

Reale und “halluzinierte” Gesichter -
Gesichtsdetektion bei schneller
Reizpräsentation

BACHELORARBEIT

Tilman Abresch
Philipps-Universität Marburg
Fachbereich Physik
AG Neurophysik

Erstgutachter: Prof. Dr. Wolfgang Einhäuser-Treyer
Zweitgutachter: Prof. Dr. Frank Bremmer
Betreuung: Marius 't Hart

Eingereicht am 23. Februar 2010

Inhaltsverzeichnis

1	Zusammenfassung	3
2	Einleitung und Zielformulierung	4
2.1	Gesichtsdetektion beim Menschen	4
2.2	Gesichtsdetektion mittels Viola-Jones Algorithmus	6
2.3	Menschliche Gesichtsdetektion und Viola-Jones: existieren Zusammenhänge?	10
2.4	Zielformulierung	11
3	Methoden	12
3.1	Generelles zu beiden Versuchen	12
3.1.1	Auswahl der Stimuli	12
3.1.2	Prinzipieller Versuchsaufbau und -ablauf	13
3.2	Versuch 1	13
3.3	Versuch 2	14
4	Ergebnisse	16
4.1	Versuch 1	16
4.1.1	Entscheidungen	16
4.1.2	Reaktionszeiten	17
4.1.3	Erfolg der einzelnen Bilder	17
4.2	Versuch 2	20
4.2.1	Entscheidungen	20
4.2.2	Reaktionszeiten	21
4.2.3	Vergleich zwischen gedrehten und nicht-gedrehten halluzinierten Gesichtern	21
5	Fazit und Diskussion	26

Abbildungsverzeichnis

1	Beispielstimuli von Thorpe et al. (2006)	5
2	Haarähnliche Eigenschaften	7
3	Das Integralbild	8
4	Kaskadierte Klassifizierung	9
5	Saliency Map Model von Cerf et al.	10
6	Anforderungen an Bildausschnitte	13
7	Prinzipieller Versuchsablauf	14
8	Entscheidungen der Versuchspersonen in Exp. 1	16
9	Reaktionszeiten für verschiedene Entscheidungen in Exp. 1	18
10	Erfolg der einzelnen Bilder in Exp. 1 und Auswahl für Exp. 2	19
11	Entscheidungen der Versuchspersonen in Exp. 2	20
12	Reaktionszeiten für verschiedene Entscheidungen in Exp. 2	22
13	Reaktionszeiten der realen Gesichter für versch. Entscheidungen	23
14	Reaktionszeiten der halluzinierten Gesichter für versch. Entscheidungen	24
15	Erfolgsrate der halluzinierten Gesichter abhängig von ihrer Orientierung	25

1 Zusammenfassung

Gesichter stellen für den Menschen besondere visuelle Stimuli dar und können außergewöhnlich schnell und zuverlässig als solche erkannt werden. Nach welchen Kriterien diese schnelle Gesichtsdetektion beim Menschen funktioniert ist jedoch noch nicht verstanden.

Eine der schnellsten Möglichkeiten maschinell Gesichter zu detektieren stellt ein auf Gesichter trainierter Viola-Jones Objektdetektionsalgorithmus dar. Dieser basiert auf Kontrasten zwischen verschiedenen Rechtecken, auf sog. "Haarähnlichen Eigenschaften".

Vergangene Studien deuten darauf hin, dass Fehldetektionen des Viola-Jones Algorithmus vom Menschen eher für Gesichter gehalten werden könnten als andere Objekte.

Ziel dieser Bachelorarbeit ist eine Klärung ebendieses Sachverhaltes. Es wird der Antwort auf die Frage nachgegangen, ob es möglich ist, dass der Mensch zur schnellen Gesichtsdetektion ebenfalls Haarähnliche Eigenschaften prüft.

In zwei Experimenten mussten Versuchspersonen nach kurzer Präsentation zweier Stimuli (20ms) möglichst schnell eine Augenbewegung zu dem Bild machen, welches (für sie) ein Gesicht enthielt.

Fehldetektionen des Viola-Jones Algorithmus ("halluzinierte Gesichter") lenkten dabei in einem ersten Experiment mehr die Aufmerksamkeit auf sich als andere Objekte. Weiterhin lenkten sie mehr von realen Gesichtern ab, als Bilder ohne Fehldetektion. Trotzdem fehlen ihnen wichtige spezifische Eigenschaften realer Gesichter, wie sich in Experiment 2 gezeigt hat. Hier wurden reale und halluzinierte Gesichter im Original sowie mit 90° Drehung als Stimuli eingesetzt. Während reale Gesichter nach ihrer Drehung nicht mehr als solche wahrgenommen wurden, bewirkte die Drehung der halluzinierten Gesichter keine signifikante Änderung.

Zusammenfassend zeigt diese Arbeit, dass Fehldetektionen des Viola-Jones Algorithmus auch von Versuchspersonen bevorzugt als Gesichter wahrgenommen werden. Ob der Grund hierfür jedoch in den Haarähnlichen Eigenschaften selbst zu suchen ist, oder aber in damit korrelierten Eigenschaften, konnte nicht abschließend geklärt werden.

2 Einleitung und Zielformulierung

Die vorliegende Arbeit beschäftigt sich mit der Gesichtsdetektion beim Menschen. Explizit sei erwähnt, dass hiermit die Frage, *ob* sich in einer Umgebung bzw. einem Bild ein Gesicht befindet, gemeint ist, nicht die Frage der Erkennung, *wer* diese Person ist.

Untersucht werden mögliche Parallelen zwischen der menschlichen Gesichtsdetektion und der Gesichtsdetektion eines von Paul Viola und Michael Jones im Jahre 2001 entwickelten Algorithmus' [Viola & Jones, 2001]. Hierzu beschäftigt sich die weitere Einleitung zunächst mit der Frage, ob Gesichter für den Menschen spezielle visuelle Stimuli darstellen. Hierauf folgend wird der Viola-Jones-Algorithmus näher erläutert werden, um anschließend zu klären, warum Gemeinsamkeiten zur menschlichen Gesichtsdetektion bestehen könnten.

2.1 Gesichtsdetektion beim Menschen

Die Idee, Gesichter könnten für den Menschen spezielle visuelle Stimuli darstellen, ist sehr naheliegend. Im Umfeld des Menschen sind Gesichter sehr häufige Stimuli, denen eine besondere Bedeutung zukommt: als soziale Wesen sind Menschen seit jeher darauf angewiesen, zunächst die Existenz anderer Menschen zu registrieren und schließlich auch Gesichter den entsprechenden Personen zuzuordnen. Zudem enthalten Gesichter viele Informationen, die von Bedeutung sein können, wie bspw. Stimmung, Alter und Geschlecht.

Bereits 1982 wurden gesichtssensitive Neuronen im temporalen Cortex von Rhesusaffen gefunden [Perrett et al., 1982]. Hierbei waren einige Zellen signifikant aktiver bei der Präsentation von Photographien oder Skizzen von Gesichtsausschnitten in normaler Anordnung (also als intakte Gesichter), als in durchmischter Anordnung von Gesichtsteilen. Andere Zellen in visuellen Arealen des temporalen Cortex zeigten in ähnlichen Versuchen gleiche Aktivitäten bei normaler und vertauschter Anordnung einzelner Gesichtsteile [Rolls et al., 1994]. Dies legt die Hypothese nahe, dass einige Zellen dieser Region auf verschiedene Eigenschaften spezialisiert sind (teilweise auf Eigenschaften in Gesichtgröße, teilweise auf kleinere), wobei die Position innerhalb des Gesichtes mitunter eine Rolle spielt.

Eine wesentliche Rolle in der Verarbeitung von Gesichtern beim Menschen spielt die Fusiform Face Area (FFA), welche sich auf der ventralen Oberfläche des Fusiform Gyrius im Temporallappen befindet. Dieses Areal ist bei der Präsentation von Gesichtern signifikant aktiver als bei Vergleichsstimuli [Kanwisher et al., 1997]. Auch wenn einige Untersuchungen zeigen, dass nicht ausschließlich Gesichter dort verarbeitet werden [Gauthier et al., 2000], so scheint dieses Areal doch wesentlich zu sein für die Verarbeitung von Gesichtern in normaler Orientierung [Moscovitch et al., 1997].

An die Frage, ob Gesichter spezielle visuelle Stimuli sind, schließt sich nahtlos die Frage an, ob sie auch besonders schnell detektiert werden können. Dies scheint der Fall zu sein, wenngleich nicht abschließend geklärt ist, ob die Ursache bei holistischen Eigenschaften von Gesichtern, High-Level Factors, zu suchen ist, wie Hershler & Hochstein auf Grund einer Untersuchung des "pop-out" Effektes vorschlagen [Hershler & Hochstein, 2004], oder die Ursache eher bei Low-Level Eigenschaften (in dem Fall Informationen, die in der Amplitude der Fouriertransformation des Bildes enthalten sind) zu suchen ist [VanRullen, 2005].

Einige weitere Untersuchungen zur schnellen Gesichtsdetektion basierten auf erzwungenen Entscheidungen [Thorpe et al., 2006]. Hier wurden den Versuchspersonen zwei Bilder auf einem Bildschirm gezeigt, jeweils eines links und eines rechts. Es enthielt immer ein Bild ein Gesicht, während auf dem anderen ein Gegenstand wie z.B. ein Baum oder eine Laterne zu sehen war (siehe Abb. 1). Die Aufgabe der Versuchspersonen bestand darin, so schnell zum Gesicht zu blicken, wie es ihnen möglich war; die Reaktion wurde über die Augenbewegung gemessen. Erstaunlich war eine Erfolgsrate von 94% bei einer durchschnittlichen Reaktionszeit von 147ms, was sehr wenig Zeit zur visuellen Signalverarbeitung lässt. Ähnliche Versuche, bei denen zu Tierbildern geblickt werden musste, zeigten Reaktionszeiten von etwa 230ms [Kirchner & Thorpe, 2005].

Menschliche Gesichter können also sehr schnell und präzise verarbeitet und detektiert werden.



Abbildung 1: Beispielstimuli aus dem Experiment von Thorpe et al. (2006). Quelle: [Thorpe et al., 2006]

2.2 Gesichtsdetektion mittels Viola-Jones Algorithmus

Beim Viola-Jones Algorithmus handelt es sich um einen für Objektdetektion entwickelten Algorithmus, welcher an Hand von Testbildern auf die Erkennung von Gesichtern trainiert wurde [Viola & Jones, 2001]. Das Besondere an ihm ist seine Schnelligkeit: abhängig von der Auflösung der zu untersuchenden Bilder ist eine Gesichtsdetektion fast in Echtzeit möglich. Eine der häufigsten Anwendungen ist die automatisierte Gesichtsfokussierung wie sie in vielen handelsüblichen Digitalkameras vorkommt.

Beim Viola-Jones Algorithmus werden lediglich die Luminanzinformationen des Bildes benötigt, er verwendet also ausschließlich Graustufenbilder. Die schnelle Detektion basiert auf vier grundlegenden Prinzipien:

- *Den Haarähnlichen Eigenschaften* - auf Kontrasten basierende Rechteckeigenschaften, die in verschiedenen Größen an verschiedenen Stellen des Bildes getestet werden.
- *Dem Integralbild* - einer Bildaufbereitung, die ein schnelles Testen der Haarähnlichen Eigenschaften ermöglicht.
- *Dem Boosting-Algorithmus AdaBoost* - ein Algorithmus, der die besten Haarähnlichen Eigenschaften auswählt und geschickt kombiniert.
- *Der Kaskadierten Klassifizierung* - einem System, mit möglichst wenig Rechenaufwand möglichst viele Nicht-Gesichter auszuschließen.

Haarähnliche Eigenschaften

Die Gesichtsdetektion des Viola-Jones Algorithmus erfolgt durch Testen von Haarähnlichen Eigenschaften. Diese verdanken ihren Namen der Ähnlichkeit zu Haar-Wavelets, welche nach dem ungarischen Mathematiker Alfréd Haar benannt sind. Haarähnliche Eigenschaften sind auf Kontrasten basierende Rechteckeigenschaften, wie sie links in Abb. 2 zu sehen sind. Ihre Lokalisation innerhalb der Detektorfläche ist hierbei Teil der Eigenschaft.

Soll nun eine derartige Eigenschaft in einem Bild nachgeprüft werden, so wird die Summe der Pixelwerte in den entsprechenden Rechtecken betrachtet und verglichen. Überschreitet die Differenz zwischen den dunkel markierten und weiß gelassenen Rechtecken einen festgelegten Schwellenwert, so gilt die Eigenschaft als zutreffend. In Abb. 2 rechts sieht man zwei zutreffende Rechteckeigenschaften: die Augenpartie ist dunkler als die Wangenpartie weiter unten; der Nasenrücken ist hell, verglichen mit den dunkleren Augen.

Wollte man nun ein Bild auf sämtliche mögliche Haarähnliche Eigenschaften testen, so würde dies extrem viel Rechenaufwand bedeuten. Grund hierfür ist einmal, dass bei einer Detektorgröße von ca. 24x24 Einheiten bereits eine Anzahl von ca. 50 000 möglichen Haarähnlichen Eigenschaften existiert. Weiterhin will man sämtliche Eigenschaften nicht

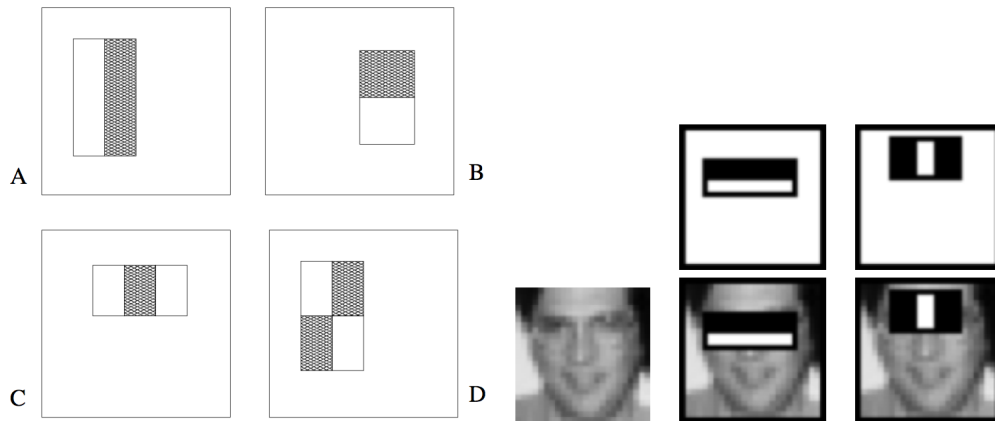


Abbildung 2: Beispiele Haarähnlicher Eigenschaften und Haarähnliche Eigenschaften bezogen auf ein Gesicht. Quelle: [Viola & Jones, 2001]

nur bezogen auf das ganze Bild testen, sondern in vielen Größen an sämtlichen möglichen Stellen - man spricht hier von sog. "Unterbildern", die getestet werden.

Ein erster Schritt, diesen Rechenaufwand zu verkleinern, stellt das Integralbild dar, welches im Folgenden erläutert wird.

Das Integralbild

Das Integralbild enthält an jeder Stelle des Bildes nicht den Helligkeitswert des Pixels (Farben spielen hier keine Rolle, da wie erwähnt nur Graustufen von Bedeutung sind), sondern die Summe über alle Pixel, die sich in dem Rechteck links und über ihm befinden (Schraffierter Bereich, Abb. 3 links). Der Wert ii eines Pixels berechnet sich also wie in Gl. 1 dargestellt.

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y') \quad (1)$$

Will man nun die Summe der Pixel in einem Rechteck des Bildes wissen, so benötigt man lediglich die vier Werte der Eckpunkte im Integralbild. Die Summe über alle Pixel in Rechteck D in Abb. 3 rechts ergibt sich folglich wie in Gl. 2.

$$D = (x_4, y_4) - (x_2, y_2) - (x_3, y_3) + (x_1, y_1) \quad (2)$$

Wie man sieht, können Haarähnliche Eigenschaften durch das Integralbild sehr einfach und mit wenig Rechenaufwand nachgeprüft werden.

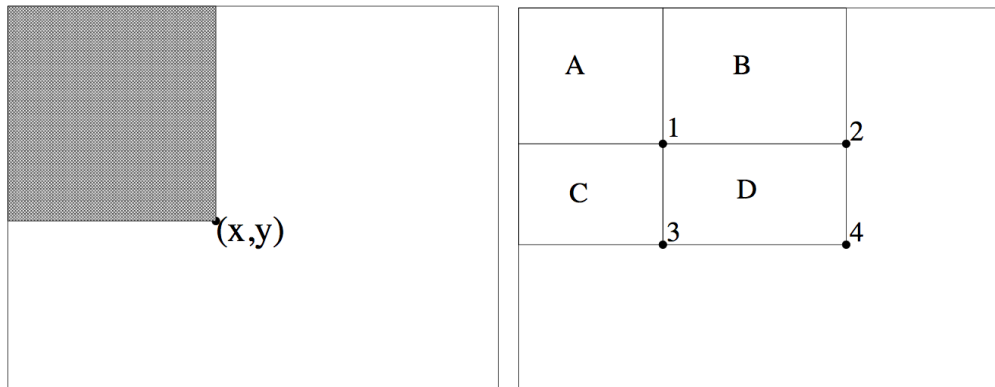


Abbildung 3: Konstruktion des Integralbildes und Veranschaulichung der Berechnung eines Rechteckes. Quelle: [Viola & Jones, 2001]

AdaBoost Maschinlernalgorithmus

Trotz der schnellen Nachprüfbarkeit führt die Menge an möglichen Eigenschaften weiterhin zu einem nicht vertretbaren Rechenaufwand. Es hat sich jedoch herausgestellt, dass eine deutlich kleinere Anzahl an Haarähnlichen Eigenschaften ausreicht, um zuverlässig Gesichter zu detektieren.

Hierzu werden anhand von Testbildern mit Gesichtern (bspw. 1000 Testbilder) Eigenschaften gefunden, deren Zutreffen Gesichter besser lokalisieren kann als der Zufall (z.B. in 51% der Fälle).

Eine leichte Abwandlung des Maschinlernalgorithmusses AdaBoost kombiniert nun jeweils mehrere dieser Eigenschaften, so genannte “schwache Klassifikatoren”, zu einem “starken Klassifikator”. Die Schwellenwerte der schwachen Klassifikatoren und der Gesamtschwellenwert eines starken Klassifikators werden auf der Basis vorheriger Klassifikatorbildungen optimiert. Dies geschieht, in dem jenen Testfällen, in denen vorherige Klassifikatoren falsch lagen, ein höheres Gewicht beigemessen wird. Wichtig ist bei der Klassifikatorenbildung in erster Linie, dass sämtliche Gesichter erkannt werden. Eine hohe false-positive Rate (von bspw. 40%) ist nebensächlich.

Die Anzahl der schwachen Klassifikatoren in einem starken Klassifikator ist variabel. So enthält z.B. der kleinste starke Klassifikator der standardmäßig in OpenCV¹ implementierten Gesichtsdetektion lediglich die beiden in Abb. 2 rechts dargestellten Eigenschaften.

¹OpenCV ist eine quelloffene Programmibibliothek für C und C++ im Bereich Bildverarbeitung und maschinelles Sehen.

Kaskadierte Klassifizierung

Nun verfügen wir über eine Reihe starker Klassifikatoren, die auf Grund des Integralbildes schnell die entsprechenden Haarähnlichen Eigenschaften in allen Unterbildern² überprüfen können.

Eine weitere Geschwindigkeitssteigerung erreichen wir durch Kaskadierte Klassifizierung. Hierzu werden die starken Klassifikatoren nach Größe sortiert, also nach der Zahl der schwachen Klassifikatoren aus denen sie zusammengesetzt sind. Sämtliche Unterbilder durchlaufen nun den ersten starken Klassifikator ("1" in Abb. 4), wobei Unterbilder, die von diesem nicht als Gesicht erkannt werden, komplett aussortiert werden. Die Aussortierung des ersten Klassifikators betrifft üblicherweise um die 60% der Unterbilder ohne Gesicht. Nur die Unterbilder, welche den ersten starken Klassifikator erfolgreich durchlaufen haben, erreichen Klassifikator "2", der nach dem selben Prinzip agiert (dies stellt kein Problem dar, da die starken Klassifikatoren explizit darauf ausgelegt waren, keine echten Gesichter zu verwerfen). Die Anzahl der zu untersuchenden Unterbilder nimmt dadurch rapide ab, während die Klassifikatoren mit der Zeit komplexer werden. Die Anzahl der verwendeten starken Klassifikatoren ist eine Abwägung zwischen Rechenaufwand und Genauigkeit der Gesichtsdetektion. Üblich ist eine Anzahl von etwa 200 starken Klassifikatoren.

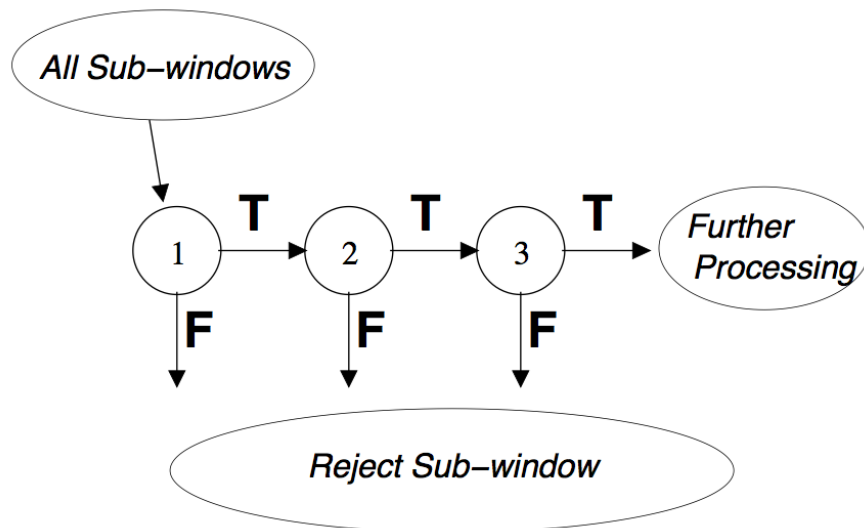


Abbildung 4: Kaskadierte Klassifizierung. Quelle: [Viola & Jones, 2001]

²Um nicht sämtliche Unterbilder untersuchen zu müssen, wird die Größe des zu betrachtenden Bildausschnittes üblicherweise schrittweise um den Faktor 1,1 vergrößert.

2.3 Menschliche Gesichtsdetektion und Viola-Jones: existieren Zusammenhänge?

Nachdem wir nun wissen, wie der bisher für seine Genauigkeit schnellste Gesichtsdetektionsalgorithmus funktioniert und dass der Mensch speziell Gesichter schneller verarbeiten kann als viele andere Stimuli, stellt sich nun die Frage, in wieweit vielleicht auch der Mensch zur schnellen Detektion auf die Erkennung Haarähnlicher Eigenschaften zurückgreift.

Interessante Hinweise hierzu ergeben sich aus Experimenten zu Saliency Map Models [Cerf et al., 2008]. Dies sind Modelle zur Aufmerksamkeitsvorhersage (siehe hierzu auch [Koch & Ullman, 1985] sowie [Itti & Koch, 2005]), welche hier verwendet wurden, um Fixationen von Versuchspersonen vorherzusagen. Das übliche Saliency Map Model, welches auf Low-Level Features, wie in dem Fall Kontrasten in Farbe, Intensität und Orientierung basiert, wurde ergänzt um Detektionen des Viola-Jones Algorithmus' (siehe Abb. 5). Das kombinierte Saliency Map Model war dem ursprünglichen, ausschließlich auf Low-Level Features basierten, Saliency Map Model weit überlegen; die Qualität der Fixationsvorhersage stieg deutlich.

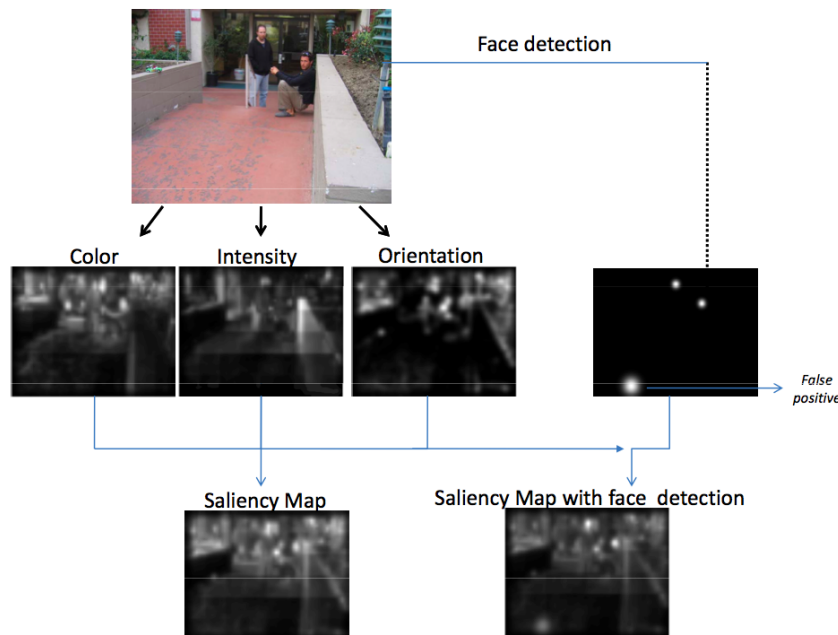


Abbildung 5: Erweitertes Saliency Map Model. Ein auf Farbe, Intensität und Orientierung basierendes Saliency Map Model wird um eine auf Gesichtsdetektion basierende Komponente ergänzt. Quelle: [Cerf et al., 2008]

Ein erstaunliches Ergebnis war, dass bei Bildern ohne Gesichter, bei denen es Fehldetektionen des Viola-Jones Algorithmus gab, das kombinierte Saliency Map Model trotzdem weiterhin die Fixationen gut vorhersagen konnte. Dies legt die Frage nahe, ob in diesem

Versuch die Stellen der Fehldetektionen auf den ersten Blick von den Versuchspersonen ebenfalls für Gesichter gehalten wurden.

Weitere Hinweise hierzu ergeben sich aus Untersuchungen zu Kopf- und Augenbewegungen [Einhäuser et al., 2009]. Hier ergab sich, dass bei freier Bewegung in natürlichen Umgebungen die Versuchspersonen blickzentriert auf Gesichter sahen. Fehldetektionen des Viola-Jones Algorithmus wurden ebenfalls durch Kopfbewegung grob zentriert, nicht jedoch genauer durch Augenbewegungen - im Gegensatz zu realen Gesichtern, wo dies der Fall war. Dies legt Nahe, dass die Fehldetektionen “auf den ersten Blick” für Gesichter gehalten wurden, beim genaueren Hinsehen jedoch nicht mehr, weshalb die anschließende Augenbewegung zur Zentrierung nicht erfolgte.

Die Ergebnisse beider Studien sind Hinweise darauf, dass der Mensch zur schnellen Gesichtsdetektion Haarähnliche Eigenschaften prüfen könnte.

2.4 Zielformulierung

Als zentrales Ziel der Forschung können also folgende Fragen festgehalten werden:

- Werden Fehldetektionen des Viola-Jones Algorithmus’ vom Menschen auf den ersten Blick eher für Gesichter gehalten als andere Objekte?
- Zeigen also Gesichtsdetektion mittels Viola-Jones und schnelle Gesichtsdetektion beim Menschen ein ähnliches Verhalten?
- Kann es sein, dass der Mensch zur schnellen Gesichtsdetektion Haarähnliche Eigenschaften prüft?

3 Methoden

3.1 Generelles zu beiden Versuchen

3.1.1 Auswahl der Stimuli

Sämtliche Bilder, die den Probanden gezeigt wurden, sind Bilder aus realen Umgebungen. Es handelt sich um Ausschnitte aus Einzelbildern (Frames), welche Filmen der EyeSeeCam entnommen wurden. Die EyeSeeCam Aufnahmen stammen aus früheren Forschungsprojekten [Schumann et al., 2008]. Die EyeSeeCam ist ein Kamerasystem, welches auf dem Kopf montiert wird. Anhand der Analyse der Augenbewegung richtet sich die Kopfkamera stets auf das Blickfeld, welches von der Trägerin³ anvisiert wird.

Es wird zwischen drei Kategorien von Bildern unterschieden:

- *Reale Gesichter* sind echte Gesichter, die sowohl vom Menschen als auch vom Viola-Jones Algorithmus⁴ für Gesichter gehalten werden.
- *Halluzinierte Gesichter* sind keine echten Gesichter, werden aber vom Viola-Jones Algorithmus für Gesichter gehalten; vom Menschen bei genauerer Betrachtung nicht.
- *Objekte* sind keine echten Gesichter und werden weder vom Menschen noch vom Viola-Jones Algorithmus für Gesichter gehalten.

Damit Einzelbilder aus einem EyeSeeCam-Film entnommen wurden, mussten einige Kriterien erfüllt werden (siehe auch Abb. 6):

- Der Detektionsbereich musste sich im mittleren Neuntel des Bildes befinden. Hierdurch sollte sichergestellt werden, dass die Trägerin der EyeSeeCam zur Aufnahmezeit tatsächlich in Richtung des Gesichtes/Objektes geblickt hat und dieses nicht nur im peripheren Sichtbereich aufgetreten ist.
- Zwischen zwei ausgewählten Einzelbildern musste mindestens eine Sekunde (entspricht 30 Einzelbildern) liegen, um nicht zu viele ähnliche Bilder auszuwählen.
- Bei realen und halluzinierten Gesichtern durfte es pro Einzelbild nur eine Detektion gegeben haben. Es sollte hierdurch ausgeschlossen werden, dass ein Gesicht/Objekt ausgewählt wird, welches eventuell gar nicht die Aufmerksamkeit der EyeSeeCam Trägerin auf sich gezogen hatte. Bei Objekten durfte es gar keine Detektion gegeben haben.
- Im Falle einer Detektion, die allen bisherigen Kriterien entsprach, musste der Detektionsbereich mindestens 50px x 50px groß sein, damit eine ausreichende Auflösung für die Versuche gegeben war.

³Wie am FB Physik üblich, wird aus Gründen der Lesbarkeit ausschließlich die weibliche Form verwendet. Sofern nicht anders vermerkt sind immer beide Geschlechter gemeint.

⁴Es wurde die standardmäßig in OpenCV implementierte Version verwendet.

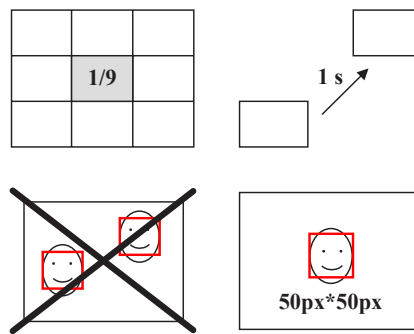


Abbildung 6: Anforderungen, die Bildausschnitte erfüllen mussten um in den Experimenten Verwendung zu finden. Die oberen beiden Anforderungen mussten sämtliche Bilder erfüllen, die unteren nur reale und halluzinierte Gesichter.

3.1.2 Prinzipieller Versuchsaufbau und -ablauf

Den Versuchspersonen wurden auf einem Bildschirm in 67,5 cm Entfernung zwei Bilder gezeigt, welche jeweils 16cm x 16cm groß waren, was einem Sehwinkel von je $14^\circ \times 14^\circ$ entspricht. Der Abstand beider Bilder betrug ca. $2,8^\circ$ Sehwinkel.

Nach einer kurzen Fixationszeit wurden den Versuchspersonen beide Bilder gleichzeitig für 20ms gezeigt. Anschließend wurden zwei Zufallsmuster ("Masken") an den Stellen eingeblendet, wo vorher die Bilder gewesen waren. Die Masken waren links und rechts gleich und bestanden aus dem Mittel aller verwendeten Bilder, versehen mit einer randomisierten Phasenverschiebung in der Fourier-Transformation. Nach Einsetzen einer Augenbewegung oder Blinzeln verblieben die Masken noch weitere 500ms, bevor eine einsekündige Pause vor dem nächsten Trial einsetzte. Der Ablauf ist in Abb. 7 noch einmal dargestellt. Aufbau und Ablauf sind ähnlich dem Versuch von Kirchner & Thorpe 2005 [Kirchner & Thorpe, 2005].

Die Versuchspersonen bekamen die Aufgabe, so schnell wie möglich von beiden Bildern dasjenige anzublicken, welches ein Gesicht enthält.

Mit einer Infrarotkamera (Eyelink-2000 eyetracking system, SR Research Ltd., Mississauga, ON, Canada) wurde die Augenbewegung der Versuchspersonen registriert, um später die erste vollzogene Sakkade daraufhin zu betrachten, wie schnell zu welchem Bild geblickt wurde.

3.2 Versuch 1

In Versuch 1 wurden pseudorandomisierte Kombinationen aus Bildern aller drei Kategorien verwendet. Jede Kategorie verfügte über 50 Bilder; es wurden je 500 Kombinationen zwischen Bildern verschiedener Kategorien gebildet. Insgesamt entstanden so 500 Trials

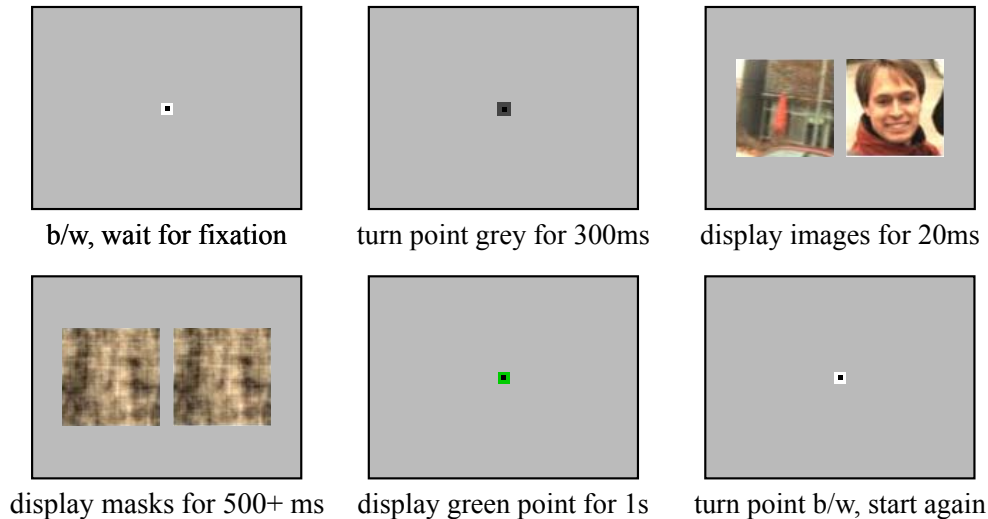


Abbildung 7: Der prinzipielle Versuchsablauf. Zunächst wird ein schwarz-weißer Fixationspunkt gezeigt, der sich bei ausreichender Fixation grau färbt und nach 300ms von den beiden Bildern abgelöst wird, welche für 20ms gezeigt werden. Den Bildern folgen Masken, welche nach ausreichender Augenbewegung oder Blinzeln noch 500ms verbleiben. Anschließend wird für 1s ein grüner Fixationspunkt eingeblendet, bevor der Ablauf von Neuem startet.

x 3 = 1500 Trials, die den Versuchspersonen in 15 Blöcken à 100 Trials gezeigt wurden. Als Versuchspersonen wurden 4 Frauen und 4 Männer im Alter von 18 bis 25 Jahren untersucht.

Die zentrale Frage in Versuch 1 war, ob halluzinierte Gesichter mehr von realen Gesichtern ablenken, also bessere Distraktoren sind, als Objekte.

3.3 Versuch 2

Für Versuch 2 wurden die 10 erfolgreichsten halluzinierten Gesichter aus Versuch 1 entnommen, sowie 10 unterschiedlich erfolgreiche reale Gesichter, die im Erfolgsspektrum gleichmäßig verteilt waren. Die Bildauswahl findet sich in Abb. 10; sie ist durch eine rote Umrandung gekennzeichnet. Diese 20 Bilder wurden sowohl im Original verwendet, sowie zusätzlich mit 90° Drehung.

Die Drehung wurde gewählt, da es verschiedene Hinweise gibt, dass gedrehte Gesichter vom Menschen schlechter erkannt werden. So befinden sich wichtige Informationen über Gesichter vermehrt in horizontalen Bildinformationen [Dakin & Watt, 2009], welche nicht invariant gegenüber einer Drehung von 90° sind. Weiterhin findet die Verarbeitung von Gesichtern normaler Orientierung vermutlich in anderen Hirnregionen statt als jene von gedrehten Gesichtern [Moscovitch et al., 1997].

Reale und "halluzinierte" Gesichter - Gesichtsdetektion bei schneller Reizpräsentation

Gezeigt wurden sämtliche mögliche Kombinationen aller vier Kategorien, was $10 \times 10 \times 4 = 400$ Trials pro Versuchsperson ergab. Diese wurden in 4 Blöcken à 100 Trials gezeigt. Versuchspersonen waren in diesem Versuch 4 Frauen und 4 Männer im Alter von 19 bis 31 Jahren.

Die zentrale Frage in Versuch 2 war, wie die Drehung der Bilder die Erkennungsleistung der Gesichter bzw. die Distraktionsleistung der halluzinierten Gesichter beeinflusst.

4 Ergebnisse

4.1 Versuch 1

4.1.1 Entscheidungen

In Abb. 8 sind die Ergebnisse von Versuch 1 bezüglich der Entscheidungen der Versuchspersonen dargestellt. Auf der x-Achse sehen wir die drei Kombinationen der Kategorien der Bilder; auf der y-Achse ist aufgetragen in wieviel Prozent der Fälle sich für die auf der x-Achse zuerst genannte Kategorie entschieden wurde. Die schwarzen Vierecke stellen die einzelnen Versuchspersonen dar, während die roten Kreise den Durchschnitt aller Versuchspersonen markieren. Dieser ist als numerische Angabe auch unter der Kennzeichnung der x-Achse vermerkt. Die angegebenen Signifikanzlevel in sämtlichen Graphiken entsprechen der standardmäßigen Konvention⁵.

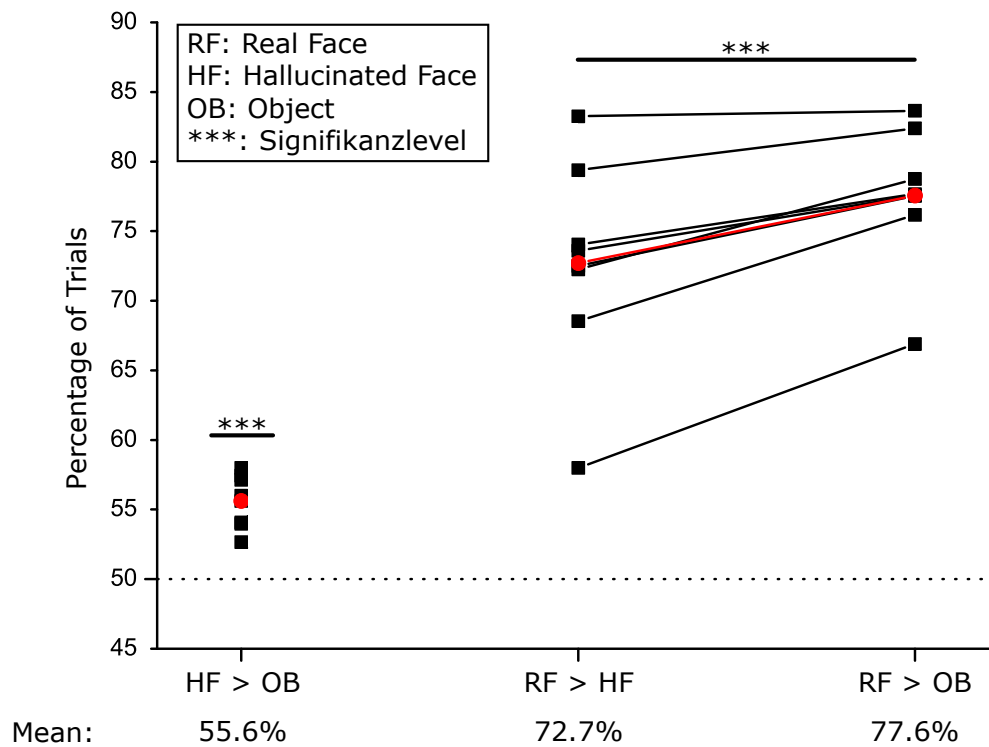


Abbildung 8: Entscheidungen der Versuchspersonen in Exp. 1. Die schwarzen Vierecke stellen die verschiedenen Versuchspersonen dar, während die roten Kreise den Mittelwert angeben.

⁵Für den P-Wert gilt dementsprechend: ***: $p < 0,001$, **: $p < 0,01$.

Wie man sieht, entschieden sich die Versuchspersonen bei Entscheidungen zwischen halluziniertem Gesicht und Objekt in 55,6% der Fälle für das halluzinierte Gesicht. Die Verteilung über die verschiedenen Versuchspersonen unterscheidet sich signifikant vom Zufall (50%) [$t(7) = 8.37$, $p < 0.0001$], wie ein Einstichproben t-Test zeigt. Halluzinierte Gesichter wurden also signifikant häufiger für Gesichter gehalten als Objekte.

In Entscheidungen zwischen realen Gesichtern und halluzinierten Gesichtern wurde sich in 72,7% der Fälle für das reale Gesicht entschieden. Betrachtet man nun die Entscheidungen zwischen realen Gesichtern und Objekten, so zeigt sich, dass in diesem Fall nicht nur die durchschnittliche Erfolgsrate (77,6%) größer ist, als im vorherigen Fall, sondern auch, dass sich sämtliche Versuchspersonen verbessern. Die Verbindungslinien in Abb. 8 verknüpfen jeweils die Erfolgsrate der Versuchspersonen in den beiden rechten Kategorien. Die Erfolgsraten dieser beiden Kategorien unterscheidet sich signifikant [$t(7) = -5.04$, $p = .0007$], was durch einen Zweistichproben t-Test geprüft wurde. Man kann also festhalten, dass die halluzinierten Gesichter bessere Distraktoren darstellen als die Objekte.

4.1.2 Reaktionszeiten

In Abb. 9 findet sich eine Übersicht über die Reaktionszeiten. Es zeigt sich, dass in den verschiedenen Kategorien von Entscheidungen gerade die Mediane nur unwesentlich variieren. Die mittleren Reaktionszeiten unterscheiden sich etwas mehr, aber auch dies bewegt sich in einem Bereich von maximal 30ms.

Da die Reaktionszeiten nicht normalverteilt sind (siehe Abb. 9), werden Wilcoxon-Tests angewendet, um signifikante Unterschiede zwischen den Medianen zu finden. Es zeigt sich, dass Entscheidungen, reale Gesichter auszuwählen immer signifikant schneller von Statten gingen als Entscheidungen, Objekte auszuwählen (jeweils unabhängig von der Alternative, die es bei der Entscheidung gab; sämtliche p-Werte $< .03$). Entscheidungen, reale Gesichter (wieder unabhängig von der Alternative bei der Auswahl) auszuwählen wurden außerdem signifikant schneller getroffen, als halluzinierte Gesichter Objekten vorzuziehen (alle p-Werte $< .005$). Sämtliche nicht erwähnte Kombinationen zeigten keinen signifikanten Unterschied ($p > .05$).

4.1.3 Erfolg der einzelnen Bilder

In Abb. 10 finden sich sämtliche Bilder, sortiert nach ihrem Erfolg in Experiment 1. Erfolg meint hier die Häufigkeit, in der diese Bilder von den Versuchspersonen ausgewählt wurden. Es lässt sich feststellen, dass dunklere Bilder zumindest tendenziell schlechter abschnitten als hellere Bilder, auch wenn es einige Ausnahmen gibt.

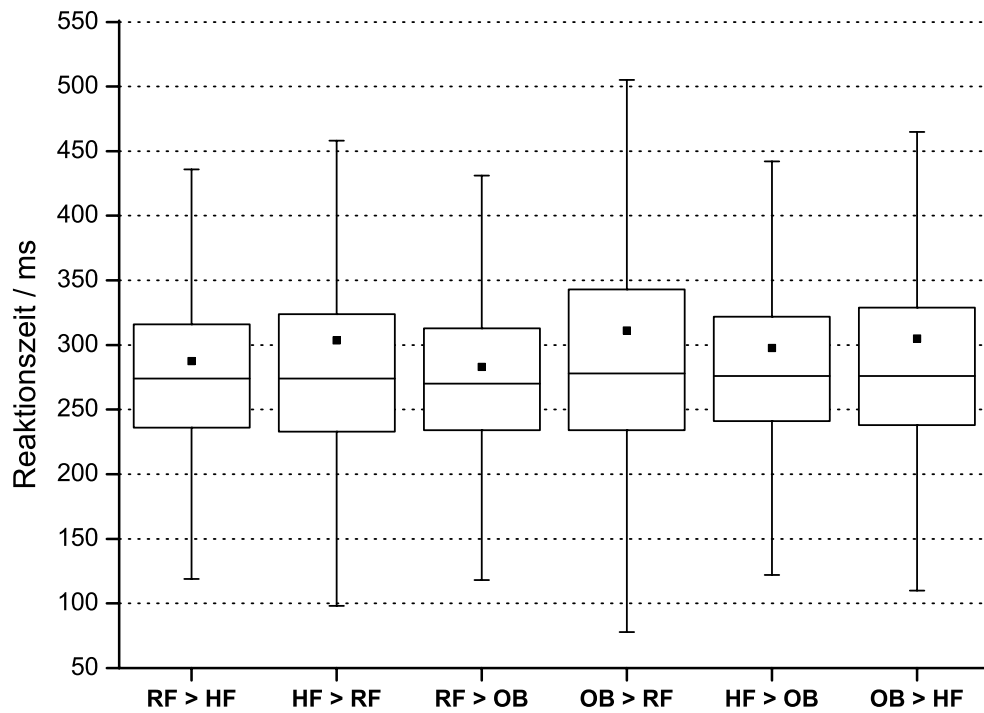


Abbildung 9: Reaktionszeiten für verschiedene Entscheidungen in Exp. 1. Die Länge der Box entspricht dem Interquartilsabstand (IQR). Die enthaltene Linie stellt den Median dar, während die schwarzen Punkte das arithmetische Mittel kennzeichnen. Die Whisker sind nach der Tukey'schen Definition konstruiert (Beschränkung auf 1,5 IQR). Für die x-Achse gilt: RF: Real Face, HF: Hallucinated Face, OB: Object.



Abbildung 10: Sämtliche verwendete Stimuli (Reihenfolge: reale Gesichter, halluzinierte Gesichter, Objekte), sortiert nach ihrem Erfolg innerhalb ihrer Kategorie in Versuch 1 (zeilenweise). Die rot eingerahmten Bilder stellen die Auswahl für Versuch 2 dar.

4.2 Versuch 2

4.2.1 Entscheidungen

In Abb. 11 finden sich die Ergebnisse bezüglich der Entscheidungen in Versuch 2.

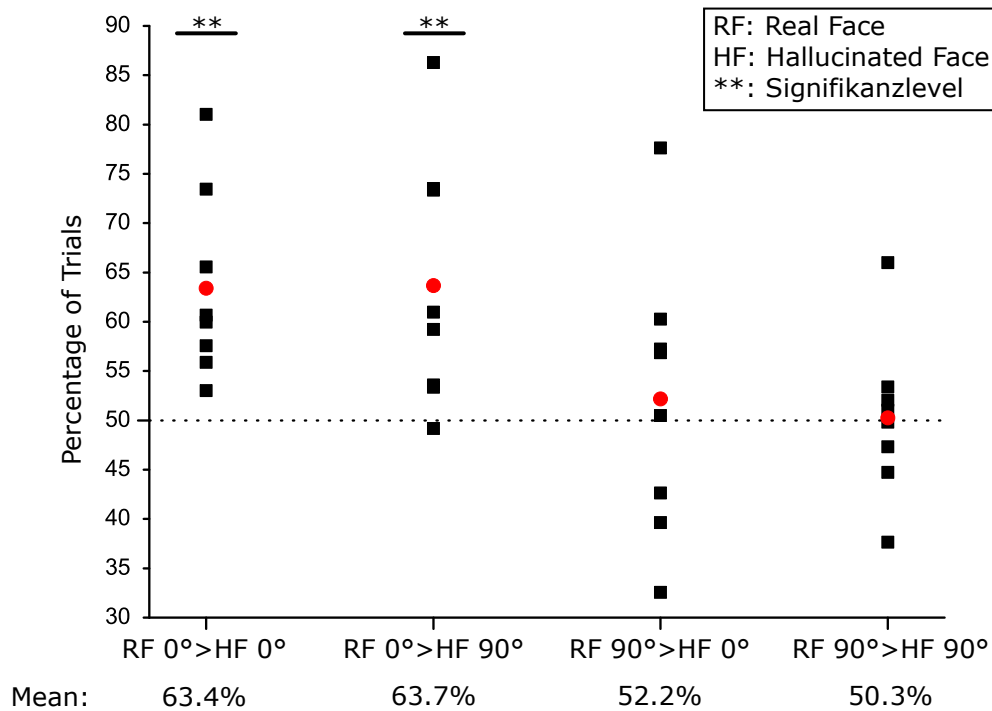


Abbildung 11: Entscheidungen der Versuchspersonen in Exp. 2. Die schwarzen Vierecke stellen die verschiedenen Versuchspersonen dar, während die roten Kreise den Mittelwert angeben.

Mussten sich die Versuchspersonen zwischen ungedrehten realen und ungedrehten halluzinierten Gesichtern entscheiden, so wählten sie in 63,4% der Fälle die realen Gesichter. Die Verteilung über die Versuchspersonen unterscheidet sich signifikant vom Zufall [$t(7) = 3.98$, $p = .003$].

Wurden den Versuchspersonen ungedrehte reale Gesichter und gedrehte halluzinierte Gesichter gezeigt, so ändert sich die Erkennungsleistung der realen Gesichter (im Mittel 63,7%) nicht signifikant [$t(7) = -0.15$, $p = .44$], wenngleich die Streuung der Versuchspersonen zunimmt. Auch diese Verteilung unterscheidet sich signifikant vom Zufall [$t(7) = 3.02$, $p < 0.01$].

In Entscheidungen mit gedrehtem realen Gesicht, konnte dieses nicht signifikant besser oder schlechter identifiziert werden, als dies per Zufall möglich gewesen wäre - sowohl bei ungedrehtem [$t(7) = .43$, $p = .34$], als auch bei gedrehtem [$t(7) = .09$, $p = .47$] halluzinierten Gesicht. Zwischen den beiden Bedingungen mit ungedrehtem (52,2%) und

gedrehtem (50,3%) halluzinierten Gesicht ist kein signifikanter Unterschied festzustellen [$t(7) = .53$, $p = .30$], wenngleich im zweiten Fall die Streuung der Versuchspersonen geringfügig abnimmt.

Es hat sich also gezeigt, dass die Drehung der realen Gesichter ihre Erkennbarkeit derart senkt, dass sich diese nicht mehr signifikant vom Zufall unterscheidet, während die Drehung der halluzinierten Gesichter keinen signifikanten Effekt auf die Entscheidung der Versuchspersonen hatte.

4.2.2 Reaktionszeiten

In Abb. 12 finden sich die mittleren Reaktionszeiten sowie deren Mediane. Auch hier sind die Mediane relativ gleichbleibend, während die Reaktionszeiten stärker variieren als in Versuch 1.

Die einzigen signifikanten Unterschiede (Wilcoxon-Test) zwischen den Medianen ergeben sich zwischen den Fällen "RF 0° > HF 0°", Median: 290ms, und "RF 90° > HF 90°", Median: 312ms, [$z = 2.44$, $p = .015$], sowie zwischen den Fällen "RF 0° > HF 90°", Median: 298ms, und "RF 90° > HF 90°", Median: 312ms, [$z = 2.18$, $p = .03$]. Entscheidungen, ein gedrehtes halluziniertes Gesicht einem gedrehten realen Gesicht vorzuziehen, dauerten also signifikant länger als Entscheidungen, ein ungedrehtes reales Gesicht einem halluzinierten Gesicht gleich welcher Orientierung vorzuziehen.

In Abb. 13 und 14 finden sich die Reaktionszeiten der einzelnen realen und halluzinierten Gesichter, aufgespalten in Fälle, in denen die Bilder ausgewählt wurden und in denen die Bilder nicht ausgewählt wurden, sowie in die verschiedenen Orientierungen.

Es lässt sich kein eindeutiges "Muster" in den Reaktionszeiten für die einzelnen Bilder finden. Die Reaktionszeiten waren bei den erfolgreicher realen Gesichtern geringfügig kürzer, wenn das jeweilige Gesicht gewählt wurde (unabhängig von der Orientierung). Die auftretenden Effekte sind jedoch recht klein. Weitere systematische Besonderheiten in größerem Ausmaß sind nicht vorhanden.

4.2.3 Vergleich zwischen gedrehten und nicht-gedrehten halluzinierten Gesichtern

Betrachten wir nun noch den Erfolg der halluzinierten Gesichter abhängig von ihrer Orientierung in Entscheidungen mit nicht-gedrehtem realen Gesicht. In Abb. 15 ist ihre Erfolgsrate bei den beiden Orientierungen aufgetragen. Es ist zu sehen, dass einige Bilder in einer bestimmten Orientierung erfolgreicher waren, es allerdings keine extremen Ausreißer gibt, die in einer speziellen Orientierung deutlich eher oder deutlich weniger für Gesichter gehalten wurden.

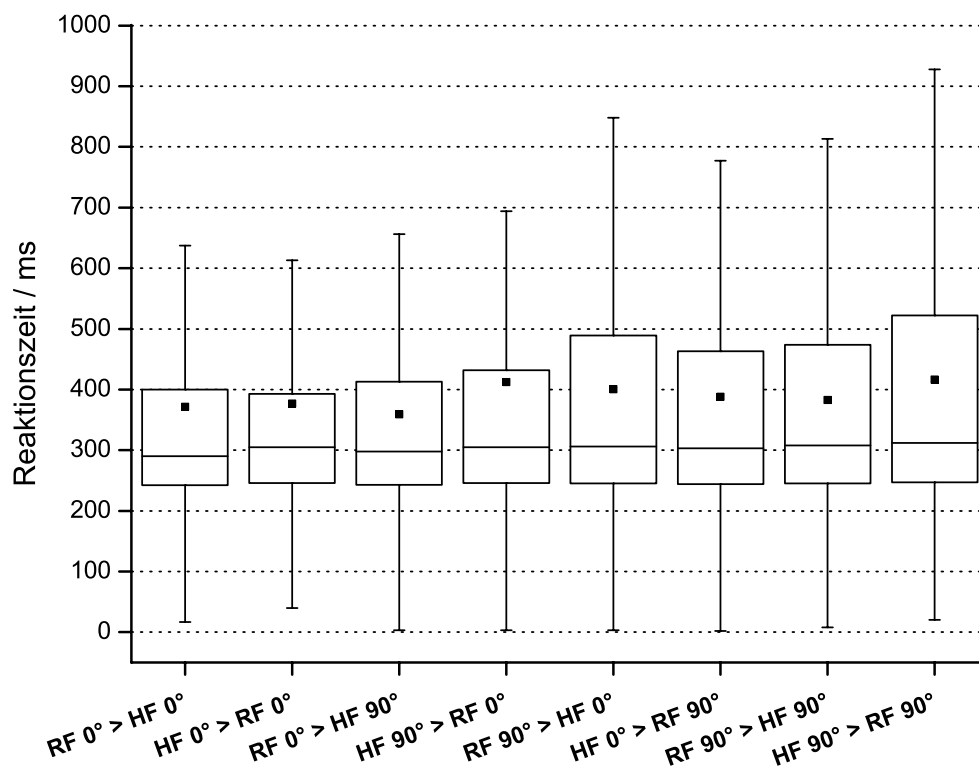


Abbildung 12: Reaktionszeiten für verschiedene Entscheidungen in Exp. 2. Die Länge der Box entspricht dem Interquartilsabstand (IQR). Die enthaltene Linie stellt den Median dar, während die schwarzen Punkte das arithmetische Mittel kennzeichnen. Die Whisker sind nach der Tukey'schen Definition konstruiert (Beschränkung auf 1,5 IQR). Für die x-Achse gilt: RF: Real Face, HF: Hallucinated Face, OB: Object.

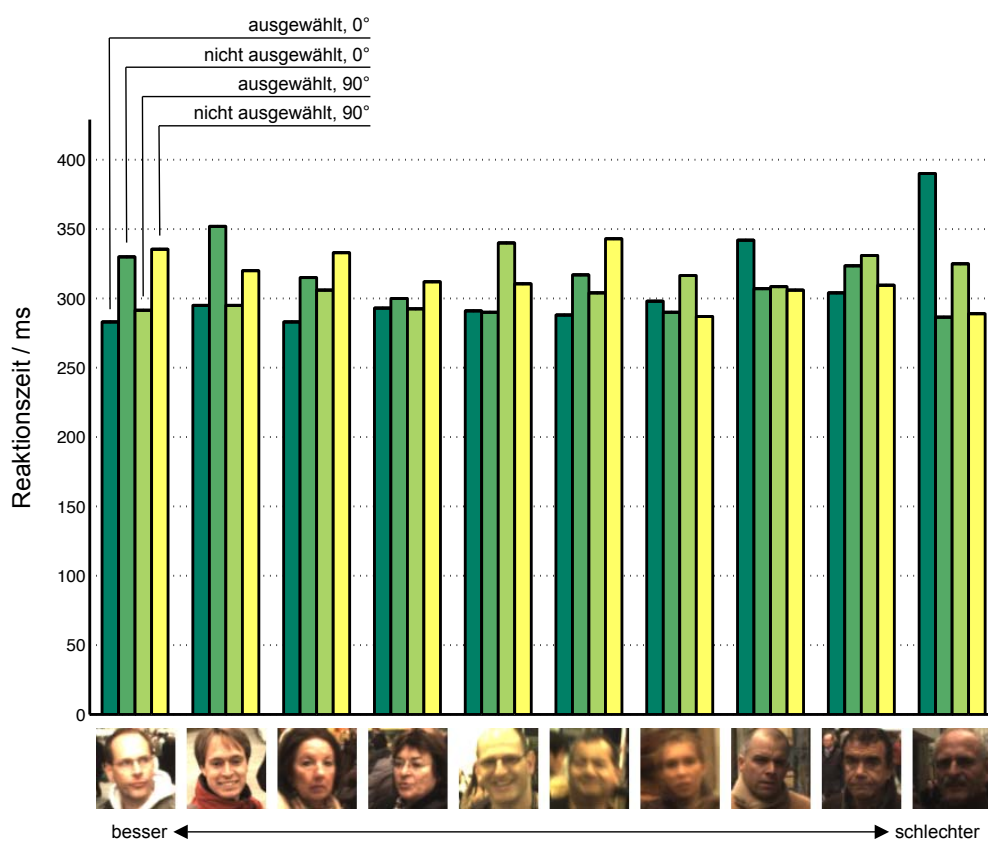


Abbildung 13: Reaktionszeiten der realen Gesichter für versch. Entscheidungen. Die Sortierung entlang der x-Achse erfolgt anhand der Erfolgsrate des Bildes in Exp. 2 (in beiden Orientierungen).

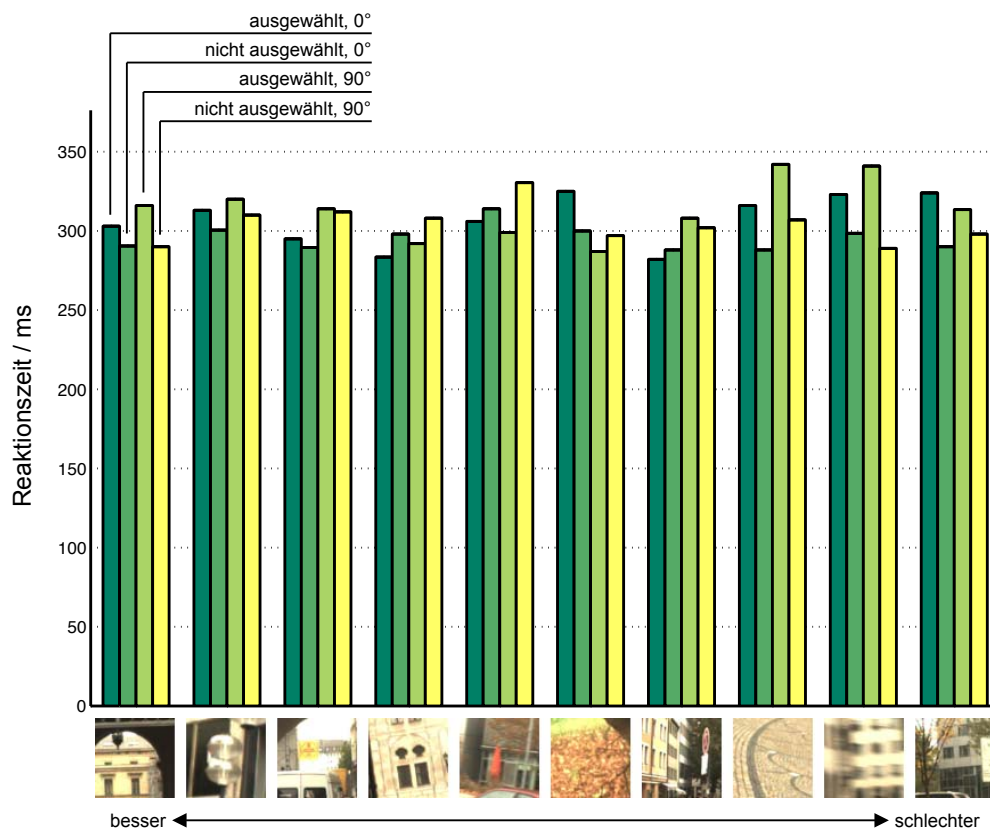


Abbildung 14: Reaktionszeiten der halluzinierten Gesichter für versch. Entscheidungen. Die Sortierung entlang der x-Achse erfolgt anhand der Erfolgsrate des Bildes in Exp. 2 (in beiden Orientierungen).

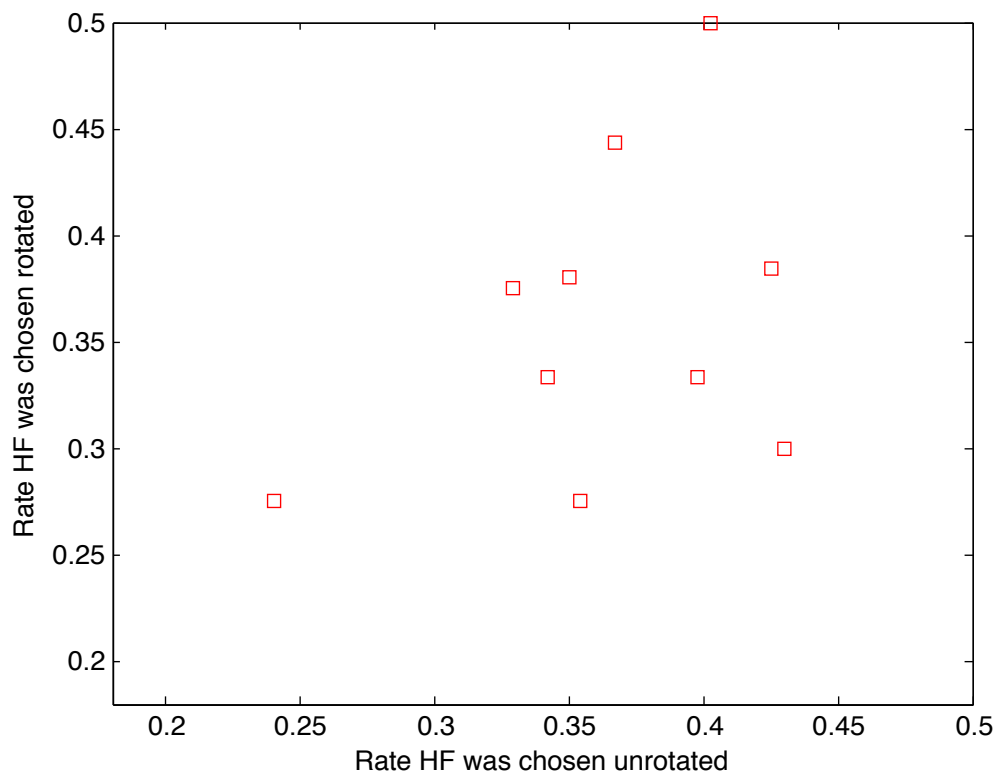


Abbildung 15: Erfolgsrate der halluzinierten Gesichter abhängig von ihrer Orientierung (in Entscheidungen mit ungedrehten realen Gesichtern). Die 10 roten Quadrate stellen die 10 verschiedenen halluzinierten Gesichter dar.

5 Fazit und Diskussion

Die Sonderstellung der halluzinierten Gesichter in der menschlichen Wahrnehmung, die auf Grund früherer Forschungen [Cerf et al., 2008] [Einhäuser et al., 2009] vermutet wurde, hat sich bestätigt. Wie aus den Ergebnissen in Versuch 1 (siehe 4.1.1) hervorgeht, lenkten halluzinierte Gesichter signifikant besser von realen Gesichtern ab als Objekte. Offensichtlich wurden sie von den Versuchspersonen im Vergleich zu den Objekten "auf den ersten Blick" eher für Gesichter gehalten. Betrachtet man jedoch den Erfolg der einzelnen Bilder (siehe Abb. 10), so stellt sich die Frage, ob Low-Level Features wie Luminanz, Kontrast oder Farbe ebenfalls eine Rolle im Entscheidungsprozess der Versuchspersonen gespielt haben.

Diese Vermutung scheint sich in Versuch 2 für die realen Gesichter nicht zu bestätigen. Durch die Drehung der Bilder ändert sich nichts an Kontrast, Luminanz oder Farbe. Die realen Gesichter wurden trotzdem gar nicht mehr als solche erkannt (siehe 4.2.1), ihre gute Erkennbarkeit in ungedrehtem Zustand lässt sich also nicht alleine auf Low-Level Features zurückführen. Vielmehr verfügen die realen Gesichter anscheinend über Eigenschaften, welche nicht drehungsvariant sind und die besondere Erkennungsleistung ermöglichen. Dies steht im Einklang mit der besonderen Rolle, die horizontale Informationen in Gesichtern haben [Dakin & Watt, 2009], sowie der Erkenntnis, dass Gesichter normaler Orientierung im menschlichen Gehirn speziell verarbeitet werden [Moscovitch et al., 1997].

Die etwas schlechtere Erkennungsleistung der realen Gesichter in Versuch 2, verglichen mit Versuch 1, erstaunt nicht, da in Versuch 2 die besten 10 halluzinierten Gesichter Verwendung fanden, während die Auswahl an realen Gesichtern breiter gefächert war.

Betrachtet man die Ergebnisse von Versuch 2 in Bezug auf die halluzinierten Gesichter, so könnten Low-Level Features eine Rolle gespielt haben - allerdings auch andere drehungsinvariante Eigenschaften. Die drehungsanfällige(n) Eigenschaft(en), welche die realen Gesichter so schnell detektierbar machen, besitzen sie jedoch nicht. Anders lässt sich ihre unveränderte Leistung nach ihrer Drehung nicht erklären.

Aus den gemessenen Reaktionszeiten ergeben sich keine maßgeblichen weiteren Erkenntnisse. Für die im Vergleich zu ähnlichen Experimenten [Thorpe et al., 2006] langen Reaktionszeiten lassen sich zwei Begründungen finden. Einmal stammen die hier verwendeten Bilder der realen Gesichter aus realen Umgebungen und sind somit keine optimalen Portraits und damit evtl. schwerer zu detektieren. Weiterhin war der Präsentationszeitraum (20ms) extrem kurz im Vergleich zu 400ms [Thorpe et al., 2006]. Es stellt eine erstaunliche Leistung des menschlichen Gehirns dar, nach einer derart kurzen Reizpräsentation eine Entscheidung in der erreichten Genauigkeit und den erreichten Reaktionszeiten zu fällen.

Die zentrale Frage ist nun, ob anhand dieser Ergebnisse ausgeschlossen werden kann, dass der Mensch zur schnellen Gesichtsdetektion Haarähnliche Eigenschaften prüft. Meiner Ansicht nach ist dies aus verschiedenen Gründen nicht der Fall:

- Unklar bleibt, wie gut die halluzinierten Gesichter vom Viola-Jones Algorithmus erkannt wurden. Da es nicht möglich ist, eine Art “Gesichtsartigkeit” für eine Detektion ausgeben zu lassen, wäre es durchaus möglich, dass bspw. alle realen Gesichter sehr gut erkannt wurden, die halluzinierten Gesichter jedoch geradeso.
- Haarähnliche Eigenschaften (zumindest ein spezielles, auf nicht-gedrehte Gesichter spezialisiertes Set) sind nicht drehungsinvariant. So erkennt der Viola-Jones Algorithmus weder die gedrehten realen, noch die gedrehten halluzinierten Bilder.

Möglich ist natürlich, dass der Mensch zur Gesichtsdetektion verschiedene Kriterien anwendet - drehungsinvariante und solche, die ihre Eigenschaften mit der Drehung verlieren. Sind die drehungsinvarianten Eigenschaften bei den halluzinierten und realen Gesichtern gleichmäßig verteilt, während die anderen Eigenschaften nur bei realen Gesichtern vorhanden sind, lässt sich das Ergebnis von Versuch 2 - insbesondere die Gleichwertigkeit von gedrehten realen Gesichtern und beliebig orientierten halluzinierten Gesichtern - gut erklären. Fraglich bliebe dann allerdings, warum die halluzinierten Gesichter durch die orientierungsgebundenen Viola-Jones Detektoren selektiert wurden.

Zur Klärung der formulierten Forschungsziele sind also weitere Untersuchungen notwendig, für die die durchgeführten Versuche eine interessante Basis darstellen.

Denkbar wäre bspw. Versuch 1 mit Stimuli, deren Low-Level Features angeglichen wurden (oder gezielt teilweise angeglichen und teilweise nicht angeglichen wurden) oder ein um Objekte erweiterter Versuch 2. Auch die Untersuchung des “pop-out” Effektes [Hershler & Hochstein, 2004] [VanRullen, 2005] von gesichtstypischen Haarähnlichen Eigenschaften wäre eine interessante Option.

Literatur

- [Cerf et al., 2008] Cerf, M.; Harel, J.; Einhäuser, W.; Koch, C. (2008). Predicting human gaze using low-level saliency combined with face detection. *Advances in Neural Information Processing (NIPS) 2007*, 20:241-248.
- [Dakin & Watt, 2009] Dakin, S. C. & Watt, R. J. (2009). Biological "bar codes" in human faces. *Journal of Vision*, 9(4):2, 1-10, <http://journalofvision.org/9/4/2/>, doi:10.1167/9.4.2.
- [Einhäuser et al., 2009] Einhäuser, W.; Schumann, F.; Vockeroth, J.; Bartl, K.; Cerf, M.; Harel, J.; Schneider, E.; König, P. (2009). Distinct Roles for Eye and Head Movements in Selecting Salient Image Parts during Natural Exploration. *Annals of the New York Academy of Sciences*, 1164:188-193.
- [Gauthier et al., 2000] Gauthier, I.; Skudlarski, P.; Gore, J.C.; Anderson, A.W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, 2000 Feb;3(2):191-7.
- [Hershler & Hochstein, 2004] Hershler, Orit & Hochstein, Shaul (2004). At first sight: A high-level pop out effect for faces. *Vision Research*, 45 (2005), 1707-1724.
- [Itti & Koch, 2005] Itti, Laurent & Koch, Christof (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, 40 (2000), 1489-1506.
- [Kanwisher et al., 1997] Kanwisher, N.; McDermott, J.; Chun, M.M. (1997). The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for Face Perception. *The Journal of Neuroscience*, June 1, 1997, 17(11):4302-4311.
- [Kirchner & Thorpe, 2005] Kirchner, Holle & Thorpe, Simon J. (2005). Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Vision Research*, 46 (2006), 1762-1776.
- [Koch & Ullman, 1985] Koch, C. & Ullman, S. (1985). Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology*, 1985, 4:219-227.
- [Moscovitch et al., 1997] Moscovitch, M.; Winocur, G.; Behrmann, M. (1997). What is special about face recognition? Nineteen experiments on a person with visual object agnosia and dyslexia but normal face recognition. *Journal of Cognitive Neuroscience*, Fall 1997, Vol. 9, No. 5, Pages 555-604.
- [Perrett et al., 1982] Perrett, D.I.; Rolls, E.T.; Caan, W. (1982). Visual Neurones Responsive to Faces in the Monkey Temporal Cortex. *Experimental Brain Research*, 1982, 47: 329-342.
- [Rolls et al., 1994] Rolls, E.T.; Tovee, M.J.; Purcell, D.G.; Stewart, A.L.; Azzopardi, P. (1994). The responses of neurons in the temporal cortex of primates, and face identification and detection. *Experimental Brain Research*, 1994, 101:473-484.

- [Schumann et al., 2008] Schumann, F.; Einhäuser-Treyer, W.; Vockeroth, J.; Bartl, K.; Schneider, E.; König, P. (2008). Salient features in gaze-aligned recordings of human visual input during free exploration of natural environments. *Journal of Vision*, 8(14):12, 1-17, <http://journalofvision.org/8/14/12/>, doi:10.1167/8.14.12.
- [Thorpe et al., 2006] Thorpe, Simon J.; Crouzet, Sébastien; Kirchner, Holle & Fabre-Thorpe, Michèle (2006). Ultra rapid face detection in natural images: Implications for computation in the visual system. *Centre de Recherche Cerveau & Cognition*, UMR 5549, Université Paul Sabatier, Toulouse, France.
- [VanRullen, 2005] VanRullen, Rufin (2005). On second glance: Still no high-level pop-out effect for faces. *Vision Research*, 46 (2006), 3017-3027.
- [Viola & Jones, 2001] Viola, Paul & Jones, Michael (2001). Robust Real-time Object Detection. *Second International Workshop on Statistical and Computational Theories of Vision - Modelling, Learning, Computing, and Sampling*, Vancouver, Canada, July 13, 2001.

Danksagung

Für die gute Zusammenarbeit und Betreuung möchte ich mich an dieser Stelle ganz herzlich bei Marius 't Hart und Wolfgang Einhäuser-Treyer bedanken, die durch ihre Unterstützung und ihre weiterführenden Ratschläge ganz erheblich zum Erfolg dieser Arbeit beigetragen haben.

Bei meinen Eltern möchte ich mich unter anderem für die ermöglichte und gelassene Freiheit bedanken.

Anton Riske, der einige Unterstützung, v.a. in der Auswahl der Stimuli, geleistet hat, möchte ich hier ebenfalls danken.

Weiterhin gilt mein Dank allen Lektorinnen und Lektoren, sowie Christine Roth und Marnix Naber für die gute Arbeitsatmosphäre in Zimmer 4.

Erklärung

Hiermit erkläre ich, dass ich meine Bachelorarbeit zum Thema:

**Reale und “halluzinierte” Gesichter - Gesichtsdetektion
bei schneller Reizpräsentation**

selbständig verfasst sowie alle wesentlichen
Quellen und Hilfsmittel angegeben habe.

Name, Vorname: **Abresch, Tilman**

Ort, Datum

Unterschrift

Einverständniserklärung

Ich erkläre mich damit einverstanden, dass die vorliegende Arbeit

Reale und “halluzinierte” Gesichter - Gesichtsdetektion bei schneller Reizpräsentation

in Bibliotheken allgemein zugänglich gemacht wird. Dazu gehört, dass sie

- von der Bibliothek der Einrichtung, in der ich meine Arbeit angefertigt habe, zur Benutzung in ihren Räumen bereit gehalten wird,
- in konventionellen und maschinenlesbaren Katalogen, Verzeichnissen und Datenbanken verzeichnet wird,
- der UB für die lokale Benutzung und für Fernleihzwecke zur Verfügung steht,
- im Rahmen der urheberrechtlichen Bestimmungen für Kopierzwecke genutzt werden kann.

Marburg,

Unterschrift des Autors

Unterschrift des betreuenden Hochschullehrers