrum der Bilder erfasst und auf Diskrepanz hin verglichen.
5. Kontextbasierte Maßeinheit: Diese Messung erfolgt mittels Kontextwahrscheinlichkeiten. Für die statistischen Informationen werden die Nachbarschaftspixel herangezogen. Änderungen bei den Kontextwahrscheinlichkeiten könnten Hinweise auf eine Störung im Bild sein.

6. HVS basierte Maßeinheit: Das menschlich visuelle System ist viel zu komplex um es im Gesamten verstehen bzw. übernehmen zu können, doch wenn vereinfachte Modelle übernommen werden, korrelieren die subjektiven mit den objektiven Messungen meist besser. [ASS02]

Alle objektiven Verfahren dienen dazu, automatisiert die Bildqualität zu berechnen. Viele der existierenden Ähnlichkeitsmessungen können einfach berechnet werden. Insbesondere ist es wichtig, dass die objektiven Bildqualitätsmessungen mit der subjektiven Qualitätsmeinung korrelieren, da sonst nicht von der optimalen Qualität eines Bildes gesprochen werden kann. Dies motiviert die Verwendung von HSV-basierten Methoden. [ASS02]

Eine Studie von S. Grgic, M. Grgic und M. Mrak [GGM04] verdeutlicht die Notwendigkeit HSV-basierter Systeme. Für die Bewertung werden komprimierte Bilder herangezogen und objektive sowie subjektive Wertungen erfasst. Bei der Auswertung gab es jedoch keine Übereinstimmung zwischen der subjektiven Wertung und den pixel-basierten Fehlermaßen des MSE (mittlere quadratische Fehler) oder PSNR (Peak Signal to Noise Ratio). Lediglich beim PQS (Picture Quality Scale) gab es eine zufrieden stellende Korrelation zu der subjektiven Wertung. Die PQS-Methode setzt Eigenschaften ein, die für das HVS relevant sind. [GGM04]

Dennoch sollte erwähnt werden, dass einfache Übereinstimmungsmaße für einfache Verzerrungen akzeptable Ergebnisse bringen. Um eine hohe Korrelation zwischen der wahrnehmungsbasierten visuellen Qualität und der objektiven Qualitätsmessung zu erreichen, sollten unbedingt HVS basierte Qualitätsmessungen bzw. Eigenschaften des HVS miteinbezogen werden. [SPA07]

## 5.2. HVS basierte Systeme

Nun soll auf ein HVS basiertes System näher eingegangen werden. Dieses Verfahren arbeitet unter Einfluss des menschlichen visuellen Systems und greift dementsprechend einige Funktionen des HVS auf.

### ***Structural Similarity Index (SSIM)***

Der SSIM Index führt seine Berechnungen auf die strukturellen Informationen eines Bildes zurück. Die strukturellen Informationen spiegeln die Struktur der Objekte in einem Bild

wieder und sind abhängig von der Helligkeit und dem Kontrast. Das menschliche visuelle System würde ebenfalls diese strukturellen Informationen einem Bild entnehmen.

Für dieses Verfahren werden nun die strukturellen Unterschiede zwischen einem Referenz- und einem Testbild berechnet. Für die Berechnung der Bildqualität werden die Helligkeit, der Kontrast sowie die strukturelle Verzerrung vom Referenz- sowie vom Testbild berechnet. Für die Bemessung der Helligkeit und des Kontrasts wird eine einfache Abtastungsstatistik von Bildbereichen herangezogen. Anschließend werden Varianzen sowie Kovarianzen für die Helligkeit und den Kontrast in jedem kleinen Fenster des Referenz- und Testbilds erstellt. Die SSIM-Differenzen stellen eine einfache Bildstatistik zwischen der Helligkeit, Kontrast und Struktur dar und werden zu einem Wert zusammengefasst. Die Kombination dieser drei Werte zeigt die Bildqualität an. Die schematische Darstellung des SSIM-Verfahrens wird in Abbildung 8 nochmals verdeutlicht. [SB07]

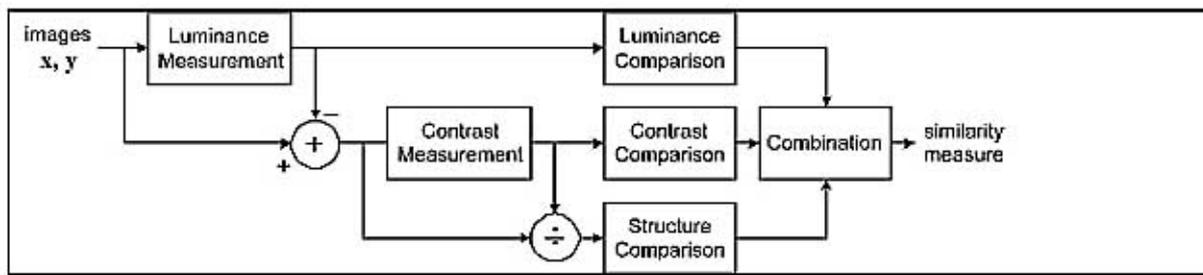


Abbildung 8: Schema des SSIM-Verfahrens: Abbildung aus [SB07]

### 5.3. Gradientmetriken

Eine Erweiterung des SSIM-Index stellt der *Gradient-based Structural Similarity-Index* (GSSIM) dar. Dieser Index untersucht insbesondere die Kanteninformationen eines Bildes. Das menschliche Auge reagiert sehr sensitiv auf Kanten- und Konturinformationen, denn diese Strukturinformationen werden zum Erkennen einer Szene benötigt und stellen daher die wichtigsten Informationen einer Bildstruktur dar. Zum Erkennen der Kanten können die Gradienten eines Bildes eingesetzt werden. [CYX06]

Ein Gradient verfügt über eine Stärke sowie Richtung. Diese Angaben sind ausreichend für die Ermittlung von Kanten innerhalb eines Bildes. Ein weiteres Einsatzgebiet des Gradienten ist die Entdeckung von Kontrastveränderungen. Dafür muss die gesamte Region untersucht werden. Das Gradientenprofil spiegelt die Form und die Schärfe eines Bildes wieder. Somit

kann durch die Ermittlung eines Gradientenfeldes der Unterschied zwischen einer niedrigen und hohen Auflösung eines Bildes festgestellt werden. [SSXS08]

#### 5.4. Verbundenheitsmetriken

Wie bereits im Kapitel Gestaltpsychologie und im Paper [LKS04] festgestellt wurde, spielt Verbundenheit innerhalb eines Objekts eine entscheidende Rolle bei der menschlichen Wahrnehmung. *We observed that for a binary document image, the distance between two pixels plays a major role in their mutual interference perceived by human eyes [LKS04].* Einzelne Pixel können zu einer Formveränderung im Bild führen. Diese Änderungen können sich durch Unverbundenheitsstellen sowie eine unebene Kante äußern.

Um ein Binärbild auf einzelne Pixelfehler hin zu untersuchen, kann das Verfahren der *Distance-Reciprocal Distortion Method (DRDM)* eingesetzt werden. Auf einem schwarz-weißen Eingabebild sind lediglich Buchstaben enthalten, da hierbei einzelne Pixel noch störender auf die menschliche Wahrnehmung wirken als bei natürlichen Bildern. Ein Hauptfaktor bei Formveränderungen ist, dass diese Störungen sofort ins Blickfeld des Beobachters geraten und weiterhin fokussiert werden.

Umso näher zwei Pixel beieinander sind, umso eher wird eine Änderung eines dieser Pixel wahrgenommen. Ein diagonales Nachbarpixel entfernt sich noch weiter als ein horizontales bzw. vertikales Nachbarpixel. Für die DRD-Methode werden nun ein Originalbild sowie ein neu reproduziertes Bild benötigt. Entsprechend der zuvor erwähnten Beobachtungen wird eine gewichtete Matrix erstellt, wobei das Zentrum den Wert 0 erhält. Anschließend wird die Differenz der beiden Bildmatrizen erstellt. Die fehlerhaften Pixelpositionen werden aufsummiert und ergeben den DRD-Fehler. Um dieses Fehlermaß zu testen, haben mehrere Beobachter eine Bildqualitätswertung in einem Labor unter gleichen Bedingungen abgegeben. Der Mittelwert der subjektiven Wertung korreliert sehr stark mit dem DRD-Fehler. Wodurch die Annahme bestätigt werden kann, dass einzelne unverbundene Pixel die menschliche Wahrnehmung stören. [LKS04]

#### 5.5. Evaluierung von Segmentierungen

Die Evaluierung von Segmentierungen kann subjektiv, mit und ohne Ground Truth oder durch das Erlernen einer Evaluierungsfunktion erfolgen [PV08]. Bei einer Segmentierung wird ein

Bild in mehrere Teile zerteilt, das bedeutet, dass für jeden Gegenstand oder jede Region ein Segment erstellt wird oder das Vorder- und Hintergrund zwei Segmente darstellen.

Die Fehlermaße für die Evaluierung von Segmentierungen können wie folgt kategorisiert werden:

- Raumdifferenz
- Grenzabstimmung
- informationsbasierte Berechnung
- nicht-parametrische Tests

Die Raumdifferenz misst die Abweichung zwischen den berechneten Bildsegmenten bzw. den Grad der Überschneidung von zwei Clustern. Bei der Grenzabstimmung wird die Übereinstimmung der Segmentierungsgrenze zwischen der berechneten Lösung und der Referenzlösung mittels Kantenstatistik gemessen. Bei der informationsbasierten Variante wird der Informationsgehalt der Segmentierung berechnet. Dies erfolgt durch die Eigenschaften von zwei Pixels innerhalb eines Segments. Das Informationsmaß kann mittels eines Klassifikators berechnet werden. Ein nicht-parametrisches Verfahren würde beispielsweise Pixelpaare abzählen, deren Segmentierungszugehörigkeit in der berechneten Lösung mit der Referenzlösung übereinstimmt. [UPH07]

In den nächsten Abschnitten wird eine Evaluierungsmöglichkeit mit einem Klassifikator und einem berechneten Index vorgestellt. Beide Varianten benötigen Bilder, die zuvor segmentiert wurden und dadurch entsprechend ihre Segmentierungszugehörigkeit bzw. Labels erhalten haben. Dies bedeutet, dass jedes Segment, das gefunden wurde, mit einem bestimmten Label gekennzeichnet wurde.

#### 5.5.1. Evaluierung mittels Klassifikators

*We formulate the computational problem of segmentation as a classification between "good" segmentations and "bad" segmentation [RM03].* Eine gute Segmentierung sollte aus klaren Begrenzungen bestehen und das Prinzip von Gleichartigkeit beinhalten, das sich in zwei Punkte teilt:

- Zum einen sind Elemente innerhalb einer Region gleich, dies inkludiert, dass die Helligkeit, die Textur usw. identisch sind.
- Zum anderen sind Elemente von verschiedenen Regionen nicht gleich. Deshalb sind auch die Helligkeit und die Textur nicht identisch, die Kantenstärke an der Regionsgrenze ist jedoch hoch.

Um Regionen entsprechend segmentieren zu können, müssen Eigenschaften gefunden werden, die die einzelnen Segmente zu Gruppen zusammenfassen. Beispiele für diese Eigenschaften wären die Texturähnlichkeit, die Helligkeit oder die Kantenstärke. [RM03]

Die Evaluierung erfolgt nun mittels eines binären Klassifikators. Dieser kann zwei Werte zurückliefern, 1 bei einer erfolgreichen Segmentierung und 0 bei einem schlechten Ergebnis. Dafür wird zu Beginn ein Klassifikator für eine gute wiedererkennende Segmentierung trainiert. Die Trainingsdaten bestehen dabei aus Segmentierungsergebnissen, welche manuell in gute und schlechte Segmente eingeteilt wurden. Für das Klassifikationstraining müssen nun passende Eigenschaften wie beispielsweise die Gradienten eines Bildes ausgewählt werden. Diese Eigenschaften werden sowohl am Originalbild als auch bei seinen Segmentierungen ermittelt. Nach erfolgreichem Training kann das Verfahren mittels der Testdaten durchgeführt werden und die Testbilder werden auf Segmentierungsfehler hin untersucht. [PV08]

### 5.5.2. Evaluierung mittels Normalized Probabilistic Rand (NPR) Index

Der Normalized Probabilistic Rand (NPR) Index führt einen quantitativen Vergleich von Bildsegmentierungsalgorithmen durch. Bei den Referenzbildern und deren Segmentierungen handelt es sich um Ground Truth Daten.

Der Probabilistic Rand (PR) Index vergleicht Segmentierungen mit Referenzlösungen durch die Gewichtung von Pixelpaaren. Für die Berechnung wird ein Set von Bildsegmentierungen benötigt. Damit sind Bilder gemeint, bei denen jedes Pixel eine Segmentierungszugehörigkeit erhalten hat. Dieses Testset wird nun mit dem manuell labelisierten Referenzset verglichen. Der Index kann nun wie folgt angegeben werden:

$$PR(S_{test}, \{S_k\}) = \frac{1}{\binom{N}{2}} \sum_{\substack{ij \\ i < j}} [c_{ij} p_{ij} + (1 - c_{ij})(1 - p_{ij})] \quad (12)$$

$S_{\text{test}}$  ist das zu testende Pixel, das mit allen Pixels aus  $S_k$  verglichen wird. Das Set von  $\{p_{ij}\}$  enthält alle Pixelpaare.  $c_{ij}$  kann die Werte  $[0,1]$  annehmen. Bei 0 gibt es keine Übereinstimmung der Labels und bei 1 wäre die Segmentierung der beiden Pixel identisch. Der Index wird über alle möglichen Paare der  $N$  Pixel berechnet. Dies ist sehr aufwändig und die Berechnung dadurch komplex. Deshalb kann eine Reduzierung durch die Nutzung einer Abtastungsschätzfunktion eingeführt werden. Diese ermittelt die Wahrscheinlichkeit, ob zwei Pixel das gleiche Label aufweisen. Anschließend wird eine Normalisierung des PR Indexes durchgeführt, welche aus dem PR Index den Normalized Probabilistic Rand (NPR) Index macht. Beim Testdurchlauf des NPR Index können nun die segmentierten Bilder auf erfolgreiche Segmentierung hin bewertet werden. [UPH07]

## 5.6. Evaluierung bei dieser Studie

In den vorigen Abschnitten wurden einige Evaluierungsmöglichkeiten vorgestellt, doch nicht alle eignen sich für die Evaluierung von Image Matting Resultaten. Die auftretenden Fehler bei den erstellten Bildern der nachfolgenden Benutzerstudie sind Artefakte in den unterschiedlichsten Bereichen. Diese Artefakte können verschiedenste Ausprägungen annehmen, oft sind diese fleckig oder leicht verschwommen. Die komplette Struktur eines Bildes ändert sich dadurch meist nur geringfügig, weshalb der SSIM-Index hier als Übereinstimmungsmaß nicht in Frage kommt. Auch eine vollständige Überprüfung, ob die Segmentierung erfolgreich war, ist hier nicht hinreichend, da die Art des Fehlers für die Studie entscheidender ist.

Einfache Übereinstimmungsmessungen können daher zur Qualitätsbewertung beitragen. Weiters könnten die Verbundenheit oder der Gradientenverlauf ein wesentliches Kriterium für die visuelle Qualität von Image Matting Resultaten sein.

## 6. Benutzerstudie

Image Matting spielt eine sehr bedeutende Rolle in der Bildverarbeitung, weshalb viel Zeit in die Forschung und Entwicklung von Algorithmen investiert wurde. Diese Vielzahl an verschiedenen Algorithmen bringt jedoch Alpha Mattes von unterschiedlicher Qualität hervor. Entscheidend bei der Entwicklung ist nicht nur die Vielzahl, sondern die Bildqualität der resultierenden Alpha Mattes. Um nun zu einer Einigung über die Qualität eines Bildes zu gelangen, werden im Rahmen dieser Studie Probanden befragt und automatisierte Fehlermaße getestet. Bisher wurden zur Evaluierung von Image Matting Resultaten einfache pixelbasierte Metriken wie MAD oder MSE eingesetzt. Ziel dieser Arbeit ist es nun festzustellen, inwieweit diese Fehlermetriken mit der menschlichen visuellen Wahrnehmung übereinstimmen.

Der Mensch spielt bei der Bildqualitätsbeurteilung eine besondere Rolle, da das menschliche Auge Kriterien zur Beurteilung heranzieht, die entscheidender sein können als die rechnerische Übereinstimmung zum Originalbild. Als Hilfsmittel kann hier eventuell die Gestaltpsychologie mit ihren Richtlinien eingesetzt werden, wie beispielsweise der Faktor der Verbundenheit [G02]. Auch die Objektwahrnehmung, wie unter Kapitel 4.3.2 bereits beschrieben, ist hier von enormer Wichtigkeit, denn nur mit ihrer Hilfe ist das Erkennen der Objekte vor unterschiedlichen Hintergründen möglich.

Das Ziel der Benutzerstudie ist die Feststellung der Bildqualität von Matting Resultaten durch Probanden sowie automatisierte Berechnungen. Somit gliedert sich die Studie in drei Segmente: in die Vorstudie, die eigentliche Befragung sowie in den Vergleich zu den automatisierten Berechnungen. Die Vorstudie zielt darauf ab Probleme aufzeigen, die bei einer Befragung auftreten könnten, um diese später bei der Studie zu verhindern. Die Befragung der Probanden besteht aus der Beurteilung des visuellen Aussehens der vorgelegten Bilder. Letztendlich werden die Ergebnisse der automatisierten Berechnungen mit der visuellen Beurteilung der Probanden verglichen und auf Übereinstimmung hin untersucht.

## 6.1. Erstellung der Bilder

Um Matting Resultate zu erzeugen, die in der Benutzerstudie bewertet werden, wurden mehrere Matting-Algorithmen auf 31 Bilder mit bekannter Alpha Matte (Ground truth) angewendet<sup>1</sup>. Die verwendeten Algorithmen sind:

- Bayesian Matting [CCSS01]
- Closed Form Solution [LLW06]
- Easy Matting [GCLDP06]
- Random Walker Matting [GSAW05]
- Robust Matting [WC07\_R]

Die Funktionsweise dieser fünf Algorithmen wurde bereits unter Kapitel 3 näher erläutert.

Mithilfe dieser Verfahren wurden zu jedem der 31 zur Verfügung gestellten Bilder Alpha Mattes erstellt. Aus diesen resultierenden Alpha Mattes wurden in weiterer Folge spezielle Bildausschnitte für die Studie herausgesucht. Das Auswahlverfahren wird in den Kapiteln 6.2.2 und 6.4.2 näher erläutert. Die zusammengesetzten Bilder aus Vordergrundobjekt und neuem Hintergrund wurden mit dem Verfahren von Levin, Lischinski und Weiss [LLW06] erstellt.

Die Ground Truth Alpha Mattes wurden mithilfe von Triangulierung [SB96] erstellt, eine nähere Erläuterung ist unter Kapitel 3.1 zu finden. Diese Alpha Mattes entsprechen der Referenz und dienen dem Vergleich mit den automatisierten Berechnungen, da diese das optimale Resultat darstellen.

Den Probanden wurden sowohl bei dieser Studie als auch bei der Vorstudie die Composites - also die Zusammensetzung aus Vordergrundobjekt und neuem Hintergrund - und nicht die Alpha Mattes gezeigt, da es unklar war, ob von den Probanden die Qualität der abstrakten Alpha Mattes beurteilt werden kann.

---

<sup>1</sup> Das Institut für Softwaretechnik und Interaktive Systeme, Arbeitsgruppe für Interaktive Multimediale Systeme, der Technischen Universität Wien, hat für diese Studie im Rahmen der Diplomarbeit eine Datenbank mit 31 Bildern, den zugehörigen Ground Truth Alpha Mattes, Bildqualitätsberechnungen sowie Mattingalgorithmen zur Verfügung gestellt.

Der neue Hintergrund ist einheitlich aus einer konstanten Farbe, wobei jedes Vordergrundobjekt vor vier unterschiedlichen Hintergrundfarben gezeigt wurde. Bei der Vorstudie wurde die Auswahl der Farben für jedes Beispielbild zufällig und unterschiedlich gewählt. Die Farbe Weiß wurde sowohl bei der Vorstudie als auch bei Hauptstudie als Hintergrundfarbe eingesetzt. Für die Hauptstudie wurden einheitlich die vier Farben Weiß, Rot, Grün und Blau gewählt.

## 6.2. Vorstudie

Vor der eigentlichen Probandenbefragung wurde eine Vorstudie mit nur wenigen Personen durchgeführt. Diese Vorstudie hat den Zweck eventuelle Probleme aufzuzeigen, um die Hauptstudie zu verbessern und das Auftreten von Schwierigkeiten bei der späteren Durchführung zu verhindern. Deshalb sollten bei der Vorstudie unterschiedliche Befragungsvarianten, Bildkategorien, Bildausschnitte, Zeitdauer usw. getestet werden.

### 6.2.1. Einteilung der Bilder in Kategorien

Bei der Vorstudie erfolgt die Einteilung der Bilder in vier Kategorien mit zusätzlichen Unterkategorien. Die ersten drei Kategorien ergeben sich durch die drei Regionen der Trimap: den Vorder-, Hintergrund und die gemischte Region. Dabei weist der Vordergrund immer den Wert 1 auf, der Hintergrund den Wert 0 und die gemischte Region enthält Werte zwischen 0 und 1. Die letzte Kategorie befasst sich mit der Verbundenheit des Vordergrundobjektes und ist durch die Gestaltpsychologie motiviert. Wie bereits erwähnt, werden zum Vergleich der neu erstellten Alpha Mattes die Ground Truth Alpha Mattes als Referenz eingesetzt. Nun werden die vier Bildkategorien näher erläutert:

1. Der Vordergrund hat im Referenzbild den Alpha-Wert 1. Wenn der Alpha-Wert im neu erstellten Bild jedoch nicht 1 ist, gibt es:
  - eine Unterschätzung von Alpha, also der neue Alpha-Wert ist kleiner 1 und größer als 0 oder
  - der Alpha-Wert ist 0.

In anderen Worten ausgedrückt, der Hintergrund ist beim Objekt durchscheinend oder das Vordergrundobjekt ist teilweise oder sogar komplett durch den Hintergrund ersetzt worden.

2. Der Hintergrund hat im Referenzbild den Alpha-Wert 0. Wenn nun der Alpha-Wert im zu vergleichenden Bild größer 0 ist, erscheinen Artefakte im Hintergrund.
3. Die gemischte Region befindet sich zwischen dem eindeutigen Vorder- und Hintergrund, somit ist der wahre Alpha-Wert größer 0 bzw. kleiner 1. Drei Unterkategorien ergeben sich nun bei den zu vergleichenden Bildern:
  - Eine Unterschätzung von Alpha, das heißt, der geschätzte Alpha-Wert des neuen Bildes ist geringer als der Referenz-Alpha-Wert.
  - Eine Überschätzung von Alpha, hier ist der neu geschätzte Alpha-Wert größer als der Referenz-Alpha-Wert.
  - Der Alpha-Wert des neuen Bildes ist 0.

Ein typisches Beispiel für die gemischte Region sind Haare. Die reproduzierten Haare werden nun heller, kontrastreicher oder verschwinden.

4. Die vierte Kategorie bezieht sich nicht auf die drei Regionen der Trimap, sondern betrachtet die Verbundenheit der Vordergrundobjekte. Dabei soll diese Kategorie zeigen, dass die Verbundenheit enorm wichtig ist und dass möglicherweise durchtrennte Objekte einen Störfaktor darstellen können. Auch die Richtlinien der Gestaltpsychologie besagen, dass nur verbundene Objekte eine Einheit [G02] bilden können.

#### 6.2.2. Auswahl der Bildausschnitte für die Vorstudie

Für die Vorstudie wurden aus den berechneten Bildern zu den zuvor erwähnten Kategorien Ausschnitte ausgewählt. Diese Bildausschnitte sind meist eher klein und weisen nur einen bestimmten Fehlerfall auf. Dies ist notwendig, damit sich die Probanden nur auf diesen einen Fehler konzentrieren und nicht durch andere Faktoren beeinflusst werden. Auch für die Auswertung ist dies erforderlich, da sonst nicht nachvollziehbar wäre, welcher Fehler nun zu der schlechten bzw. guten Beurteilung geführt hat. Die Bildausschnitte wurden zusätzlich nachbearbeitet, sodass jeder Ausschnitt tatsächlich nur einen bestimmten Fehler aufweist. Jedoch können diese Fehler aufgrund der unterschiedlichen Funktionsweise der Algorithmen verschiedene Ausprägungen annehmen. Um noch mehr Resultate zu erhalten, wurden Interpolationen von den fehlerhaften Bildausschnitten erstellt.

Bei der Vorstudie sind 18 Bildausschnitte entsprechend der Kategorien unter Kapitel 6.2.1 ausgewählt worden. Alle Beispiele beinhalten unterschiedlichste und ganz beliebige Objekte, wie etwa Stofftiere und Pflanzen. Jedoch ist zu erwähnen, dass bei der Vorstudie lediglich die Resultate von vier Algorithmen (Closed Form Solution, Easy Matting, Random Walk Matting und Robust Matting) eingesetzt wurden.

### ***Kategorie: Vordergrund***

Die Beispiele der Kategorie Vordergrund zeigen Objekte, die Flecken, Schattierungen und Transparenz im Vordergrund aufweisen. Diese Artefakte ergeben sich durch die Unterschätzung der Alpha-Werte. In Abbildung 9 ist ein Beispiel dieser Kategorie ersichtlich. Die Bilder werden aus Platzgründen nur vor der Hintergrundfarbe Weiß gezeigt.



Abbildung 9: Vorstudie: Kategorie Vordergrund, links: Originalausschnitt, restliche Ausschnitte: Resultate unterschiedlicher Algorithmen vor weißem Hintergrund

### ***Kategorie: Hintergrund***

Die Beispiele der Kategorie Hintergrund zeigen weitere beliebige Objekte, die jedoch Artefakte im Hintergrund aufweisen, da die Alpha-Werte in diesen Regionen zu hoch bzw. größer null sind. Die Flecken bzw. Fehler nehmen dabei die unterschiedlichsten Ausprägungen an. Ein Beispiel ist in Abbildung 10 zu sehen.



Abbildung 10: Vorstudie: Kategorie Hintergrund, links: Originalausschnitt, restliche Ausschnitte: Resultate unterschiedlicher Algorithmen vor weißem Hintergrund.

### ***Kategorie: Gemischte Region***

Für die gemischten Regionen wurden überwiegend Haare sowie Plastiksäcke als Beispiele ausgewählt. Haare eignen sich deshalb, da diese sich genau zwischen der Vorder- und Hintergrundregion befinden. Die Segmentierung ist deshalb so schwierig, da es keine

abgeschlossene Kante gibt. Der Plastiksack ist ebenfalls gut geeignet, da dieser durchscheinend ist und der Alpha-Wert somit größer 0 bzw. kleiner 1 ist. Ein Beispiel ist in Abbildung 11 ersichtlich.



Abbildung 11: Vorstudie: Kategorie Gemischte Region, links: Originalausschnitt, restliche Ausschnitte: Resultate unterschiedlicher Algorithmen vor weißem Hintergrund

### ***Kategorie: Verbundenheit***

Die letzte Kategorie behandelt die Verbundenheit der Objekte. In Abbildung 12 ist ein Beispiel zu sehen, das unterschiedliche Variationen eines Blumenstiels zeigt. In einem der Resultate ist der Stiel verschwommen, in einem anderen nur mehr zum Teil erhalten und somit fleckig oder komplett verschwunden.



Abbildung 12: Vorstudie: Kategorie Verbundenheit, links: Originalausschnitt, restliche Ausschnitte: Resultate unterschiedlicher Algorithmen vor weißem Hintergrund

Die Probanden werden nun die Aufgabe haben zu entscheiden, welche dieser Lösungen einem realistischen Aussehen entsprechen und so ein optimales Resultat darstellen.

### 6.2.3. Planung

Am Beginn der Befragung soll den Probanden erklärt werden, worum es sich in dieser Studie handelt bzw. wie der Ablauf erfolgt. Bei der Vorstudie selbst werden drei unterschiedliche Varianten der Befragung getestet, doch immer sind das Originalbild sowie der Originalbildausschnitt zu sehen. Das Originalbild soll der Orientierung dienen. Am Ende werden noch einige Fragen zum Ablauf und den Bildausschnitten an die Probanden gerichtet.

Bei jeder Variante der Befragung sollen die Probanden eine Begründung über seine/ihre Entscheidung abgeben. Die erste Variante der Befragung ist, dass jeweils nur zwei Resultate gezeigt werden und entschieden werden soll, welches der beiden davon besser bzw. schlechter

ist. Wenn eine Entscheidung zwischen zwei Bildern für die Probanden nicht möglich ist, sollen sie das erwähnen, anstelle sich für eines der Bilder zu entscheiden. Diese Konfliktsituation könnte ein Zeichen dafür sein, dass sich dieser Bildausschnitt eventuell nicht für die Studie eignet. Jede Kombination der vier Algorithmen wird gezeigt, bei vier Resultaten ergibt dies sechs Kombinationen. Die zweite Form der Befragung stellt einen Vergleich zwischen Interpolationen dar. Dies bedeutet, dass zwischen einem optimalen und fehlerhaften Resultat interpoliert wird. Die Probanden sollen dabei entscheiden, ab wann der Fehler zu störend und somit das Bild unbrauchbar wird. Die dritte Befragungsvariante zeigt vier Resultate gleichzeitig und die Beurteilung erfolgt mittels Schulnoten.

Die Probanden werden zusätzlich aufgefordert, die Größe der Bildausschnitte zu beurteilen, um diese eventuell zu verkleinern bzw. zu vergrößern. Letztendlich gibt es noch Abschlussfragen, die wie folgt lauten:

- Welche Hintergrundfarben haben Sie bevorzugt?
- Oder haben Sie nur auf eine bestimmte Hintergrundfarbe geachtet und welche?
- Haben Sie in Streitfällen noch andere Farben zur Entscheidungsfindung herangezogen und welche?
- Haben Sie auf größeren Bildausschnitten auch in anderen Bildregionen Fehler gesucht?

#### 6.2.4. Durchführung

Zur Vorstudie wurden vier Personen eingeladen, bestehend aus einer Frau und drei Männern im Alter von 27 bis 47 Jahren. Wie bereits erwähnt, bekamen sie eine kurze Einleitung über den Ablauf der Studie. Die Vorstudie hatte vor allem den Zweck des Informationsgewinns. Deshalb könnte jede Bemerkung, die während der Vorstudie von den Probanden geäußert wird, ein Hinweis auf eine Verbesserung der Studie sein. Aufgrund dessen wurden die Probanden sehr genau beobachtet und sämtliche Kommentare mit notiert.

Insgesamt bekamen die Probanden 31 Beispiele zu sehen. Diese Beispiele beinhalteten die vier Kategorien in beliebiger Reihenfolge sowie die Resultate der vier Algorithmen. Die Reihenfolge der Bildausschnitte war für alle Probanden identisch. Die große Auswahl der Beispiele war wichtig um später einen guten Überblick über alle Bildausschnitte zu erhalten.

Denn nur dadurch war es möglich zu erfahren, welche Ausschnitte sich gut eignen und welche eher nicht, da eine Entscheidungsfindung für die Probanden nicht möglich war.

Für die Dauer der Vorstudie wurde eine Stunde eingeplant. Dabei war der Studienleiterin sehr wohl bewusst, dass eine Stunde für die Durchführung einer Studie zu lange ist. Doch die vier ausgewählten Personen sind darauf vorbereitet worden, dass möglichst viele Informationen benötigt werden, um die Hauptstudie infolgedessen erfolgsversprechend vorzubereiten. Weiters sind die Probanden aufgefordert worden "laut zu denken", da jede Bemerkung eine wichtige Information enthalten könnte. Die Vorstudie hat in einem kleinen Raum stattgefunden, wo die Probanden ungestört tätig werden konnten. Die Bildausschnitte wurden auf einem etwa 30 Zoll großen Bildschirm gezeigt.

#### 6.2.5. Ergebnisse der Vorstudie

Die Vorstudie hat gezeigt, dass einige Änderungen sowohl bei der Durchführung als auch bei den Bildausschnitten unbedingt erforderlich waren. Die Auswertung der Vorstudie lieferte wertvolle Verbesserungsvorschläge. Diese Änderungen waren insbesondere für die nachfolgende Auswertung der Hauptstudie entscheidend.

Erstens hat die Vorstudie ergeben, dass bei der Einleitung zusätzlich ganz genau definiert werden muss, wofür die Bilder eingesetzt werden sollen, denn je nach Anwendungsgebiet könnte ein anderes Bild gewählt werden. Deshalb ist eine Aufgabenstellung wie beispielsweise, dass das neu erstellte Bild das Titelblatt eines bekannten Hochglanzmagazins ist, unbedingt erforderlich. Weiters muss unbedingt erwähnt werden, dass die gezeigten Resultate nicht nachbearbeitet werden können, auch wenn dies bei manchen Bildausschnitten leicht möglich wäre. Letztlich ist noch zu erwähnen, dass die neuen Bilder möglichst realistisch erscheinen sollen. Diese drei Punkte sind so wichtig, da bei der Vorstudie von den Probanden mehrmals gefragt bzw. darauf hingewiesen wurde, dass dieses und jenes Resultat das bessere Beispiel wäre, wenn kleine Änderungen bei bestimmten Bildausschnitten durchgeführt werden würden. Solche Fragen sollen jedoch ausgeschlossen werden. Die gezeigten Resultate liegen in der Endversion für die Entscheidung vor.

Die Dauer der Vorstudie war viel zu lang, da gegen Ende die Probanden müde und unkonzentriert wurden. Die Studie sollte deshalb um die Hälfte der Zeit reduziert und auf maximal 30 Minuten beschränkt werden.

Drei Befragungsvarianten wurden bei der Vorstudie getestet. Die erste Variante hatte zwei Resultate zur Auswahl, wo lediglich entschieden werden sollte, welches der Resultate besser bzw. schlechter sei. Diese Variante wurde von den Probanden als gut angenommen und stellte kein Problem dar. Allerdings ist diese Befragungsform durch die Vielzahl an Kombinationsmöglichkeiten ein großer Aufwand. Die zweite Variante befasste sich mit Interpolationen und es sollte entschieden werden, ab wann ein Fehler sichtbar sei. Diese Variante hat sich als nicht erfolgsversprechend erwiesen, da der Fehler meist bald zu sehen war und die restlichen Bildausschnitte kaum mehr beachtet wurden. Die letzte Form stellte einen Vergleich von vier Resultaten dar und die Beurteilung erfolgte mittels Schulnoten. Bei dieser Art der Befragung konnten die Probanden leicht eine Rangfolge abgeben, doch insbesondere die Vergabe von Schulnoten war problematisch. Zusätzlich wurde erwähnt, dass das Zeigen von mehreren Resultaten gleichzeitig einen größeren Überblick verschaffen würde, hingegen handelt man bei zwei Resultaten lediglich intuitiv.

Die Größe der Bildausschnitte wurde als gut empfunden und sollte auch nicht geändert werden, da bei einer Verkleinerung der Gesamteindruck verloren gehen und bei einer Vergrößerung der Fehler wiederum in den Hintergrund treten würde.

Bei der Frage, welche Hintergrundfarbe bevorzugt wurde, wurde von den Probanden keine eindeutige Präferenz genannt. Die Farbe Weiß als Hintergrund war für alle Probanden sehr wichtig und entscheidend. Bei unklaren Ausschnitten wurden meist alle Hintergrundfarben in Betracht gezogen.

Die Vorstudie hat für die Bildausschnitte ergeben, dass sich manche sehr gut eignen, manche keine Ergebnisse bringen und manche zu eindeutig sind. Die eindeutigen Ergebnisse sollten zufrieden stellend sein, doch meist war der Erkenntnisgewinn gleich null. Die Bildausschnitte, die keine Ergebnisse brachten, waren sich meist zu ähnlich, sodass keine Unterschiede festgestellt werden konnten. Somit konnten Bilder, die zu eindeutig waren bzw. keine Ergebnisse brachten, nicht für die Studie in dieser Form übernommen werden.

Die Vordergrundkategorie zeigt Fehler von unterschiedlichen Ausprägungen an und bestimmte Fehler haben die Probanden mehr gestört als andere. Deshalb sollten in der Hauptstudie Kategorien gefunden werden, die auf diese Art der Fehler vermehrt eingehen.

Die Artefakte im Hintergrund haben auf die Probanden immer störend gewirkt. Die gemischte Region stellte bei einigen Beispielen ein Problem dar, da die Unterschiede nicht deutlich genug zu sehen waren. Ansonsten muss ebenfalls vermehrt auf die Art des Fehlers eingegangen werden. Die Kategorie Verbundenheit konnte soweit bestätigt werden, da unverbundene Objekte tatsächlich als störend empfunden wurden. Jedoch wurde dies nur anhand von zwei Beispielen getestet.

Zu jedem Bildausschnitt wurde den Probanden ein Übersichtsbild gezeigt, indem gekennzeichnet ist, von wo der Bildausschnitt stammt. Doch dieses Übersichtsbild wurde kaum beachtet, da es sich im oberen Bildfeld des Bildschirms befunden hat.

### 6.3. Automatisierte Berechnungen

Die Bildqualität soll im Zuge dieser Studie sowohl mithilfe der menschlichen Wahrnehmung als auch mit automatisierten Berechnungen ermittelt werden. Diese automatisierten Berechnungen wurden mittels MATLAB erstellt. Dafür wurden Berechnungen hinzugezogen, welche in bisherigen Evaluierungen für Image Matting verwendet wurden und die pixelweise vorgehen, wie beispielsweise die Summe der absoluten Differenz.

Für die Eingabe der Berechnungen wird jeweils die Ground Truth Alpha Matte sowie eine reproduzierte Alpha Matte benötigt. Die Ground Truth Alpha Matte dient der Referenz. Diese beiden Bilder werden bei jeder Berechnung verglichen und der jeweilige Differenzwert bzw. Fehler zum Referenzbild wird ausgegeben.

#### ***Differenzbilder:***

Um den genauen Unterschied zwischen dem Referenzbild und dem neu erstellten Bild zu sehen, werden Differenzbilder erstellt. Diese sollen anhand eines Grauwertbildes zeigen, wo sich die unterschiedlichen Regionen befinden. Dafür wird pixelweise die Differenz der beiden Bilder erstellt und anschließend visualisiert.

#### ***Summe der absoluten Differenz:***

Um die Übereinstimmung zwischen dem Referenzbild und dem neu erstellten Bild zu ermitteln, wird die Summe der absoluten Differenz herangezogen (SAD – Sum of Absolute Difference). Für das Ergebnis wird die Differenz pixelweise zwischen den beiden Bildern aufsummiert. Die resultierende Summe gibt den Fehler zum Referenzbild an.

### ***Mittelwert der absoluten Differenz:***

Bei diesem Fehlermaß wird der durchschnittliche Abstand der Beträge ermittelt [T05]. Die Berechnung erfolgt ähnlich wie bei der Summe der absoluten Differenz, nur dass zusätzlich zur Differenzermittlung auch der Mittelwert erstellt wird (MAD – Mean Absolute Distance).

### ***Mittlerer quadratischer Fehler:***

Ein weiteres Übereinstimmungsmaß ist der mittlere quadratische Fehler (MSE – Mean Square Error). Dieser Wert kann auch als einfaches Maß für die Größe der Abweichungen herangezogen werden [S02]. Wenn der Fehlerwert nun eher gering ist, so ist die Übereinstimmung zwischen dem Referenzbild und dem zusätzlichen Bild eher groß. [T05]

## 6.4. Studie

Nach Beendigung der Vorstudie sowie der Erstellung der automatisierten Berechnungen konnte die Hauptstudie nun durchgeführt werden. Der Informationsgewinn durch die Vorstudie war unerlässlich für die Vorbereitung der Studie. Für die Durchführung kommen folgende Änderungen zum Tragen:

- die genaue Definition der Aufgabenstellung am Beginn der Studie,
- die Fehlerkategorien,
- die Bildausschnitte,
- die Befragungsform und
- die Zeitdauer.

Die genaue Definition der Aufgabenstellung ist wichtig, da alle Probanden unter der gleichen Bedingung entscheiden und sich nicht durch eventuelles Vorwissen beeinflussen lassen sollen.

Für die Studie sind neue Kategorien notwendig, da die Vorstudie gezeigt hat, dass die Ergebnisse bei vielen Bildausschnitten nicht brauchbar waren. Dieses Problem entstand durch Konfliktsituationen bei der Bildqualitätsbeurteilung durch die Probanden, da sich viele Bilder zu ähnlich waren oder der Erkenntnisgewinn zu gering war. Deshalb müssen sowohl die Kategorien als auch die Bildausschnitte überarbeitet werden. Die Kategorien werden nun entsprechend der visuellen Wahrnehmung des Menschen angepasst.

Bei der Vorstudie wurden drei verschiedene Befragungsformen getestet, wobei sich nur zwei als gut erwiesen haben. Deshalb ist wohl eine Kombination aus diesen beiden Varianten am erfolgversprechendsten.

Ein Punkt, der bei der Studie nicht außer Acht gelassen werden darf, ist die Zeitdauer. Eine Studie darf nicht zu lange dauern, da die Probanden mit der Zeit unruhig und unkonzentriert werden, was die Ergebnisse negativ beeinflussen könnte. Die Dauer der Vorstudie war viel zu lang und deshalb sollte die Studie gekürzt werden.

#### 6.4.1. Kategorien

Die neuen Kategorien sind durch die Vorstudie sowie die Gestaltpsychologie motiviert und passen sich vermehrt der menschlichen visuellen Wahrnehmung an. Die Artefakte in den Bildausschnitten werden nun entsprechend dieser Kategorien ausgewählt. Bei der Studie wird es lediglich drei Kategorien geben, die nun näher erläutert werden:

1. Verbundenheit: Die Kategorie Verbundenheit wurde bereits bei der Vorstudie eingesetzt und hat sich als vielversprechend erwiesen. Da jedoch dies nur anhand von zwei Beispielbildern getestet wurde, wird diese Kategorie neuerlich überprüft. Dieser Punkt ist durch die Gestaltpsychologie motiviert und zeigt, dass die Verbundenheit von Objekten ein wichtiger Faktor ist, da nur ein verbundenes Objekt eine Einheit bilden kann [G02]. Hingegen wirken durchtrennte Objekte oder einzelne Flecken unruhig und störend.
2. Gradienten: In diesem Punkt geht es um die weichen und harten Übergänge im Vordergrundobjekt. Die betroffenen Bereiche sind meist der Vordergrund sowie die gemischte Region. Beispiele dafür wären Haare oder durchscheinende Bereiche wie bei einem Plastiksack. Eine Über- bzw. Unterschätzung von Alpha ist als Definition zu vage formuliert, deshalb ist die Betrachtung des Gradienten einer Bildregion besser geeignet, da mittels Gradienten Kontrastveränderungen im Bild erkannt werden können und somit auch Artefakte unterschiedlichster Ausprägungen. Die Grundregel der Gestaltpsychologie "*Das Ganze ist mehr als die Summe seiner Teile [G02]*" kann hier wohl auch nicht außer Acht gelassen werden. Die Summe des gesamten Objekts ist für den Menschen entscheidend, doch Artefakte wie Flecken und Löcher können in einem Objekt unrealistisch und störend wirken. Löcher und Flecken würden zu einem

hohen Gradienten führen. Ein geringer Gradient entsteht bei eventuell unschärferen Regionen, die in der Summe jedoch eher von den Probanden angenommen werden, da diese unauffälliger erscheinen.

3. Vorder- und Hintergrund: Bei der Vorstudie ist aufgefallen, dass Fehler im Hintergrund störend wirken. Natürlich sind kleinere Fehler vernachlässigbar, aber grundsätzlich wirken diese immer störend. Deshalb ist der Vergleich zwischen Artefakten im Vorder- und Hintergrund relevanter. Die Frage, die sich hier stellt, ist, ob der Fehler im Vorder- oder Hintergrund mehr störend wirkt. Die Fehler im Vordergrundobjekt könnten oft als Lichtreflexionen interpretiert werden, während die Fehler auf einem einfarbigen Hintergrund immer eindeutig erkennbar sind und so aufdringlicher wirken. Der Grund dafür könnte die Figur-Grund-Trennung [G02] der Gestaltpsychologie sein, die zur Erkennung der Objekte beiträgt und eine klare Trennung zwischen Objekt und Hintergrund zieht.

#### 6.4.2. Auswahl der Bildausschnitte

Die Bilder der Studie wurden der unter Kapitel 6.1 erwähnten Bilddatenbank entnommen. Die resultierenden Alpha Mattes wurden mittels der fünf zuvor näher erläuterten Algorithmen (Bayesian Matting, Closed Form Solution, Easy Matting, Random Walk Matting und Robust Matting) erstellt. Zusätzlich wurden Interpolationen von den mit Fehlern behafteten Bildausschnitten erzeugt, um die Bildauswahl noch zu vergrößern. Wie bei der Vorstudie wurden Bildausschnitte ausgewählt und keine gesamten Bilder. Diese wurden so lange nachbearbeitet, bis jeder Bildausschnitt lediglich einen bestimmten Fehler aufwies. Insgesamt gab es 30 unterschiedliche Beispiele, und je 10 davon lassen sich in eine der drei zuvor erwähnten Kategorien einteilen.

##### ***Kategorie: Verbundenheit***

Als Beispielausschnitte für die Verbundenheit wurden etwa Haarsträhnen oder Blumen ausgewählt. Die meisten Verbundenheitsfehler traten bei Vordergrundobjekten mit Haaren auf. Ein häufigeres Vorkommen ist das komplette Fehlen von Objektteilen. Einige Beispiele sind in Abbildung 13 zu sehen. Zu den Bildausschnitten mit Verbundenheitsfehlern wurden auch Bildausschnitte mit anderen Fehlerausprägungen hinzugemischt wie beispielsweise Unschärfe oder Transparenz. Der Grund dafür ist, dass gezeigt werden soll, dass die

Unverbundenheit von Objekten tatsächlich schwerwiegender ist als andere Fehlererscheinungen.



Abbildung 13: Einige Beispiele für die Unverbundenheit, links: Originalausschnitt, restliche Ausschnitte: Resultate unterschiedlicher Algorithmen vor weißem Hintergrund

### ***Kategorie: Gradient***

Die Beispiele für den Gradienten bestehen überwiegend aus Bildausschnitten mit Haaren, Haarsträhnen oder transparenten Objekten. Dabei sind weiche und härtere Übergänge zu sehen. Einige Beispiele befinden sich in Abbildung 14.

Die Auswahl der Bildausschnitte erfolgte anhand von automatisierten Berechnungen. Die Erwartung war, dass die Reihung der Summe der absoluten Differenz nicht der visuellen Wahrnehmung entsprechen wird. Dies sollte im Ranking der Benutzerstudie bestätigt werden. Denn eine große Differenz zum Referenzbild soll nicht der Entscheidungsträger sein, sondern der Gradient. Die Bildauswahl wird nochmals in der ersten Reihe der Abbildung 14 mit dem Haarbüschel verdeutlicht. Das dritte Bild (SAD: 806) weist einen geringeren SAD-Wert auf als das fünfte Bild (SAD:1215). Jedoch liegt die Vermutung nahe, dass der Mensch Bild 5 bevorzugen wird. Für welches Bild sich die Probanden entscheiden, wird die Auswertung der Studie zeigen und so diese Annahme bestätigen oder eventuell auch widerlegen.

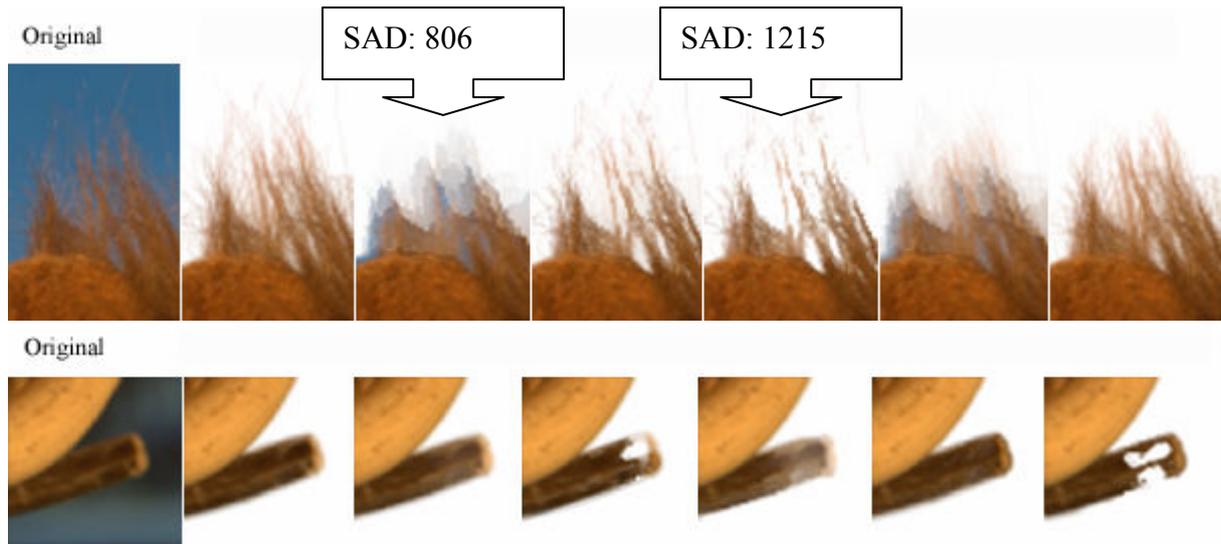


Abbildung 14: Einige Beispiele für den Gradientenfehler, links: Originalausschnitt, restliche Ausschnitte: Resultate unterschiedlicher Algorithmen vor weißem Hintergrund

### ***Kategorie: Vorder- und Hintergrund***

Nun folgen die Fehlerbewertungen im Vorder- und Hintergrund. Es gibt immer ein Vorder- und Hintergrundpaar. Um feststellen zu können, ob die Artefakte in Vorder- oder Hintergrund störender wirken, wurden die Artefakte so gewählt, dass sie den gleichen SAD-Fehler haben. Weiters sollte jedes Paar in etwa die gleiche Fläche des Fehlers haben. Die Beurteilung betrifft jedoch nicht nur ein Paar, sondern kann über alle gleichzeitig gezeigten Beispiele erfolgen. Einige Resultate sind in Abbildung 15 ersichtlich.

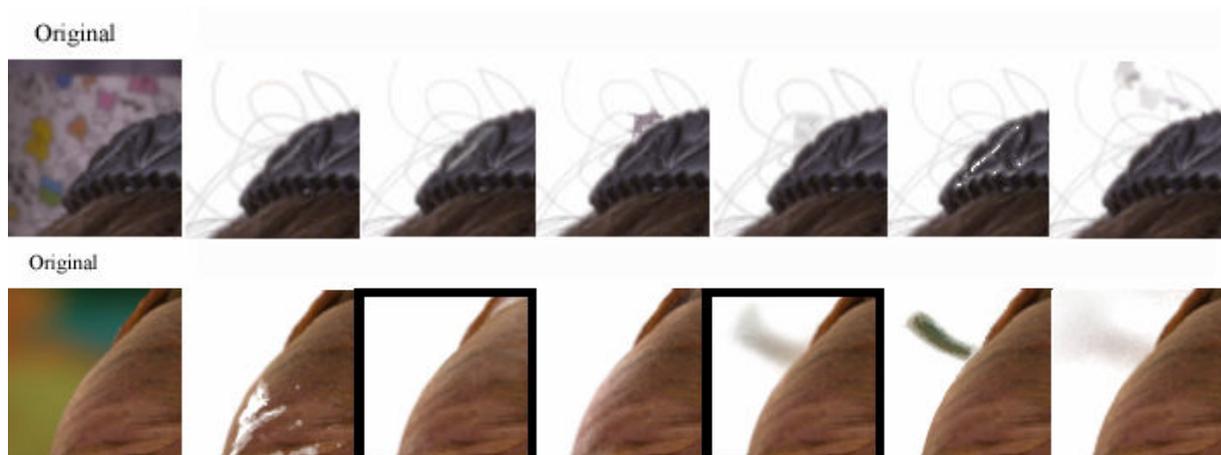


Abbildung 15: Einige Beispiele für den Vorder- und Hintergrund, links: Originalausschnitt, restliche Ausschnitte: Resultate unterschiedlicher Algorithmen vor weißem Hintergrund (die eingerahmten Bilder bilden ein Vorder- und Hintergrundpaar)

### 6.4.3. Planung

Am Beginn der Studie erhalten die Probanden eine kurze Erklärung zu Image Matting sowie eine genaue Aufgabenstellung zu dieser Studie. Anschließend wird ihnen das Setup erklärt, in dem die Beurteilung abgegeben werden soll. Zu sehen ist das Setup in Abbildung 16. Wie die Auswertung der Vorstudie ergeben hat, ist die Bewertung von Bildern durch eine Rangfolge gut möglich. Somit werden den Probanden sechs Resultate gleichzeitig gezeigt.

Ein Resultat wird immer vor vier verschiedenen Hintergrundfarben (Rot-, Grün-, Blau- und Weiß) in einer Zeile abgebildet sein. Die Farbe Weiß hat sich als gute Hintergrundfarbe herausgestellt, da dieses Bild gerne als erstes von den Probanden angesehen wird. In jeder Zeile und somit auch jedem Resultat befindet sich ganz links der Originalausschnitt, der den Vergleich erleichtern soll. An der rechten Seite befindet sich das Originalbild mit einer Markierung derjenigen Stelle, von wo der Bildausschnitt herausgelöst wurde. Das Originalbild ist relativ groß abgebildet, da bei der Vorstudie dieses kaum wahrgenommen wurde und schließlich eine Unterstützung bei der Bewertung sein soll. Insgesamt gibt es 30 unterschiedliche Bildausschnitte, die in einer zufälligen Reihenfolge den Probanden gezeigt werden.

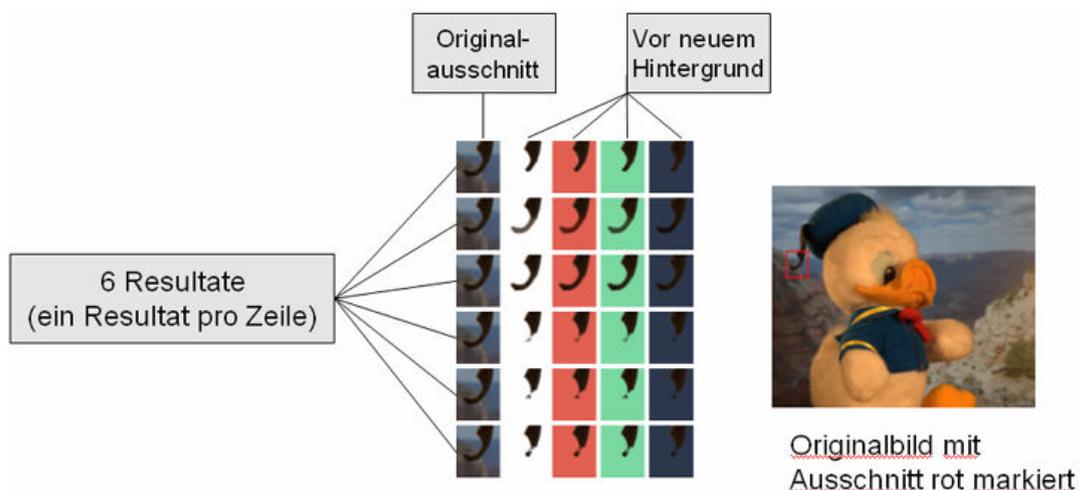


Abbildung 16: Setup für die Beurteilung der Bildausschnitte

Für die Beurteilung sollen die sechs Resultate von den Probanden entsprechend der visuellen Qualität gereiht werden. Das bedeutet, dass sich nach der Beurteilung das visuell beste Resultat an oberster Stelle und im unteren Bereich das schlechteste befinden sollte. Wobei natürlich auch die schlechteren Resultate gereiht werden sollen. In Abbildung 17 wird die Umordnung verdeutlicht. Die einzelnen Reihen können mithilfe der Maus verschoben werden. Auch wenn die einzelnen Resultate vor den unterschiedlichen Hintergrundfarben

möglicherweise verschieden gut erscheinen, muss die Bewertung immer über die gesamte Reihe getroffen und Prioritäten gesetzt werden.

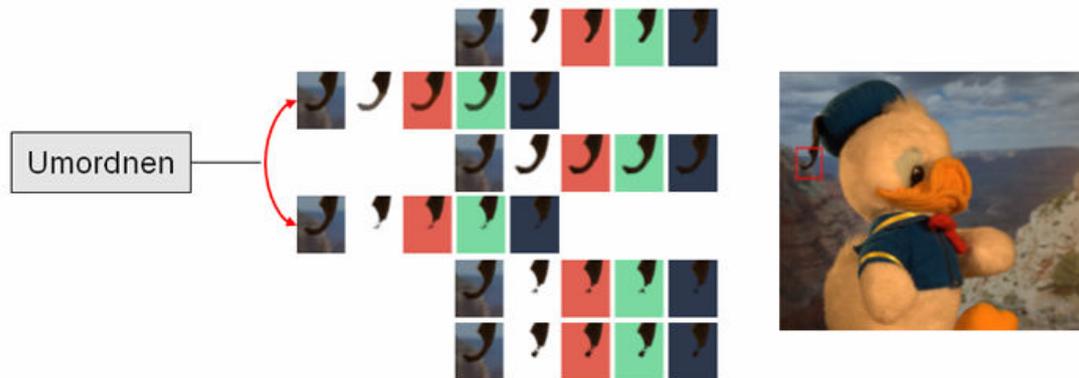


Abbildung 17: Setup zur Beurteilung der Bildausschnitte mit dem Hinweis, dass die Resultate zeilenweise umgeordnet werden können.

Wenn die Reihung abgeschlossen ist, kann der Benutzer das nächste Testbild abrufen. Es ist auch möglich, ein bereits gesehenes Bild nochmals anzusehen und Entscheidungen zu revidieren.

In der Vorstudie gab es etliche Bildausschnitte, bei denen die Probanden zu keiner Entscheidung gelangt sind, deshalb wird darauf geachtet, dass solche Bildausschnitte in dieser Form nicht mehr übernommen werden. Jedoch kann auch hier nicht ausgeschlossen werden, dass die Probanden in eine Konfliktsituation kommen und sich zwischen zwei Bildausschnitten nicht entscheiden können. Daher wird den Probanden die Möglichkeit gegeben Entscheidungsschwierigkeiten anzugeben, da diese beiden Bildausschnitte bei der Auswertung den gleichen Rang erhalten würden. Weiters werden die Probanden aufgefordert eine Begründung abzugeben, wenn die Beurteilung problematisch ist.

Da bei der Vorstudie die Zeitdauer ein großes Thema war, wird jetzt darauf geachtet, dass die Probanden etwa 30 Minuten für die komplette Studie einschließlich der Einführung am Beginn benötigen. Wobei die Probanden nicht unter Druck gesetzt werden und entsprechend ihrer Geschwindigkeit handeln sollen. Die Zeitdauer ist somit nicht beschränkt. Da die Studie nicht länger als 30 Minuten dauern sollte, sind lediglich 30 Bildausschnitte gewählt worden. Bei den Probanden wird ein Lerneffekt bei der Bildqualitätsbeurteilung erwartet, der die Bewertung der einzelnen Bilder beschleunigen wird.

Die Studie findet in einem kleinen, abgeschlossenen Raum statt. Dieser Raum ist sehr ruhig, sodass die Probanden nicht gestört oder abgelenkt werden und die volle Konzentration auf die Studie gelegt werden kann. Der Arbeitsplatz besteht hierbei aus einem Schreibtisch, einem Sessel sowie einem Bildschirm mit etwa 30 Zoll. Die Studie findet für alle teilnehmenden Probanden in diesem Raum statt, sodass die Bedingungen und auch die Bildausschnitte, da sie immer auf dem gleichen Bildschirm gezeigt werden, identisch sind.

#### 6.4.4. Durchführung

Für die Bildqualitätsbeurteilung wurden 17 Personen im Alter von 24 bis 67 Jahren eingeladen. Neun der Personen sind weiblich und acht männlich. Das Durchschnittsalter dieser 17 Personen beträgt 36,4 Jahre. Die Kandidaten weisen unterschiedliches Vorwissen sowie Werdegänge auf, sind jedoch keine Experten im Bereich Bildverarbeitung.

Am Beginn der Studie erhalten die Probanden Einweisungen bzw. Erklärungen, die im Folgenden beschrieben werden. Die Aufgabenstellung ist, dass der neue Bildausschnitt auf ein Titelblatt eines Hochglanzmagazins kommt und die möglichen Hintergrundfarben sind Rot, Blau, Grün oder Weiß. Die Resultate sollten dabei möglichst realistisch erscheinen. Für die Auswahl des Titelblatts stehen nur die gezeigten Resultate zur Verfügung.

Weiters werden die Probanden darauf hingewiesen, dass keine Nachbearbeitung möglich ist. Eventuell auftretenden Fragen nach kleinen Änderungen oder Ausbesserungen an den Bildausschnitten sollte – wie die Vorstudie gezeigt hat – jedenfalls vorgegriffen werden. Diese Einschränkung ist notwendig, da die Bedingungen für alle Probanden gleich sein sollen. Weiters gibt es eine kurze Erläuterung zu dem Setup für die Beurteilung der Bildausschnitte.

Während der Studie könnten Konfliktsituationen bzw. Entscheidungsschwierigkeiten zwischen den verschiedensten Bildausschnitten auftreten. Der erste Konfliktfall ist, dass eine Entscheidung aufgrund von Ähnlichkeit zweier Bildausschnitte nicht getroffen werden kann. Der zweite Konfliktfall entsteht bei Bildausschnitten, die sehr unterschiedlich sind, aber der Fehler bei beiden derart schwerwiegend ist, dass ebenfalls keine Entscheidung getroffen werden kann. Wenn eine der beiden Konfliktsituationen für die Probanden entsteht, sollten sie darauf hinweisen.

## 7. Auswertung der Ergebnisse

Die Auswertung erfolgt ähnlich wie bei Gijssenij, Gevers, Lucassen in [GGL08]. Zu Beginn wird der durchschnittliche Proband ermittelt. Dafür wird der Mittelwert der Probandenbewertung für jeden Bildausschnitt berechnet. Der durchschnittliche Proband stellt eine robuste subjektive Auswertung dar. Dieses resultierende Ranking dient dem Vergleich zu den einzelnen Probanden sowie den automatisierten Berechnungen und wird folglich mit Mittelproband bzw. Mittelperson angegeben. [GGL08]

Zur Berechnung der Korrelation zwischen der berechneten Mittelperson und den einzelnen Kandidaten wird der Kendall'sche Korrelationskoeffizient für stetige Merkmale herangezogen [HE07]. Die Berechnung dieses Korrelationskoeffizienten basiert auf Rangzahlen, diese entsprechen den Resultaten, die sich bei der Durchführung der Studie ergeben haben. Die resultierenden Werte liegen zwischen -1 und 1, während eine 1 eine Übereinstimmung bedeutet. Die Berechnung des Korrelationskoeffizienten erfolgt mittels MATLAB und der vordefinierten Funktion *corr*. Durch den Kendall'schen Korrelationskoeffizienten können Übereinstimmungen sowie Ausreißer ermittelt werden. Die statistischen Ausreißer werden in der nachfolgenden Auswertung näher betrachtet, da die Beweggründe des Ranking verschieden zu den anderen Probanden sind.

Zusätzlich wird der Kendall'sche Korrelationskoeffizient zwischen den automatisierten Berechnungen sowie den Probanden bzw. dem Mittelprobanden ermittelt. Denn ein Ziel dieser Studie ist es zu zeigen, dass bisher verwendete Berechnungen nicht ausreichen um Verbundenheit oder Gradienten zu beschreiben.

Bei der folgenden Auswertung der einzelnen Bildausschnitte werden jeweils die Resultate in der Reihenfolge der Wertung des Mittelprobanden sowie die zugehörigen automatisierten Berechnungen angegeben. Die Probanden haben bei der Durchführung der Studie jedes Bild vor vier Hintergrundfarben zu sehen bekommen, hier wird aufgrund der Platzersparnis jedes Resultat nur vor der Hintergrundfarbe Weiß gezeigt. Ausgenommen sind zwei Bildausschnitte, da diese vor weißem Hintergrund schwer zu erkennen und deshalb vor blauem Hintergrund dargestellt werden.

## 7.1. Analyse der Einzelergebnisse

Die anschließenden Unterkapitel enthalten die genaue Analyse der einzelnen Bilder sowie die automatisierten Berechnungen.

### 7.1.1. Kategorie: Verbundenheit

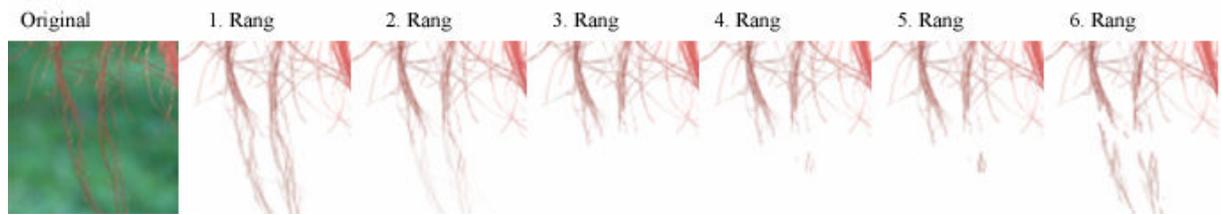
Diese Kategorie beschäftigt sich mit der Verbundenheit der Objekte. Aufgrund dessen weisen die Bildausschnitte vor allem Verbundenheitsfehler auf, was bedeutet, dass Objektteile unterbrochen oder zerrissen sind. Um zu zeigen, dass die Verbundenheit von Objekten entscheidend ist, werden gleichzeitig Bildausschnitte gezeigt, die diverse andere Fehler aufweisen wie beispielsweise Transparenz. Weiters werden Beispiele vorgelegt, bei denen Teile des Objektes fehlen, denn die Probanden sollen entscheiden, ob Objekte eventuell realistischer erscheinen, wenn Teile fehlen als wenn sie unterbrochen sind. Ein Beispiel dafür wäre eine verkürzte Haarsträhne.

Für alle Bildausschnitte wurden Berechnungen wie die absolute Differenz, der Mittelwert der absoluten Differenz sowie der mittlere quadratische Fehler ermittelt. Auf den folgenden Seiten werden nun die Ergebnisse der einzelnen Bildausschnitte näher erläutert.

#### ***Bild: Haarsträhne***

Bei diesem Bildausschnitt handelt es sich um zwei Haarsträhnen. Dabei sind die Haare teilweise unterbrochen, blasser und nur zum Teil erhalten. Die Probanden sollten hier entscheiden, ob die Durchtrennungen der Haare, komplettes Fehlen oder andere Fehler wie beispielsweise Transparenz schwerwiegender für das realistische Aussehen sind.

Das Ranking des Mittelprobanden sowie die zugehörigen errechneten Werte sind in Abbildung 18 zu sehen. Der Großteil der Probanden hat hier als Beweggrund der Entscheidung die Verbundenheit gewählt, das heißt, die Bildausschnitte mit den verbundenen Haarsträhnen wurden besser bewertet als die unverbundenen Haare. Auf Rang 3 wurde ein Bildausschnitt gewählt, bei dem die Hälfte der Haarsträhnen fehlt, jedoch entspricht die gekürzte Haarsträhne der Realität.



SAD	109	243	340	331	324	218
MAD	0,84	1,87	2,62	2,55	2,5	1,68
MSE	0,08	0,36	0,69	0,66	0,65	0,30

Abbildung 18: Rangfolge des Mittelprobanden vom Bild Haarsträhne sowie die zugehörigen berechneten Werte

Die unverbundenen Haare wurden in drei unterschiedlichen Variationen gezeigt, diese bestanden aus kleinen Resthaaren sowie aus der Durchtrennung der gesamten Haarsträhnen. Der Großteil der Probanden hat entschieden, dass die kleinen Resthaare weniger störend bzw. weniger auffallend wirken, als wenn eine gesamte Haarsträhne durchtrennt wird. Lediglich einer kleinen Gruppe von drei Personen war die Wiedererkennung zum Originalbild wichtiger und hat deshalb die durchtrennte Haarsträhne besser bewertet. Die kleinen Haarreste hingegen erschienen ihnen wie Flecken im Bild.

Bei näherer Betrachtung der automatisierten Berechnungen ist erkennbar, dass der SAD-, der MAD- sowie der MSE-Wert nicht der Reihenfolge der Probandenmeinung entspricht, die Korrelation zwischen der Mittelperson und den automatisierten Berechnungen liegt bei 0,07. Beispielsweise ist der SAD-Wert und somit die Differenz vom Bildausschnitt mit Rang 3 zum Originalbild sehr hoch, doch die menschliche visuelle Meinung entscheidet hier nach anderen Kriterien. Dieses Kriterium behandelt die Verbundenheit von Objekten und der geringe Informationsverlust wird hingenommen.

***Bild: Lasche eines Hutes 1***

Dieses Bild zeigt die Lasche eines Hutes, der zugehörige Bildausschnitt ist in Abbildung 19 zu sehen. Diese Lasche ist unterbrochen, weist Transparenzen auf oder ist nur zum Teil erhalten.

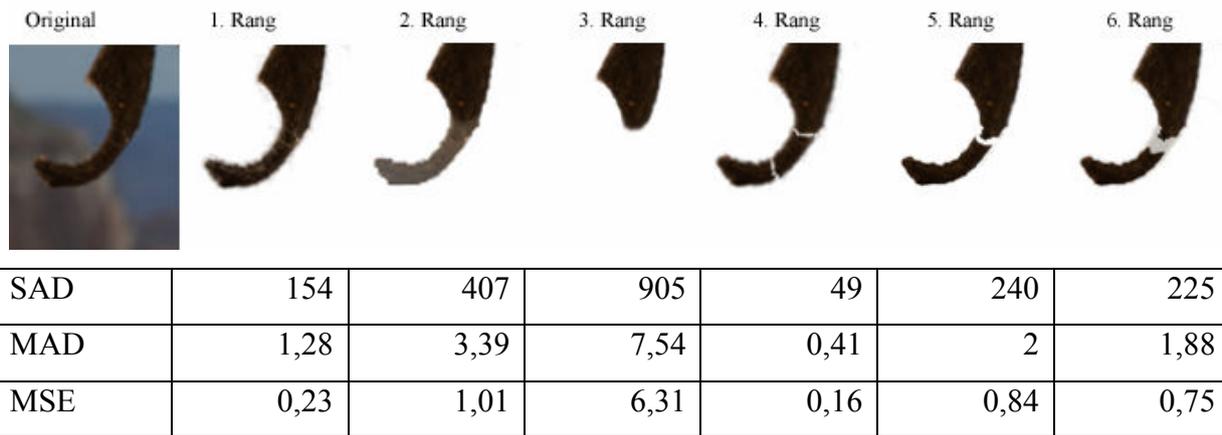


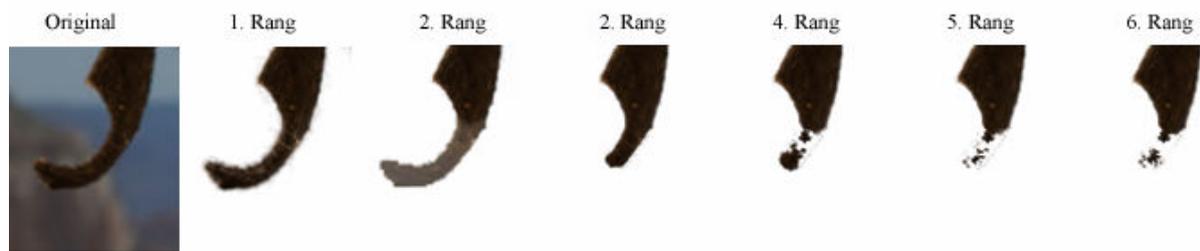
Abbildung 19: Rangfolge des Mittelprobanden vom Bild Lasche eines Hutes 1 sowie die zugehörigen berechneten Werte

Auch hier war der Entscheidungsgrund der Probanden die Verbundenheit und das verkürzte Objekt wurde erneut in das Mittelfeld gewählt. Jedoch hat eine Gruppe von sechs Personen den Bildausschnitt von Rang 3 auf den letzten Platz gesetzt, da dieser enorme Informationsverlust einen schwerwiegender Fehler darstellt und deshalb das Objekt kaum mehr erkennbar ist. Zwei Personen haben als Entscheidungsgrund zur Verbundenheit die Struktur des Objektes - also die Form und Farbe des ursprünglichen Objektes - genannt und so den Bildausschnitt von Rang 5 vor den Bildausschnitt von Rang 4 gesetzt. Eine weitere Person hat die Transparenzen des Objektes in die Entscheidungsfindung einfließen lassen und den Bildausschnitt von Rang 2 auf 3 gesetzt. Die Unverbundenheit wurde trotz allen Einflüssen als schwerstwiegend eingestuft.

Der SAD-Wert entspricht nicht der Reihenfolge der Probandenmeinung. Denn beispielsweise die Unterscheidung vom Bildausschnitt von Rang 4 ist sehr gering zum Originalbild, doch die menschliche visuelle Wahrnehmung entscheidet hier nach dem Kriterium der Verbundenheit und reiht deshalb dieses Resultat nur auf Platz 4. Dies spiegelt sich auch in der Korrelation mit dem Mittelprobanden wieder, denn der Korrelationskoeffizient zwischen der Mittelperson und SAD, MAD sowie MSE liegt bei -0,07.

### ***Bild: Lasche eines Hutes 2***

Hierbei handelt es sich um den gleichen Bildausschnitt wie im vorigen Abschnitt. Jedoch weisen die Bildausschnitte zum Teil andere Fehler auf. Zu sehen ist dies in Abbildung 20.



SAD	154	407	610	802	789	721
MAD	1,28	3,39	5,08	6,68	6,57	6,01
MSE	0,23	1,01	3,96	5,34	5,25	4,7

Abbildung 20: Rangfolge des Mittelprobanden vom Bild Lasche eines Hutes 2 sowie die zugehörigen berechneten Werte

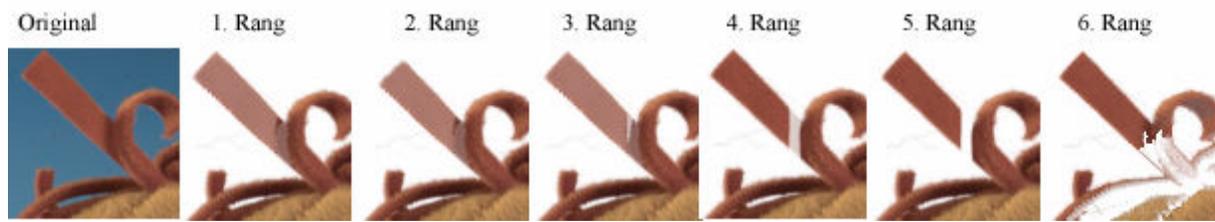
Bei diesem Bildausschnitt ist die Übereinstimmung nicht so eindeutig wie bei den zwei vorigen Bildausschnitten, doch auch hier wird die Verbundenheit als entscheidend gewertet. Sieben Personen haben die Struktur und Kontur als Entscheidungsgrund mit einfließen lassen und so den Bildausschnitt, bei dem lediglich das Endstück fehlt, auf Rang 2 oder sogar 1 gesetzt. Hingegen wurden Resultate, die im Ganzen vorhanden sind, schlechter bewertet, da diese Transparenzen aufweisen.

Die Bildausschnitte mit den unverbundenen Objektteilen wurden auf den 4., 5. und 6. Rang gewählt, jedoch gibt es bei der Reihung keine eindeutige Übereinstimmung. Eine Gruppe von 12 Personen wertet den großen Fleck von Bildausschnitt auf Rang 4 als besser, weil dieser als Teil des Originalobjekts wahrgenommen werden kann. Die kleinen Flecken hingegen wirken nur unrealistisch und störend. Eine zweite Gruppe von fünf Personen ist der Meinung, dass ein großer Fleck aufgrund seiner Dominanz mehr störend ist.

In diesem Beispiel stimmt der Mittelproband mit den automatisierten Berechnungen zu 0,86 überein. Die Reihung der automatisierten Berechnungen ist bei allen drei Varianten identisch.

### ***Bild: Haupt des Elefanten***

Bei diesem Bildausschnitt treten Durchtrennungen sowie Transparenzen in unterschiedlichen Ausprägungen bei den Haupthaaren eines Stoffelefanten auf, wie in Abbildung 21 zu sehen ist. Hier soll insbesondere gezeigt werden, ob die Transparenz eines Bildes oder die Verbundenheit mehr Einfluss auf das realistische Aussehen eines Objektes hat.



SAD	232	312	243	83	111	1211
MAD	3,13	4,22	3,29	1,12	1,5	16,38
MSE	0,7	1,87	0,81	0,72	1,28	13

Abbildung 21: Rangfolge des Mittelprobanden vom Bild Haupt des Elefanten sowie die zugehörigen berechneten Werte

Der Großteil der Probanden hat sich hier einstimmig für die Verbundenheit des Objektes entschieden. Die Rangabstufung erfolgt nach der Breite der Durchtrennungen. Die ersten beiden Bildausschnitte unterscheiden sich nur durch eine geringe Verkürzung, die den meisten Probanden auf den ersten Blick nicht aufgefallen wäre. Lediglich eine Person hat die Transparenz des Vordergrundobjektes vor die Unverbundenheit gesetzt.

Die Übereinstimmung bei diesem Bildausschnitt unter den Probanden lag hier bei 0,91. Die Korrelation zu den automatisierten Berechnungen ist jedoch nicht so eindeutig. Der Korrelationskoeffizient vom Mittelprobanden und dem SAD sowie MAD liegt bei 0,07. Die Übereinstimmung zum MSE liegt bei 0,47.

### ***Bild: Blumenstiel***

Bei diesem Bildausschnitt handelt es sich um eine Blume, in deren Mitte sich viele kleine Stiele befinden. Es weisen jedoch nur zwei dieser Stiele Artefakte auf, der Rest des Bildes wurde nachbearbeitet und ist bei allen Bildausschnitten identisch. Die reproduzierten Resultate weisen diverse Unterbrechungen sowie Transparenzen auf (Abbildung 22).



SAD	53	128	82	122	22	112
MAD	0,47	1,15	0,74	1,1	0,2	1
MSE	0,19	0,57	0,37	0,55	0,11	0,46

Abbildung 22: Rangfolge des Mittelprobanden vom Bild Blumenstiel sowie die zugehörigen berechneten Werte

Auch bei diesem Bild stand als Entscheidungsgrund die Verbundenheit im Vordergrund. Zwei Personen machen dabei eine Ausnahme, da sie die Struktur in den Vordergrund stellten. Insbesondere eine dieser beiden Personen hat alleine die Struktur des Blumenstiels entscheiden lassen. Die Unverbundenheit wurde nicht so schwerwiegend bewertet, da dies - laut diesem Probanden - bei diesem Motiv vernachlässigbar ist. Die zweite Person, die die Struktur in den Vordergrund gestellt hat, hat die Verbundenheit jedoch nicht außer Acht gelassen. Die Anzahl der Flecken sowie deren Ausprägung waren sehr wohl entscheidend. Zwei weitere Personen haben den Informationsgehalt ebenso wie die Verbundenheit betrachtet, wodurch der Bildausschnitt von Platz 5 nach vor gereiht wurde.

Jedoch ist an dieser Stelle zusätzlich zu erwähnen, dass mehrere Probanden gemeint hätten, dass die Flecken bei diesem Motiv nicht so störend wirken, da sich die einzelnen Punkte verlieren würden. Für die Bewertung war dies etwas problematischer, da die Entscheidungen nicht ganz einfach getroffen werden konnten.

Die Korrelation zwischen der Mittelperson und dem SAD, MAD sowie MSE beträgt  $-0,14$ . Die automatisierten Berechnungen stimmen hier nicht annähernd mit der menschlichen visuellen Wahrnehmung überein, da Unverbundenheitsstellen schwerwiegender gewertet werden als beispielsweise Transparenz.

***Bild: Krallen eines Stofflöwen***

Bei diesem Bildausschnitt wurde die Krallen eines Stofftieres bzw. Löwen näher betrachtet (Abbildung 23). Die Krallen weisen Transparenzen, Durchtrennungen sowie das fast komplette Fehlen der Krallen auf.

Bei diesem Bildausschnitt steht die Verbundenheit eindeutig vor allen anderen Fehlerausprägungen. Doch die extreme Blässe wirkt insbesondere für zwei Personen sehr unrealistisch, sodass diese beiden den Bildausschnitt von Rang 3 auf 4 nach unten gereiht haben. Die unterbrochene Krallen ist für den Großteil der Probanden noch besser als die beiden Bildausschnitte, bei denen die Krallen nur mehr als einzelne Linie repräsentiert wird (in Abbildung 23 Rang 5 und 6). Bei der Krallen als Linie ist wiederum die Verbundenheit entscheidend. Die geschwungene Linie - die vom Stofflöwen weggeführt - könnte in der

Vorstellung eine kleine Haarsträhne oder Ähnliches darstellen, während die Linie neben dem Stofftier nur ein störendes Artefakt sein kann.

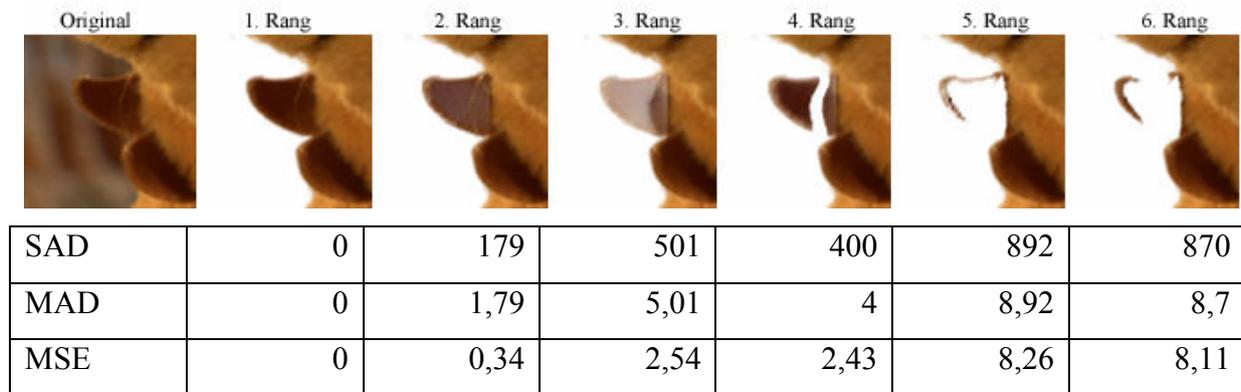


Abbildung 23: Rangfolge des Mittelprobanden vom Bild Kralle sowie die zugehörigen berechneten Werte

Der Korrelationskoeffizient zwischen Mittelperson und den automatisierten Berechnungen ist 0,73.

### ***Bild: Ananas - Blattschopf***

Dieses Beispiel zeigt eine Ananas, wobei die Farben von Vorder- und Hintergrund sehr kräftig und insgesamt kontrastreicher wirken als die zuvor gezeigten Bildausschnitte. Der Originalbildausschnitt in Abbildung 24 verdeutlicht dies. Bei diesem Beispiel treten bei einem Ananasblatt Durchtrennungen, Löcher sowie Verkürzungen auf.

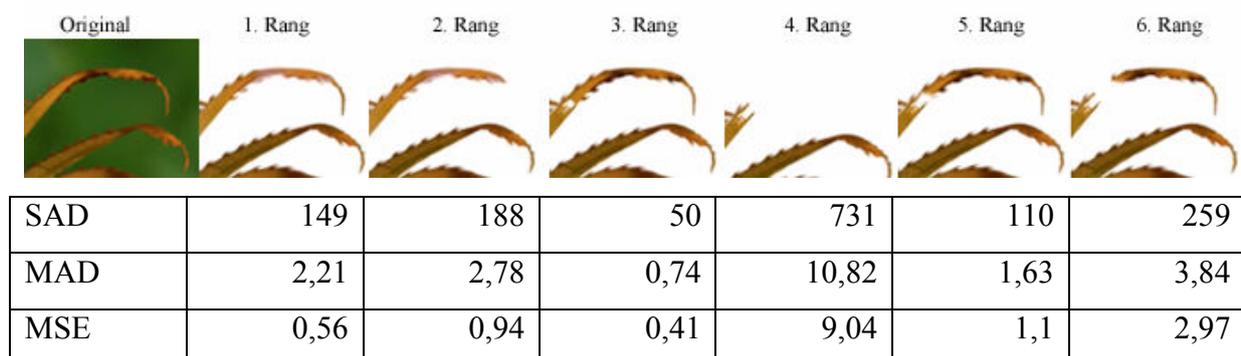


Abbildung 24: Rangfolge des Mittelprobanden vom Bild Ananas sowie die zugehörigen berechneten Werte

Auch beim letzten Beispiel der Verbundenheit ist an der Rangfolge des Mittelprobanden zu erkennen, dass für die menschliche visuelle Wahrnehmung verbundene Objekte enorm wichtig sind. Das Ranking war hier nicht eindeutig, denn insbesondere für eine Person stand erneut die Struktur und nicht die Verbundenheit im Vordergrund. Trotzdem wurden die

Bildausschnitte mit den unverbundenen Blättern auch von dieser Person nach hinten gereiht. Der Bildausschnitt von Rang 3 zeigt ein Loch im Ananasblatt, dies war für neun Personen ein realistisches Bild, da solche Löcher auch in der Natur vorkommen können. Der Bildausschnitt von Rang 4 wurde hingegen sehr unterschiedlich bewertet. Während es für die einen ein abgerissenes Blatt darstellte, ist es für die anderen der reine Informationsverlust. Allerdings wurde dieser Bildausschnitt von Rang 4 meist besser bewertet als der Ausschnitt von Rang 6, da hier das Ananasblatt ein komplett unrealistisches Aussehen annimmt.

Der Korrelationskoeffizient zwischen Mittelperson und SAD sowie MAD ergibt 0,2 und zu MSE liegt dieser bei 0,47.

Weitere Beispiele mit ähnlichen Fehlerausprägungen haben gezeigt, dass die verbundenen Objekte immer an den Anfang und die Bildausschnitte mit unverbundenen Stellen absteigend nach der Breite und Anzahl der Durchtrennungen gereiht werden. Transparenz stellt ebenfalls einen schwerwiegenden Fehler dar, der in die Entscheidung über die visuelle Qualität mit einfließt, da das Objekt unrealistische Züge erhält. Damit wird gezeigt, dass die Struktur und Kontur eines Objektes wesentlich sind. Die Probanden haben je nach Motiv entschieden, welche Fehlerausprägung eher akzeptiert werden kann. Insbesondere bei kleinen Pixelfehlern in sehr detailreichen Motiven konnte die Verbundenheit vernachlässigt werden. Doch überwiegend war die Verbundenheit hauptsächlicher Entscheidungsträger, da die Objekte sonst meist unnatürlich wirkten.

Eine weitere Fehlerausprägung bei den Vordergrundobjekten stellt der Informationsverlust dar, der jedoch sehr unterschiedlich bewertet wurde. Zum einen sind kleine Verluste kaum wahrgenommen worden und zum anderen haben große Reduktionen zu unterschiedlichen Meinungen geführt. Für eine Gruppe stellte der Informationsverlust kein Problem dar, denn Verluste sind jedenfalls besser als wenn Objektteile unverbunden sind. Für die andere Gruppe war das Motiv als solches nicht mehr erkennbar und somit erfuhr es eine Ablehnung.

#### 7.1.2. Kategorie: Gradient

Diese Kategorie behandelt die Gradienten innerhalb eines Vordergrundobjektes. Dabei soll gezeigt werden, dass ein hoher Gradient innerhalb eines Objektes mehr stören kann als ein niedriger. Anders ausgedrückt: Flecken bzw. Löcher im Vordergrundobjekt wirken oft störender, als wenn das Objekt konstant blasser bzw. transparenter wird. Um dieses Problem

zu verdeutlichen, werden zehn Bildausschnitte mit den unterschiedlichsten Fehlerausprägungen gezeigt. Auch für diese Bildausschnitte wurden automatisierte Berechnungen mittels MATLAB erstellt. Diese Berechnungen sind die absolute Differenz, der Mittelwert der absoluten Differenz sowie der mittlere quadratische Fehler.

**Bild: Kleid**

Bei diesem Bild ist ein bunt gemustertes Kleid einer Puppe zu sehen, ersichtlich in Abbildung 25. Jedoch weist der Stoff des Kleides Fehler wie Löcher und Transparenzen in unterschiedlichen Ausprägungen auf.



Abbildung 25: Rangfolge des Mittelprobanden vom Bild Kleid sowie die zugehörigen berechneten Werte

Die Rangfolge der Probandenbewertung ist in Abbildung 25 ersichtlich. Der Großteil aller Probanden hat die Flecken im Kleid störender empfunden als die Unschärfe bzw. die Transparenz. Die Löcher entsprachen nicht der Norm und wurden somit eher nach hinten gereiht. Allerdings kann auch Unschärfe störend wirken, wodurch kleine Flecken von den Probanden eher akzeptiert wurden. Insbesondere bei sehr auffälligen Flecken wie bei den Bildausschnitten von Rang 5 und 6 wird die Blässe bzw. die Transparenz bevorzugt. Vier Personen haben den Bildausschnitt von Rang 5 auf den letzten Platz gesetzt, da die Übergänge der Flecken noch etwas härter sind als beim Bildausschnitt von Rang 6.

Die automatisierten Berechnungen wie SAD und MAD korrelieren mit der Mittelperson mit einem Betrag von 0,47. MSE liegt dabei etwas höher bei 0,6.

**Bild: Stroh**

Bei diesem Bildausschnitt wird ein kleiner Teil eines Strohballens gezeigt, zu sehen ist dies in Abbildung 26. Der Bildausschnitt weist Transparenzen sowie Löcher auf. Die Transparenzen

steigern sich in der Stärke und die Löcher werden in zwei unterschiedlichen Variationen gezeigt.

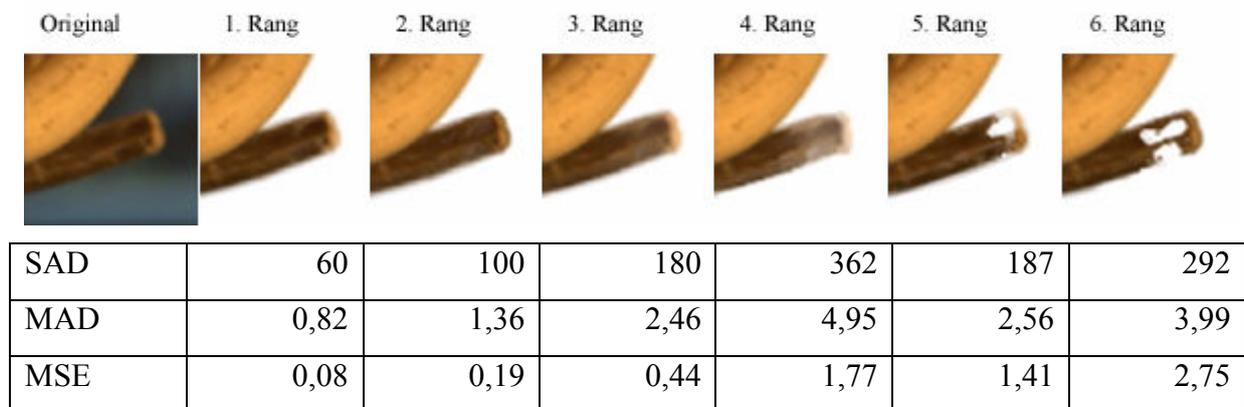


Abbildung 26: Rangfolge des Mittelprobanden vom Bild Stroh sowie die zugehörigen berechneten Werte

Bei der Rangfolge des Mittelprobanden weist das Vordergrundobjekt am Beginn sehr kleine, fast unscheinbare Flecken auf, diese gehen in Transparenzen über, bis am Ende deutliche Löcher enthalten sind. Diese Reihung wird in Abbildung 26 verdeutlicht. Bei der Probandenbewertung gab es eine kleine Unstimmigkeit zwischen den ersten beiden Bildausschnitten, da einige Probanden den Unterschied kaum erkennen konnten. Lediglich einen Ausreißer gab es bei diesem Beispiel, dieser stellte die Struktur des Objektes in den Vordergrund und bezeichnete die Unschärfe als störenden Faktor. Dies bedeutet, dass der Proband den Bildausschnitt von Rang 4 auf den letzten Platz und der Ausschnitt von Rang 6 auf den 4. Platz gereiht hat. Der Grund für diese Reihung war, dass beim Bildausschnitt von Rang 6 die Kanten deutlicher und die Spitze des Stiels besser erkennbar waren. Die Unschärfe vom Bildausschnitt von Rang 4 war hingegen sehr störend und nicht annähernd mit dem Original vergleichbar.

Die Übereinstimmung unter den Probanden lag bei diesem Beispiel bei 0,94. Die Korrelationskoeffizienten zwischen der Mittelperson und dem SAD sowie MAD liegt bei 0,71. Die Korrelation zum MSE liegt erneut etwas höher bei 0,86.

### ***Bild: Elefantenrüssel***

Bei diesem Bildausschnitt wird der Rüssel eines Stoffelefanten gezeigt. Innerhalb des Rüssels treten Fehlerausprägungen wie Transparenz, Löcher und Unschärfe auf.

Die Bewertung des Mittelprobanden ist in Abbildung 27 ersichtlich. Die Rangfolge der Fehler beginnt bei leichter Unschärfe, geht über in Transparenzen und endet bei massiven Löchern. Beim ersten Bildausschnitt waren sich alle Teilnehmer einig. Doch die zunehmende Unschärfe haben nicht alle Probanden hingenommen und stattdessen den Bildausschnitt von Rang 3 vorgezogen. Eine Gruppe von fünf Personen haben den Bildausschnitt von Rang 4 besser bewertet als den Bildausschnitt von Rang 3, da die hellen Stellen besser aussehen als kleine eindeutige Löcher. Eine weitere Gruppe von fünf Personen hat den Bildausschnitt von Rang 4 schlechter bewertet, da dieser unrealistisch, fremd und ausgeschnitten aussieht. Die Struktur ist bei diesem Bildausschnitt kaum mehr erkennbar und so sind die Bildausschnitte von Rang 5 und 6 besser, auch wenn Löcher enthalten sind.

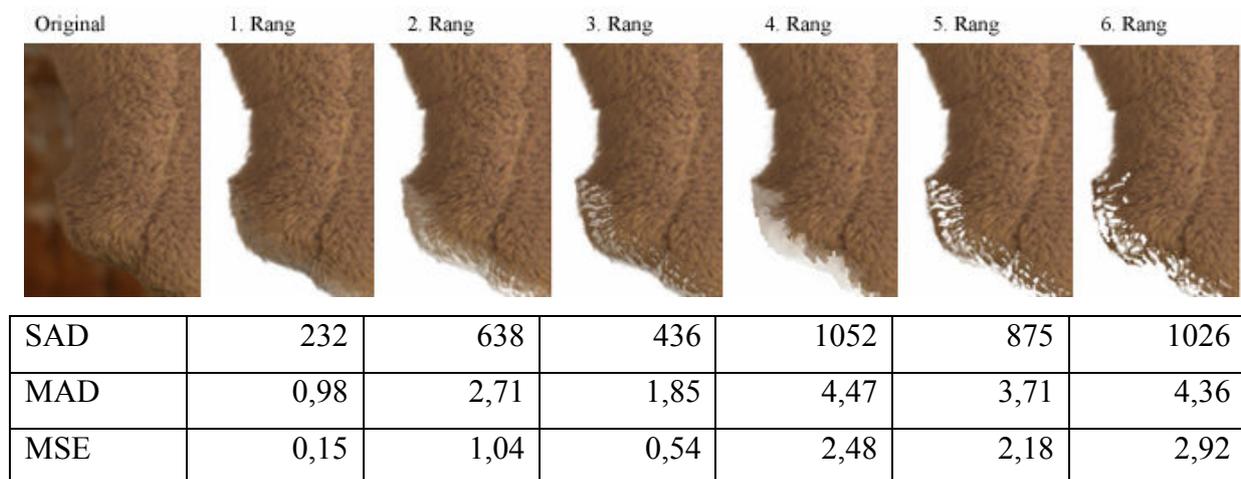


Abbildung 27: Rangfolge des Mittelprobanden vom Bild Elefantenrüssel sowie die zugehörigen berechneten Werte

Der Korrelationskoeffizient zwischen SAD sowie MAD und der Mittelperson liegt bei 0,6. MSE liegt bei 0,73. Die drei letzten Bilder zeigen, dass der Gradient entscheidender ist als beispielsweise der SAD-Wert. Der SAD-Fehler beim Bildausschnitt von Rang 4 weist den höchsten Wert auf, jedoch wurde dies nicht so von den Probanden gesehen, da eindeutige Löcher, die auf einen hohen Gradienten hinweisen würden, störender sind.

### ***Bild: Haarbüschel***

Das Bild zeigt Haare, die von einer Stoffgiraffe stammen. Bei der Reproduzierung des Vordergrundobjektes auf einen neuen Hintergrund treten Schatten, Verkürzungen sowie Reduzierungen auf. Die Fehlerwerte sowie die Rangfolge der Probanden sind in Abbildung 28 zu sehen.

Die Probanden haben bei diesem Beispiel annähernd einstimmig gewertet, sodass die Übereinstimmung bei 0,91 liegt. Trotzdem gibt es wenige Ausreißer. Bei den ersten beiden Bildausschnitten wurde der Fehler bzw. der Unterschied kaum erkannt, die verkürzten Haare wirken somit nicht als störend und sind kaum aufgefallen. Ein ähnliches Problem ist bei den Bildausschnitten von Rang 3 und 4 aufgetreten. Doch da der Bildausschnitt von Rang 3 über mehr Details verfügt, wurde dieser vom Großteil der Probanden auch auf diesen Platz gereiht. Lediglich zwei Personen hatten den Bildausschnitt von Rang 5 besser gereiht, da dieser Ausschnitt über mehr Details bzw. Informationen verfügt. Die Platzierung von Rang 6 wurde mit einer Ausnahme einstimmig entschieden.

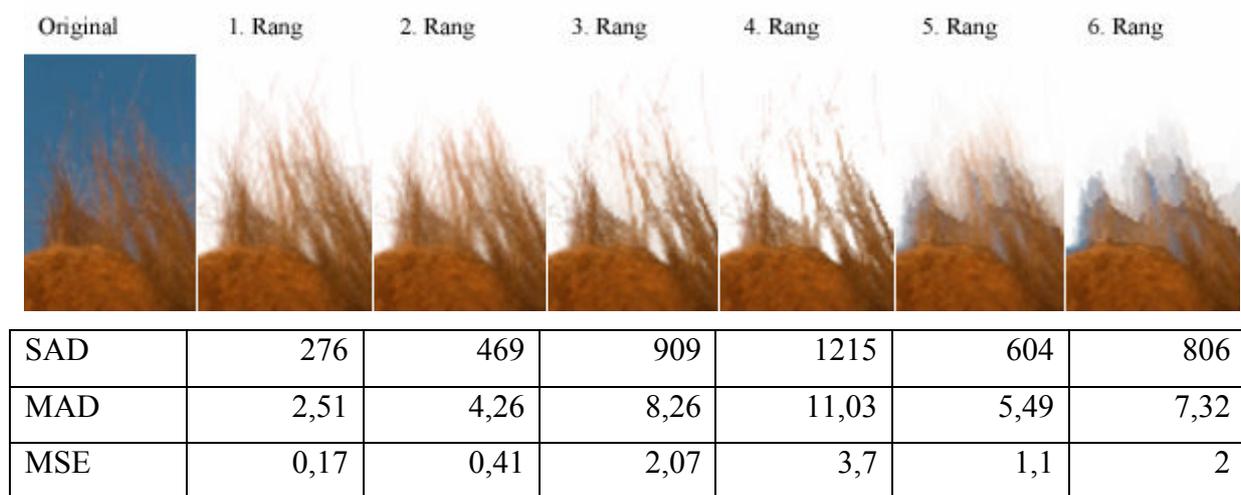
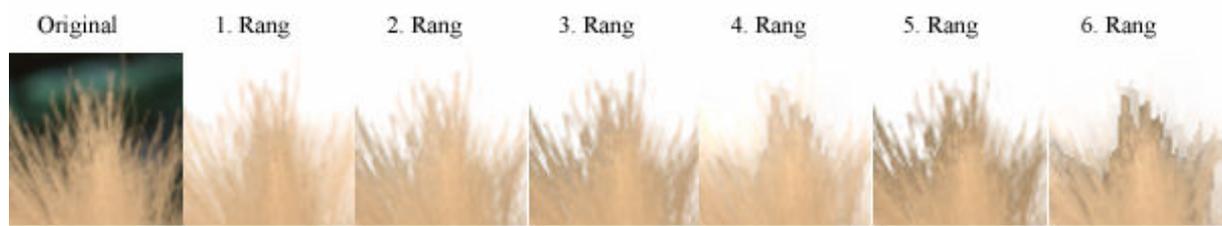


Abbildung 28: Rangfolge des Mittelprobanden vom Bild Haarbüschel sowie die zugehörigen berechneten Werte

Die Mittelperson stimmt mit 0,47 mit den automatisierten Berechnungen überein. Die Problematik der automatisierten Berechnungen entsteht hier insbesondere bei den beiden letzten Beispielen, da diese bessere Resultate erhielten als die Bildausschnitte von Rang 3 und 4. Jedoch weisen diese Bilder eindeutige Schatten auf, die von den Probanden abgelehnt wurden.

### ***Bild: Haare***

Dieses Beispiel zeigt die Problematik eines Haarbüschels, das zu einem Stoffschaf gehört. Die Haare weisen dabei harte Übergänge auf oder wirken weich gezeichnet. Zum Teil sind die Haare dabei noch recht gut erhalten, manchmal verschwinden sie oder werden an den Rändern dunkler, da die Segmentierung von Vorder- und Hintergrund nicht optimal funktioniert hat. Die Resultate sind in Abbildung 29 zu sehen.



SAD	541	436	651	356	866	706
MAD	7,32	5,9	8,8	4,82	11,71	9,55
MSE	1,24	0,74	1,66	0,52	2,94	2,08

Abbildung 29: Rangfolge des Mittelprobanden vom Bild Haare sowie die zugehörigen berechneten Werte

Bei diesem Bildausschnitt gab es keine Einigkeit sowie unterschiedliche Motivationen unter den Probanden. Für einige Probanden waren weiche Übergänge bei den Haaren sehr wichtig, wodurch die Bildausschnitte von Rang 1, 2 und 4 nach vorne gereiht wurden. Auch Details und Struktur standen für einige Personen im Vordergrund, wodurch der Bildausschnitt von Rang 3 auf den ersten Platz gereiht wurde. Einigkeit gab es lediglich bei der Bewertung des letzten Platzes, ausgenommen sind davon zwei Personen, die den 5. und 6. Rang vertauscht haben.

Wie bereits erwähnt, ist der SAD-Wert nicht ausreichend, um das Aussehen von Bildern zu beurteilen. Dies verdeutlicht das Bildpaar von Abbildung 29. Bei näherer Betrachtung des Bildpaares von Rang 3 und 4 ist gleich ersichtlich, dass der SAD-Wert des Bildes von Rang 3 höher ist als vom 4. Rang. Jedoch fällt die subjektive Wertung umgekehrt aus.

Bei weiteren Bildausschnitten mit Haaren ist aufgefallen, dass weiche Übergänge bevorzugt werden. Auch Farbwiedergabe und Details sind wichtige Entscheidungsträger. Bei all diesen Beispielen konnte jedoch keine komplette Übereinstimmung unter den Probanden festgestellt werden, da durch die Vielzahl an Details bei einem solchen Bild oft die Unterscheidung schwierig war, wodurch gefühlsmäßig bewertet wurde. Enorme Informationsverluste, harte Übergänge und Schatten im Hintergrund des Vordergrundobjektes wurden eher abgelehnt.

### ***Bild: Plastiksack 1***

Dieses Beispiel behandelt einen durchscheinenden Plastiksack und zeigt daher ein halbtransparentes Objekt, bei dem der Hintergrund zu sehen ist. In diesem Fall wird das Vordergrundobjekt vor blauem Hintergrund gezeigt, da vor dem weißen Hintergrund nicht alle Resultate gut genug sichtbar sind. In Abbildung 30 sind die Bildausschnitte und deren

Rangfolge zu sehen. Der reproduzierte Plastiksack weist stärkere und auch schwächere Kanten auf, zusätzlich ergeben sich Flecken, die zum Teil störend wirken oder kaum auffallen.

Bei diesem Bildausschnitt gab es unter den Probanden keine Einigkeit. Bei diesem und auch den folgenden Beispielen - die ebenfalls einen Ausschnitt des Plastiksackes zeigen – hatten die Probanden Probleme eine Entscheidung über die Wertung zu treffen. Die Resultate unterscheiden sich teilweise nur in Kleinigkeiten, wie beispielsweise der Farbintensität. Trotzdem haben alle Probanden eine Wertung abgegeben, doch nicht immer konnten sie eine zusätzliche Begründung abgeben, da die Beurteilung rein gefühlsmäßig war.

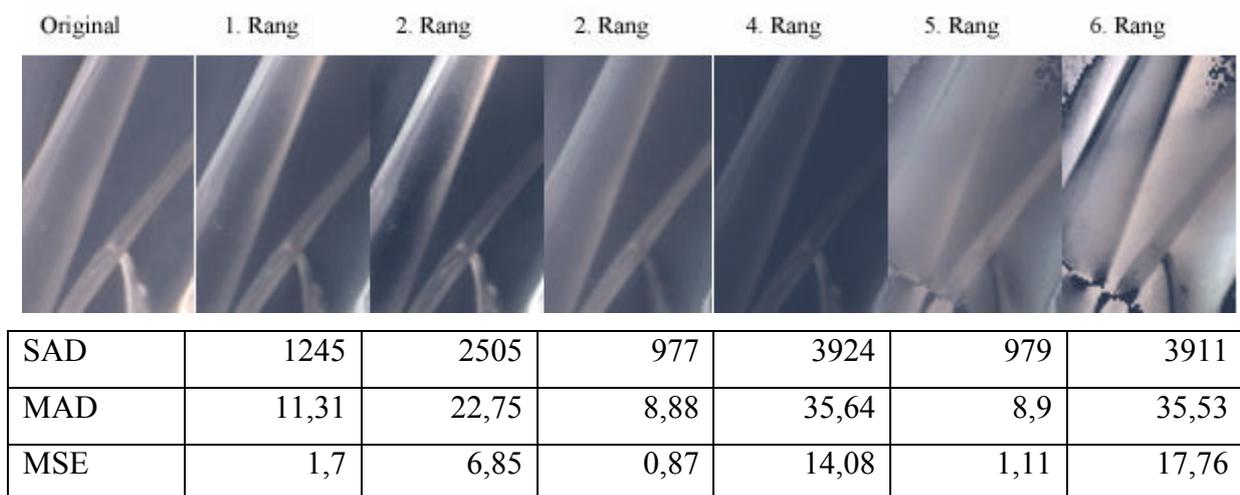


Abbildung 30: Rangfolge des Mittelprobanden vom Bild Plastiksack 1 sowie die zugehörigen berechneten Werte

Insbesondere die Wertung der Bildausschnitte von Rang 1 bis 3 ist einem Großteil der Probanden schwer gefallen, da sich diese sehr ähnlich sind. Eine Gruppe von sechs Probanden war der Meinung, dass der Bildausschnitt von Rang 2 das beste Resultat darstellt, obwohl dieser nicht dem Original entspricht. Entscheidend waren hier der Kontrast des Bildes sowie die klaren Kanten, wodurch der Plastiksack am besten als dieser erkennbar ist. Die schlechte Bewertung des Bildausschnitts von Rang 4 ergibt sich aufgrund der Blässe und des geringen Informationsgehalts des Bildes. Vier Personen haben dieses Resultat noch schlechter bewertet, da wenige Artefakte wie im Bildausschnitt von Rang 5 besser sind, da immerhin der Plastiksack als dieser noch erkennbar ist. Einigkeit gab es schließlich beim schlechtesten reproduzierten Vordergrundobjekt.

Zwischen den automatisierten Berechnungen wie dem SAD sowie MAD und dem Mittelprobanden gibt es ebenfalls keine Übereinstimmung, der Korrelationskoeffizient liegt lediglich bei 0,29. Zwischen Mittelperson und MSE gibt es eine Übereinstimmung von 0,41.

**Bild: Plastiksack 2**

Dieses Gradientenbeispiel zeigt erneut den Plastiksack. Auch hier sind Übergänge sowie Transparenzen zu erkennen. Die Resultate von diesem Bild – zu sehen in Abbildung 31 - werden ebenfalls auf blauem Hintergrund gezeigt, da wieder ein Vordergrundobjekt enthalten ist, das vor weißem Hintergrund kaum sichtbar wäre.

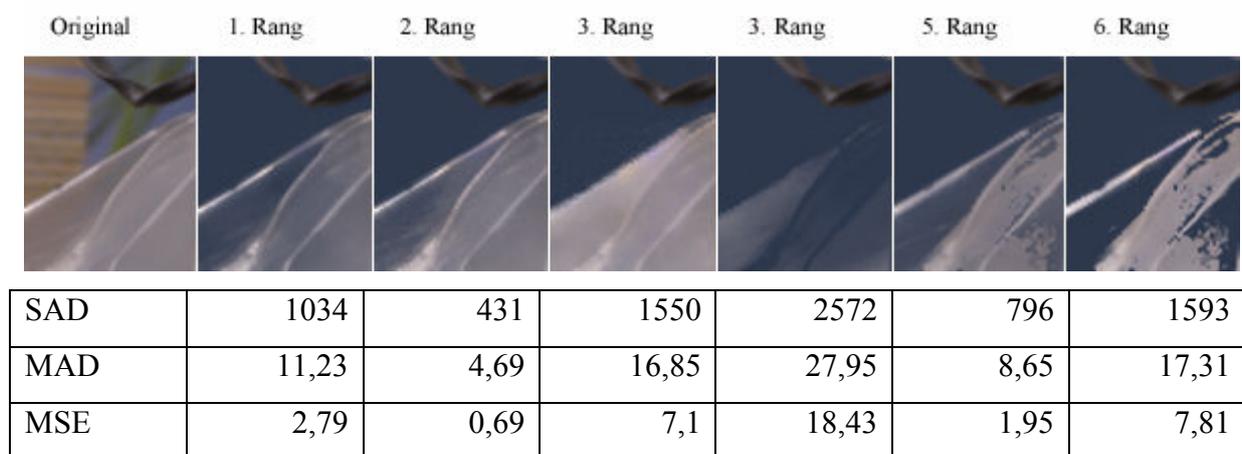


Abbildung 31: Rangfolge des Mittelprobanden vom Bild Plastiksack 2 sowie die zugehörigen berechneten Werte

Das Thema Plastiksack verursacht Entscheidungsprobleme bei den Probanden und das sollte auch bei diesem Beispiel nicht anders sein. Bei diesem Bildausschnitt ist keine Übereinstimmung unter den Probanden vorhanden. Die ersten beiden Bilder weisen zu wenige Unterschiede auf, während die anderen Bilder nur schwerwiegende Fehler haben. Der Bildausschnitt von Rang 3 war zu wenig durchscheinend, doch für einen Probanden war es das beste Resultat aufgrund der Struktur. Insbesondere zwei Probanden haben nach Transparenz entschieden, da ein Plastiksack durchscheinend sein sollte. Das schlechteste Resultat wurde nicht von allen als solches bewertet, denn für eine Probandin hat der Sack das Äußere eines nassen Plastiksacks angenommen, wodurch das Objekt ein sehr realistisches Aussehen erhält.

Bei den automatisierten Berechnungen sowie der Mittelperson gibt es nur eine Übereinstimmung von 0,29.

Auch bei einem weiteren Beispiel mit einem Plastiksack waren die Entscheidungsgründe die Farbtreue, leichte Transparenz sowie keine zu auffälligen Fehler. Mit zu auffälligen Fehlern sind Artefakte gemeint, die offensichtlich nicht zu dem Objekt gehören. Farbtreue war ebenfalls ein wichtiges Kriterium, da nur ein farbechtes Beispiel als gut bewertet werden kann. Die Transparenz stellte bei all diesen Beispielen eine Erschwerung der Entscheidung dar, denn ein Plastiksack muss durchscheinend sein. Die Meinungen darüber, wie viel Transparenz noch natürlich ist, gehen jedoch auseinander. Weiters wurden die Konturen des Plastiksacks in die Bewertung miteinbezogen, denn nur wenn diese eindeutig und klar erkennbar waren, kann das Objekt einen guten Rang erhalten.

### 7.1.3. Kategorie: Vordergrund-Hintergrund-Problem

Diese Kategorie ist durch die Vorstudie motiviert, bei der zu erkennen war, dass Fehler im Hintergrund immer als störend empfunden werden. Infolgedessen wird der Hinter- mit dem Vordergrund verglichen, um zu untersuchen, ob Fehler im Hintergrund tatsächlich störender wirken als Fehler im Vordergrund. Im Vordergrund können sich Artefakte oft besser tarnen, da sie eventuell wie Lichteffekte erscheinen, hingegen stechen Artefakte im Hintergrund immer hervor. Bei dieser Kategorie gibt es jeweils zwei Bildausschnitte, bei denen ein Fehler im Vorder- und einer im Hintergrund auftritt und dabei den gleichen SAD-Wert aufweisen.

#### ***Bild: Fell eines Stoffelefanten***

Beim ersten Beispiel des Vorder- und Hintergrundproblems wird ein Fellstück eines Stofftieres gezeigt. Wie bereits erwähnt, gibt es immer ein Vorder- und Hintergrundpaar, das den gleichen SAD-Fehler aufweist, jedoch müssen die Fehler weder die gleiche Struktur noch Form oder Farbe haben. Die Artefakte an sich können wie bei den vorigen Beispielen die unterschiedlichsten Ausprägungen annehmen.

Die Reihung der Probanden ist in Abbildung 32 zu sehen und verdeutlicht, dass die Bildausschnitte mit den Artefakten im Vordergrund besser bewertet wurden. Die Artefakte bzw. Flecken im Hintergrund haben bei diesem Beispiel sehr störend gewirkt und wurden somit ans Ende gereiht. Die Übereinstimmung unter den Probanden liegt bei 0,86 und ist damit sehr hoch. Ein Proband hat allerdings eine komplett andere Reihung abgegeben und zwar hat er die Ränge 2 und 3 mit 4 und 5 getauscht, da für ihn das Objekt und dessen Struktur im Vordergrund stehen. Aufgrund dieser Sichtweise sollte das Vordergrundobjekt keine Artefakte beinhalten und die Störungen im Hintergrund werden dabei ignoriert. Eine

weitere Probandin hat den Bildausschnitt von Rang 3 an den 5. Platz gestellt, da die weißen Punkte sowie die weißen Schattierungen zu unnatürlich auf sie gewirkt haben. Ein anderer Proband hat stattdessen den 2. und 3. Rang vertauscht, da der Bildausschnitt von Rang 3 für ihn schärfer und kontrastreicher gewirkt hat.

Original	1. Rang	2. Rang	3. Rang	4. Rang	5. Rang	6. Rang
						
SAD	112 (VG)	190 (VG)	114 (VG)	190 (HG)	112 (HG)	114 (HG)

Abbildung 32: Rangfolge des Mittelprobanden vom Bild Fell sowie die zugehörigen berechneten Werte, VG steht für Artefakte im Vordergrund und HG für Artefakte im Hintergrund

### ***Bild: Jacke eines Stofftieres***

Bei diesem Bildausschnitt ist das Ende einer Stoffjacke zu sehen, ersichtlich in Abbildung 33. Die Stoffjacke gehört zu einem kleinen Bären. Zum Teil weisen die Resultate leichte Schattierungen auf und zum anderen sind bei der Segmentierung Artefakte aus dem Hintergrund übernommen worden. Die Segmentierung hat hier nicht optimal funktioniert, da sich der Hintergrund nur geringfügig vom Vordergrund unterscheidet.

Original	1. Rang	2. Rang	3. Rang	4. Rang	5. Rang	6. Rang
						
SAD	53 (VG)	53 (HG)	112 (VG)	88 (HG)	88 (VG)	112 (HG)

Abbildung 33: Rangfolge des Mittelprobanden vom Bild Jacke sowie die zugehörigen berechneten Werte, VG steht für Artefakte im Vordergrund und HG für Artefakte im Hintergrund

Bei diesen Bildausschnitten gab es keine eindeutige Übereinstimmung unter den Probanden wie beim vorigen Beispiel. Eine Gruppe von sechs Personen hat sich erneut für den Vordergrund ausgesprochen. Der Rest hat hingegen entsprechend der Fehlerintensität bewertet. Das bedeutet, dass der Vordergrund meist bevorzugt wurde, außer wenn die Fehler zu schwerwiegend wurden, wie beim Resultat mit Rangnummer 5. Ein Proband hat wieder die Artefakte im Hintergrund besser bewertet, da die Struktur und Kontur des

Vordergrundobjektes für das Aussehen eines Bildes wichtiger sind. Ein weiterer Proband hat das Artefakt des Bildausschnitts von Rang 6 als zugehörig eingestuft und deshalb auf Platz 2 gereiht. Die restliche Bewertung erfolgte nach dem Ausmaß der Artefakte.

Bei diesem Beispiel kann die Annahme über das Vorder- und Hintergrundproblem nicht eindeutig bestätigt werden, da die Teilnehmer eher nach der Fehlerintensität entschieden haben.

**Bild: Blume**

Hierbei handelt es sich um die Außenseite einer Blume, die wiederum leichte sowie stärkere Artefakte aufweist. Bei der Segmentierung sind Teile des Hintergrundes übernommen worden, da der Hinter- und Vordergrund in einer kleinen Region die gleiche Farbe haben. Erneut gibt es jeweils ein Vorder- und Hintergrundpaar, das allerdings unterschiedliche Fehler aufweist, zu sehen ist dies in Abbildung 34.

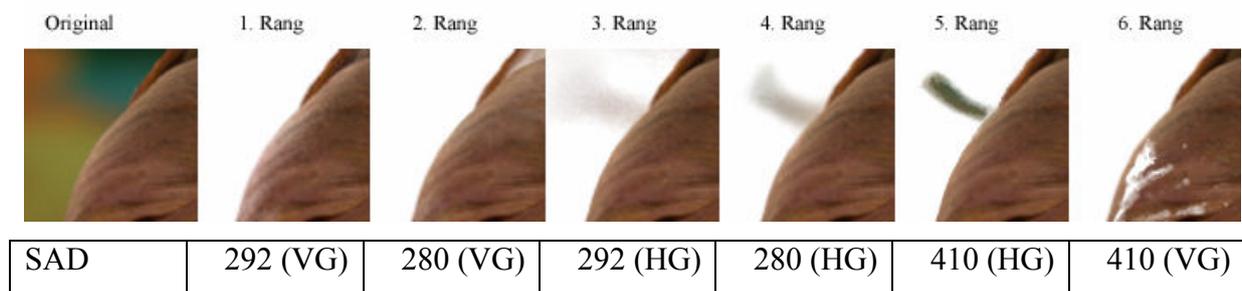


Abbildung 34: Rangfolge des Mittelprobanden vom Bild Blume sowie die zugehörigen berechneten Werte, VG steht für Artefakte im Vordergrund und HG für Artefakte im Hintergrund

Bei diesem Bildausschnitt gibt es keine einstimmige Meinung unter den Probanden, jedoch wurden die Artefakte im Vordergrund weniger bemängelt. Die kleinen Farbveränderungen der ersten beiden Bildausschnitte sind sehr unauffällig und treten daher in den Hintergrund. Ausgenommen davon wird der Bildausschnitt, bei dem das Vordergrundobjekt Löcher aufweist. Diese sind zu schwerwiegend, als dass sie übersehen werden könnten. Somit ist die schlechte Bewertung auf die Auffälligkeit der Löcher zurückzuführen. Eine Gruppe von drei Personen hat den Bildausschnitt von Rang 6 nach vor gereiht, da die Fehler im Hintergrund jedenfalls störender empfunden wurden als die Löcher im Vordergrund. Eine Probandin hat das Bild von Rang 5 auf den 3. Platz gesetzt, da dieses Artefakt zur Blume gehören könnte, für die Schatten hingegen gibt es keine natürliche Erklärung.

Bei diesem Bildausschnitt ist die bevorzugte Wertung des Vordergrundes nicht mehr eindeutig. Der Fehler im Vordergrundobjekt von Rang 6 ist zu offensichtlich und unnatürlich, wodurch dieser Bildausschnitt von der Mehrheit der Probanden abgelehnt wurde. Bei den weniger intensiven Fehlern in der Vorder- oder Hintergrundregion konnte die Annahme bestätigt werden, dass die Fehler im Vordergrundobjekt weniger aufdringlich sind.

**Bild: Hut**

Dieser Bildausschnitt zeigt den Hut einer Figur. Dabei sind auf der Hutoberfläche Schattierungen und kleine Flecken zu erkennen. Im Hintergrund treten Artefakte in Form eines Streifens sowie eines intensiven Flecks auf.

Bei der Betrachtung der Rangfolge des Mittelprobanden in Abbildung 35 ist deutlich ersichtlich, dass die Fehler im Vordergrundobjekt als weniger störend gelten. Drei Personen schätzten den Bildausschnitt des 3. Rangs schlechter ein, da die kleinen intensiven Flecken zu auffällig sind. Zwei weitere Personen reichten diesen Bildausschnitt nach vorne, da das Vordergrundobjekt dem Original aufgrund der Farbtreue ähnlicher sei. Ein Proband reichte die Bildausschnitte mit den Hintergrundartefakten nach vorne, da ihm das korrekte Objekt wichtiger war. Der Fleck im Hintergrund von Bildausschnitt mit Rangnummer 5 wurde sehr unterschiedlich bewertet. Für zehn Personen stellte der intensive Fleck eine geringere Störung dar als die Streifen der beiden anderen Bildausschnitte. Diese Wertung dürfte mit der Größe des Artefakts zusammenhängen. Der Rest der Probanden reiht den Bildausschnitt von Rang 5 jedoch an den 5. und 6. Platz.

Original	1. Rang	2. Rang	3. Rang	4. Rang	5. Rang	6. Rang
						
SAD	191 (VG)	287 (VG)	134 (VG)	134 (HG)	191 (HG)	287 (HG)

Abbildung 35: Rangfolge des Mittelprobanden vom Bild Hut sowie die zugehörigen berechneten Werte, VG steht für Artefakte im Vordergrund und HG für Artefakte im Hintergrund

Die Artefakte im Vordergrund sind wieder weniger störend als die Artefakte im Hintergrund. Lediglich eine Person hat ein völlige andere Reihung und somit Meinung abgegeben.

### **Bild: Kürbis**

Dieser Bildausschnitt beinhaltet einen Kürbis, zu sehen ist dies in Abbildung 36. Die Oberfläche des Kürbis weist leichte Schattierungen sowie kleine Löcher auf. Im Hintergrund sind Schatten sowie ein Punktnebel zu erkennen.

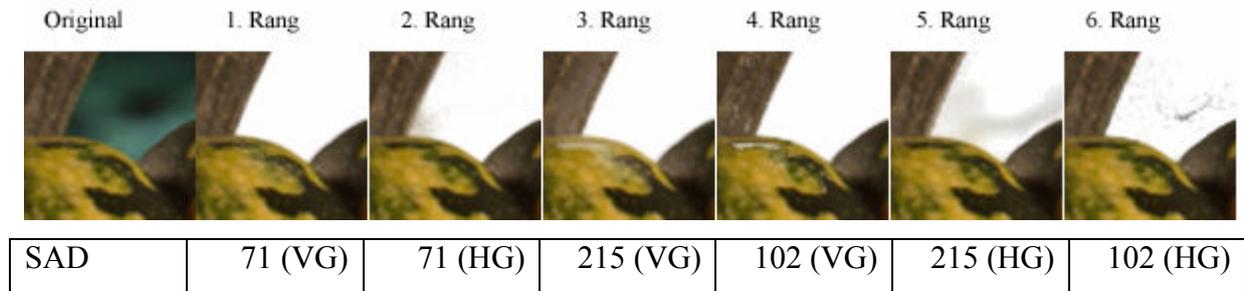


Abbildung 36: Rangfolge des Mittelprobanden vom Bild Kürbis sowie die zugehörigen berechneten Werte, VG steht für Artefakte im Vordergrund und HG für Artefakte im Hintergrund

Die Reihung des Mittelprobanden ist in Abbildung 36 ersichtlich. Der Bildausschnitt vom 1. Rang wurde vom Großteil der Probanden an diese Stelle gewählt. Die weitere Reihung von Rang 2 bis 4 wurde nicht einstimmig gewählt. Eine Gruppe von Probanden hat die kleinen Schatten des Bildausschnitts von Rang 2 kaum gestört, während die anderen diesen Schatten bereits als unnatürlich abgewiesen haben. Eine weitere Problematik stellte sich bei den Bildausschnitten von Rang 3 und 4, da die Kürbisoberfläche Flecken bzw. Schattierungen aufweist. Meist wurden die Schattierungen besser bewertet, da es sich um Lichtreflexionen handeln könnte, die Löcher hingegen sehen unnatürlich aus. Bei den letzten beiden Beispielen gab es erneut unterschiedliche Meinungen. Zum einen waren die Schatten weniger störend als die einzelnen Pixelfehler. Zum anderen war der Schatten im Hintergrund zu großflächig, wodurch dieser abgelehnt wurde.

Bei den vorigen Beschreibungen ist zu erkennen, dass wiederum die Fehler im Vordergrund besser bewertet wurden. Doch bei diesem Bildausschnitt hat sich eine weitere kleine Gruppe herauskristallisiert, diese hat nämlich die Rangfolge umgedreht. Das bedeutet, dass die Bildausschnitte mit den Hintergrundartefakten besser bewertet wurden. Zum einen liegt dies daran, dass der Struktur der Kürbisoberfläche ein höherer Stellenwert zugeschrieben wurde und zum anderen haben die Fehler am Kürbis für sie unnatürlich ausgesehen.

Bei diesem Beispiel wird verdeutlicht, dass die Bildausschnitte mit den Vordergrundartefakten nicht immer besser bewertet werden, sondern dass die Artefakte an sich eine wichtige Rolle bei der Beurteilung spielen und nicht alleine der Standort.

**Bild: Pullover**

Beim diesem Bild wird ein Pullover eines Stoffbären gezeigt. Hier soll die Problematik einer klein gemusterten Oberfläche näher betrachtet werden. Erneut gibt es Artefakte sowohl im Vorder- als auch im Hintergrund, doch da es sich um einen gestrickten Pullover handelt, könnten die Artefakte im Hintergrund eventuell realistische Erklärungen finden. Die Resultate sind in Abbildung 37 ersichtlich.

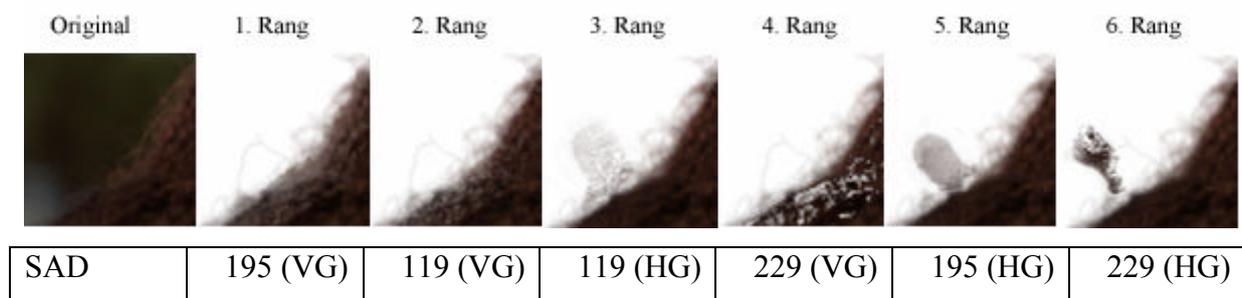


Abbildung 37: Rangfolge des Mittelprobanden vom Bild Pullover sowie die zugehörigen berechneten Werte, VG steht für Artefakte im Vordergrund und HG für Artefakte im Hintergrund

Die Rangfolge des Mittelprobanden zeigt eine Bevorzugung der Artefakte im Vordergrund. Die ersten beiden Bildausschnitte wurden meist auch an diese Stelle gewählt, jedoch von einigen Probanden in umgekehrter Reihenfolge, da der Bildausschnitt von Rang 1 eine leichte Unschärfe aufweist. Beim dritten Bildausschnitt mit Vordergrundartefakten ist die Reihung jedoch nicht so eindeutig, denn einige Probanden haben diesen an den dritten und viele auch an den letzten Platz gesetzt, da die Fehler und somit der Informationsverlust zu groß war. Die Bildausschnitte mit den Hintergrundartefakten wurden eher schlecht bewertet, da die Fehler eindeutige Störungen sind und nicht zu dem Pullover hinzugehören. Ausgenommen sind zwei Personen, diese haben den Bildausschnitt von Rang 6 besser bewertet, da dieses zusätzliche Artefakt sehr wohl eine Schlaufe des Pullovers sein könnte. Der Rest der Probanden hat jedoch keine Erklärung zu den Hintergrundartefakten gefunden, wodurch diese Bilder auch die schlechte Wertung erhalten haben.

Bei diesem Bildausschnitt haben einige Probanden nach der Intensität des Fehlers bzw. Artefakts geurteilt und nicht nach dem Standort des Fehlers.

### **Bild: Haarspange**

Das letzte Beispiel zeigt eine Haarspange eines Puppenkopfs, zu sehen ist dies in Abbildung 38. Dabei weist die Haarspange kleine Löcher auf und weiters treten kleine Artefakte in den einzelnen herumfliegenden Haaren auf. Die Artefakte im Hintergrund stammen von der bunten Tapete im Originalbild.

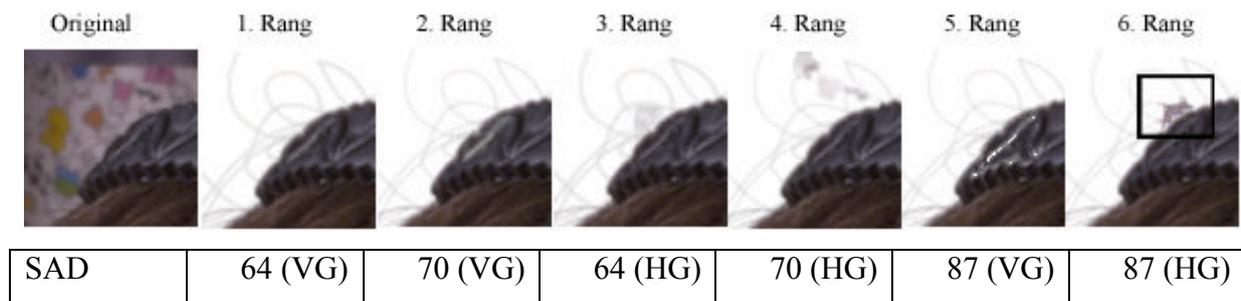


Abbildung 38: Rangfolge des Mittelprobanden vom Bild Haarspange sowie die zugehörigen berechneten Werte, VG steht für Artefakte im Vordergrund und HG für Artefakte im Hintergrund

Auch beim letzten Bild des Vorder- und Hintergrundproblems gibt es keine Einigkeit unter den Probanden. Der Mittelproband zeigt eine Bevorzugung der Vordergrundartefakte. Die kleinen Löcher des Bildausschnitts mit Rangnummer 5 wurden zum Teil an den letzten Platz gesetzt, da diese zu störend wirkten und somit nicht angenommen wurden. Eine kleine Gruppe von drei Personen hat den Bildausschnitten mit den Hintergrundartefakten den Vorzug gegeben, da diese nicht sehr schwerwiegend sind. Die Fehler auf der Haarspange hingegen wirkten für viele Probanden als sehr auffällig und nicht der Realität entsprechend. Doch der Großteil der Probanden hat auch hier die Bildausschnitte mit den Vordergrundartefakten besser bewertet. Beispielsweise wurde das Artefakt vom Bildausschnitt mit Rangnummer 6 in der eingerahmten Region als "Maus" bezeichnet und dies ist in der Realität wohl wirklich nicht möglich.

Wenn weitere Beispiele mit ähnlichen Fehlerausprägungen näher betrachtet werden, dann ist klar ersichtlich, dass nicht alleine der Standort sondern die Fehlerintensität entscheidend ist. Dazu muss erwähnt werden, dass kleinere Fehler in Vordergrundobjekten tatsächlich oft weniger auffällig sind als im Hintergrund, wodurch diese eine bessere Wertung erfahren. Doch wenn Artefakte in Vordergrundobjekten zu großflächig oder intensiv werden, ist die Entscheidung nicht mehr offensichtlich. Im Gegenteil, solche Fehler können noch auffälliger erscheinen. Bei etlichen Beispielen war es möglich, für die Hintergrundartefakte Erklärungen

für eine Zugehörigkeit zu finden, dadurch haben diese Bildausschnitte ebenfalls eine bessere Wertung bekommen. Für wenige Probanden war immer das Objekt im Mittelpunkt des Interesses, wodurch die Fehler im Hintergrund immer ignoriert wurden, da alleine die Struktur eines Objektes gezählt hat.

## 7.2. Schlussfolgerungen der Studie

Zuerst ist zu erwähnen, dass die Studie im Gesamten betrachtet sehr erfolgreich abgelaufen ist. Die Probanden haben äußerst wertvolle Informationen geliefert. Die Studie hat auch gezeigt, dass die menschliche Wahrnehmung subjektiv ist und nicht so leicht vorausgesagt werden kann. Trotzdem wurden bestimmte im Vorfeld getroffene Annahmen bestätigt oder auch widerlegt.

Die Verbundenheit von Objekten ist für die menschliche visuelle Wahrnehmung wesentlich, eine sehr große Anzahl an Probanden hat dieses Kriterium als die wichtigste Eigenschaft eines Objektes eingestuft. Die Übereinstimmung der Probanden war hier sehr hoch. Zudem hat sich bestätigt, dass die bisher verwendeten Fehlermetriken wie SAD oder MSE nicht mit der menschlichen Wahrnehmung in der Präsenz von unverbundenen Stellen korrespondieren. In anderen Worten kann der SAD Unverbundenheit nicht richtig erklären. Dies spiegelt sich auch in der Korrelation zwischen den automatisierten Berechnungen und dem Mittelprobanden wieder, da der SAD sowie MAD bei 0,27 und der MSE bei 0,33 liegt. Der SAD sowie MAD weisen immer die gleiche Rangfolge und somit auch den gleichen Korrelationskoeffizienten auf, wodurch in weiterer Folge nur mehr vom SAD gesprochen wird.

Bei einem Großteil der Bildausschnitte wurde die getroffene Annahme über den Gradientenverlauf bestätigt. Die Übereinstimmung der Probanden war hier erneut sehr hoch. Ein hoher Gradient in der Alpha Matte weist oft auf störende Flecken hin, die auch von den Probanden ablehnt wurden. Ein niedriger Gradient hingegen weist auf einen Verlauf im Vordergrundobjekt hin, der als weniger störend betrachtet wird. Dieser Verlauf verursacht häufig Unschärfe im Objekt, die jedoch visuell weniger störend wirkt als auffällige Flecken bzw. Löcher. Auch hier konnte festgestellt werden, dass die automatisierten Berechnungen wie der SAD oder MSE nicht mit der visuellen Wahrnehmung übereinstimmen. Der Korrelationskoeffizient zwischen Mittelperson und SAD liegt bei 0,44 und zu MSE bei 0,50.

Die Kategorie des Vorder- und Hintergrundproblems hat keine Einigkeit unter den Probanden hervorgebracht. Stattdessen haben sich zwei Benutzergruppen bzw. zwei Richtungen herauskristallisiert. Eine dieser Gruppen interpretiert die Artefakte im Vordergrund um, beispielsweise als Lichtreflexionen und würde so die vordefinierte Annahme bestätigen. Eine andere Gruppe jedoch wertet nach der Struktur des Vordergrundobjektes, Störungen im Hintergrund werden als gegeben angenommen und damit ignoriert. Es kann somit zusammengefasst werden, dass die ursprüngliche Annahme zu dieser Kategorie nicht bestätigt werden kann.

Daten	SAD	MSE
Verbundenheit	0,27	0,33
Gradient	0,44	0,50
Vorder- und Hintergrund	0,49	0,55

Abbildung 39: Tabelle mit Übereinstimmungsmaßen

Zusammenfassend kann gesagt werden, dass zu allen Bildausschnitten die Übereinstimmungsberechnungen wie die Summe der absoluten Differenz sowie der mittlere quadratische Fehler keine hohe Korrelation zur Mittelperson haben, wie in der Tabelle von Abbildung 39 ersichtlich ist. Deshalb eignen sich diese Bewertungsmaße nicht für die Feststellung der visuellen Qualität in der Präsenz von Unverbundenheit oder Störungen des Gradientenverlaufs.

## 8. Zusammenfassung

Image Matting beschreibt ein Bild als eine Kombination aus Vorder- und Hintergrund. Das Mischungsverhältnis zwischen Vorder- und Hintergrund wird dabei durch eine Alpha Matte definiert. Durch Berechnungen dieser Alpha Matte kann in weiterer Folge das Vordergrundobjekt aus einem beliebigen Hintergrund herausgelöst und auf einem neuen wieder eingesetzt werden. Um die Qualität von verschiedenen Matting Algorithmen quantitativ vergleichen zu können ist eine Referenzlösung erforderlich. Diese kann mit Triangulierung berechnet werden.

Diese Arbeit beschäftigt sich mit der Frage, inwieweit einfache pixelbasierte Fehlermetriken die visuelle Qualität von Matting Resultaten beschreiben können. Dazu wurde eine Benutzerstudie durchgeführt. Davor war zum einen das Wissen über das menschliche visuelle System sowie zum anderen das Aufzeigen von objektiven Evaluierungsmöglichkeiten notwendig. Das Verständnis über die menschliche Wahrnehmung ist bedeutsam, da sonst nicht verstanden werden kann, warum bestimmte Bilder von besserer visueller Qualität sind als andere. Aufgrund der Wichtigkeit der visuellen Qualität eines Bildes sind bereits eine Vielzahl automatisierter Evaluierungsvarianten entwickelt worden, einige davon wurden im Rahmen dieser Arbeit kurz erläutert.

Der Ablauf der Studie teilt sich in die subjektive und objektive Wertung. Dafür wurden zuvor drei Bildkategorien erstellt gemäß der Verbundenheit, des Gradientenverlaufs und des Vorder- sowie Hintergrunds. Bei der Verbundenheit sollte überprüft werden, ob unverbundene Objekte tatsächlich schlechter bewertet werden, da diese nicht der Realität entsprechen. Beim Gradientenverlauf ging es darum, ob eindeutige Löcher bzw. Flecken in einem Bild als störender empfunden werden als wenn weiche Übergänge zu sehen sind. Flecken weisen dabei einen hohen Gradienten auf und die weichen Übergänge einen niedrigen Wert. Bei einer Vorstudie konnte bereits festgestellt werden, dass Hintergrundartefakte immer störend wirken. Deshalb war der Vergleich zwischen Vorder- und Hintergrundartefakten von Interesse und es sollte festgestellt werden, ob Artefakte im Hintergrund störender wirken als solche im Vordergrund.

Für die subjektive Bewertung wurden 17 Personen eingeladen. Dreißig Bildausschnitte wurden von diesen beurteilt und nach visueller Qualität gereiht. Die Bedingungen waren für alle Kandidaten gleich. Für die objektive Bewertung wurden Berechnungen wie die Summe der absoluten Differenz, der Mittelwert der absoluten Differenz sowie der mittlere quadratische Fehler erstellt.

Bei der Auswertung konnte nun festgestellt werden, dass die menschliche Wahrnehmung bestimmten Kriterien folgt, aber natürlich auch subjektiv ist. Die Annahme über die Verbundenheit sowie den Gradientenverlauf konnte bestätigt werden. Unverbundene Objekte wirken unrealistisch und erhielten dadurch keine gute Wertung. Die pixelbasierten automatisierten Bewertungsmaße hingegen konnten keine gute Übereinstimmung mit den Probanden erzielen.

Die Aussage, dass Artefakte im Vordergrund weniger störend sind, konnte in dieser Studie nicht bestätigt werden. Wobei erwähnt werden soll, dass es hier zwei Meinungsgruppen von Probanden gab. Eine Gruppe hätte die Annahme bestätigt. Der zweiten Gruppe hingegen war nur die Struktur des Vordergrundobjekts wichtig, wodurch die Artefakte im Hintergrund hingenommen wurden.

Ein Thema für zukünftige Arbeiten wäre die Weiterentwicklung von automatisierten HVS-basierten Systemen. Diese Arbeit hat verdeutlicht, dass pixelbasierte Bewertungsmaße nicht ausreichend sind für die Feststellung der visuellen Qualität von Bildern. Deshalb sind automatisierte Qualitätsberechnungen notwendig, die vermehrt auf die menschliche visuelle Wahrnehmung eingehen. Da der Mensch unbewusst viele Kriterien in seine Beurteilung mit einbezieht, sollte sich ein neues Bewertungsmaß ebenso wenig auf ein Kriterium beschränken und viel mehr unterschiedlichste Eigenschaften eines Bildes, wie beispielsweise die Struktur und Verbundenheit, miteinander verbinden.

## 9. Literaturverzeichnis

[A06] Ivan D. Artamonow, *Optische Täuschungen*, Verlag Haari Deutsch, 2006

[ASS02] Ismail Avcibas, Bülent Sankur and Khalid Sayood: Statistical Evaluation of Image Quality Measures, *Journal of Electronic Imaging* 11(2), pp. 206-223, 2002

[CCSS01] Yung-Yu Chuang, Brian Curless, David H. Salesin and Richard Szeliski: A Bayesian Approach to Digital Matting. *IEEE Computer Vision and Pattern Recognition*, Vol. II, 264-271, 2001

[CYX06] Guan-Hao Chen, Chun-Ling Yang and Sheng-Li Xie: Gradient-Based Structural Similarity for Image Quality Assessment, Image Processing, *IEEE International Conference on Image Processing*, pp. 2929-2932, 2006

[G02] E. Bruce Goldstein, *Wahrnehmungspsychologie*, Spektrum Akademischer Verlag, Heidelberg, Berlin. Zweite deutsche Auflage herausgegeben von Manfred Ritter, 2002

[GCLDP06] Yu Guan, Wei Cheny, Xiao Liang, Zi'ang Ding and Qunsheng Peng: Easy Matting: A Stroke Based Approach for Continuous Image Matting. In *Eurographics*, Vol. 25, pp.567-576, 2006

[GGL08] Arjan Gijsenij, Theo Gevers and Marcel P. Lucassen: A Perceptual Comparison of Distance Measures for Color Constancy Algorithms. *European Conference on Computer Vision*, Vol. 5302, pp. 208-221, 2008

[GGM04] Sonja Grgic, Mislav Grgic and Marta Mrak: Reliability of Objective Picture Quality Measures, *Journal of Electrical Engineering*, Vol. 55, No. 1-2, pp. 3-10, 2004

[GSAW05] Leo Grady, Thomas Schiwietz, Shmuel Aharon and Rüdiger Westermann: Random Walks for Interactive Alpha Matting, *Visualization, Imaging and Image Processing*, pp. 423 – 429, 2005

- [HE07] Joachim Hartung und Bärbel Elpelt: Multivariate Statistik: Lehr- und Handbuch der angewandten Statistik, Veröffentlicht von Oldenbourg Wissenschaftsverlag, 2007
- [JK05] Olivier Juan and Renaud Keriven: Trimap Segmentation for Fast and User-friendly Alpha Matting. In *Variational, Geometric, and Level Set Methods in Computer Vision*, pp. 186-197, 2005
- [LKS04] Haiping Lu, Alex C. Kot and Yun Q. Shi: Distance-Reciprocal Distortion Measure for Binary Document Images, *IEEE Signal Processing Letters*, Vol. 11, No. 2, 2004
- [LLW06] Anat Levin, Dani Lischinski and Yair Weiss: A Closed Form Solution to Natural Image Matting. *IEEE Computer Vision and Pattern Recognition*, pp. 61- 68, 2006
- [PV08] Bo Peng and Olga Veksler: Parameter Selection for Graph Cut Based Image Segmentation, *British Machine Vision Conference*, 10pp., 2008
- [R02] Austin Roorda: Human Visual System – Image Formation, *Encyclopedia of Imaging Science and Technology*, Vol. 1, pp. 539-557, 2002
- [RM03] Xiaofeng Ren and Jitendra Malik: Learning a Classification Model for Segmentation. *Proceedings of Ninth IEEE International Conference on Computer Vision*, Vol.1, pp. 10 – 17, 2003
- [RT00] Mark A. Ruzon and Carlo Tomasi: Alpha Estimation in Natural Images. In *CVPR 2000*, pp. 18-25, 2000
- [S02] Werner A. Stahel: Statistische Datenanalyse: Eine Einführung für Naturwissenschaftler, Veröffentlicht von Vieweg und Teubner Verlag, 4. Auflage, 2002
- [SB96] Alvy Ray Smith and Jim Blinn: Blue Screen Matting. *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, pp. 259-268, 1996
- [SB07] Kalpana Seshadrinathan and Alan C. Bovik: New Vistas in Image and Video Quality Assessment, *Human Vision and Electronic Imaging XII*, Vol. 6492, 2007

- [SPA07] Eric A. Silva, Karen Panetta and Sos S. Azaian: Quantifying Image Similarity Using Measure of Enhancement by Entropy, *Mobile Multimedia/Image Processing for Military and Security Applications*, Vol. 6579, 2007
- [SSXS08] Jian Sun (Xi'an Jiaotong University), Jian Sun (Microsoft Research Asia), Zongben Xu and Heung-Yeung Shum: Image Super-Resolution using Gradient Profile Prior, *Computer Vision and Pattern Recognition*, pp. 1-8, 2008
- [T05] Klaus D. Tönnies: Grundlagen der Bildverarbeitung, Verlag Pearson Studium, 2005
- [UPH07] Ranjith Unnikrishnan, Caroline Pantofaru and Martial Hebert: Toward Objective Evaluation of Image Segmentation Algorithms, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 29, No. 6, pp. 929-944, 2007
- [WC07] Jue Wang and Michael Cohen: Image and Video Matting: A Survey. In *Foundations and Trends in Computer Graphics and Vision*, Vol. 3, No. 2, 2007
- [WC07\_R] Jue Wang and Michael F. Cohen: Optimized Color Sampling for Robust Matting. *IEEE Computer Vision and Pattern Recognition*, pp. 1-8, 2007
- [WAC07] Jue Wang, Maneesh Agrawala and Michael F. Cohen: Soft Scissors: An Interactive Tool for Realtime High Quality Matting. *SIGGRAPH*, 6pp, 2007
- [WSB03] Zhou Wang, Hamid R. Sheikh and Alan C. Bovik: Objective Video Quality Assessment, Chapter 41 in *The Handbook of Video Databases: Design and Applications*, (B. Furht and O. Marqure, ed.), *CRC Press*, pp. 1041-1078, 2003

## 10. Abbildungsverzeichnis

Abbildung 1, S. 7: Bildausschnitt mit pixelbasierter Qualitätsmessung

Quelle: Im Rahmen der Diplomarbeit selbst erstellt.

Abbildung 2, S. 9: Image Matting

Quelle: Im Rahmen der Diplomarbeit selbst erstellt.

Abbildung 3, S. 11: Trimap links: Exakt spezifizierte Trimap; Trimap rechts: Grob spezifizierte Trimap

Quelle: Im Rahmen der Diplomarbeit selbst erstellt.

Abbildung 4, S. 14: Bayesche Verfahren

Quelle: Chuang, Curless, Salesin and Szeliski: A Bayesian Approach to Digital Matting [CCSS01]

Abbildung 5, S. 14: Bayesche Verfahren: Verteilungen der optimalen  $F$ ,  $B$  und  $\alpha$  Parameter

Quelle: Chuang, Curless, Salesin and Szeliski: A Bayesian Approach to Digital Matting [CCSS01]

Abbildung 6, S. 16: Vorder- und Hintergrundrekonstruktion

Quelle: Levin, Lischinski and Weiss: A Closed Form Solution to Natural Image Matting [LLW06]

Abbildung 7, S. 21: Schematisches Diagramm des menschlichen visuellen Systems

Quelle: Wang, Sheikh and Bovik: Objective Video Quality Assessment [WSB03]

Abbildung 8, S. 31: Schema des SSIM-Verfahrens

Quelle: Seshadrinathan and Bovik: New Vistas in Image and Video Quality Assessment [SB07]

Abbildung 9, S. 40: Vorstudie: Kategorie Vordergrund

Quelle: Im Rahmen der Diplomarbeit selbst erstellt.

Abbildung 10, S. 40: Vorstudie: Kategorie Hintergrund

Quelle: Im Rahmen der Diplomarbeit selbst erstellt.

Abbildung 11, S. 41: Vorstudie: Kategorie gemischte Region

Quelle: Im Rahmen der Diplomarbeit selbst erstellt.

Abbildung 12, S. 41: Vorstudie: Kategorie Verbundenheit

Quelle: Im Rahmen der Diplomarbeit selbst erstellt.

Abbildung 13, S. 49: Einige Beispiele für die Unverbundenheit

Quelle: Im Rahmen der Diplomarbeit selbst erstellt.

Abbildung 14, S. 50: Einige Beispiele für den Gradientenfehler

Quelle: Im Rahmen der Diplomarbeit selbst erstellt.

Abbildung 15, S. 50: Einige Beispiele für den Vorder- und Hintergrund

Quelle: Im Rahmen der Diplomarbeit selbst erstellt.

Abbildung 16, S. 51: Setup für die Beurteilung der Bildausschnitte

Quelle: Im Rahmen der Diplomarbeit selbst erstellt.

Abbildung 17, S. 52: Setup zur Beurteilung der Bildausschnitte mit dem Hinweis, dass die Resultate zeilenweise umgeordnet werden können

Quelle: Im Rahmen der Diplomarbeit selbst erstellt.

Die folgenden Abbildungen enthalten jeweils die Rangfolge des Mittelprobanden sowie die zugehörigen berechneten Werte und wurden im Rahmen der Diplomarbeit von der Verfasserin erstellt:

Abbildung 18, S. 56: Bild Haarsträhne

Abbildung 19, S. 57: Bild Lasche eines Hutes 1

Abbildung 20, S. 58: Bild Lasche eines Hutes 2

Abbildung 21, S. 59: Bild Haupt des Elefanten

Abbildung 22, S. 59: Bild Blumenstiel

Abbildung 23, S. 61: Bild Krallen eines Stofflöwen  
Abbildung 24, S. 61: Bild Ananas  
Abbildung 25, S. 63: Bild Kleid  
Abbildung 26, S. 64: Bild Stroh  
Abbildung 27, S. 65: Bild Elefantenrüssel  
Abbildung 28, S. 66: Bild Haarbüschel  
Abbildung 29, S. 67: Bild Haare  
Abbildung 30, S. 68: Bild Plastiksack 1  
Abbildung 31, S. 69: Bild Plastiksack 2  
Abbildung 32, S. 71: Bild Fell eines Stoffelefanten  
Abbildung 33, S. 71: Bild Jacke eines Stofftieres  
Abbildung 34, S. 72: Bild Blume  
Abbildung 35, S. 73: Bild Hut  
Abbildung 36, S. 74: Bild Kürbis  
Abbildung 37, S. 75: Bild Pullover  
Abbildung 38, S. 76: Bild Haarspange  
Abbildung 39, S. 78: Bild Tabelle mit Übereinstimmungsmessungen