

Das Fourierspektrum von Gesichtsbildern  
in Photographie und Kunst  
und dessen Einfluss auf die Gesichtswahrnehmung

Dissertation  
zur Erlangung des akademischen Grades  
doctor medicinae (Dr. med.)

**vorgelegt dem Rat der Medizinischen Fakultät  
der Friedrich-Schiller-Universität Jena**

**von Marko Blickhan  
geboren am 29. 05. 1985 in Boston (USA)**

Gutachter:

1. Prof. Dr. med Dr. rer. nat. habil. Christoph Redies, FSU Jena
2. PD Dr. habil. Holger Wiese, FSU Jena
3. Prof. Dr. Ulrich Ansorge, Universität Wien

Tag der öffentlichen Verteidigung:

03. Juli 2012

### Abkürzungsverzeichnis

CGL	Corpus geniculatum laterale
cp	cycles per image (Schwingungen pro Bild)
EEG	Elektroenzephalografie
fMRT	funktionelle Magnetresonanztomographie
HSF	high spatial frequency (hohe Ortsfrequenzen)
LSF	low spatial frequency (tiefe Ortsfrequenzen)
V1	primäre Sehrinde

## Inhaltsverzeichnis

<b>1</b>	<b>EINLEITUNG .....</b>	<b>1</b>
<b>1.1</b>	<b>Visuelles System und Ortsfrequenz-Filtermodell.....</b>	<b>2</b>
1.1.1	Sehbahn .....	2
1.1.2	Ganglienzellen und rezeptive Felder.....	3
1.1.3	Ortsfrequenz-Filtermodell des visuellen Systems.....	5
<b>1.2</b>	<b>Fourierspektrum und Kodierung natürlicher Szenen.....</b>	<b>7</b>
1.2.1	Fourierspektrum von natürlichen Szenen.....	7
1.2.2	Effiziente Kodierung von natürlichen Szenen .....	9
<b>1.3</b>	<b>Neuroästhetik.....</b>	<b>10</b>
1.3.1	Hintergrund .....	10
1.3.2	Fourierspektrum von ästhetischen Bildern.....	11
<b>1.4</b>	<b>Gesichtswahrnehmung.....</b>	<b>13</b>
1.4.1	Hintergrund .....	13
1.4.2	Gesichtsblindheit.....	15
1.4.3	Fourierspektrum von Gesichtsfotografien.....	16
<b>2</b>	<b>ZIELE DER ARBEIT .....</b>	<b>20</b>
<b>3</b>	<b>PUBLIZIERTE ORIGINALARBEITEN .....</b>	<b>22</b>
<b>3.1</b>	<b>Artists portray human faces with the Fourier statistics of complex natural scenes. Redies C, Hänisch J, Blickhan M, Denzler J. Network. 18(3):235-48. 2007.....</b>	<b>22</b>
<b>3.2</b>	<b>1/f<sup>p</sup> Characteristics of the Fourier power spectrum affects ERP correlates of face learning and recognition. Blickhan M, Kaufmann JM, Denzler J, Schweinberger SR, Redies C. Biological Psychology. 88(2-3):204-14. 2011. . .....</b>	<b>23</b>

<b>4</b>	<b>DISKUSSION</b> .....	<b>24</b>
4.1	Kritische Betrachtung der Methodik .....	25
4.2	Fourierspektrum von Porträts und Gesichtern.....	26
4.3	Fourierspektrum und Gesichterlernen .....	27
4.4	Bedeutung der Arbeit für die Gesichterforschung.....	28
4.5	Bedeutung der Arbeit für die Neuroästhetik.....	30
<b>5</b>	<b>SCHLUSSFOLGERUNGEN</b> .....	<b>32</b>
<b>6</b>	<b>LITERATUR- UND QUELLENVERZEICHNIS</b> .....	<b>33</b>
<b>7</b>	<b>ANHANG</b> .....	<b>46</b>
7.1	Fourieranalyse .....	46
7.2	EEG- und fMRT-Korrelate der Gesichterwahrnehmung.....	48
7.3	Lebenslauf .....	50
7.4	Danksagung.....	52
7.5	Eigenanteil an den Publikationen .....	53
7.6	Ehrenwörtliche Erklärung .....	54

## Zusammenfassung

**Hintergrund:** Das visuelle System ist an die Wahrnehmung natürlicher Szenen angepasst und kann diese effizient kodieren. Ein gemeinsames Merkmal natürlicher Szenen ist das Abfallen des doppellogarithmischen Fourierspektrums (Amplitudenspektrum), mit einem Anstieg von minus zwei ( $1/f^2$ -Eigenschaften). Dies wird durch hohe örtliche Korrelation und durch Skaleninvarianz von natürlichen Szenen hervorgerufen. Untersuchungen von unserer Arbeitsgruppe unter Leitung von C. Redies und J. Denzler und von der amerikanischen Arbeitsgruppe von D. Graham und D. Field zeigten, dass ästhetische gemalte Bilder ebenfalls  $1/f^2$ -Eigenschaften haben. Die Ergebnisse wurden als Hinweis interpretiert, dass Künstler ihre Werke an Eigenschaften des visuellen Systems anpassen. Bisher wurden ästhetische Bilder genutzt, die teilweise auch natürliche Szenen abbilden. Wie bilden Künstler aber Objekte ab, die andere Amplitudenspektren als natürliche Szenen besitzen? Das zeigen wir in unserer ersten Veröffentlichung anhand von Gesichtsfotografien, welche einen steileren Anstieg im Amplitudenspektrum besitzen.

Der Einfluss des Fourierspektrums auf die Gesichterwahrnehmung wurde bisher mit bandpass-gefilterten Stimuli untersucht, welche nur einen Teilbereich des Frequenzspektrums enthalten. In verschiedenen Studien wurde der mittlere Frequenzbereich als besonders wichtig hervorgehoben. Bandpass-gefilterte Stimuli unterscheiden sich deutlich in ihrer spektralen Energie, weil das Fourierspektrum mit zunehmender Frequenz abfällt. Bisher wurden noch keine Studien durchgeführt, in welchen der Einfluss des Anstieges des Fourierspektrums auf die Gesichterwahrnehmung erforscht wird. In meiner Arbeit konnte ich durch Manipulation des Anstiegs konfundierende Störvariablen wie unterschiedliche Energie von bandpass-gefilterten Stimuli vermeiden. Des Weiteren konnte durch diese Manipulation der relative Anteil von groben und feinen Strukturen im Bild verändert werden. In unserer zweiten Veröffentlichung zeigten wir in einem Lernexperiment, wie das Erlernen und Erkennen unbekannter Gesichter durch Manipulation von  $1/f^p$ -Eigenschaften des Fourierspektrums beeinflusst wurde.

**Methodik:** Für die erste Veröffentlichung wurden 300 gemalte ästhetische Porträts von namhaften Künstlern digitalisiert. Fotografien von Gesichtern erhielten wir von der Yale- und der „AR Face“-Datenbank. Der Anstieg des Fourierspektrums von Porträts und Gesichtsfotografien wurde, wie in zuvor durchgeführten Studien unserer Arbeitsgruppe, mit Matlab ermittelt und verglichen.

In der zweiten Veröffentlichung erstellten wir zwei Gruppen von Gesichtsfotografien mit veränderten  $1/f^p$ -Eigenschaften: Zum einen Gesichter mit steilerem Anstieg, wodurch grobe Strukturen verstärkt und feine abgeschwächt wurden; zum anderen Gesichter mit flacherem Anstieg und  $1/f^2$ -Eigenschaften mit reduzierten groben und verstärkten feinen Strukturen. In einem Gesichter-Lernexperiment wurden Verhaltensdaten und Elektroenzephalografie-Korrelate der Gesichterwahrnehmung untersucht.

**Ergebnisse und Schlussfolgerungen:** Unsere erste Studie zeigte, dass ästhetische gemalte Porträts  $1/f^2$ -Eigenschaften haben, die denen natürlicher Szenen ähnlich sind und sich in dieser Hinsicht deutlich von Gesichtsfotografien unterscheiden. Die Unabhängigkeit der  $1/f^2$ -Eigenschaften der ästhetischen Porträts von Ursprungsland, Jahrhundert und verwendeter Technik war ein weiterer Hinweis, dass Künstler ihre Abbildungen an Kodierungsmechanismen des visuellen Systems anpassen und nicht die Eigenschaften der Objekte abbilden, welche diese natürlicherweise besitzen.

In dem Gesichter-Lernexperiment der zweiten Studie konnten Fotos mit verstärkten groben Strukturen und verminderten feinen Strukturen schlechter gelernt werden. Es zeigten sich langsamere Reaktionszeiten und im Elektroenzephalogramm verminderte neurophysiologische Korrelate der Gesichterwahrnehmung (N170, N250 und central positivity). Im Gegensatz dazu konnten Gesichtsfotos mit reduzierten groben und verstärkten feinen Strukturen sowie  $1/f^2$ -Eigenschaften, die gemalten Porträts und natürlichen Szenen ähnlich sind, leichter gelernt werden und es fanden sich größere neurophysiologische Korrelate der Gesichterwahrnehmung (N170, N250 und central positivity). Unbekannte Gesichter mit steilem Anstieg, welche mehr Energie im mittleren Frequenzbereich hatten, wurden schlechter gelernt als Gesichter mit flachem Anstieg. Dies widerspricht bisherigen Annahmen, dass der mittlere Frequenzbereich am wichtigsten für die Erkennung von Gesichtern sei.

Es wird daraus gefolgert, dass  $1/f^2$  Eigenschaften, wie sie bei natürlichen Szenen auftreten, das Erlernen und Erkennen von Gesichtern erleichtern können.

# 1 Einleitung

In der Neuroästhetik werden neurobiologische Grundlagen für die Wahrnehmung ästhetischer Stimuli untersucht. Insbesondere gibt es einen Zusammenhang zwischen Funktionsweisen des visuellen Systems und den Eigenschaften grafischer Kunstwerke. Das erste Experiment (siehe Kapitel 3.1) der vorliegenden Arbeit beschäftigte sich mit Eigenschaften im Fourierspektrum von ästhetischen gemalten Porträts, welche die Kodierung im visuellen System erleichtern könnten. Das zweite Experiment (siehe Kapitel 3.2) untersuchte die Auswirkungen dieser Eigenschaften im Fourierspektrum auf die Wahrnehmung von Gesichtern. Diese Arbeit liefert einen Baustein für den Zusammenhang zwischen elementaren Funktionsweisen des visuellen Systems und höheren kognitiven Funktionen am Beispiel der Gesichtserkennung, welche ihrerseits eine fundamentale Bedeutung für die soziale Interaktion darstellt.

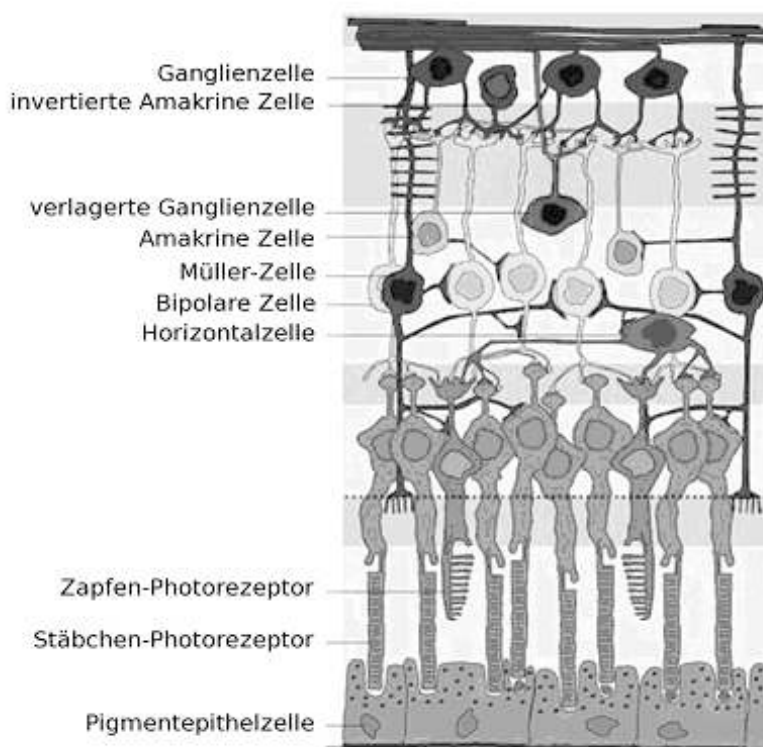
In diesem Abschnitt erörtere ich zunächst den Zusammenhang zwischen der Physiologie des visuellen Systems und Ortsfrequenz- und Fourieranalysen. Insbesondere wird die Anpassung des visuellen Systems an die Wahrnehmung natürlicher Szenen erläutert und Gemeinsamkeiten der Eigenschaften im Fourierspektrum natürlicher Szenen und ästhetischer Bilder diskutiert. Des Weiteren stelle ich die Bedeutung des Fourierspektrums für die Neuroästhetik und für die Gesichtswahrnehmung dar.



## 1.1 Visuelles System und Ortsfrequenz-Filtermodell

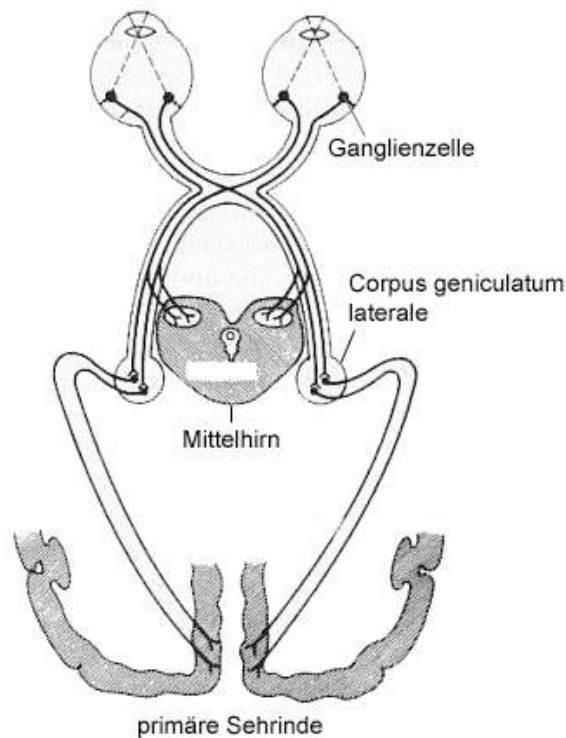
### 1.1.1 Sehbahn

Bereits in der Retina (siehe Abb. 1) werden visuelle Informationen mit unterschiedlicher örtlicher Auflösung verarbeitet. Zum einen legen Dichte und Größe der Fotorezeptoren fest, mit welcher Auflösung unterschiedliche Lichtsignale wahrgenommen werden. Zum anderen kommt es durch die Verschaltung zwischen Fotorezeptoren und nachfolgenden Zellschichten zu unterschiedlicher Ortsauflösung.



**Abb. 1 Zellulärer Aufbau der Retina. Ganglienzellen bilden das erste Neuron der Sehbahn. Horizontalzellen sind an der Verschaltung der rezeptiven Felder beteiligt. Modifiziert von J. Etzrodt nach P. Hartmann de.wikipedia.org 2009 (GFDL-Lizenz)**

In der Fovea centralis, dem Bereich des schärfsten Sehens erfolgt eine 1:1-Verschaltung von einem Fotorezeptor auf eine Ganglienzelle (Prasad und Galetta 2011). In der Peripherie werden mehrere Fotorezeptoren mit einer Ganglienzelle verschaltet (siehe Kapitel 1.1.2). Ganglienzellen projizieren mit einer nahezu 1:1-Umschaltung auf Zellen des Corpus geniculatum laterale (CGL, siehe Abb. 2), welches retinotopisch organisiert ist.

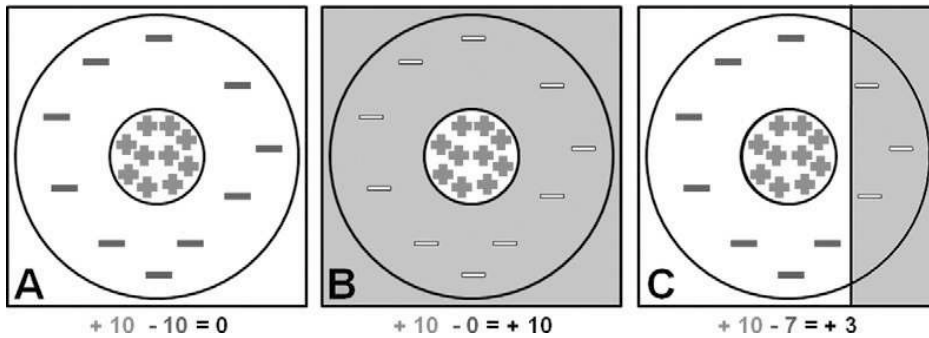


**Abb. 2 Sehbahn. Neuronaler Verlauf vom Auge zum Gehirn. Eine neuronale Umschaltung erfolgt in den Ganglienzellen der Retina, im CGL und in der primären Sehrinde. Modifiziert nach Guhr (2002)**

Die Afferenzen aus der Retina bilden nur einen kleinen Teil der synaptischen Verschaltung im CGL. Die Mehrheit der Afferenzen wird aus dem Nucleus reticularis thalami und aus dem visuellen Kortex erhalten. Dadurch werden visuelle Informationen moduliert, bevor sie zum Großhirn gelangen. Vom CGL projiziert die Sehstrahlung zur primären Sehrinde (V1) im Okzipitallappen. Auch in V1 bleibt die retinotopische Gliederung erhalten, wobei es zu einem starken Anstieg der Neuronenzahl und einer Überrepräsentierung der Fovea centralis kommt.

### 1.1.2 Ganglienzellen und rezeptive Felder

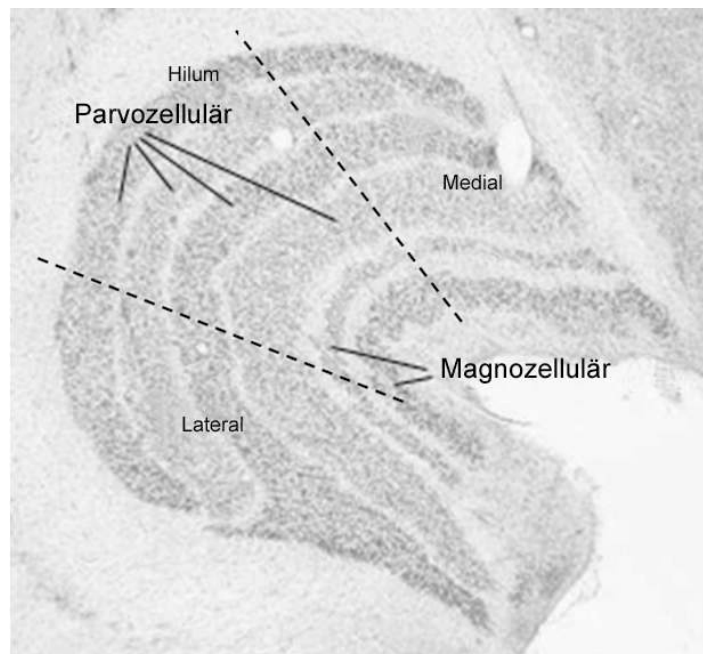
In der Retina konvergieren  $10^8$  Fotorezeptoren auf  $10^6$  Ganglienzellen. Dabei erhält eine Ganglienzelle visuelle Informationen von zahlreichen Fotorezeptoren, die zu einem rezeptiven Feld zusammengefasst werden (Prasad und Galetta 2011). Rezeptive Felder bestehen aus Zentrum und Peripherie (siehe Abb. 3), welche durch laterale Inhibition der Horizontalzellen (siehe Abb. 1) antagonistisch verschaltet sind.



**Abb. 3** Rezeptives Feld einer „On“-Zentrum Ganglienzelle. (A) Bei Beleuchtung des gesamten rezeptiven Feldes gleichen sich neuronale Hemmung und Erregung aus. (B) Beleuchtung des Zentrums führt zu starker neuronaler Erregung. (C) Beleuchtung des Zentrums und eines Teils der Peripherie führt zu geringer neuronaler Erregung. Aus Prasad und Galetta (2011)

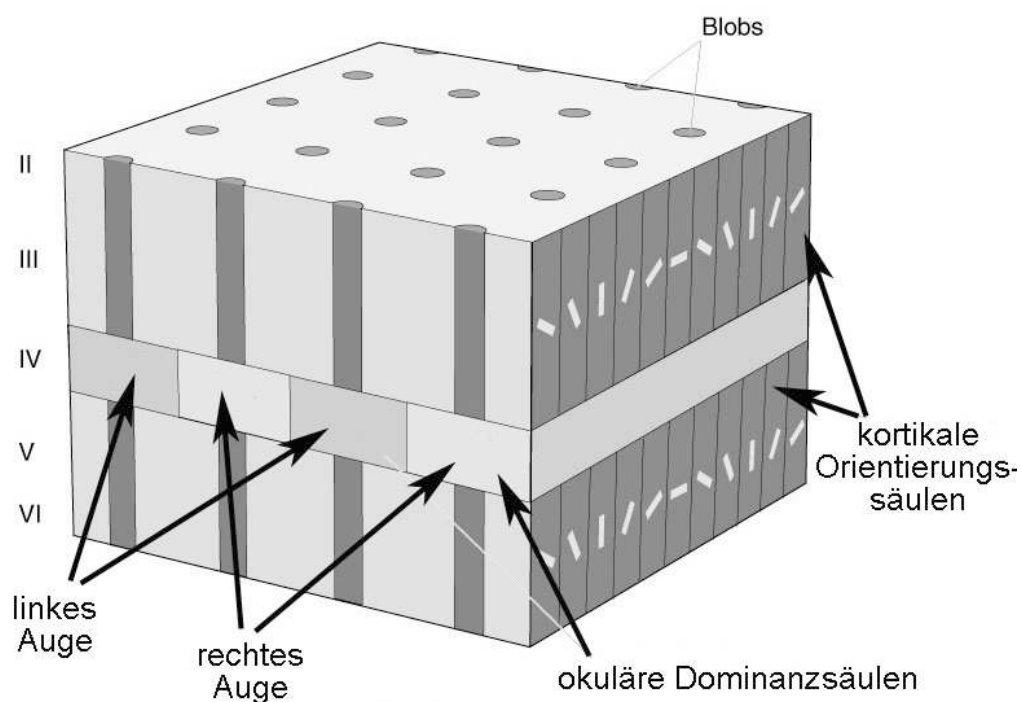
Nach anatomisch-physiologischen Gesichtspunkten lassen sich hauptsächlich zwei Ganglienzelltypen unterscheiden, die in unterschiedlichen Signalwegen zu finden sind:

Im magnozellularen Signalweg projizieren Parasol-Ganglienzellen, die große rezeptive Felder und damit eine geringe Ortsauflösung besitzen, auf magnozellige Schichten im CGL (siehe Abb. 4). Von dort gelangen diese Signale zu der Schicht 4Ca der primären Sehrinde. Dieser Signalweg dient der Wahrnehmung von Bewegung, grober räumlicher Auflösung und groben Strukturen.



**Abb. 4** Parvozelluläre und magnozellige Schichten des Corpus geniculatum laterale. Modifiziert nach Prasad und Galetta (2011)

Der parvozelluläre Signalweg führt zu großer Sehschärfe und ermöglicht die Wahrnehmung von Farben, detaillierter räumlicher Auflösung und feinen Strukturen. Dies wird durch Midget-Zellen ermöglicht, die im Gegensatz zu Parasol-Zellen kleine rezeptive Felder besitzen. Dieser Signalweg projiziert auf die Schicht 4C $\beta$  der primären Sehrinde. Man geht davon aus, dass V1 aus rezeptiven Feldern besteht, die in Orientierungssäulen angeordnet sind (Hubel und Wiesel 1962). Sie reagieren auf die Ausrichtung einer Linie oder einer Kante (siehe Abb. 5). Des Weiteren können Zellen in V1 durch Informationen außerhalb des eigenen rezeptiven Feldes beeinflusst werden (siehe Kapitel 1.1.3, De Valois und De Valois 1980).



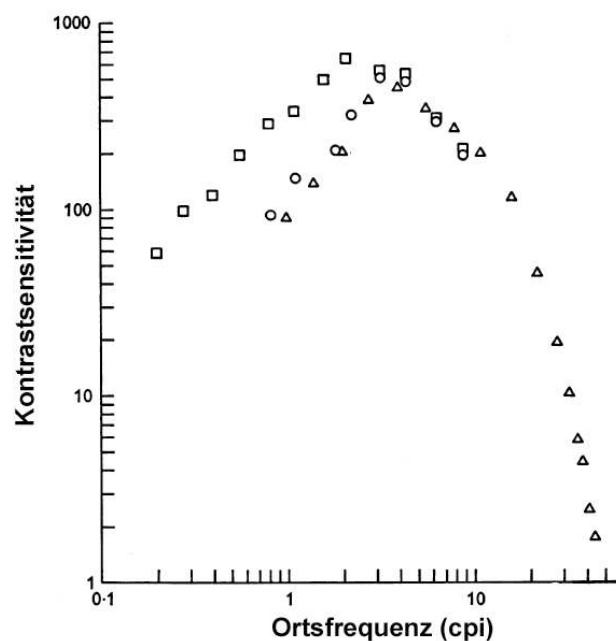
**Abb. 5** Kortikale Säulen in V1 reagieren auf eine bestimmte räumliche Orientierung. Modifiziert von H. G. Hinghofer-Szalkay nach Bear et al. (2007)

### 1.1.3 Ortsfrequenz-Filtermodell des visuellen Systems

Die örtliche Auflösung von Fotorezeptoren und ihre Organisation zu rezeptiven Feldern mit unterschiedlicher Größe führen dazu, dass visuelle Informationen zunächst in Form eines passiven Ortsfrequenzfilters verarbeitet werden (Westheimer 2001).

Erste psychologische Experimente lieferten Hinweise, dass das visuelle System Informationen ähnlich einer Fouriertransformation (siehe Kapitel 7.1) in verschiedene

Ortsfrequenzen zerlegt (Campbell und Robson 1968). In diesem Experiment wurden Gitter verschiedener Frequenz mit unterschiedlich hoher Kontrastschwelle wahrgenommen (siehe Abb. 6). Die Autoren schlussfolgerten, dass es Kanäle für unterschiedliche Ortsfrequenzen gibt, welche für eine bestimmte Bandbreite selektiv sind. In Studien mit invasiver Mikroelektrodenableitung konnte die Ortsfrequenz-Bandbreite von Neuronen in V1 direkt bestimmt werden (Maffei und Fiorenti 1973). Diese Ergebnisse wurden von De Valois zu einem Ortsfrequenz-Filtermodell zusammengefasst, dem zufolge visuelle Informationen in V1 mit einem linearen Ortsfrequenzfilter in ihre Ortsfrequenz-Komponenten zerlegt werden (De Valois und De Valois 1980).



**Abb. 6 Kontrastsensitivität für unterschiedliche Ortsfrequenzen. Die Kontrastschwelle steigt zunächst mit steigender Ortsfrequenz und sinkt anschließend wieder ab. □ 57 cm Abstand und 10° Öffnungswinkel. ○ 57cm Abstand und 2° Öffnungswinkel. ▲ 285 cm Abstand und 2° Öffnungswinkel. cpi (cycles per image) Schwingungen pro Bild. Modifiziert nach Campbell und Robson (1968)**

In V1 entspricht die Orientierung des rezeptiven Feldes der Orientierung der Ortsfrequenz-Komponente. Die Größe des erregenden und hemmenden Bereichs des rezeptiven Feldes bedingt die Ortsfrequenz (Barlow und Olshausen 2004). Diese rezeptiven Felder werden auch durch parallel zum eigenen Feld verlaufende Felder beeinflusst (De Valois und De Valois 1980, Geisler 2008). Damit hat man ein lineares Modell für das rezeptive Feld, das man zum Beispiel mit Hilfe der Fouriertransformation analysieren und

modellieren kann. Weitere Untersuchungen zeigten jedoch zahlreiche nichtlineare Phänomene (Carandini et al. 2005). Demnach müssen Modelle für visuelle Informationsverarbeitung sowohl lineare als auch nichtlineare Phänomene berücksichtigen (Graham 2011, Carandini et al. 2005).

## **1.2 Fourierspektrum und Kodierung natürlicher Szenen**

### **1.2.1 Fourierspektrum von natürlichen Szenen**

Fotographien natürlicher Szenen werden oft zur Untersuchung des visuellen Systems genutzt, da sie komplexe und natürliche Reize darstellen. Natürliche Szenen werden hier definiert als Landschaftsaufnahmen, welche keine von Menschen hergestellten Objekte und keine Nahaufnahmen von Pflanzen oder Tieren enthalten. Nach Simoncelli und Olshausen (2001) ist das sensorische System sowohl durch Evolution also auch Entwicklung an die Wahrnehmung der natürlichen Umgebung angepasst. Insbesondere soll das visuelle System natürliche Szenen mit minimaler Redundanz repräsentieren können (Field 1987, Willmore et al. 2011, Tolhurst et al. 2009). Daraus ergibt sich erstens die Frage nach den Mechanismen dieser redundanten Repräsentation (siehe Kapitel 1.2.2) und zweitens welche statistischen Eigenschaften natürlicher Szenen das visuelle System zur redundanten Repräsentation nutzen kann (siehe unten in diesem Kapitel).

Natürliche Szenen zeichnen sich durch hohe örtliche Korrelation der Intensität aus (Simoncelli und Olshausen 2001). Das heißt, dass die Intensität der Pixel um einen bestimmten Ort sehr ähnlich ist und bei zunehmendem Abstand von diesem Ort weniger ähnlich ist (siehe Abb. 7). Die Korrelation der Intensität an zwei Punkten ist damit nur von ihrem Abstand abhängig, jedoch unabhängig von der absoluten Position. Die örtliche Korrelation bewirkt, dass das eindimensionale doppellogarithmische Fourierpowerspektrum (siehe Kapitel 7.1) natürlicher Szenen bei steigender Frequenz mit einem Anstieg von  $p=-2$  abfällt ( $1/f^2$ -Eigenschaften, siehe Abb. 8, Burton und Moorhead 1987, Field 1987, Tolhurst et al. 1992, Ruderman und Bialek 1994, Simoncelli und Olshausen 2001). Man geht davon aus, dass diese Eigenschaft außerdem durch Skaleninvarianz unserer visuellen Umgebung hervorgerufen wird. Skaleninvarianz bedeutet, dass bei Ein- und Auszoomen die statistischen Bildeigenschaften und damit der Anstieg des Fourierspektrums des jeweiligen Bildes gleich bleiben.

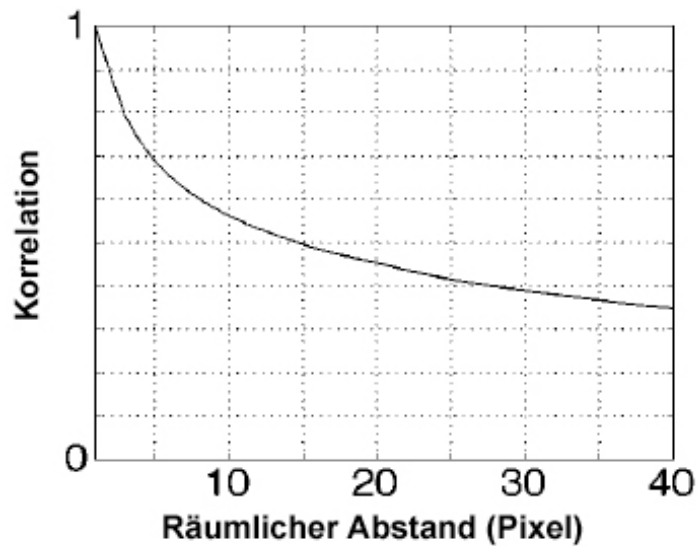


Abb. 7 Örtliche Korrelation bei natürlichen Szenen. Geringer räumlicher Abstand führt zu hoher Korrelation. Großer räumlicher Abstand führt zu geringer Korrelation. Modifiziert nach Simoncelli und Olshausen (2001)

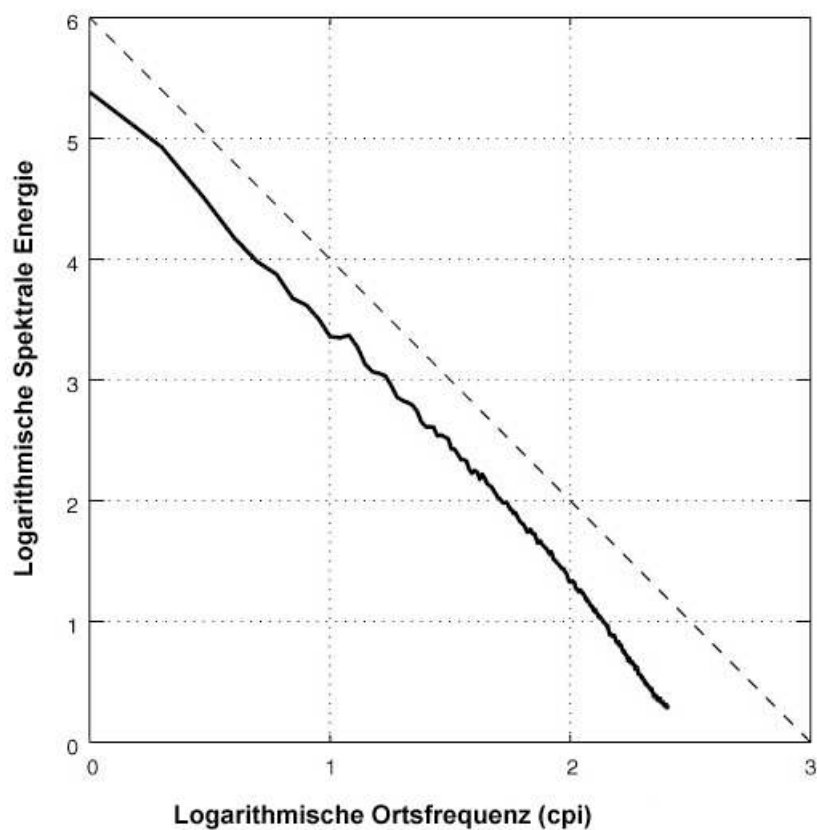


Abb. 8 Eindimensionales Fourier-spektrum. Die Gestrichelte Linie zeigt einen idealen Anstieg von -2. Die durchgezogene Linie zeigt die gemittelten Fourierspektren mehrerer natürlicher Szenen. Modifiziert nach Simoncelli und Olshausen (2001)

Neben der örtlichen Korrelation gibt es weitere gemeinsame Eigenschaften natürlicher Szenen wie das Orts-Zeit-Verhältnis (van Hateren und Ruderman 1998), die Farbstatistik (Webster und Mollon 1997, Hoyer und Hyvärinen 2000) und das Kontrast-Intensitäts-Verhältnis (Brady und Field 2000, Frazor und Geisler 2006). Das visuelle System kann diese statistischen Eigenschaften zur effizienten Kodierung nutzen (siehe Kapitel 1.2.2). In dieser Arbeit soll der Einfluss des Anstiegs des Fourierspektrums auf das visuelle System erörtert werden.

### **1.2.2 Effiziente Kodierung von natürlichen Szenen**

Natürliche Szenen können vom visuellen System effizient codiert werden (Graham und Field 2009). Dabei unterscheidet man die repräsentative, die metabolische und die Lern-Effizienz im visuellen System.

*Repräsentative Effizienz* beschreibt die Eigenschaften von Neuronen, mit geringer Anzahl und geringem Aufwand Stimuli zu verarbeiten. Hierzu werden Korrelationen der Stimuli ausgenutzt. So können beispielsweise bei der Verarbeitung natürlicher Szenen Informationen mit reduzierter Dimension repräsentiert werden, das heißt, dass weniger Neurone nötig sind oder ein reduziertes Antwortmuster vorliegt. Eine weitere Eigenschaft natürlicher Szenen im Gegensatz zu weißem Rauschen (siehe 1.3.2) ist, dass eine bestimmte Anordnung linearer Filter zu Antworten führen kann, welche keine Normalverteilung aufweisen. Eine starke Abweichung von der Normalverteilung führt zu einem unabhängigen Code (Graham und Field 2009). Bei einem unabhängigen Code repräsentieren möglichst wenige Neuronen die gleichen Informationen, so dass eine große Anzahl der Neuronen inaktiv und nur wenige aktiv sind. Diese Eigenschaft der Kodierung wird auch als „sparse coding“ bezeichnet. Neuronale Netze des primären visuellen Kortex sind an Redundanz von natürlichen Szenen angepasst und können diese „sparse“ repräsentieren (Hoyer und Hyvärinen 2000, Taylor et al. 2000, Wachtler et al. 2001, Lee et al. 2002, Caywood et al. 2004, Olshausen und Field 1996). „Sparseness“ konnte ebenfalls bei Untersuchung von Neuronen im visuellen Cortex von Frettchen (Tolhurst et al. 2009) und Primaten gezeigt werden (Vinje und Gallant 2000, Vinje und Gallant 2002, Willmore und Tolhurst 2001, David et al. 2004, Willmore et al. 2011).

*Metabolische Effizienz* beschreibt die Kosten der Kodierung, wie sie z.B. durch Aktionspotentiale hervorgerufen werden. Nach Attwell und Laughlin (2001) und Lennie



(2003) führen metabolische Kosten im Vergleich zu verfügbaren metabolischen Reserven dazu, dass die Frequenz von Aktionspotentialen auf unter 1Hz beschränkt ist. Demnach wird Effizienz im visuellen System sowohl nach repräsentativen als auch metabolischen Gründen durch die reduzierte Zahl aktiver Neurone und reduzierte Aktionspotentialraten charakterisiert.

*Lern-Effizienz* beschreibt die Eigenschaft des Nervensystems, dass es durch Entwicklung und Lernen entsteht. Dadurch kann es an verschiedene Aufgaben angepasst werden. Um zu Lernen müssen Neurone zum einen selektiv für bestimmte Eigenschaften sein und zum anderen dürfen sie nicht von Eigenschaften, die sich in irrelevanter Art verändern, beeinflusst werden. Das Nervensystem muss also über veränderbare Größen generalisieren können, da ein Gesicht z. B. selten in der gleichen Beleuchtung wahrgenommen wird. Ein streng „sparser“ Code würde ein Neuron für jedes Objekt entwickeln, auch Grossmutter-Zell-Hypothese genannt (Barlow 1972). Allerdings würde das visuelle System dafür zu viele Objekterkennungs-Neurone benötigen (Gross 2002), sodass Objekterkennungsstrategien in ihrer „sparseness“ variieren müssen (Riesenhuber und Poggio 2000). Lern-Effizienz führt zu Strategien, welche von metabolischer und repräsentativer Effizienz abweichen können. Hierbei kommt es zu einem komplexen Zusammenspiel von angeborenen und erlernten Eigenschaften des Nervensystems.

## **1.3 Neuroästhetik**

### **1.3.1 Hintergrund**

Es gibt verschiedene Vorstellungen über die Grundlagen ästhetischer Wahrnehmung. Dabei werden zum einen soziokulturelle Faktoren in den Vordergrund gestellt (Danto 1981, Goodman 1976) und zum anderen universelle Aspekte betrachtet, welche unabhängig von soziokulturellen Faktoren für alle Menschen gleich sind (Schelling 1907, Burke 1757, Hume 1826, Kant 1790, Paul 1988, Adorno 1970). Der Begriff „ästhetische Bilder“ wird von mir wie bei Redies (2008) zur Unterscheidung von kontemporären nicht-ästhetischen Kunstwerken genutzt. Nicht-ästhetische Kunst soll eher durch soziokulturelle Faktoren, Philosophie und Psychologie auf den Betrachter wirken und hat nicht den Anspruch ästhetische Wahrnehmung hervorzurufen.

Universelle Aspekte ästhetischer Wahrnehmung werden in jüngster Zeit von Neurowissenschaftlern untersucht und es wird ergründet, wie ästhetische Erfahrungen mit

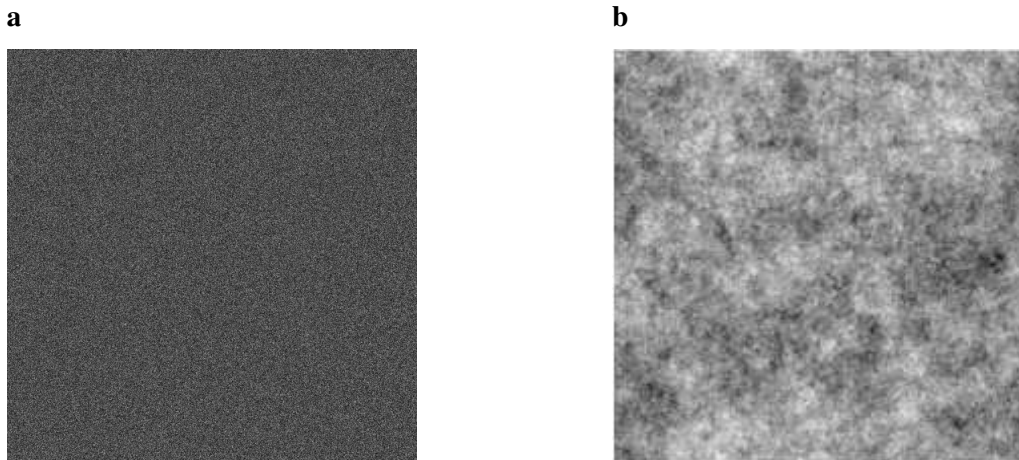
physiologischen Funktionen des Gehirns zusammenhängen (Zeki 1999, Livingstone 2002). Verschiedene Kunstrichtungen nutzen bestimmte Eigenschaften des visuellen Systems aus, die jedoch meist auf eine spezielle Kunstrichtung beschränkt sind und nicht für verschiedene Richtungen verallgemeinert werden können (Goguen und Myin 2004). Zwischen der Ästhetik von natürlichen Szenen und ästhetischen gemalten Bildern wird auf philosophischer Seite ein Zusammenhang postuliert (Adorno 1970, Hegel 1833, Kant 1790, Koppe 2004, Paul 1988, Paul 1998). Visuelle Kunstwerke und natürliche Szenen müssen zunächst vom visuellen System, welches an die Wahrnehmung natürlicher Stimuli angepasst ist (siehe Kapitel 1.2.2), verarbeitet werden. Künstler haben die Möglichkeit, Werke an ihr visuelles System anzupassen (Graham und Redies 2010, Zeki 1999). Unsere eigenen Untersuchungen (Redies et al. 2007) sowie die einer amerikanischen Arbeitsgruppe (Graham und Field 2007) haben gemeinsame statistische Eigenschaften bei gemalten ästhetischen Bildern und bei natürlichen Szenen nachgewiesen (siehe Kapitel 1.3.2).

### **1.3.2 Fourierspektrum von ästhetischen Bildern**

In zwei unabhängigen Untersuchungen von ästhetischen Bildern konnte gezeigt werden, dass deren Fourierspektren denen von natürlichen Szenen ähneln (Graham und Field 2007, Redies et al. 2007). Dabei wurde für ästhetische Bilder ein Anstieg von  $-2$  im eindimensional gemittelten doppellogarithmischen Fourierspektrum gefunden ( $1/f^2$ -Eigenschaften, siehe Kapitel 7.1) und geschlossen, dass ästhetische Bilder auch skaleninvariante Eigenschaften aufweisen. Dies konnte sowohl für ästhetische Bilder der westlichen Hemisphäre als auch der östlichen Hemisphäre gezeigt werden. In diesen Studien wurden nur ästhetische Bilder untersucht, da die Wahrnehmung nicht-ästhetischer Bilder wie zum Beispiel in Zeitgenössischer Kunst eher durch soziokulturelle Aspekte beeinflusst wird (Danto 1981, Goodman 1976, Redies 2007).

Eine Sonderform der Skaleninvarianz sind Fraktale, die per definitionem auch im Ortsbereich selbstähnlich sind. Es gibt eine Präferenz für fraktale Eigenschaften sowohl bei natürlichen Szenen als auch bei Computer- und von Menschenhand generierten Objekten (Aks und Sprott 1996, Spehar et al. 2003, Hagerhall et al. 2004, Sprott 1993). Solche Eigenschaften konnten beispielsweise für Gemälde von Jackson Pollock gezeigt werden (Taylor et al. 1999, Taylor 2002). Fraktale Eigenschaften sind jedoch keinesfalls ausreichend

für ästhetische Wahrnehmung. Computergenierte selbstähnliche Muster können nicht ästhetisch sein (Lee et al. 2001, Ruderman 1997).



**Abb. 9 (a) Weißes Rauschen mit flachem Fourierspektrum. Modifiziert nach Bautsch M. de.wikipedia.org 2006. (b)  $1/f^2$  Rauschen. Aus Simoncelli und Olshausen (2001)**

Die statistischen Ähnlichkeiten zwischen natürlichen Szenen und ästhetischen Bildern sind ein Zeichen dafür, dass Künstler ihre Werke an das visuelle System anpassen (Graham und Redies 2010, Zeki 1999). Des Weiteren könnten diese Eigenschaften mit der „Induzierung“ von ästhetischem Empfinden assoziiert sein (Redies 2007). Diesbezüglich konnte gezeigt werden, dass Bilder mit Fourierspektren welche von natürlichen Szenen abweichen, als unangenehm wahrgenommen werden (Fernandez und Wilkins 2008, O'Hare und Hibbard 2011, Juricevic et al. 2010).  $1/f^2$ -Eigenschaften sind jedoch nicht ausreichend, um ästhetisches Empfinden hervorzurufen. Beispielsweise kann man ein zweidimensionales weißes Rauschsignal (flaches Fourierspektrum) mit einem zweidimensionalen  $1/f^2$ -Rauschsignal vergleichen (siehe Abb. 9). Das  $1/f^2$ -Rauschsignal hat keinerlei Ähnlichkeit mit einer natürlichen Szene oder einem ästhetischen Bild und wird nicht als ästhetisch wahrgenommen.

Die ästhetischen Bilder, welche in den beiden oben genannten Studien (Graham und Field 2007, Redies et al. 2007) untersucht wurden, bildeten zu einem Teil natürliche Szenen ab. Dies wäre eine einfache Erklärung für die ähnlichen Fourierspektren von ästhetischen Bildern und natürlichen Szenen. Gesichtsfotografien besitzen dagegen einen steileren Anstieg im Fourierspektrum als natürliche Szenen (Bosworth et al. 2006, Torralba und Oliva 2003). Daher ist der Vergleich der Fourierspektren von Gesichtsfotografien und gemalten

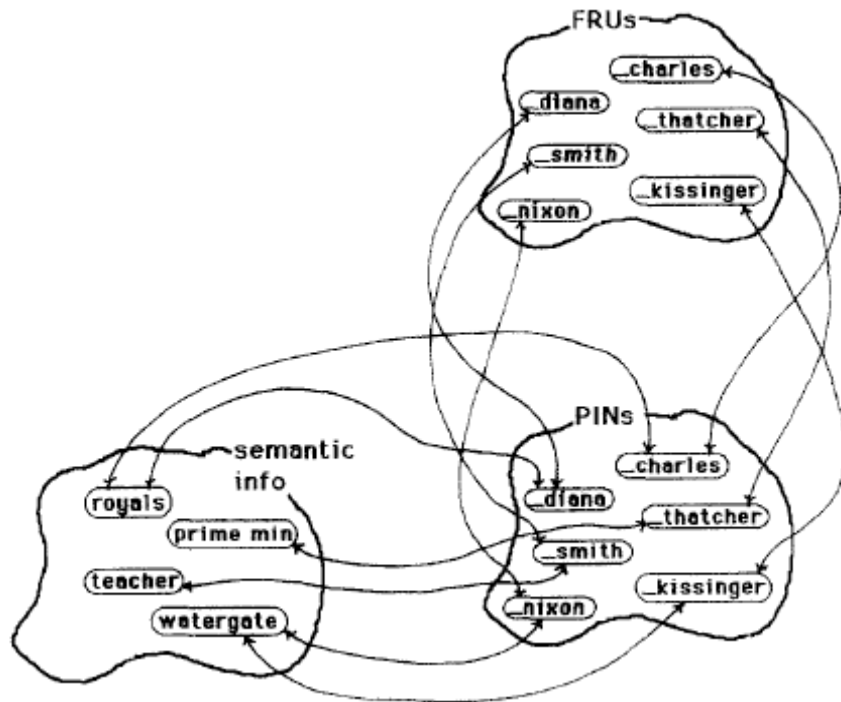
ästhetischen Porträts besonders interessant. Wir formulierten die Hypothese, dass gemalte ästhetische Porträts ähnliche Fourierspektren wie andere ästhetische Bilder und natürliche Szenen haben und sich darin von Gesichtsfotografien unterscheiden.

## **1.4 Gesichtswahrnehmung**

### **1.4.1 Hintergrund**

Gesichterwahrnehmung ist ein wichtiger Bestandteil der alltäglichen sozialen Interaktion. Auf der einen Seite kann man die Identität von Personen anhand des Gesichtes erkennen und auf der anderen Seite werden zahlreiche Informationen z.B. über Emotionen, Alter und Gesundheit anhand des Gesichtes wahrgenommen. In der Gesichtswahrnehmungsforschung werden drei Ansätze unterschieden: kognitiver, neurophysiologischer und psychophysikalischer Ansatz (Ruiz-Soler und Beltran 2006).

Im *kognitiven* Ansatz wurden mit funktionellen Modellen die verschiedenen Stadien der Gesichterwahrnehmung beschrieben. Eine wichtige Grundlage war das Modell von Bruce und Young (1986), welches die folgende Schritte beinhaltet: Zunächst kommt es zu einer bildlichen Kodierung der visuellen Informationen, die durch strukturelle Kodierung zu einer holistischen Repräsentation des Gesichtes verarbeitet werden und mit „face recognition units“ im visuellen System verglichen werden können. Wenn holistische Repräsentation und „face recognition unit“ übereinstimmen, wird ein Gesicht erkannt und auf zusätzliche semantische und episodische Information über die Person, zu der das Gesicht gehört, zugegriffen. Die semantischen Informationen werden in „person identity nodes“ kodiert. Schließlich wird ermöglicht, dass auch der Namen der Person zugeordnet werden kann. Dieses Modell konnte jedoch einige Phänomene beim Zugriff auf semantische Informationen wie z.B. semantisches Priming nicht erklären (Valentine et al. 1996). Beim semantischen Priming kann ein Gesicht leichter erkannt werden, wenn zuvor ein semantisch zugehöriger anderer Stimulus gezeigt wird. Das Modell von Bruce und Young wurde zu einem interaktiven Aktivationsmodell (IAC-Modell- „interactive activation and competition“) weiterentwickelt und modifiziert (Burton et al. 1999). Hierbei gibt es unterschiedliche funktionelle Neuronen-gruppen wie z.B. „face recognition unit“, „person identity nodes“ und „semantic information unit“, welche sich gegenseitig aktivieren können. Nach Burton spielen die „person identity nodes“ eine zentrale Rolle bei der Erkennung von Personen (siehe Abb. 10).



**Abb. 10 Interaktives Aktionsmodell (IAC-Modell).** FRU- „face recognition unit“. PIN-„person identity node“. Aus Burton et al. (1990)

Im Unterschied zu dem zuvor beschriebenen Modell von Bruce enthalten die „person identity nodes“ keine semantischen Informationen, sondern sind die Knotenpunkte, die aktiviert werden müssen, um Zugang zu weiteren semantischen Informationen zu erhalten („semantic information units“). Demnach werden Gesichter zunächst über „face recognition units“ wahrgenommen und auf dem Niveau der „person identity nodes“ einer Person zugeordnet. Dadurch können verschiedene Informationsquellen wie z.B. Gesicht, Stimme, Name oder Beruf auf dem Niveau der „person identity nodes“ konvergieren und zu der Erkennung einer Person beitragen.

In *neurophysiologischen* Studien wurden physiologische Korrelate der Gesichterwahrnehmung mit Hilfe des EEGs und der funktionellen Magnetresonanztomographie (fMRT) erforscht (siehe Kapitel 7.2). Diese Studien konnten den zeitlichen und räumlichen Verlauf der zuvor beschriebenen kognitiven Modelle näher erklären (Schweinberger und Burton 2003, Gobbini und Haxby 2007). An der Gesichterwahrnehmung ist ein Netzwerk verschiedener Gehirnareale beteiligt, das aus dem Sulcus temporalis superior, der „fusiform face area“ und der „occipital face area“ besteht (siehe Kapitel 7.2). Bei einer Schädigung

dieser Areale kann es zu einer Gesichtsblindheit (Prosopagnosie) kommen (siehe Kapitel 1.4.2).

Im *psychophysikalischen* Ansatz wurde die Bedeutung des Fourierspektrums für die Gesichterwahrnehmung untersucht (siehe Kapitel 1.4.3).

## **1.4.2 Gesichtsblindheit**

Gesichtsblindheit (Prosopagnosie) bezeichnet die Unfähigkeit, bekannte Personen anhand von Gesichtern zu erkennen und unbekannte Gesichter zu erlernen, wobei die Wahrnehmung anderer Gesichtsmerkmale wie Ausdruck, Emotionen und Blickrichtung nicht gestört ist (Bodamer 1947, Grüter et al. 2008). Man unterscheidet eine erworbene und eine kongenitale Prosopagnosie.

Bei der erworbenen Prosopagnosie führt eine strukturelle Schädigung des Gehirns, zum Beispiel ein Trauma oder ein Schlaganfall, zu der gestörten Gesichterwahrnehmung. Durch die Schädigung kommt es häufig zu anderen kognitiven Defiziten wie visuelle Agnosie, Farbblindheit, Quadrantenanopsie oder gestörte Objekterkennung (Grüter et al. 2008, Kennerknecht et al. 2006). Einzelfälle, bei denen nur die Gesichtserkennung und nicht die Objekterkennung betroffen ist, geben Hinweise darauf, dass es Gesichter-spezifische Gehirnregionen und Funktionen gibt. So konnten Busigny et al. (2010) einen Patienten in ihrer Studie untersuchen, der ähnliche Objekte im Gegensatz zu Gesichtern problemlos unterscheiden konnte.

Die kongenitale Prosopagnosie ist angeboren und nicht mit einer strukturellen Schädigung des Gehirns verbunden (De Haan 1999, Grüter et al. 2008, Kennerknecht et al. 2006). Bei Betroffenen von dieser Krankheit kann eine gestörte Objekt- und Szenenerkennung vorliegen, während Farbblindheit und Quadrantenanopsie nicht assoziiert sind.

Sowohl für die kongenitale als auch die erworbene Prosopagnosie konnte gezeigt werden, dass die holistische Gesichtswahrnehmung gestört ist (Busigny und Rossion 2010, Stollhoff et al. 2010), sodass die Gesichtsmerkmale einzeln analysiert werden (Orban de Xivry et al. 2008).

### 1.4.3 Fourierspektrum von Gesichtsfotografien

In der ersten Studie dieser Arbeit (siehe Kapitel 3.1) wurde gezeigt, in welcher Art und Weise Künstler  $1/f^p$ -Eigenschaften von Gesichtern verändern. Anschließend untersuchten wir, welchen Einfluss die Veränderung des Fourierspektrums auf die Gesichterwahrnehmung hat. Dabei wurden  $1/f^2$ -Eigenschaften von ästhetischen gemalten Porträts mit Kodierungsmechanismen im visuellen System in Zusammenhang gebracht. Wir spekulierten, dass diese Eigenschaften das Erlernen und Erkennen von Gesichtern beeinflussen könnten.

Aufgrund der Erkenntnis, dass das visuelle System Informationen in Ortsfrequenz-Komponenten zerlegt (siehe Kapitel 1.1.3), untersuchten Studien des psychophysikalischen Ansatzes, welcher Anteil des Frequenzspektrums zur Gesichtererkennung nötig ist (Ruiz-Soler und Beltran 2006). Es wurden meist Bandpass-gefilterte Stimuli genutzt, die nur einen bestimmten Anteil des Frequenzspektrums enthalten (siehe Abb. 11). Soweit uns bekannt, ist unsere Studie (siehe Kapitel 3.2) die erste, die den Einfluss des Anstiegs des Fourierspektrums auf die Gesichtswahrnehmung untersucht.



**Abb. 11 Beispiele für frequenzgefilterte Gesichter: (a) unmodifiziertes Gesicht. (b) Tiefpass- bzw. LSF-Gesichter enthalten nur grobe Strukturen. (c) Hochpass- bzw. HSF-Gesichter enthalten nur feine Strukturen. Aus Ruiz-Soler und Beltran (2006)**

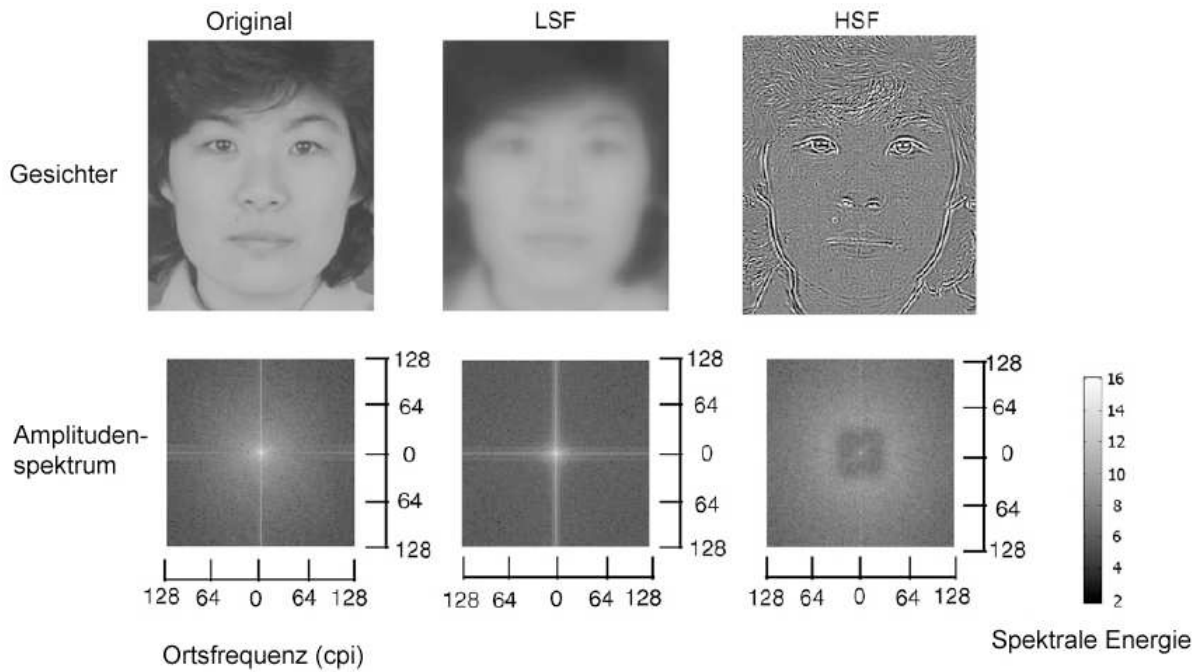
In Verhaltensstudien konnte gezeigt werden, dass der mittlere Frequenzbereich von 8 bis 16 Schwingungen pro Bild (cpi) am Wichtigsten für die Gesichtererkennung ist (Costen et al. 1994, Fiorentini et al. 1983, Morrison und Schyns 2001, Näsänen 1999, Collin et al. 2004). Trotzdem erreichten Gesichter aus den anderen Frequenzbereichen noch sehr gute Erkennungsraten, sodass Gesichtererkennung auch außerhalb des optimalen Frequenzbereichs gut möglich ist.

Neurophysiologische Studien mit fMRT und EEG, welche die Wahrnehmung von Bandpass-gefilterten Gesichtern untersuchen, zeigten zum Teil widersprüchliche Ergebnisse für Hoch- (HSF) und Tiefpass- (LSF) gefilterte Stimuli. Vuilleumier et al. (2003) fanden mittels fMRT, dass ungefilterte und HSF-Gesichter eine höhere Aktivität als LSF-Gesichter in der „fusiform face area“ hervorrufen und dass LSF-Gesichter bei wiederholter Repräsentation zu geringerem Abfall des Signals in der „fusiform face area“ (Wiederholungseffekt) führten. Die Autoren argumentierten, dass LSF-Gesichter schwächere Repräsentationen in der „fusiform face area“ zur Folge haben. Hingegen fanden Goffaux et al. (2003) in einem Gesichter-Lernexperiment höhere Erkennungsraten für LSF- als für HSF-Gesichter und keine Unterschiede für die N170-Komponente im EEG (siehe Kapitel 7.2). Andere Studien mit evozierten Potentialen zeigten größere Amplituden der N170 für LSF- als für HSF-Gesichter (Halit et al. 2006) oder für HSF- als für LSF-Gesichter (Hsiao et al. 2005, Nakashima et al. 2008) oder gleiche N170-Amplituden für HSF- und für LSF-Gesichter (Flevaris et al. 2008). Die verschiedenen Ergebnisse lassen sich zum einen durch unterschiedliche Stimulus Manipulation (zum Beispiel Kontrast, Helligkeit, Frequenzgrenzen der Filter) und zum anderen durch unterschiedliche Aufgaben (zum Beispiel passive Wahrnehmung, Gesichtserkennung, Lernexperiment) erklären. Unterschiedliche Frequenzbereiche können je nach Nützlichkeit flexibel für die Aufgabe verwendet werden (Morrison und Schyns 2001, Loftus und Harley 2004, Schyns und Oliva 1999).

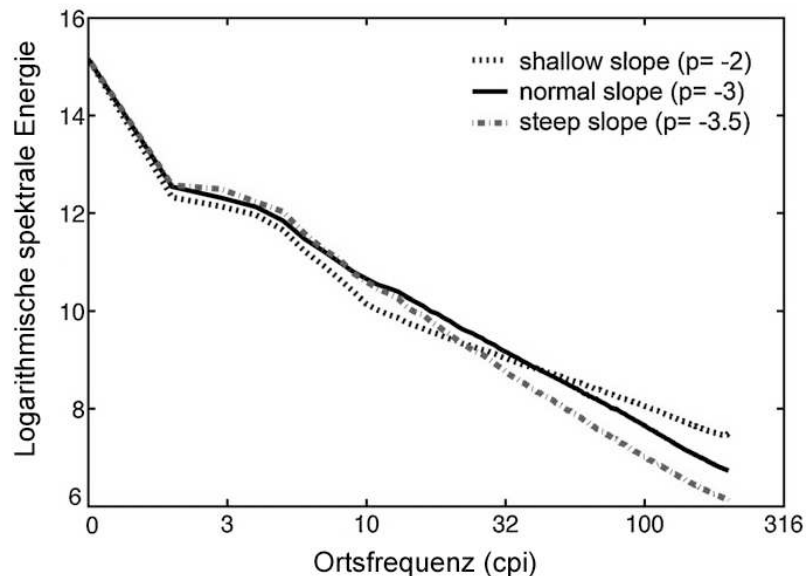
Da HSF-Gesichter wesentlich weniger spektrale Energie als LSF-Gesichter haben, erhöhten Nakashima et al. (2008) die spektrale Energie von HSF-Gesichtern (siehe Abb. 12) und fanden eine größere N170 für HSF- als für unmanipulierte Gesichter.

Um diese konfundierenden Störvariablen zu vermeiden, haben wir den Anstieg des Fourierspektrums von Gesichtern ohne die Anwendung traditioneller Bandpassfilter manipuliert (siehe Abb. 13).





**Abb. 12** Bandpass-gefilterte Stimuli mit ihren Fourierspektren. LSF-Gesichter enthalten spektrale Energie im Zentrum des Amplitudenspektrums. HSF-Gesichter enthalten Energie in der Peripherie des Amplitudenspektrums. Beim Vergleich des Amplitudenspektrums von Original- und HSF-Gesichtern sieht man, dass bei den HSF-Gesichtern nicht nur ein Hochpassfilter angewendet wurde, sondern auch die spektrale Energie in der Peripherie erhöht wurde. Modifiziert nach Nakashima et al. (2008)



**Abb. 13** Gemittelte Fourierspektren der Stimuli unserer Studie.  $p$  bezeichnet den Anstieg. SHALLOW SLOPE Gesichter haben einen flacheren Anstieg. STEEP SLOPE Gesichter besitzen einen steileren Anstieg.

Wir haben zum einen STEEP SLOPE Gesichter mit steilerem Anstieg, verstärkten groben und abgeschwächten feinen Strukturen hergestellt (siehe Abb. 14). Zum anderen stellten wir SHALLOW SLOPE Gesichter her mit flacherem Anstieg ( $1/f^2$ -Eigenschaften), verstärkten feinen und abgeschwächten groben Strukturen. SHALLOW SLOPE Gesichter besitzen verstärkte feine und abgeschwächte grobe Strukturen. STEEP SLOPE Gesichter enthalten verstärkte grobe und abgeschwächte feine Strukturen. Dadurch enthalten unsere Stimuli das ganze Frequenzspektrum und nur der relative Anteil von groben und feinen Strukturen ändert sich. Dies hat den Vorteil, dass sich die Stimuli in der gesamten spektralen Energie nicht unterscheiden und die manipulierten Gesichter realistischer wirken und weniger Artefakte enthalten. Wir konnten SHALLOW SLOPE Gesichter mit  $1/f^2$ -Eigenschaften erzeugen, die natürlichen Szenen und ästhetischen Bildern darin ähnlich sind. In einem Gesichter-Lernexperiment (siehe Kapitel 3.2) sollte gezeigt werden, wie dies durch den Anstieg des Fourierspektrums beeinflusst wird. Wir nahmen an, dass Gesichter mit  $1/f^2$ -Eigenschaften und damit verstärkten feinen Strukturen und abgeschwächten groben Strukturen besser gelernt werden können.



**Abb. 14 Beispiele für Gesichter mit manipuliertem Anstieg im Fourierspektrum. SHALLOW SLOPE Gesichter. STEEP SLOPE Gesichter besitzen abgeschwächte feine und verstärkte grobe Strukturen.**

## 2 Ziele der Arbeit

Die Untersuchung der statistischen Eigenschaften ästhetischer Bilder dient zum einen der Erforschung der Grundlagen ästhetischer Wahrnehmung und zum anderen der Erforschung von Funktionsweisen des visuellen Systems.

Der erste Teil der Arbeit beinhaltete deskriptive Analysen des Anstiegs der Fourierspektren ( $1/f^p$ -Eigenschaften) von Gesichtsfotografien und gemalten ästhetischen Porträts. In diesem Projekt sollten universelle statistische Eigenschaften von ästhetischen Bildern gefunden und untersucht werden. Die vorliegende Arbeit schloss sich an Studien unserer Arbeitsgruppe und einer amerikanischen Arbeitsgruppe an, welche zeigten, dass ästhetische gemalte Bilder ähnliche  $1/f^2$ -Eigenschaften im Fourierspektrum besitzen wie Fotografien von natürlichen Szenen (Redies et al. 2007, Graham und Field 2007). Da sich in der Kategorie ästhetische Bilder auch natürliche Szenen befanden, führte dies eventuell zu einer Ähnlichkeit der Fourierspektren. In unserer Studie sollte herausgefunden werden, mit welchen  $1/f^p$ -Eigenschaften Künstler Objekte darstellen, die andere Fourierspektren als natürliche Szenen besitzen. Dazu wurden Gesichtsfotografien untersucht, welche einen steileren Anstieg im Fourierspektrum aufweisen. In Zusammenarbeit mit dem Lehrstuhl für Digitale Bildverarbeitung von Prof. Dr. J. Denzler wurden  $1/f^p$ -Eigenschaften der Fourierspektren von Gesichtsfotografien und gemalten ästhetischen Porträts ermittelt und verglichen. Wir fanden heraus, dass ästhetische gemalte Porträts ähnliche  $1/f^2$ -Eigenschaften wie natürliche Szenen haben und sich von Gesichtsfotografien unterscheiden.

Im zweiten Teil der Arbeit sollte die durch den Künstler hervorgerufene Veränderung des Fourierspektrums auf Fotos von Gesichtern übertragen und anschließend deren Wahrnehmung untersucht werden. Damit schloss sich diese Arbeit an eine große Anzahl an Studien an, die den Einfluss des Frequenzspektrums auf die Gesichterwahrnehmung erforschen (siehe Kapitel 1.4, Ruiz-Soler und Beltran 2006). Soweit mir bekannt ist, stellt meine Arbeit die erste Studie dar, die Gesichter mit verändertem Anstieg des Fourierspektrums erstellt hat und nicht wie sonst üblich Bandpass-gefilterte Stimuli. Unsere Annahme war, dass Gesichter mit ähnlichem Fourierspektrum wie natürliche Szenen ebenfalls effizienter codiert werden können und in einem Lernexperiment besser abschneiden. In einer gemeinsamen experimentellen Studie mit dem Lehrstuhl für Allgemeine Psychologie von Prof. Dr. S. Schweinberger wurde der Einfluss von  $1/f^p$ -Eigenschaften auf das Erlernen

und Erkennen unbekannter Gesichter untersucht. Insbesondere durch die Erfassung von EEG-Korrelaten der Gesichtswahrnehmung konnten wir zeigen, wie die Veränderung des relativen Anteils an groben und feinen Strukturen das Gesichterlernen beeinflusst.

### **3 Publierte Originalarbeiten**

#### **3.1 *Artists portray human faces with the Fourier statistics of complex natural scenes.***

***Redies C, Hänisch J, Blickhan M, Denzler J.  
Network. 18(3):235-48. 2007.***

## Artists portray human faces with the Fourier statistics of complex natural scenes

CHRISTOPH REDIES<sup>1</sup>, JAN HÄNISCH<sup>2</sup>, MARKO BLICKHAN<sup>1</sup>,  
& JOACHIM DENZLER<sup>2</sup>

<sup>1</sup>*Institute of Anatomy I, School of Medicine, Friedrich Schiller University, D-07740 Jena, Germany, and* <sup>2</sup>*Department of Computer Science, Friedrich Schiller University, D-07740 Jena, Germany*

*(Received 2 May 2007; revised 14 July 2007; accepted 14 July 2007)*

### Abstract

When artists portray human faces, they generally endow their portraits with properties that render the faces esthetically more pleasing. To obtain insight into the changes introduced by artists, we compared Fourier power spectra in photographs of faces and in portraits by artists. Our analysis was restricted to a large set of monochrome or lightly colored portraits from various Western cultures and revealed a paradoxical result. Although face photographs are not scale-invariant, artists draw human faces with statistical properties that deviate from the face photographs and approximate the scale-invariant, fractal-like properties of complex natural scenes. This result cannot be explained by systematic differences in the complexity of patterns surrounding the faces or by reproduction artifacts. In particular, a moderate change in gamma gradation has little influence on the results. Moreover, the scale-invariant rendering of faces in artists' portraits was found to be independent of cultural variables, such as century of origin or artistic techniques. We suggest that artists have implicit knowledge of image statistics and prefer natural scene statistics (or some other rules associated with them) in their creations. Fractal-like statistics have been demonstrated previously in other forms of visual art and may be a general attribute of esthetic visual stimuli.

**Keywords:** *visual stimuli, natural scenes, psychophysics, emotional processing*

## Introduction

The fundamental nature of esthetic judgment remains unknown, despite attempts by artists, philosophers and psychologists to define universal principles that characterize what makes art esthetically pleasing to human observers. Several scholars in the field have argued that all humans share the same concept of beauty (Burke 1757; Hume 1757; Kant 1790; Schelling 1907; Kandinsky 1912; Adorno 1970; Paul 1988) and some have concluded that biological factors must be taken into account in order to explain esthetic experience (Burke 1757; Paul 1988). More recently, in the emerging field of neuroesthetics, neuroscientists speculated that esthetic experience is a product of brain function and is closely linked to perceptual processes (Rentschler et al. 1988; Gregory et al. 1995; Werner and Ratliff 1999; Zeki 1999; Livingstone 2002; Cavanagh 2005). Following this general idea, we hypothesized that esthetic art is a phenomenon of resonance between the artist's visual system and his creations (Redies 2008). In our model, this resonant state of neural activity is purposefully induced by the artist through a constant feedback between the work of art being created and the artist's visual system.

In a search for possible neuronal mechanisms that are linked to esthetic perception, we previously measured Fourier statistics in graphic art from diverse periods and countries of the Western hemisphere (Redies et al. 2007). Results showed that, on average, artists create their works of art with fractal-like statistical properties, independent of the cultural variables present in the set of images analyzed. These fractal-like properties are reflected in a  $1/f^2$  Fourier power spectrum (or  $1/f$  amplitude spectrum;  $f$ : spatial frequency) and imply that works of graphic art display scale invariance. Similar fractal-like statistical properties have been demonstrated for natural scenes (Burton and Moorhead 1987; Field 1987; Tolhurst et al. 1992; Ruderman and Bialek 1994; Ruderman 1997; Simoncelli and Olshausen 2001; Olshausen and Field 2004).

Fractal structure was previously detected in the abstract paintings by Jackson Pollock (Taylor et al. 1999), and image statistics similar to those of natural scenes have been found also in a set of color paintings from diverse Western and Asian cultures (Graham and Field 2007). Moreover, human observers show a general preference for fractal-like structures in landscape silhouettes (Hagerhall et al. 2004). We proposed that this similarity between natural scenes and esthetic visual art relates to the fact that both types of stimuli can be perceived as beautiful by human observers (Redies et al. 2007; Redies 2008).

In the present study, we examined a favorite subject matter of artists, human faces. Photographic images of human faces do not display fractal-like, scale-invariant statistics and the slope of the curve in the log-log plot of spectral power (amplitude squared) vs. spatial frequency is steeper than for natural scenes (Torralba and Oliva 2003; Bosworth et al. 2006). We asked whether artists render human faces with the same statistics as photographs of faces. Our results for a large set of graphic art of Western provenance show that this is not the case. Paradoxically, artists portrait human faces with scale-invariant Fourier statistics that are characteristic of complex natural scenes. This finding suggests that artists might have implicit knowledge of complex scenes statistics (or of unknown rules associated with complex scene statistics) and prefer these statistics or rules in their creations.

## Material and methods

### *Image data*

Two photographic face databases (1, 2), a natural scene database (3) and two databases containing portraits by artists (4, 5) were analyzed.

- (1) The Yale face database B (Georghiades et al. 2001) consists of monochrome images of 10 people that were photographed with 9 different poses under 64 illumination conditions in front of a simple laboratory or office background. Original images were  $640 \times 480$  pixels.
- (2) The AR face database (Martinez and Benavente 1998) contains color images of 126 people with different facial expressions, illumination conditions and occlusions, photographed on a uniformly bright background. Image size was  $768 \times 576$  pixels. Images were converted to grayscale values.

Centered passport-type details of  $480 \times 480$  pixels (Yale face database) or  $576 \times 576$  pixels (AR face database) were cut from each image for analysis. Examples are shown in Figure 1D–I.

- (3) For comparison, images from the Groningen natural scene database (van Hateren and van der Schaaf 1998) were analyzed. The same dataset of 208 images analyzed previously (Redies et al. 2007) was used. Centered details of  $1024 \times 1024$  pixels were cut from the original monochrome images of  $1536 \times 1024$  pixels. Examples are shown in Figure 1A–C.
- (4) A database of 306 portraits by artists was generated. Reproductions were digitized from various art books by a calibrated scanner (Perfection 3200 Photo, Seiko, Epson Corporation, Nagano, Japan). No compression or image enhancement algorithms were applied. Images were scanned in 8-bit grayscale at a resolution of at least 1024 pixels width and length. The database consisted of monochrome or lightly colored (washed) works on paper (graphic art). The portraits represented various cultural backgrounds from the Western hemisphere and were created by artists from different countries and centuries, employing different techniques (Table I).
- (5) Using the same scanning procedure, calibrated scans were obtained from reproductions of colored portraits (oil paintings) that originated from a cultural background similar to that of the monochrome portraits. Color images were converted to grayscale using the YIQ transform where luminance is expressed as the sum of the weighted contributions from the RGB channels (relative weights: R, 0.3; G, 0.59; B, 0.11), as previously done in another study of colored art images (Graham and Field 2007).

The scanner was calibrated for gamma gradation with the IT8 target printed on reflective paper (LaserSoft Imaging, Kiel, Germany). The target displayed 24 gray values of measured luminances. A grayscale conversion table was generated that allowed transformation of all monochrome scans to linearized gray scale values. For color scans, the scanner was gamma calibrated with the same target using the SilverFast Ai Professional Scan Software, version 6.5 (LaserSoft Imaging).





Figure 1. Examples of the images analyzed. (A–C) Examples from the Groningen database of natural scenes (van Hateren and van der Schaaf 1998). (D–F) Examples from the Yale face database B (Georghiades et al. 2001). (G–I) Examples from the AR face database (Martinez and Benavente 1998). (K–L) Examples of padded images of monochrome portraits by artists (K, drawing by Martin Schongauer, about 1465; L, drawing (self-portrait) by Caspar David Friedrich 1820; and P, drawing by Julius Schnorr von Carolsfeld 1817). (N–P) Details displaying the face with an eye distance similar to that of the photographic faces in D–I. Slope constants and deviations from the fitted line (in parentheses) were: A,  $-1.93$  (0.005); B,  $-1.73$  (0.002); C,  $-2.00$  (0.008); D,  $-3.37$  (0.003); E,  $-3.31$  (0.004); F,  $-3.28$  (0.001); G,  $-3.50$  (0.048); H,  $-3.69$  (0.005); I,  $-3.68$  (0.124); K,  $-1.84$  (0.008); L,  $-1.96$  (0.046); M,  $-2.30$  (0.024); N,  $-1.84$  (0.031); O,  $-1.87$  (0.043); and P,  $-2.09$  (0.041). The images shown in K–M were reproduced with permission from “Das Berliner Kupferstichkabinett”, Akademischer Verlag, Berlin, 1994 (inventory numbers: K, 976-1; L, 916-2; and M, 5212; © Staatliche Museen zu Berlin, Kupferstichkabinett).

Table I. Slopes of the fitted line for portraits by artists (details), calculated separately for different cultural and other variables.

	Slope (mean $\pm$ SD)	<i>n</i>
All	-2.12 $\pm$ 0.30	306
Background		
Homogeneous	-2.11 $\pm$ 0.29	195
Complex	-2.13 $\pm$ 0.31	111
Headdress		
No	-2.11 $\pm$ 0.30	188
Yes	-2.13 $\pm$ 0.29	118
Gender		
Child	-2.09 $\pm$ 0.25	26
Women	-2.14 $\pm$ 0.30	46
Man, without beard	-2.10 $\pm$ 0.30	145
Man, with beard	-2.13 $\pm$ 0.32	89
View		
Front	-2.11 $\pm$ 0.30	253
Side	-2.15 $\pm$ 0.28	53
Century		
15th Century	-1.95 $\pm$ 0.16	20
16th Century	-2.10 $\pm$ 0.24	89
17th Century	-2.05 $\pm$ 0.36	34
18th Century	-2.18 $\pm$ 0.16	18
19th Century	-2.16 $\pm$ 0.37	50
20th Century	-2.16 $\pm$ 0.32	95
Country		
Italy	-2.14 $\pm$ 0.27	57
Flanders	-1.87 $\pm$ 0.26	34
France	-2.24 $\pm$ 0.37	45
Germany	-2.12 $\pm$ 0.26	150
Other countries	-2.14 $\pm$ 0.32	20
Techniques		
Etching	-2.04 $\pm$ 0.33	50
Engraving	-2.08 $\pm$ 0.24	17
Lithograph	-2.20 $\pm$ 0.23	27
Woodcut	-2.37 $\pm$ 0.44	13
Charcoal, chalk	-2.16 $\pm$ 0.26	100
Pencil, silver point	-2.02 $\pm$ 0.23	59
Pen drawing	-2.05 $\pm$ 0.33	31
Brush drawing	-2.32 $\pm$ 0.37	9

Note: Values are means  $\pm$  SD. *n*: number of images analyzed in each category.

The reproductions chosen for analysis were of relatively large size and high quality and displayed works of art with no or only minor defects (paper cuts, stains, folds etc.). In all portraits, faces covered a large part of the image.

The artistic portrait database was analyzed in two different formats. First, as described previously, the scanned images were padded according to square ones by adding a uniform border with a gray value equal to the average gray value in the image (Redies et al. 2007). Examples are shown in Figure 1K–M.

Second, square details of the portraits were generated showing face, neck and shoulders of the portrayed persons at a magnification comparable to that of the photographic face databases (Figure 1N–P). For normalization, eye distance was

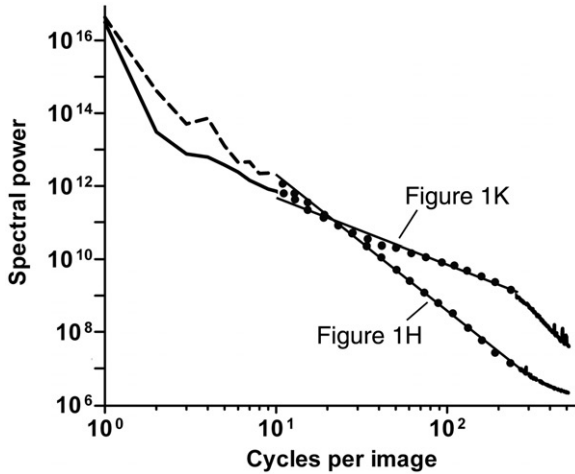


Figure 2. Example of the Fourier spectral analysis. In the log–log plane, Fourier power (amplitude squared) was plotted as a function of spectral frequency. A line was fitted to values that were binned at regular logarithmic intervals between 10 to 256 cycles per image (dots). The dashed and solid lines represent results for the images displayed in Figure 1H and K, respectively. Slopes and deviations from the fitted line are  $-3.69$  and  $0.005$  (for Figure 1H) and  $-1.84$  and  $0.007$  (for Figure 1K).

measured (front views) or estimated on the basis of the distance between eyes and the mouth (side views).

### *Image analysis*

Image analysis was carried out using Matlab as described previously (Redies et al. 2007). Briefly, each input image from the test sets of different dimensions was resized to  $1024 \times 1024$  pixels by bicubic interpolation. After transforming each image into the frequency domain using Fast Fourier Transform, the rotational average of the power spectrum was computed for each frequency. Power spectrum (amplitude squared) and frequency were analyzed in the log–log plane (Figure 2). Next, a least squares fit of a line to the log–log power spectrum was performed by fitting data points that were binned at regular intervals. Only the frequency range between 10 and 256 cycles per image was used for the fitting. This restriction minimized the effect of artifacts in our analysis, for example artifacts due to low pass filtering, rectangular sampling, raster screen or noise in the images. The result for each image is the slope of the line and the deviation of the data points from that line, calculated as the sum of the squares of the deviations of the data points, divided by the number of data points.

In total, we analyzed five different data sets, consisting of natural scenes (208 images), photographic images of faces (Yale face database B, 5776 images; AR face database, 3313 images), monochrome portraits by artists (306 images) and colored oil portraits converted to grayscale values (141 images).

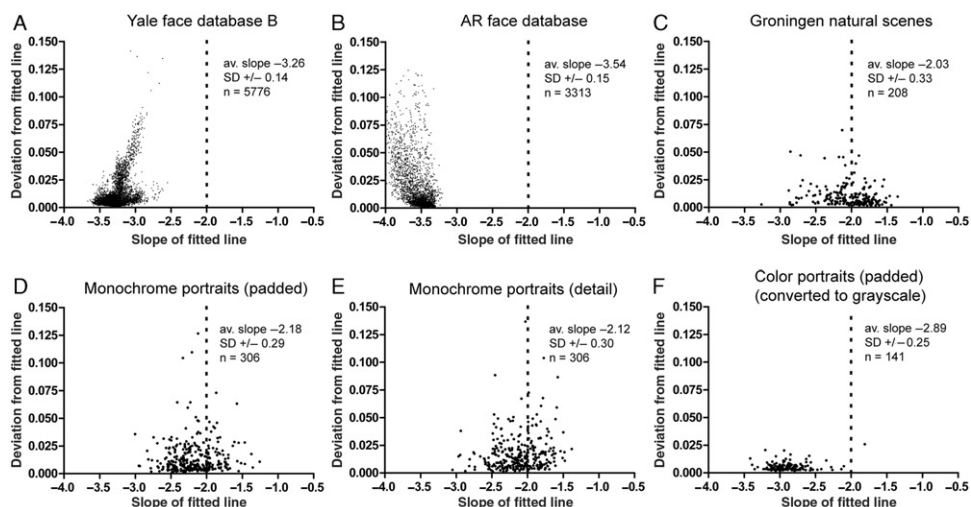


Figure 3. Results of the Fourier spectral analysis. Each dot in the scatter diagrams represents the slope of the fitted line for one image and the deviation of the measured data points from the fitted line for that image. Data shown in C are from Redies et al. (2007). Av. slope, average slope for the set of images.

## Results

In Figure 2, Fourier spectral power of two representative images from the databases is plotted as a function of spectral frequency. In the log–log plane, the binned data points deviate only slightly from the straight fitted line, within the frequency range analyzed. However, the two fitted lines differ in their slope. The fitted line of the face photograph is steeper (slope of  $-3.69$ ) than that of the artist's rendering of a human face (slope of  $-1.84$ ). A slope constant of about  $-2$  (or  $-1$  if spectral amplitude instead of power is plotted) indicates that the image has scale-invariant or fractal-like properties, as previously shown for natural (complex) scenes (Burton and Moorhead 1987; Field 1987; Tolhurst et al. 1992; Ruderman and Bialek 1994; Ruderman 1997; Simoncelli and Olshausen 2001; Olshausen and Field 2004). Close-up views of simple objects generally result in steeper slopes (Torralba and Oliva 2003; Bosworth et al. 2006; Redies et al. 2007).

Figure 3 shows scatter diagrams with the slope of each image plotted on the  $X$  axis and the deviations from the fitted lines plotted on the  $Y$  axis, for each of the image datasets analyzed. The majority of images can be fitted well by a straight line, as indicated by the small deviations of the data points from the fitted line. The average slopes are  $-3.26$  for the Yale face database (Figure 3A) and  $-3.54$  for the AR face database (Figure 3B). This difference is probably due to the office background in the Yale face database. After replacing this background by a white background in 30 randomly selected images from the Yale face database, the slope became more negative for all images; the average slope for the 30 images shifted significantly from  $-3.28$  ( $\pm 0.12$  SD) to  $-3.57$  ( $\pm 0.15$  SD;  $p < 0.0001$ , paired  $t$ -test).

For natural scenes (Figure 3C) and monochrome portraits (Figure 3D), slopes were significantly higher than for the face photographs ( $-2.03$  and  $-2.18$ , respectively; nonparametric statistical analysis by Kruskal–Wallis test with

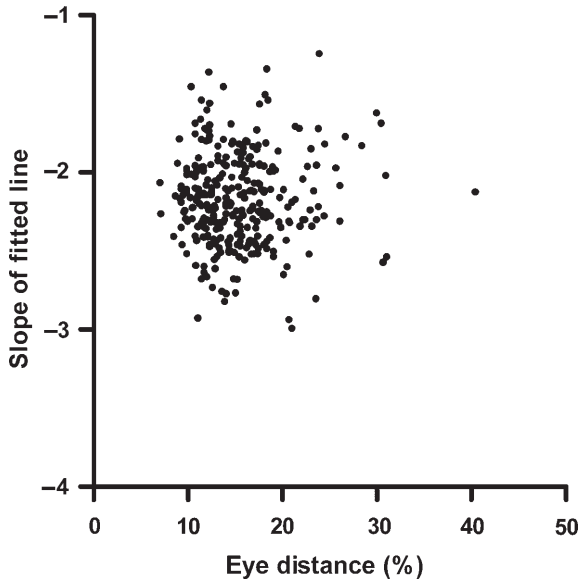


Figure 4. Slope of the fitted lines plotted as a function of eye distance for the 306 padded images of monochrome portraits. Eye distance was expressed in percent of the image dimension.

Dunn's multiple comparison post-test,  $p < 0.001$ ). For color (oil) portraits that were converted to monochrome images, the average slope was  $-2.89$  (Figure 3F), which is significantly more negative than the slope for natural scenes or monochrome portraits ( $p < 0.001$ ) and significantly more positive than the slope for the two-face-photograph databases ( $p < 0.001$ ).

The difference between the slopes of monochrome portraits and face photographs may be due to the fact that, in some of the portraits, faces were viewed from a larger distance and were embedded in complex scenes. As an index of face size in the padded portraits, the eye distance was expressed as a percentage of image dimension. Average eye distance was  $19.6\%$  ( $\pm 1.9$  SD) in the Yale face database, and  $19.7\%$  ( $\pm 1.4$  SD) in the AR face database, compared to  $15.5\%$  ( $\pm 4.6$  SD) in the portrait database. Figure 4 shows the dependency of the measured slope constants on the eye distance for the monochrome portraits. The two variables did not significantly correlate with each other (Spearman correlation coefficient  $r = -0.003$ ). We repeated our analysis for details of the portraits, which were enlarged in size so as to match approximately the size of the photographed faces. For the portrait details (Figure 1N–P), average eye distance was  $20.3\%$  ( $\pm 5.8$  SD). The mean slope for this dataset was  $-2.12$  ( $\pm 0.30$  SD; Figure 3E), which is close to the average slope of the padded portraits ( $-2.18$ ; Figure 3D).

For the face details, there were only small or no significant differences in the average slope constants between faces painted on homogeneous vs. complex background, between persons portrayed with and without headdress, between faces of children, women, and men with and without beards, or between front views and

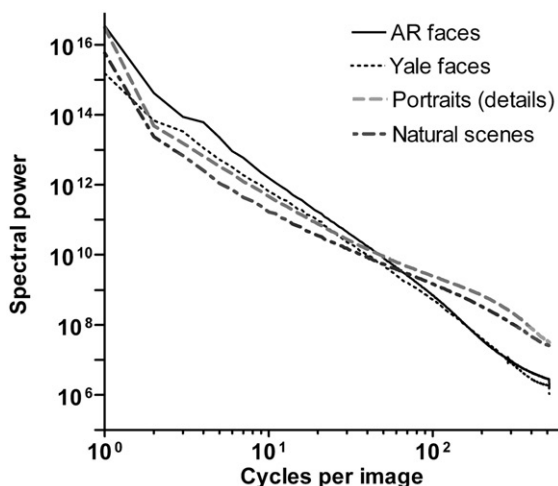


Figure 5. Average curves for the different categories of face images and natural scenes.

side views of faces (Table I). Also, cultural variables, such as techniques, centuries and country of origin had only a small or no significant influence on the slopes (Table I).

In Figure 5, the logarithmic average of all spectral power curves for the different image categories is plotted as a function of spectral frequency in the log–log space. The curves for natural scenes and portraits by artists are shallower than those for photographs of faces.

The scanner used for digitizing the reproductions of portraits from art books was calibrated for linearized conversion of color and brightness into pixel values (see ‘Materials and methods’). However, we cannot control for gamma gradation during reproduction in art books. We therefore asked what effect moderate degrees of gamma gradation have on the slopes measured by us. Figure 6 shows that the effect of gamma values between 0.25 and 4 is minor.

## Discussion

### *Methodological considerations*

Our analysis reveals that artists endow human faces with image statistical properties similar to those of complex natural scenes (Burton and Moorhead 1987; Field 1987; Tolhurst et al. 1992; Ruderman and Bialek 1994; Ruderman 1997; Simoncelli and Olshausen 2001; Olshausen and Field 2004). Before accepting this result, trivial explanations for our findings and experimental artifacts must be excluded. We therefore carried out control experiments, which show that the present result is unlikely to originate in reproduction artifacts and that it cannot be explained by systematic differences in the complexity of the visual patterns surrounding the faces in the portraits.

A number of artifacts might possibly influence the measurements of the slopes in the log–log plots, for example artifacts caused by reproducing art images in books.

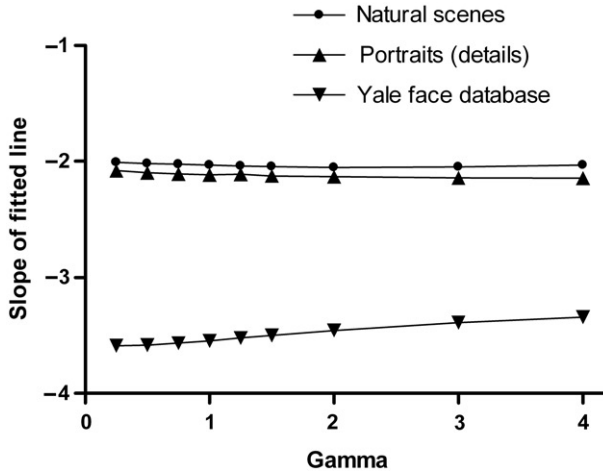


Figure 6. Average slopes of the fitted lines plotted as a function of gamma gradation applied to three different image datasets (Yale face database, natural scenes and details of portraits by artists). From the Yale face database, a subset of 300 randomly selected images was used for the analysis.

One such artifact may be nonlinearities in the transformation of color and brightness to pixel values during photography, scanning, and printing. Such nonlinearities are commonly expressed as changes in gamma gradation. Here, we demonstrate that moderate degrees of gamma gradation, which can be anticipated in the reproduction process, have only a minor effect on the values of the slope constant measured in our experiment (Figure 6). A similar robustness of the slopes has been previously reported in natural scenes for changes in contrast or in gray value offset (for a review, see Ruderman (1997)). The effect of other reproduction artifacts has been minimized by restricting the frequency range in the analysis (see ‘Materials and methods’). It is therefore unlikely that reproduction artifacts have a major effect on our results. The frequency range restriction may explain why other investigators, who did not restrict the frequencies range, obtained slightly lower values for the slopes (Tolhurst et al. 1992; Graham and Field 2007).

Secondly, the three databases of human faces differ in the complexity of the background shown in the images. Images from the AR face database show the face, neck and shoulders of each person on a uniform bright background. Images from the Yale face database show similar body parts in front of an office background, resulting in less negative slope values. Images from the art portrait database generated by us depict persons or faces at variable distances and with backgrounds of different degrees of complexity. It is thus possible that the higher slope values reflect a higher complexity of the rest of the image rather than of the face. This possibility, however, was excluded by normalizing the eye distance in the portraits to those of the photographic faces. Moreover, we did not observe any difference in the slopes between faces portrayed on a complex background and faces portrayed on a homogeneous background (Table I).

Thirdly, artists often portray humans with elaborate accessories, such as fancy hats, which represent complex visual stimuli and may also result in higher slope constants. However, slopes of portraits with and without headdress were not



significantly different from each other (Table I). Also, the absence or presence of beards, which may also induce complexity in the portraits, did not influence the results (Table I).

*A paradoxical shift of image statistics in artists' portraits*

Our results suggest that artists have an implicit knowledge of image statistics and tend to shift the statistics of human faces in their portraits toward the fractal-like statistics of complex natural scenes. As a result, artists portray human faces with statistics different from those of face photographs. This paradoxical shift demonstrates that artists do not necessarily strive to represent natural objects as they are in reality. Rather, they follow unspecified rules that call for an implementation of image statistics similar to those of complex natural scenes. A similar conclusion has been reached for biased samples of nonrepresentational (abstract) art, including oil paintings (Taylor et al. 1999; Taylor 2002; Redies et al. 2007).

The present results are in line with previous observations for a large set of graphic art of the Western hemisphere (Redies et al. 2007). This study showed that, on average, graphic art is created by artists with the fractal-like statistics of natural scenes. However, in our previous study, we did not compare the statistics of art images and their natural counterparts and most works of art included in our previous study depicted complex scenes.

*Sampling bias and the universality of image statistics in art*

The artistic portraits analyzed here represent a biased sample of art images. First, we demonstrate fractal-like properties only for monochrome portraits or portraits, which were washed with thin color and converted to monochrome images. The inclusion of the color dimension in our analysis would have complicated the analysis.

After conversion to monochrome images, fully colored portraits (color oil paintings) show Fourier spectral statistics in between those of photographed faces and natural scenes (Figure 3F). Color is an important attribute to art and adds to its esthetic appearance. It may thus come as no surprise that the luminance component of color art has different Fourier statistics than that of monochrome art. Graham and Field (2007) recently obtained Fourier statistics similar to natural scenes also for monochrome renderings of color paintings. Their biased sample of art, however, contained complex scenes and was not restricted to portraits, which may explain the difference in the results.

Another bias stems from the fact that we selected works of art from well-known artists that have been preserved in prestigious museums. We assume that the esthetic value of these works of art is an important reason why they have been conserved, in some cases over many centuries. Due to this bias, conclusions about the image statistics of art reached in the present study likely apply only to esthetic forms of art but not to other contemporary forms of nonesthetic art (see discussion in Redies 2008).



Despite these biases, our sample of graphic art contains representational art from a large variety of different cultural backgrounds within the Western hemisphere and different graphic techniques. As shown previously for a set of Western graphic art, which depicted multiple subject matters and included abstract art (Redies et al. 2007), the dependence of the slopes on the cultural variables is small, if significant at all (Table I). Similar statistics were found for the abstract drip paintings of Jackson Pollock (Taylor et al. 1999) and in a set of paintings that included a large proportion of art from the Middle East and Asia (Graham and Field 2007). The widespread occurrence of this finding in different forms of art and artistic techniques and in various human cultures is striking, but its universality in all form of esthetic art remains to be established.

### *Questions and hypotheses*

Fractal-like properties may be a general attribute of esthetic visual displays but cannot be a sufficient criterion for esthetic art for several reasons. First, computer-generated artificial images with  $1/f^2$  power statistics (Ruderman 1997; Lee et al. 2001; Olshausen and Field 2000) do not necessarily look esthetically pleasing. Second, the range of slope values measured for artistic portraits in the present study overlaps extensively with examples of image classes that are little or not at all esthetic (Redies et al. 2007). Third, there is a clear difference in the profoundness of esthetic appeal between art objects and natural scenes; these differences do not correlate with differences in the measured slopes.

If  $1/f^2$  power statistics are not sufficient to induce esthetic perception, what is the reason for artists to shift image statistics in portraits? Does this shift provide insight into the sensory principles underlying esthetic perception? In an attempt to address this question, we would like to raise the following two speculative points:

- (1) The visual system is adapted to the statistics of complex natural scenes by evolution and development (Field 1987; Olshausen and Field 1996; Parraga et al. 2000; Vinje and Gallant 2000; Simoncelli and Olshausen 2001; Hoyer and Hyvärinen 2002). In turn, artists adapt their creations to functional features intrinsic to the human visual system (Zeki 1999). The present results are compatible with the hypothesis (Redies 2008) that the functional features, to which artists induce resonance in their visual system, are related, in some unknown way, to the adaptation of the visual system to natural scenes. Following this idea, the  $1/f^2$  power statistics discovered in visual art should be thought of as a corollary of other, as of yet unidentified, principles of esthetic perception. Artists may not be able to express these statistical principles in precise, every-day language (Redies 2008). For example, Fourier analysis can hardly be carried out in the conscious human mind. Indeed, Fourier analysis is a scientific concept that most artists cannot have been aware of until the 20th century.
- (2) Alternatively, it may be argued that artists often aim to convey or emphasize particular traits of their subjects (for example, personality traits or expressed emotions). To achieve this goal in the artistic portraits, artists might use specific artistic techniques (for example, sketching with lines or fine textures) that carry more energy in the higher frequency range. However, in art images

depicting complex (natural) scenes with similar techniques, the frequency spectra of the depicted scenes did not change on average (Redies et al. 2007). Therefore, graphic art is not generally associated with an increase in higher frequencies. Moreover, as discussed above, other artistic techniques result also in art images with scale-invariant properties.

## Acknowledgements

The authors thank Aleix M. Martinez, Athos S. Georghiades and Hans van Hateren for generous permission to use the databases generated by them and to reproduce images from their databases in Figure 1D–I.

## References

- Adorno TW. 1970. *Ästhetische Theorie*. Frankfurt: Suhrkamp.
- Bosworth RG, Bartlett MS, Dobkins KR. 2006. Image statistics of American sign language: Comparison with faces and natural scenes. *J Opt Soc Amer A - Opt, Image Sci Vision* 23:2085–2096.
- Burke E. 1757. *A philosophical enquiry into the origin of our ideas of the sublime and beautiful*. London: Dodsley.
- Burton GJ, Moorhead IR. 1987. Color and spatial structure in natural scenes. *Appl Phys* 26:157–170.
- Cavanagh P. 2005. The artist as neuroscientist. *Nature* 434:301–307.
- Field DJ. 1987. Relations between the statistics of natural images and the response properties of cortical cells. *J Opt Soc Amer A - Opt Image Sci Vision* 4:2379–2394.
- Georghiades AS, Belhumeur PN, Kriegman DJ. 2001. From few to many: Illumination cone models for face recognition under variable lighting and pose. *IEEE T Pattern Anal* 23:643–660.
- Graham DJ, Field DJ. 2007. Statistical regularities of art images and natural scenes: Spectra, sparseness and nonlinearities. *Spatial Vision* (in press).
- Gregory RL, Harris J, Heard P, Rose D. 1995. *The artful eye*. Oxford: Oxford University Press.
- Hagerhall CM, Purcell T, Taylor R. 2004. Fractal dimension of landscape silhouette outlines as a predictor of landscape preference. *J Environ Psychol* 24:247–255.
- van Hateren JH, van der Schaaf A. 1998. Independent component filters of natural images compared with simple cells in primary visual cortex. *Proc R Soc B* 265:359–366.
- Hoyer PO, Hyvärinen A. 2002. Sparse coding of natural contours. *Neurocomputing* 44–46:459–466.
- Hume D. 1757. *Of the standard of taste. The philosophical works of David Hume*. London: Longman, Green.
- Kandinsky W. 1912. *Über das Geistige in der Kunst, insbesondere in der Malerei*. München: Piper.
- Kant I. 1790. *Kritik der Urteilskraft*. In: Weischedel W, editor. *Werkausgabe in zwölf Bänden* (1992). Frankfurt: Suhrkamp.
- Lee AB, Mumford D, Huang J. 2001. Occlusion models for natural images: A statistical study of a scale-invariant dead leaves model. *Int J Comput Vis* 41:35–59.
- Livingstone MS. 2002. *Vision and art: The biology of seeing*. New York: Harry N. Abrams.
- Martinez AM, Benavente R. 1998. *The AR Face Database*. CVC Technical Report, Vol. #24.
- Olshausen BA, Field DJ. 1996. Natural image statistics and efficient coding. *Network Comp Neural* 7:333–339.
- Olshausen BA, Field DJ. 2000. Vision and the coding of natural images. *Am Sci* 88:238–245.
- Olshausen BA, Field DJ. 2004. Sparse coding of sensory inputs. *Curr Opin Neurobiol* 14:481–487.
- Parraga CA, Troscianko T, Tolhurst DJ. 2000. The human visual system is optimised for processing the spatial information in natural visual images. *Curr Biol* 10:35–38.
- Paul G. 1988. Philosophical theories of beauty and scientific research on the brain. In: Rentschler I, Herzberger B, Epstein D, editors. *Beauty and the brain. Biological aspects of aesthetics*. Basel: Birkhäuser. pp 15–27.

- Redies C. 2008. A universal model of esthetic perception based on the sensory coding of natural stimuli. *Spatial Vision* (in press).
- Redies C, Hasenstein J, Denzler J. 2007. Fractal-like image statistics in visual art: Similarity to natural scenes. *Spatial Vision* (in press).
- Rentschler I, Caelli T, Maffei L. 1988. Focusing in on art. In: Rentschler I, Herzberger B, Epstein D, editors. *Beauty and the brain*. Basel: Birkhäuser. pp 181–216.
- Ruderman DL. 1997. Origins of scaling in natural images. *Vision Res* 37:3385–3398.
- Ruderman DL, Bialek W. 1994. Statistics of natural images - scaling in the woods. *Phys Rev Lett* 73:814–817.
- Schelling FWJ. 1907. *Philosophie der Kunst*. In: Schröter M, editor. *Schellings Werke*. 3rd ed. München: Beck (Reprint 1959).
- Simoncelli EP, Olshausen BA. 2001. Natural image statistics and neural representation. *Annu Rev Neurosci* 24:1193–1216.
- Taylor RP. 2002. Order in Pollack's chaos - Computer analysis is helping to explain the appeal of Jackson Pollock's paintings. *Sci Am* 287:116–121.
- Taylor RP, Micolich AP, Jonas D. 1999. Fractal analysis of Pollock's drip paintings. *Nature* 399:422.
- Tolhurst DJ, Tadmor Y, Chao T. 1992. Amplitude spectra of natural images. *Ophthal Physiol Opt* 12:229–232.
- Torralba A, Oliva A. 2003. Statistics of natural image categories. *Network* 14:391–412.
- Vinje WE, Gallant JL. 2000. Sparse coding and decorrelation in primary visual cortex during natural vision. *Science* 287:1273–1276.
- Werner JS, Ratliff F. 1999. Some origins of the lightness and darkness of colors in the visual arts and in the brain. *Techné* 9–10:61–73.
- Zeki S. 1999. Art and the brain. *J Consciousness Stud* 6–7:76–96.

**3.2  $1/f^p$  Characteristics of the Fourier power spectrum affects ERP correlates of face learning and recognition.**

**Blickhan M, Kaufmann JM, Denzler J, Schweinberger SR, Redies C.**

***Biological Psychology*. 88(2-3):204-14. 2011.**



# $1/f^p$ Characteristics of the Fourier power spectrum affects ERP correlates of face learning and recognition

M. Blickhan<sup>a</sup>, J.M. Kaufmann<sup>b,d</sup>, J. Denzler<sup>c</sup>, S.R. Schweinberger<sup>b,d</sup>, C. Redies<sup>a,\*</sup>

<sup>a</sup> Institute of Anatomy I, Jena University Hospital, School of Medicine, Friedrich Schiller University, Jena, Germany

<sup>b</sup> Department of General Psychology and Cognitive Neuroscience, Friedrich Schiller University, Jena, Germany

<sup>c</sup> Department of Computer Science, Friedrich Schiller University, Jena, Germany

<sup>d</sup> DFG Research Unit Person Perception, Friedrich Schiller University, Jena, Germany

## ARTICLE INFO

### Article history:

Received 30 November 2010

Accepted 8 August 2011

Available online 19 August 2011

### Keywords:

Face learning

Face perception

Event-related potential

N170

N250

Spatial frequency

Fourier power spectrum

Natural vision

## ABSTRACT

We investigated the influence of Fourier power spectrum ( $1/f^p$ ) characteristics on face learning while recording ERPs that are associated with the representation of faces. Two image sets with an altered  $1/f^p$  characteristics were created. The first set consisted of stimuli with a STEEP SLOPE ( $1/f^{3.5}$ ) and therefore enhanced low spatial frequencies (LSF) and attenuated high spatial frequencies (HSF). The second set consisted of stimuli with a SHALLOW SLOPE ( $1/f^2$ ), similar to complex natural scenes and artwork, resulting in enhanced HSF and attenuated LSF. Faces with a SHALLOW SLOPE elicited larger N170 and N250 amplitudes and larger old/new effects for central positivity in comparison to unmodified faces. The opposite effect was observed for faces with a STEEP SLOPE that led to slower reaction times. This result suggests that diminishing the ratio of fine detail (HSF) to coarse structures (LSF) impairs face learning, whereas increasing it facilitates neurocognitive correlates of face learning.

© 2011 Elsevier B.V. All rights reserved.

## 1. Introduction

The human visual system uses spatial filtering to extract visual information (De Valois and De Valois, 1980; Westheimer, 2001). For example, visual information in the brain is processed through two separate pathways (Livingstone and Hubel, 1988): the magnocellular stream provides coarse information (low spatial frequencies) and the parvocellular stream provides fine information (high spatial frequencies). Several studies have demonstrated that the spatial frequency profile of face images has an effect on face learning and identity recognition. Typically, bandwidth frequency-filtered faces were used to study these effects (for a recent review, see Ruiz-Soler and Beltran, 2006).

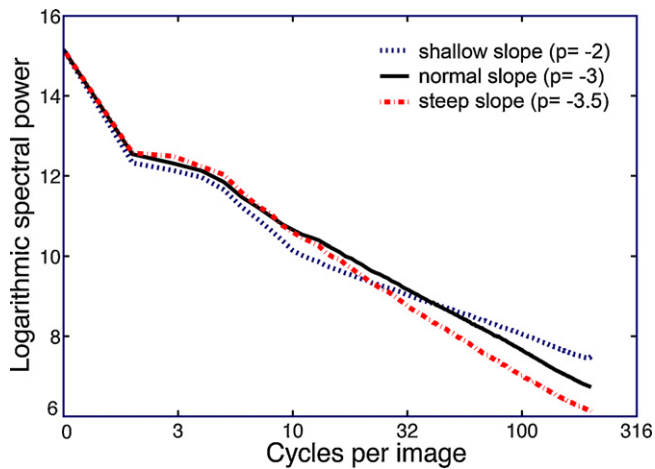
For face recognition, the importance of the frequency range between 8 and 16 cycles per image (cpi) was implicated in behavioural studies (Collin et al., 2004; Costen et al., 1994; Fiorentini et al., 1983; Morrison and Schyns, 2001; Näsänen, 1999). A similar range is preferred for face identity discrimination (Costen et al., 1996; Parker and Costen, 1999), although other frequency

ranges have also been suggested to contribute to face perception (Ginsburg, 1978; Halit et al., 2006; Harmon and Julesz, 1973).

Neurophysiological studies have also explored the processing of faces that were bandwidth frequency filtered. Results are inconsistent and favor different frequency ranges. In an fMRI study, Vuilleumier et al. (2003) measured the activation of fusiform cortex, which is thought to play an important role in face recognition (Kanwisher et al., 1997). They found a higher activation in fusiform cortex for unfiltered (broad spatial frequency, BSF) faces and high spatial frequency (HSF) faces in comparison to low spatial frequency (LSF) faces. Furthermore, they found long-lag repetition effects for HSF faces and BSF faces but not for LSF faces. They concluded that LSF faces lead to less stable face representation in fusiform cortex. In another fMRI study, Eger et al. (2004) found immediate repetition effects that were invariant to spatial frequency composition in fusiform cortex. By contrast, response adaption to repetition in occipital face regions only occurred with identical stimulus input with the same spatial frequency composition. In an event-related potential (ERP) study, Goffaux et al. (2003) found higher accuracy for LSF than for HSF faces in a familiarity task, in which bandwidth frequency-filtered faces had to be recognized. No significant differences between LSF and HSF faces were found for reaction times and amplitudes of the ERP component N170, which is widely assumed to reflect early face processing prior to face identity recognition (Eimer, 2000a,b). However there is some

\* Corresponding author at: Institute of Anatomy I, University of Jena School of Medicine, Teichgraben 7, D-07743 Jena, Germany. Tel.: +49 3641 938 511; fax: +49 3641 938 512.

E-mail address: [redies@mti.uni-jena.de](mailto:redies@mti.uni-jena.de) (C. Redies).



**Fig. 1.** Averaged Fourier spectral power of the 30 stimuli from each category. In the log-log plane, Fourier power (amplitude squared) was plotted as a function of spectral frequency (cpi, cycles per image). SHALLOW SLOPE – enhanced high spatial frequencies and diminished low spatial frequencies; NORMAL SLOPE – not manipulated; STEEP SLOPE – enhanced low spatial frequencies and diminished high spatial frequencies.  $p$ , slope of the curve. Spectral power was larger for the STEEP SLOPE than NORMAL SLOPE faces from 0 to 9 cpi and larger for STEEP SLOPE than SHALLOW SLOPE faces from 0 to 22 cpi. NORMAL SLOPE faces had larger power than SHALLOW SLOPE from 0 to 43 cpi and SHALLOW SLOPE faces had larger power above 43 cpi. A spatial frequency of 1 cpi is equivalent to 0.075 cycles per degree (cpd); 3 cpi = 0.23 cpd; 10 cpi = 0.75 cpd; 32 cpi = 2.41 cpd; 100 cpi = 7.52 cpd).

evidence that the N170 can be modulated by repeated face presentation (Caharel et al., 2009). Other ERP studies that focused on face detection or passive viewing tasks and not on identification showed larger N170 amplitudes for LSF faces than HSF faces (Halit et al., 2006), for HSF faces than LSF faces (Hsiao et al., 2005; Nakashima et al., 2008), or equal N170 amplitudes for HSF faces and LSF faces (Flevaris et al., 2008).

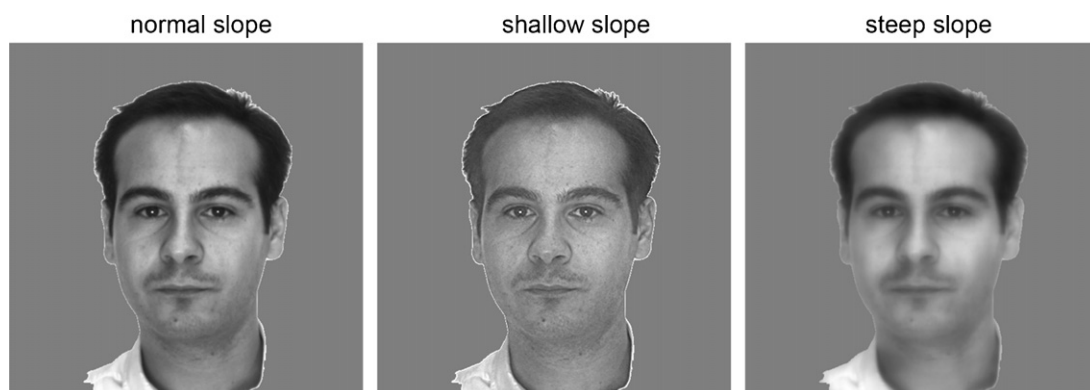
This inconsistency may be due to different tasks used in these studies (familiarity, gender and passive viewing task), or to different types of stimulus manipulation (contrast, luminance and different frequency cut offs). It has been shown that the face processing system uses frequency scales according to their diagnostic value for the respective task (Loftus and Harley, 2004; Morrison and Schyns, 2001; Schyns and Oliva, 1999).

Various ERP studies showed that bandwidth frequency-filtered faces did not elicit larger N170 amplitudes than unfiltered faces (Flevaris et al., 2008; Goffaux et al., 2003; Halit et al., 2006). Interestingly Nakashima et al. (2008) found a larger N170 for HSF faces than for BSF (unfiltered) faces and concluded that the N170 represents fine feature encoding of faces. One possible explanation for

this discrepancy is that, by equalizing the contrast in their stimulus set, Nakashima et al. (2008) increased the spectral power of the HSF faces. Typically, a face has more spectral power in the lower frequencies than in the higher frequencies (Keil, 2008; Redies et al., 2007b). As a consequence, an HSF face has less overall spectral power than an LSF face. Possibly, the larger spectral power of the HSF faces, which was increased by contrast equalization, led to the larger N170 in comparison to BSF faces.

The above problem suggests that it may be advantageous to modulate the spatial frequency profile of faces content without “traditional” bandwidth filters, in order to avoid the possible confounding effect of differences in overall spectral power of the images. The radially averaged Fourier spectral power of most natural images, including face images, falls nearly linearly with increasing spatial frequency ( $f$ ), according to a power law ( $1/f^p$ ; Keil, 2008; Redies et al., 2007b). In log-log plots of Fourier power versus spatial frequency (Fig. 1), the value  $p$  corresponds to the slope of the plotted line and provides a measure for the ratio of high to low spatial frequencies in the image. The different stimuli, which our visual system perceives under normal everyday conditions, usually vary in their spectral slope and are not bandwidth frequency filtered. By manipulating the slope value, we changed the ratio of high and low spatial frequency power without affecting the overall spectral power of the face images. These images still contain the full frequency spectrum and may result in more realistic stimuli compared to bandwidth frequency-filtered images (Fig. 2).

The image manipulation carried out by us has further perceptual implications. It is well established that, in contrast to face images, the slope value of complex natural scenes is around  $-2$  ( $1/f^2$  characteristics), or  $-1$  if Fourier amplitude is plotted instead of Fourier power (Burton and Moorhead, 1987; Tolhurst et al., 1992). This result implies that complex natural scenes are scale invariant (fractal-like), that is, when zooming in and out of a natural scene, the statistical properties of the Fourier spectral components remain constant. Moreover, it has been shown that the human visual system processes the statistical regularities of natural scenes with an efficient and sparse code (Hoyer and Hyvärinen, 2002; Olshausen and Field, 1996; Párraga et al., 2000; Vinje and Gallant, 2000, 2002). Scale invariance was also found in art images of different cultural provenance (Graham and Field, 2007; Redies et al., 2007a). Strikingly, it has been reported that artists portray human faces with the Fourier statistics of complex natural scenes, despite the fact that face photographs have significantly more negative slope values, indicative of less fine structure than natural scenes and art images (Keil, 2008; Redies et al., 2007b). A possible explanation is that artists intuitively endow their artworks with the statistical regularities that are processed optimally and efficiently in their visual system (Redies, 2007; Zeki, 1999).



**Fig. 2.** Example of the stimuli. NORMAL SLOPE – not manipulated; SHALLOW SLOPE – enhanced high spatial frequencies and diminished low spatial frequencies; STEEP SLOPE – enhanced low spatial frequencies and diminished high spatial frequencies.



The face images that were used in this study had an average slope value of  $p = -3.06$  before manipulation. We manipulated the slope of these images (Figs. 1 and 2) and created faces with a slope value of  $-2$  (SHALLOW SLOPE faces), similar to art portraits and natural scenes. These faces contain less LSF and more HSF than the faces that were not manipulated (NORMAL SLOPE faces). Therefore they comprise less coarse structures and more fine structures, which are supposed to be more important in a task that involves fine feature perception (Boeschoten et al., 2005; Goffaux et al., 2005; Schyns et al., 2002). We also created faces with a slope value of  $-3.5$  (STEEP SLOPE faces), i.e. they contain more LSF and less HSF than the NORMAL SLOPE faces and therefore they exhibit a higher ratio of coarse structures to fine structures.

While recording ERPs, we used a face learning task to study the influence of Fourier spectrum characteristics on face identification. Our reasoning was that, if images with  $1/f^2$  characteristics are more efficiently encoded by the human visual system (Párraga et al., 2000), we should expect to see an advantage for such faces in a face learning task. We assume that this advantage should stand out more clearly when learning an unfamiliar face, compared to familiar stimuli, which already have an established neural representation. In addition, the stimulus manipulation also enables us to study the contribution of the ratio of fine structures to coarse structures to face learning. To the best of our knowledge, this is the first electrophysiological study that uses a manipulation of the slope of the power spectrum, thereby avoiding confounding factors such as differences in contrast, overall spectral power or frequency cut offs.

In contrast to previous ERP studies that focused on the N170, we have also analyzed later components, which have been associated with the processing of identity specific information. During this study the following ERP components were analyzed: P100 - which is sensitive to low level stimulus conditions and may sometimes show sensitivity to early face processing (Herrmann et al., 2005; Nakashima et al., 2008); N170 - which is supposed to reflect early structural face encoding and is generally thought to be widely insensitive to face identity (Bentin and Deouell, 2000; Eimer, 2000a,b), however there is some evidence that the N170 can be modulated by repeated face presentation (Caharel et al., 2009); N250 - which is sensitive to face repetition and is larger for familiar than unfamiliar faces (Kaufmann et al., 2009; Schweinberger et al., 2002a,b; Tanaka et al., 2006); and central-parietal positivity - which has been reported to show larger old/new effects for learned than new faces and is assumed to be a neural correlate of retrieval of visual face information (Bentin and Deouell, 2000; Eimer, 2000a; Paller et al., 2000).

## 2. Methods

### 2.1. Participants

Twenty healthy participants (11 women) aged between 18 and 26 years ( $M = 22.3$ ,  $SD = 1.7$ ) contributed data. They were each paid € 12. All participants were right handed, as assessed by the Edinburgh handedness inventory (Oldfield, 1971) and had normal or corrected to normal vision.

### 2.2. Stimuli and apparatus

The stimulus set consisted of 90 face images (45 women and 45 men) that were taken from the AR face database (Martinez and Benavente, 1998) and converted to grayscale. All faces were looking straight into the camera and displayed a neutral expression. The unmodified images (NORMAL SLOPE faces) had an average power-spectrum slope of  $p = -3.06$  and a resolution of  $574 \times 574$  pixels. Our aim was to manipulate the slope of the radially averaged power spectrum (Fig. 1) of the grayscale face images. First, we created images with slope values of  $-2$  (SHALLOW SLOPE), which exhibit a higher ratio of fine structures (HSF) to coarse structures (LSF) than the NORMAL SLOPE faces. Second, we created images with slope values of  $-3.5$  (STEEP SLOPE), which exhibit a lower ratio of fine structures (HFS) to coarse structures (LSF) than the NORMAL SLOPE faces. Processing artefacts in the images (differences in mean gray levels and patchy appearance of the faces) were minimized

by not altering the first couple of cycles per image (cpi), and by creating a smooth transition between manipulated and unmanipulated power spectra, as described below.

In order to change the Fourier power spectra of the face images, each image was transformed into the frequency domain using Fast Fourier Transform. For each frequency, the rotational average of the power spectrum was computed as described by previous studies (Graham and Field, 2007; Redies et al., 2007a). For further analysis, the log-log plane of frequency and averaged power spectrum was used. The slope value of the double logarithmic power spectrum was obtained by linear regression in the range between 10 and 127 cpi. This range has been previously used to avoid possible artefacts at the extremes of the power spectrum (Redies et al., 2007a). Then a factor was calculated so that multiplication of the averaged power spectrum with the factor resulted in the desired slope value of the log-log plot. The power spectrum in the normal plane was multiplied by this factor, which leads to the desired change of the slope. A manipulation of the lower frequencies causes image distortions and therefore the first 2 cpi were not altered. The mean gray value of the images remained constant. The amplitude step at the transition between the unmodified and the modified parts of the spectrum lead to frequency filtering artefacts and the produced faces can easily be identified as manipulated. For the images with a slope value of  $-2$  (SHALLOW SLOPE faces), a smooth transition between the unmodified and the modified parts of the spectrum was therefore created by replacing the amplitude step with a sinusoidal transition function. As a result, the new power spectrum consists of 3 parts: the first part comprises the first 2 cpi that were not altered. The second part is the transition region (from 2 to 12 cpi) and the third part has a slope of  $-2$  (rest of the spectrum). For the images with a slope of  $-3.5$  (STEEP SLOPE faces), the first 5 cpi were not altered and the y-intercept was lowered to avoid an amplitude step. The modified power spectrum and the original phase spectrum were used to create the modified face image by performing an inverse Fourier transformation. The changes in the power spectrum lead to some modifications in the background of the faces, which made it easy to distinguish manipulated images from original images. To avoid this and to be sure that our modification affected the face and not just the background, the background of all images was changed into a uniform gray before and after the manipulation. Therefore the NORMAL SLOPE faces have a slightly less negative slope value ( $p = -3.06$ ) than the images from the AR face database ( $p = -3.54$ ) with a white background (Redies et al., 2007a). In the last step the contrast of the manipulated face image was adapted to the contrast of the unmodified face image, using the adjust contrast function of the Matlab computer program.

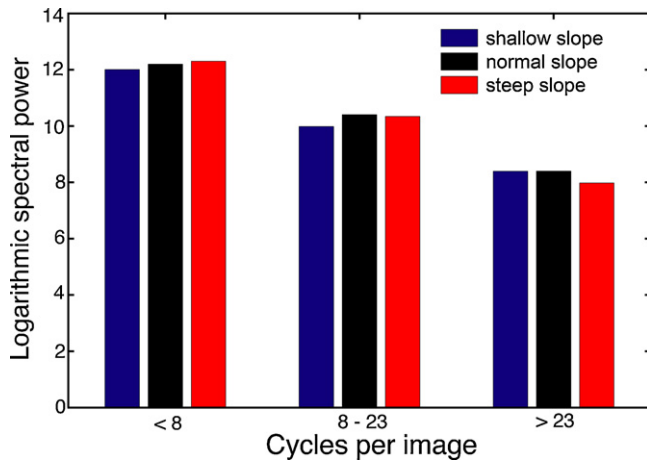
One third of the 90 images was not manipulated (NORMAL SLOPE faces; see also Fig. 2), one third was manipulated to obtain a slope of  $-2$  (SHALLOW SLOPE faces) and one third was manipulated to obtain a slope of  $-3.5$  (STEEP SLOPE faces). The mean gray value of the images was not affected by the manipulation, because the first 2 cpi were not altered. An analysis of the overall spectral power revealed no significant differences of the 3 image categories (NORMAL SLOPE, SHALLOW SLOPE and STEEP SLOPE). Nevertheless the images differed in their spectral power in certain frequency ranges. As shown by the rotationally averaged power spectrum (Fig. 1), the following Fourier spectral power relationships between the three image types were obtained: STEEP SLOPE faces > NORMAL SLOPE faces (range between 0 and 9 cpi), STEEP SLOPE faces > SHALLOW SLOPE faces (0–22 cpi), NORMAL SLOPE faces > SHALLOW SLOPE faces (0–43 cpi), and SHALLOW SLOPE faces > NORMAL SLOPE faces and STEEP SLOPE faces (above 43 cpi). For further comparison of the spectral power across the 3 image categories, we computed averaged spectral power in 3 frequency ranges: low spatial frequencies (<8 cpi), middle spatial frequencies (8–23 cpi), high spatial frequencies (>23 cpi). The cutoffs were chosen to correspond with prior studies (Flevaris et al., 2008; Halit et al., 2006; Vuilleumier et al., 2003). The averaged data was computed by averaging the rotational averaged power spectrum in each frequency range so that the percentage of change of spectral power was similar for SHALLOW SLOPE and STEEP SLOPE faces (see Figs. 3 and 4).

Images were displayed on a 15" monitor with a resolution of  $1280 \times 1024$  pixels. Using EPrime™ (Version 1.4) software, stimuli were presented on a black background at a viewing distance of 60 cm. Viewing distance was kept constant by using a chin rest. The size of the images was  $14 \text{ cm} \times 14 \text{ cm}$ , corresponding to a visual angle of  $\sim 13.3^\circ \times 13.3^\circ$ . Throughout this study, a spatial frequency of 1 cpi is therefore equivalent of 0.075 cycles per degree of visual angle (cpd).

### 2.3. Procedure

Before the experiment, participants gave informed consent and were encouraged to ask questions in case anything remained unclear at any time during the experiment. Then an electrode cap (Easycap™) was applied and participants were guided to the electrically shielded chamber.

**Learning phase.** In this phase, 45 unfamiliar face images (23 men and 22 women, 15 SHALLOW SLOPE, 15 NORMAL SLOPE and 15 STEEP SLOPE faces) were presented. Each stimulus was presented for 3200 ms and participants were instructed to memorize the faces for a following recognition test. Each face was preceded by a fixation cross presented for 500 ms and after each face a blank screen was presented for 1500 ms. Faces were shown in randomized order. In total, there were 4 learning blocks and, in each block, each face was presented once, resulting in 4 repetitions



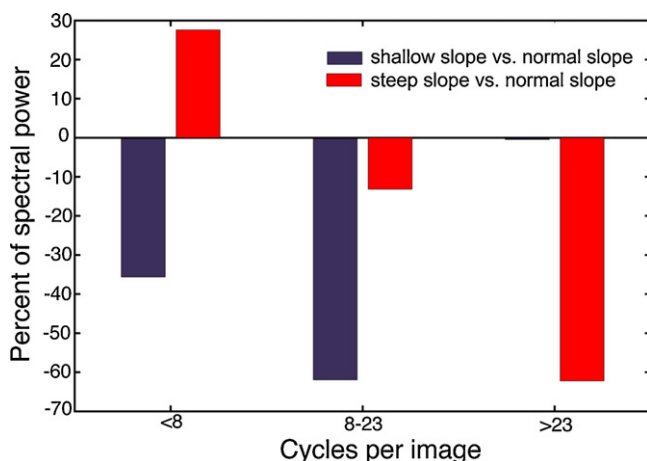
**Fig. 3.** Averaged spectral power of the 3 image categories in 3 frequency ranges: low spatial frequencies (<8 cpi), middle spatial frequencies (8–23 cpi), high spatial frequencies (>23 cpi). SHALLOW SLOPE faces had decreased spectral power in the low and middle spatial frequencies in comparison to NORMAL SLOPE faces. STEEP SLOPE stimuli had decreased spectral power in the middle and high spatial frequencies.

per face across the entire learning phase. After each block, there was a short break and participants decided when they were ready to continue.

**Test phase.** In the test phase, 39 of the 45 learned faces (i.e. the 13 SHALLOW SLOPE, 13 NORMAL SLOPE and 13 STEEP SLOPE faces shown in the learning phase) and 39 novel faces (faces not shown in the learning phase) were presented. The novel faces also consisted of 13 SHALLOW SLOPE, 13 NORMAL SLOPE and 13 STEEP SLOPE faces. Each face was presented for 2000 ms. Face stimuli were preceded by a fixation cross (500 ms) and followed by a blank screen (1500 ms). Participants were instructed to use two marked keys on a standard computer keyboard to categorize the faces as learned (shown in the learning phase) or novel (not shown in the learning phase), using the index fingers of the right and left hand. The assignment of response hand and response alternative was counterbalanced across participants.

After reading the instructions on the monitor, participants performed 24 practice trials consisting of the remaining 6 learned plus 6 additional novel faces (2 per frequency condition each). Each face was presented twice during the practice trials. These trials were not included in the data analysis. For practice trials, feedback was given regarding response speed and response accuracy.

After performing the practice trials, 3 blocks consisting of 78 faces each (39 learned, 39 novel) were presented. Within blocks all stimuli were presented in random order. Faces had to be categorized as learned or novel. Response times and accuracy were measured. Only for slow responses (responses exceeding 2000 ms) visual feedback was given (“Please respond faster”). After a maximum of 39 trials, participants were given the opportunity to have a short break. During the test phase, each face was therefore presented 3 times.



**Fig. 4.** Percentage of change of averaged spectral power for SHALLOW SLOPE and STEEP SLOPE faces in comparison to NORMAL SLOPE faces. Spectral power for the SHALLOW SLOPE faces was decreased in the low frequency range (36%) and middle frequency range (62%) and equal in the high frequency range. Spectral power for the STEEP SLOPE faces was decreased in the middle frequency range (14%) and high frequency range (62%) and increased in the low frequency range (28%).

## 2.4. Event-related potentials

**Data recording.** The electroencephalogram (EEG) was recorded in an electrically shielded room. Data was recorded with sintered Ag/AgCl electrodes mounted on an electrode cap (EasyCap™, Falk Minow Services, Herrsching-Breitbrunn, Germany) using SynAmps amplifiers (NeuroScan Labs, Sterling, VA), arranged according to the extended 10/20 system at the scalp positions Fz, Cz, Pz, Iz, Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T7, T8, P7, P8, FT9, FT10, P9, P10, PO9, PO10, F9, F10, F9', F10', TP9 and TP10. TP10 (right upper mastoid) served as initial common reference, and a forehead electrode (AFz) served as ground. Impedances were kept below 10 kΩ and were typically below 5 kΩ. The horizontal electro-oculogram (EOG) was recorded from F9' and F10' at the outer canthi of both eyes. The vertical EOG was monitored bipolarly from electrodes above and below the right eye. All signals were recorded with AC (0.05 Hz high pass, 40 Hz low pass, -6 dB attenuation, 12 dB/octave), and sampled at a rate of 250 Hz.

**Pre-processing of ERP data.** Ocular artefacts were corrected using the automatic artefact correction algorithm implemented in Besa™ (Version 5.1.8.10). Trials with non-ocular artefacts and trials with incorrect behavioural responses were discarded. ERPs were averaged separately for each channel and experimental condition. Each averaged ERP was low-pass filtered at 20 Hz with a zero phase-shift digital filter, and recalculated to average reference, excluding the vertical and horizontal EOG channels.

**Analysis of ERP data.** Only correct and artefact-free trials were included for averaging. The average numbers of correct and artefact-free trials per condition were as follows: learning phase: 28.3, 28.3, 28.6; test phase learned faces: 31.8, 33.5 and 31.5; novel faces: 35.3, 34.7 and 34.5 (for NORMAL SLOPE, SHALLOW SLOPE and STEEP SLOPE, respectively). All measures were taken relative to a 200 ms baseline preceding the target stimulus. Effects were quantified at regions of interest (ROI) that were based on maximum amplitudes of a particular component in grand mean waveforms and on previous research (Kaufmann and Schweinberger, 2008; Kaufmann et al., 2009; Schweinberger et al., 2002a,b). For the analysis of P100, the ROI included electrodes O1/O2, PO9/PO10 and Iz. The ROI for N170 encompassed electrodes P9/P10 and PO9/PO10. The ROI for N250 included P9/P10, PO9/PO10 and TP9/TP10. Central-parietal positivity between 500 and 600 ms was analyzed at electrodes C3/C4, P3/P4, Cz and Pz. For P100 and N170, peak analyses were performed. Peak latencies were determined separately for the left and right hemisphere, at the electrode pair within a particular ROI that showed maximum amplitudes. For P100, latency was additionally determined at Iz. Peak amplitudes were analyzed at all electrodes of a particular ROI at the previously identified peak latencies. For N170, in the case of strongly lateralized peaks, i.e. in the absence of a clear peak over one hemisphere, amplitudes for this hemisphere were measured at the peak latencies identified over the other hemisphere. This was necessary in the case of two participants. For N250 and central-parietal positivity, ERPs were quantified by mean amplitudes between 300–400 ms and 500–600 ms, respectively. These time windows were chosen based on visual inspection of the grand mean waveforms. For the test phase, wrong or missing responses were discarded from the analysis.

## 3. Results

For all behavioural and ERP data, where appropriate, Epsilon corrections for heterogeneity of covariances were performed with the Huynh–Feldt method (Huynh and Feldt, 1976). Multiple comparisons were not corrected because we focused on differences to the unmanipulated normal slope condition of both shallow and steep slope faces (2 comparisons). However, we added the exact *p* values for all multiple comparisons.

### 3.1. Behavioural results

With 87% hits and 94% correct rejections, accuracy was high. Response time (RT) and accuracy were analyzed by performing ANOVAs with repeated measurements on the factor frequency (SHALLOW SLOPE, NORMAL SLOPE, STEEP SLOPE; see also Table 1) and familiarity (learned faces shown in the learning phase, and novel faces not shown in the learning phase).

Accuracy was higher for the novel than the learned faces  $F(1,19) = 6.77, p < .05$ . The factor frequency had no significant effect on accuracies,  $F(2,38) = 1.16, p > .32$ , and furthermore there was no interaction between familiarity and frequency,<sup>1</sup>  $F(2,38) = 1.34, p < .27$ . The RT analysis revealed a significant effect of frequency,

<sup>1</sup> An additional analysis of hits minus false alarms with repeated measurements on frequency did not reach significance,  $F(2,38) = 1.16, p > .33$ .



**Table 1**

Reaction times and accuracies for the test phase. Results that revealed significant effects compared to NORMAL SLOPE faces are marked with \*.

Image category	Reaction Times in ms (SEM)		Accuracies in % (SEM)	
	Learned faces	Novel faces	Learned faces	Novel faces
NORMAL SLOPE	823 (16.3)	842 (27.7)	86.8 (2.5)	96.6 (1.2)
SHALLOW SLOPE	824 (14.7)	910 (28.9)*	89.1 (2.8)	92.8 (2.1)
STEEP SLOPE	874 (19.3)*	909 (28.2)*	85.4 (3.1)	93.1 (2.1)

**Table 2**

Latency of P100 and N170 in the learning phase and in the test phase (averaged across learned and novel faces). Results that revealed significant effects compared to NORMAL SLOPE faces are marked with \*.

Image category	P100 Latency in ms (SEM)		N170 Latency in ms (SEM)	
	Learning phase	Test phase	Learning phase	Test phase
NORMAL SLOPE	130 (2.0)	125 (2.4)	191 (2.2)	185 (3.4)
SHALLOW SLOPE	132 (2.1)*	131 (2.1)*	193 (2.6)	190 (3.1)*
STEEP SLOPE	132 (2.4)*	129 (2.7)*	193 (2.4)	190 (3.2)*

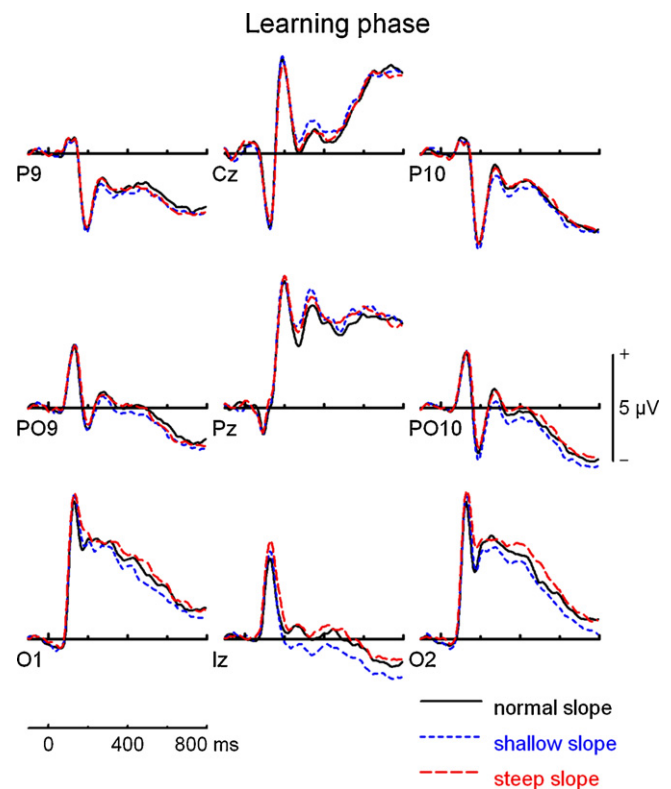
$F(2,38)=27.57$ ,  $p<.0001$ , and an interaction between familiarity and frequency,  $F(2,38)=6.49$ ,  $p<.01$ . Post-hoc comparisons for the learned faces revealed faster RTs for NORMAL SLOPE than for STEEP SLOPE faces,  $F(1,19)=17.34$ ,  $p=.001$ . No RT differences were found between SHALLOW SLOPE and NORMAL SLOPE faces,  $F(1,19)<1$ . Post-hoc comparisons for the novel faces revealed faster RTs for the NORMAL SLOPE in comparison to STEEP SLOPE and SHALLOW SLOPE faces ( $F(1,19)=30.82$ ,  $p<.0001$ , and  $F(1,19)=31.08$ ,  $p<.0001$ , respectively, see also Table 1).

### 3.2. Electrophysiological data

#### 3.2.1. Learning phase

**P100.** Peak latency was investigated by performing an ANOVA with repeated measurements on the factors electrode location (left vs. right vs. Iz) and frequency (SHALLOW SLOPE vs. NORMAL SLOPE vs. STEEP SLOPE). This analysis yielded a significant effect of the factor frequency,  $F(2,38)=5.44$ ,  $p<.01$ . Post-hoc testing revealed slightly but significantly shorter latencies for NORMAL SLOPE faces ( $M=130$  ms) compared to both SHALLOW SLOPE faces ( $M=132$  ms),  $F(1,19)=10.51$ ,  $p<.01$ , and STEEP SLOPE faces ( $M=132$  ms),  $F(1,19)=7.54$ ,  $p<.05$  (Table 2).

P100 peak amplitudes were analyzed by an ANOVA with repeated measurements on the factors electrode (O1 vs. O2 vs. Iz vs. PO9 vs. PO10) and frequency (NORMAL SLOPE vs. SHALLOW SLOPE vs. STEEP SLOPE). This analysis revealed a significant main effect of the factor frequency,  $F(2,38)=4.71$ ,  $p<.05$ , which was further modulated by a two-way interaction between frequency and electrode,  $F(8,152)=2.48$ ,  $p<.05$ . Separate comparisons showed a trend for larger<sup>2</sup> amplitudes for SHALLOW SLOPE compared to NORMAL SLOPE faces,  $F(1,19)=4.27$ ,  $p=.05$  (Table 3). This trend was not further modulated by an interaction with electrode. A comparison between NORMAL SLOPE and STEEP SLOPE faces revealed larger P100 amplitudes for STEEP SLOPE faces,  $F(1,19)=6.95$ ,  $p=.016$ . This effect was further specified by a two-way interaction between electrode and frequency,  $F(4,76)=3.37$ ,  $p=.026$ . Visual inspection of Fig. 5 suggests largest amplitude differences between NORMAL SLOPE and STEEP SLOPE faces at electrode Iz. This observation was confirmed by post-hoc comparisons for each electrode, which only revealed significant differences in P100 amplitudes between NORMAL SLOPE and STEEP SLOPE faces at electrodes Iz,  $F(1,19)=27.55$ ,  $p<.0001$ , and O2,  $F(1,19)=4.46$ ,  $p=.048$ .



**Fig. 5.** Grand-mean ERPs at selected electrodes of the regions of interest (ROI) in the learning phase.

**N170<sup>3</sup>.** Peak latency was tested by performing an ANOVA with repeated measurements on the factors hemisphere (left vs. right) and frequency (NORMAL SLOPE vs. SHALLOW SLOPE vs. STEEP SLOPE). This analysis did not yield any significant effects (Table 2).

Peak amplitudes of N170 were investigated by performing an ANOVA with repeated measurements on the factors hemisphere (left vs. right), frequency (NORMAL SLOPE vs. SHALLOW SLOPE vs. STEEP SLOPE) and site (P9/P10 vs. PO9/PO10). The ANOVA revealed larger N170 amplitudes over the right hemisphere,  $F(1,18)=14.09$ ,  $p<.01$ . The main effect of frequency was marginally significant,

<sup>2</sup> For components characterized by a positive polarity (P100 and central-parietal positivity), larger amplitudes always refer to more positive amplitudes (and analogous for the negative components N170 and N250).

<sup>3</sup> Data from one participant had to be excluded from all analyses of N170 because no clear peak could be identified.

**Table 3**

Amplitudes in  $\mu\text{V}$  in the learning phase at the following electrodes: P100 at IZ; N170 at PO9/PO10; N250 at P9/P10, PO9/PO10 and TP9/TP10; central positivity at C3/C4, P3/P4, Cz and Pz. Results that revealed significant effects compared to NORMAL SLOPE faces are marked with \*.

	P100 peak amplitude in $\mu\text{V}$ (SEM)	N170 peak amplitude in $\mu\text{V}$ (SEM)		N250 mean amplitude in $\mu\text{V}$ (SEM)	Central positivity mean amplitude in $\mu\text{V}$ (SEM)
	Iz	PO9	PO10	ROI	ROI
NORMAL SLOPE	4.3 (0.6)	-1.2 (0.5)	-2.4 (0.4)	-1.4 (0.3)	2.7 (0.4)
SHALLOW SLOPE	4.6 (0.6)*	-1.4 (0.4)*	-2.9 (0.3)*	-1.6 (0.3)*	3.1 (0.4)
STEEP SLOPE	5.1 (0.6)*	-1.1 (0.4)	-2.4 (0.4)	-1.3 (0.3)	3.0 (0.3)

$F(2,36) = 3.23, p = .05$ , as was the two-way interaction between frequency and site,  $F(2,36) = 3.06, p = .05$ .<sup>4</sup> Visual inspection of Fig. 5 suggests that frequency effects were largely limited to more posterior sites. This impression was confirmed by post-hoc ANOVAs with repeated measurements on hemisphere and frequency (NORMAL SLOPE vs. SHALLOW SLOPE and NORMAL SLOPE vs. STEEP SLOPE, respectively), only including electrodes PO9 and PO10. These analyses suggested larger N170 for SHALLOW SLOPE compared to NORMAL SLOPE faces,  $F(1,18) = 5.57, p = .03$ , whereas no significant differences were found between NORMAL SLOPE and STEEP SLOPE faces,  $F(1,18) < 1$  (Table 3).

**N250.** An ANOVA with repeated measurements on the factors frequency, hemisphere and site revealed a main effect of the factor frequency,  $F(2,38) = 6.51, p < .01$ . Visual inspection of Fig. 5 indicates that N250 was most negative for SHALLOW SLOPE faces. This impression was confirmed by additional post-hoc ANOVAs comparing SHALLOW SLOPE with NORMAL SLOPE faces,  $F(1,19) = 6.1, p = .023$ . Mean amplitudes of N250 did not differ between NORMAL SLOPE and STEEP SLOPE faces,  $F < 1$  (Table 3).

**500–600 ms.** Mean amplitudes of central-parietal positivity were analyzed by performing an ANOVA with repeated measurements on the factors frequency and electrode (C3, C4, P3, P4, Cz, Pz). There was no effect of frequency,  $F(2,38) = 1.38, p > .2$  (Table 3), and also no interaction between frequency and electrode,  $F(1,19) = 1.63, p > .1$ .

### 3.2.2. Test phase

ERPs in the test phase were analyzed by performing ANOVAs according to those applied in the learning phase. Additionally, the factor familiarity was introduced to compare effects for learned faces (shown in the learning phase) with effects for novel faces (not shown in the learning phase).

**P100.** For P100 latencies, the ANOVA yielded a main effect of frequency,  $F(2,38) = 20.49, p < .0001$ , and a two-way interaction between familiarity and electrode,  $F(2,38) = 2.31, p < .05$ . Post-hoc testing revealed that latencies were shorter for NORMAL SLOPE compared to STEEP SLOPE faces and SHALLOW SLOPE faces,  $F(1,19) = 21.37, p < .0001$  and  $F(1,19) = 37.04, p < .0001$ , respectively (Table 2).

For P100 peak amplitudes, a significant effect of frequency was found,  $F(2,38) = 6.13, p < .01$ . This effect was further modulated by a two-way interaction between frequency and electrode,  $F(8,144) = 4.0, p < .001$ . There was also an interaction between familiarity and electrode,  $F(4,76) = 4.51, p < .01$ . Post-hoc testing over separate electrodes revealed significant effects of familiarity,  $F(1,19) = 12.57, p < .01$ , and frequency,  $F(2,38) = 25.37, p < .0001$ , only at IZ and not at O1 and O2. P100 was larger for learned faces in comparison to novel faces at IZ. Post-hoc ANOVAs for this electrode showed that P100 was smaller for NORMAL SLOPE com-

pared to SHALLOW SLOPE,  $F(1,19) = 18.53, p < .0001$ , and smaller for NORMAL SLOPE compared to STEEP SLOPE faces,  $F(1,19) = 50.516, p < .0001$  (Fig. 6 and Table 4).

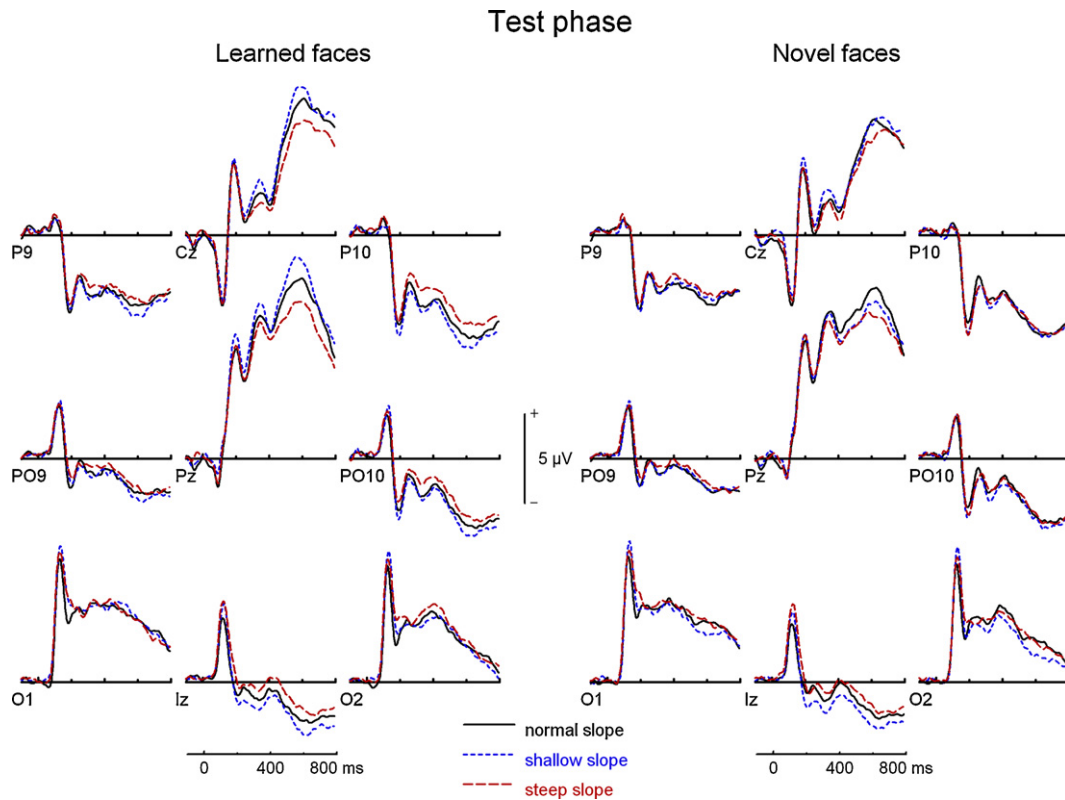
**N170.** Latencies of N170 were modulated by the factor frequency,  $F(2,36) = 15.89, p < .01$ . Separate post-hoc ANOVAs revealed shorter latencies for NORMAL SLOPE in comparison to SHALLOW SLOPE faces,  $F(1,18) = 22.53, p < .0001$ , and STEEP SLOPE faces,  $F(1,18) = 18.63, p < .0001$  (see also Table 2).

For N170 peak amplitudes, there were main effects of frequency,  $F(2,36) = 13.49, p < .0001$ , and hemisphere,  $F(1,18) = 24.79, p < .0001$ . The factor familiarity was not significant,  $F(2,38) < 1$ . Main effects were further qualified by a two-way interaction between frequency and hemisphere,  $F(2,36) = 4.81, p < .05$ , a three-way interaction between electrode, frequency and familiarity,  $F(2,36) = 6.45, p < .01$ , and a three-way interaction between hemisphere, frequency and familiarity,  $F(2,36) = 5.7, p < .01$ . Visual inspection of Fig. 6 suggested larger frequency effects over the right hemisphere, which also seemed to depend on familiarity. Separate ANOVAs for both hemispheres showed significant effects of frequency for left s,  $F(2,36) = 7.41, p < .01$ , and right hemispheric electrodes,  $F(2,36) = 11.04, p < .001$ . Over the left hemisphere N170 amplitude was smaller for STEEP SLOPE faces in comparison to NORMAL SLOPE faces,  $F(1,18) = 17.76, p = .0005$ , irrespective of familiarity. SHALLOW SLOPE faces did not differ from NORMAL SLOPE faces,  $F(1,18) < 1$ . Over the right hemisphere an interaction between frequency and familiarity,  $F(2,36) = 13.6, p < .0001$ , was found. Separate ANOVAs at right hemispheric sites for learned and novel faces revealed frequency effects both for learned,  $F(2,36) = 17.71, p < .0001$ , and for novel faces,  $F(2,36) = 7.41, p < .01$ . For the learned faces (Figs. 6 and 7), N170 amplitudes were larger for SHALLOW SLOPE faces compared to NORMAL SLOPE faces,  $F(1,18) = 10.74, p = .004$ , and STEEP SLOPE faces,  $F(1,18) = 29.67, p < .0001$ , which again had smaller N170 amplitudes than NORMAL SLOPE faces,  $F(1,18) = 8.89, p = .008$ . For the novel faces, N170 amplitudes were also larger for SHALLOW SLOPE faces compared to NORMAL SLOPE faces,  $F(1,18) = 12.71, p = .002$ . In contrast to learned faces, for novel faces, STEEP SLOPE faces evoked a larger N170 compared to NORMAL SLOPE faces,  $F(1,18) = 9.33, p = .007$  (see also Table 4).

**N250.** Mean amplitudes were influenced by the factors frequency,  $F(2,38) = 10.93, p < .0001$ , hemisphere,  $F(1,19) = 8.83, p < .01$ , and a trend for familiarity,  $F(1,19) = 3.43, p = .08$ . Furthermore, there was a two-way interaction between familiarity and frequency,  $F(2,38) = 4.31, p < .05$ . Frequency effects were only significant for learned faces,  $F(2,38) = 14.53, p < .0001$ , ( $F(2,38) < 1$  for novel faces). Post-hoc ANOVAs for the learned faces revealed more negative amplitudes for NORMAL SLOPE faces compared to STEEP SLOPE faces  $F(1,19) = 12.37, p = .002$  (Table 4). For SHALLOW SLOPE vs. NORMAL SLOPE faces, no significant difference was found  $F(1,19) < 1$  (Figs. 6 and 7). The factor familiarity reached significance for SHALLOW SLOPE faces,  $F(1,19) = 5.38, p = .032$  (Fig. 9), and NORMAL SLOPE faces,  $F(1,19) = 4.77, p = .04$ . For STEEP SLOPE faces there was no effect of familiarity,  $F(1,19) < 1$ .

**500–600 ms.** For central-parietal positivity, there were main effects of frequency,  $F(2,38) = 11.81, p < .0001$  (see also Fig. 6) and familiarity,  $F(1,19) = 46.67, p < .0001$ , with larger positivity

<sup>4</sup> We also performed an equivalent analysis with mean amplitudes (time segment  $x-y$  ms). This analysis revealed a significant main effect of frequency,  $F(2,36) = 3.60, p < .05$ , as well as a significant two-way interaction between frequency and site,  $F(2,36) = 3.66, p < .05$ .

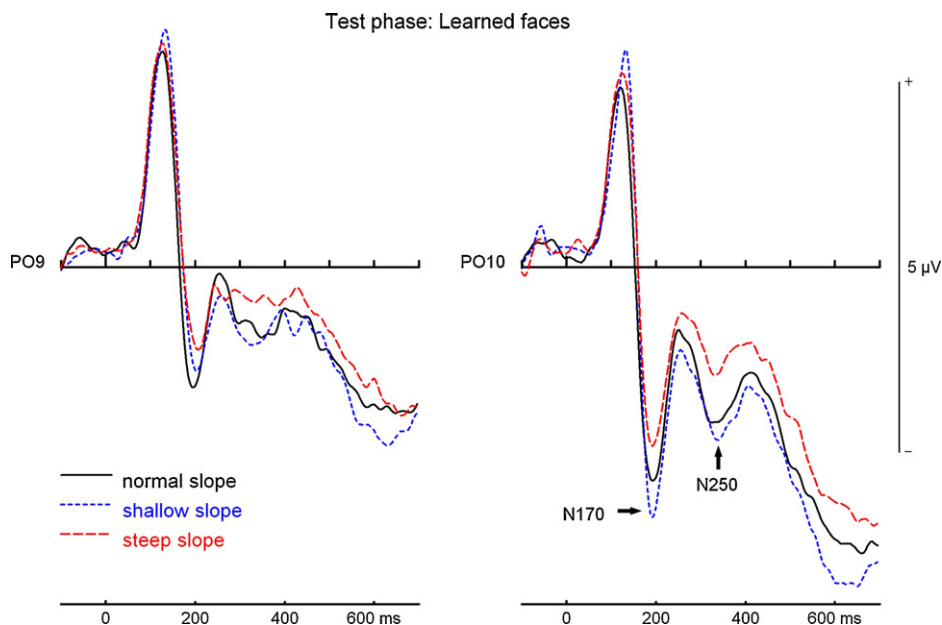


**Fig. 6.** ERPs at selected electrodes of the regions of interest (ROI) in the test phase for learned and novel faces.

for learned faces (Fig. 9). These effects were further qualified by an interaction between familiarity and frequency,  $F(2,38) = 6.45$ ,  $p < .01$ , and a three-way interaction between familiarity, frequency and electrode,  $F(10,190) = 2.17$ ,  $p < .05$ . Post-hoc ANOVAs for learned faces revealed more positive amplitudes for SHALLOW SLOPE faces compared to NORMAL SLOPE faces,  $F(1,19) = 5.82$ ,  $p = .026$ , which again were more positive than amplitudes for

STEEP SLOPE faces,  $F(1,19) = 14.40$ ,  $p = .001$  (Table 4). For the novel faces, significant differences were limited to electrodes P4 and Pz, where amplitudes were more positive for NORMAL SLOPE than for SHALLOW SLOPE and STEEP SLOPE faces,  $F(1,19) = 4.6$ ,  $p = .04$ , and  $F(1,19) = 4.49$ ,  $p = .04$ , respectively (see also Fig. 6).

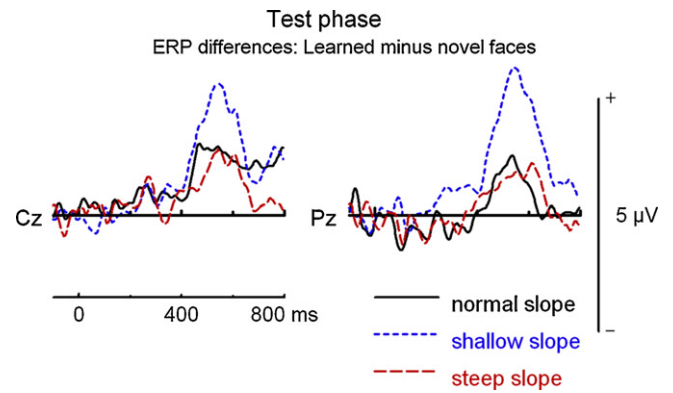
We also analyzed central-parietal positivity as an old/new effect (learned minus novel faces). The ANOVA revealed a sig-



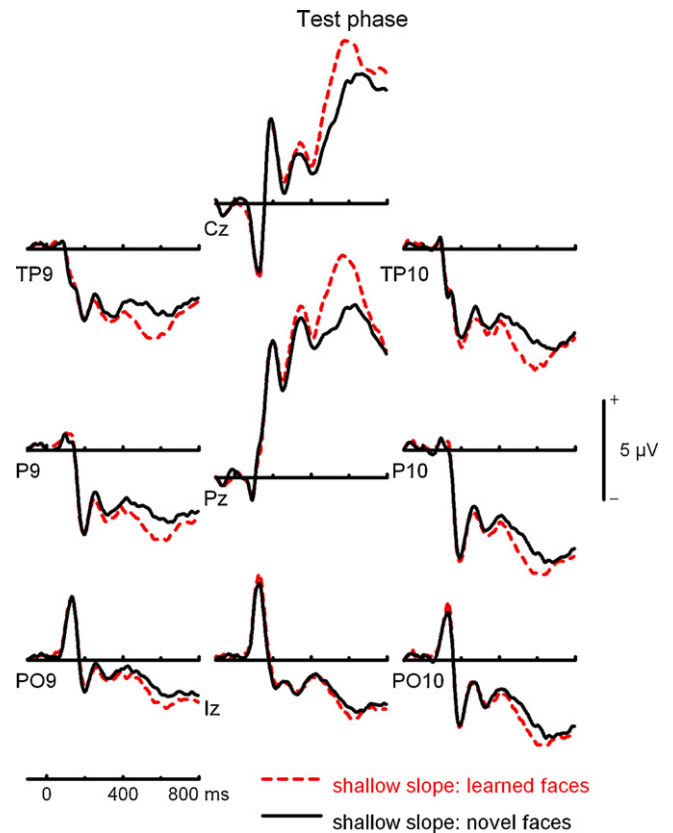
**Fig. 7.** ERPs at electrode sites PO9/PO10 in the test phase for learned faces. Note the increase of N170 absolute amplitude from STEEP SLOPE to NORMAL SLOPE to SHALLOW SLOPE faces over the right hemisphere (see arrow). N250 amplitudes were more negative for NORMAL SLOPE faces in comparison to STEEP SLOPE faces over the right hemisphere, whereas no differences between SHALLOW SLOPE and STEEP SLOPE faces were found (see arrow).

**Table 4**  
Amplitudes in  $\mu\text{V}$  in the test phase. P100 at IZ; N170 at P9/PO9 (left) and P10/PO10 (right hemisphere); N250 at P9/TP10, PO9/PO10, TP9/TP10; central positivity and old/new effect at C3/C4, P3/P4, Cz and Pz. Results that revealed significant effects compared to NORMAL SLOPE faces are marked with \*.

Image category	P100 (Iz) peak amplitude in $\mu\text{V}$ (SEM)		N170 (left) peak amplitude in $\mu\text{V}$ (SEM)		N170 (right) peak amplitude in $\mu\text{V}$ (SEM)		N250 mean amplitude in $\mu\text{V}$ (SEM)		Central positivity mean amplitude in $\mu\text{V}$ (SEM)		Old/new effect mean amplitude in $\mu\text{V}$ (SEM)	
	Learned	Novel	Learned	Novel	Learned	Novel	Learned	Novel	Learned	Novel	Learned minus novel	Learned minus novel
NORMAL SLOPE	4.1 (0.6)	3.9 (0.6)	-3.1 (0.5)	-4.1 (0.4)	-2.9 (0.3)	-2.6 (0.3)	6.9 (0.5)	5.9 (0.5)	1.0 (0.3)			
SHALLOW SLOPE	4.9 (0.6)*	4.5 (0.6)*	-3.1 (0.4)	-4.7 (0.4)*	-3.0 (0.3)	-2.7 (0.3)	7.7 (0.5)*	5.5 (0.6)	2.2 (0.3)*			
STEEP SLOPE	5.3 (0.6)*	5.1 (0.7)*	-2.7 (0.4)*	-4.6 (0.4)*	-2.4 (0.3)*	-2.5 (0.3)	5.9 (0.4)*	5.2 (0.4)	0.7 (0.3)			



**Fig. 8.** Old/new effects (learned minus novel faces) at Cz and Pz for NORMAL SLOPE, SHALLOW SLOPE and STEEP SLOPE faces.



**Fig. 9.** ERPs in the test phase for learned and novel faces averaged across frequency conditions. Larger absolute amplitudes for learned faces were observed for N250 and central positivity (old/new effect).

nificant effect of frequency,  $F(2,38)=6.45, p < .01$ , which was further modulated by an interaction between frequency and site,  $F(10,190)=2.17, p < .05$ . Subsequent comparisons suggested a larger old/new effect for SHALLOW SLOPE than for NORMAL SLOPE faces,  $F(1,19)=5.98, p = .024$ , whereas there was no difference in the old/new effect between NORMAL and STEEP SLOPE faces,  $F(1,19) < 1$  (see also Table 4 and Fig. 8).

**4. Discussion**

In the present experiment, previously unfamiliar faces had to be learned and later discriminated from novel faces. One third of the faces each were unmodified (NORMAL SLOPE) or had manipulated power spectra (SHALLOW SLOPE faces and STEEP SLOPE faces, respectively).



The behavioural results for the learned faces show longer RTs for the STEEP SLOPE faces in comparison to NORMAL SLOPE, whereas no differences between NORMAL SLOPE faces and SHALLOW SLOPE faces were found. These results indicate a disadvantage for STEEP SLOPE faces in this face identification task (Table 1). This effect was not present for accuracy, possibly due to high performance levels close to ceiling. Indeed, several participants achieved an accuracy of 100% in one or multiple conditions. In an ERP study by Goffaux et al. (2003), accuracy was lower for HSF than LSF faces and RTs did not differ between these two conditions in the familiarity task. One reason for the discrepant results could be the difference in stimulus manipulation. The bandwidth frequency-filtered faces result in HSF faces that have less spectral power and also lower contrast than LSF images (Nakashima et al., 2008). The faces that we used contain the entire frequency range and therefore the SHALLOW SLOPE faces had the same overall spectral power as the NORMAL SLOPE and STEEP SLOPE faces.

Stronger effects of the frequency manipulation were found in the ERP data. Peak analyses for the P100 component showed that its amplitude was larger and latency was longer for the manipulated than for the unmodified faces (Fig. 6 and Table 2). This early component is largely influenced by low-level stimulus content and sometimes associated with early face encoding processes (Herrmann et al., 2005). The findings are also in line with reports that atypical faces elicit a larger P100 than typical faces (Halit et al., 2000) and that a distortion of a face through inverting the eye region causes increased P100 latency (Doi et al., 2007). In this case, the longer latency and larger amplitude could be explained in terms of early face processing being more difficult for frequency manipulated stimuli. In the study by Flevaris et al. (2008), a delay of the P100 latency (but no increase in amplitude) was observed for filtered images. This absence of the amplitude effect might be due to the fact that they analyzed P100 at posterior lateral electrodes. In the present study, P100 amplitude effects were only significant at inferior occipital midline electrode IZ. Furthermore P100 was largest for STEEP SLOPE faces. Similar findings were obtained in a study by Nakashima et al. (2008) that showed increased P100 amplitudes for LSF faces and decreased amplitudes for HSF faces. In our study, this decrease of P100 amplitudes for SHALLOW SLOPE faces was not observed, possibly because, in contrast to frequency bandwidth-filtered stimuli, our SHALLOW SLOPE faces still contained low spatial frequencies. The larger P100 amplitudes for learned compared to novel faces cannot be explained by low-level stimulus content because novel and learned faces contained the same number of NORMAL SLOPE, SHALLOW SLOPE and STEEP SLOPE stimuli.

N170 amplitudes in the learning and the test phase were larger for SHALLOW SLOPE compared to NORMAL SLOPE stimuli (Figs. 5–7) and might represent the processing of fine features of a face. Furthermore N170 amplitudes for the learned faces in the test phase were smaller for STEEP SLOPE compared to NORMAL SLOPE stimuli. Similar results were found by Nakashima et al. (2008), after increasing the spectral power of the HSF stimuli while keeping the contrast normalized. In other ERP studies, smaller N170 amplitudes were found for HSF in comparison to LSF stimuli (Goffaux et al., 2003; Halit et al., 2006), but these authors used HSF faces that contained less overall spectral energy. This result suggests that the N170 effect does not just depend on the frequency range, but also on the spectral energy within this particular range. The latency of the N170 was longer for the manipulated stimuli than for the unmodified stimuli (Table 2). This difference was most likely caused by latency differences before the N170, because a comparable effect was already observed for the P100 latency (Flevaris et al., 2008).

The negative occipito-temporal N250 has been associated with the successful acquisition and activation of face representations (Kaufmann et al., 2009; Tanaka et al., 2006). In line with this

idea, the present N250 was increased for learned in comparison to novel faces for the NORMAL SLOPE and SHALLOW SLOPE conditions (Fig. 9). In the learning phase, the N250 was larger for SHALLOW SLOPE than NORMAL SLOPE faces, but in the test phase, this difference was not significant anymore (Fig. 7). One explanation for the different N250 effect would be the different task requirements in the study and test phase (Trenner et al., 2004). As another explanation, detailed information as provided by the SHALLOW SLOPE faces (Boeschoten et al., 2005) could be more important for face learning, while the face is still relatively unfamiliar. After the unfamiliar faces had been repeated four times during the learning phase, SHALLOW SLOPE and NORMAL SLOPE stimuli produced comparable amplitudes of N250 while performing memory judgements repeatedly in the test phase. However, the N250 amplitude of the learned faces decreased for STEEP SLOPE in comparison to NORMAL SLOPE faces in the test phase (Fig. 7). This decrease may reflect the fact that the STEEP SLOPE faces, which contain less spectral power above 9 cpi, cause a less reliable representation of the identity of a face than the NORMAL SLOPE faces. The N250 for novel faces was unaffected by the frequency manipulation, most likely because these faces were unfamiliar and not learned intentionally.

The central-parietal positivity associated with modality-independent person recognition (Paller et al., 2000) was larger for learned than for novel faces (Fig. 9). Whereas no frequency effects were found in the learning phase, during the test phase amplitudes of the central-parietal positivity for the learned faces were larger for SHALLOW SLOPE than NORMAL SLOPE faces and larger for NORMAL SLOPE than STEEP SLOPE faces (Fig. 6). Furthermore the old/new effect of the central-parietal positivity was larger for SHALLOW SLOPE than for NORMAL SLOPE faces, and smaller for STEEP SLOPE compared to NORMAL SLOPE faces (Fig. 8), suggesting facilitated retrieval of visual face information in the SHALLOW SLOPE condition (Paller et al., 2000). This might also reflect more reliable representations and more confident responses to faces learned as SHALLOW SLOPE versions.

Most of the frequency effects were over right hemispheric electrodes and reflected a bias for SHALLOW SLOPE faces (high spatial frequencies enhanced). This contradicts previous literature that found a right hemispheric bias towards low spatial frequencies (Sergent, 1982). However it should be noted that empirical evidence for the hemispheric spatial frequency hypothesis has been equivocal (Grabowska and Nowicka, 1996; Sergent, 1987).

This study shows that STEEP SLOPE in comparison to NORMAL SLOPE faces exhibit a disadvantage for the acquisition of new face representations. Due to low number of available images, stimuli were presented in the same version during learning and test phases. Our results, like those of previous research (e.g., Tanaka et al., 2006), may mainly reflect pictorial face learning and do not reflect generalization to recognition from other image variants. This is a limitation of our and many other studies on electrophysiological correlates of face learning (for further discussion see also Kaufmann et al., 2009). Nevertheless, using the same images at learning and test typically maximizes the effects obtained, which may be appropriate for a first study on this topic. The repeated presentation of the same stimuli causes repetition effects as a confounding factor. This confound should not have caused differences between conditions, as multiple repetitions were used equally in all conditions. Therefore we would expect similar results with one presentation during learning and test.

In contrast to previous studies, the stimuli that we used were not frequency bandwidth-filtered, but their ratio of fine detail (HSF) to coarse structures (LSF) was altered. For the STEEP SLOPE stimuli, reaction times were slower and amplitudes of N170, N250 and central-parietal positivity were decreased in the test phase. STEEP SLOPE faces contained attenuated fine detail and enhanced coarse structures. In the context of the present study, it is therefore likely

that face identification – at least at a relatively early stage of familiarization – is impaired if the ratio of fine detail to coarse structures is diminished. STEEP SLOPE faces contain more spectral power from 0 to 22 cpi in comparison to SHALLOW SLOPE stimuli and therefore, the results indicate an advantage for faces containing more spectral power above 22 cpi and a disadvantage for faces that have a more negative slope of the power spectrum (Fig. 1). These findings are in line with results from an fMRI study by Vuilleumier et al. (2003), who found that LSF faces did not lead to a lasting representation in fusiform cortex. On the other hand, these findings contradict previous research, which suggested a preference for the range of 8–16 cpi in face recognition tasks (Costen et al., 1996; Näsänen, 1999), because the SHALLOW SLOPE faces that we used had less spectral energy than STEEP SLOPE faces below 22 cpi. In contrast to frequency bandwidth filtering, the method applied here enhances or diminishes the spectral power in certain frequency ranges. Therefore, our face stimuli still contained the whole frequency range. Different stimuli, which our visual system perceives under normal everyday conditions, usually contain the whole frequency range and vary in their spectral slope and are not frequency bandwidth filtered. Our approach may thus provide insight into the role of different spatial frequencies under more natural conditions.

Furthermore, the present electrophysiological results might indicate an advantage for SHALLOW SLOPE stimuli in comparison to NORMAL SLOPE stimuli for face learning. The SHALLOW SLOPE images produced larger amplitudes of N170 and central-parietal positivity in the test phase and a larger N250 in the learning phase. The latter finding suggests that SHALLOW SLOPE faces facilitate the acquisition of a new face representation (cf. Tanaka et al., 2006). This finding could imply that face identification is improved when the ratio of fine detail to coarse structure is increased, as provided by SHALLOW SLOPE stimuli. This effect might be stronger if a face is seen for the first time, because the N250 was larger for SHALLOW SLOPE than NORMAL SLOPE faces in the learning phase but not in the test phase. The slope of the SHALLOW SLOPE stimuli was chosen to correspond to the slope that was found in photographs of complex natural scenes (around  $-2$ ;  $1/f^2$  characteristics). It has been shown that the human visual system encodes natural scenes efficiently (Simoncelli and Olshausen, 2001; Vinje and Gallant, 2000).  $1/f^2$  characteristics have also been found in art images (Graham and Field, 2007; Redies et al., 2007a), including face portraits drawn by artists (Redies et al., 2007b). This result suggested that artists may have an intuitive knowledge of image statistics and prefer statistical regularities similar to those of natural scenes in their artworks. The present electrophysiological results demonstrate that face stimuli with  $1/f^2$  characteristics result in an increase of ERP components that have been associated with the successful acquisition of new face representations, when compared to the unmodified stimuli. It is unclear whether this effect is face specific, because we did not use non-face images in this study (Scott et al., 2006; Tanaka et al., 2006). Moreover, Párraga et al. (2000) showed that performance in an object discrimination task is higher for images with  $1/f^2$  characteristics than for images that deviate from these statistics. Taken together, these results indicate that, besides natural scenes, other categories of images with  $1/f^2$  characteristics might also be encoded more efficiently by the human visual system.

## 5. Conclusions

Our results suggest more efficient face learning for stimuli with an increased ratio of fine detail to coarse structures, and less efficient face identity learning for stimuli with a diminished ratio. Of particular interest is the finding that SHALLOW SLOPE stimuli with a slope of  $-2$ , similar to natural scenes and artwork, led to an

increase of ERP components that have been associated with the successful acquisition of new face representations, in comparison to unmodified stimuli. The results suggest that enhancing fine structures (high spatial frequencies) facilitates learning the identity of an unfamiliar face.

## Acknowledgments

We thank Dr. Aleix Martinez for permission to use the AR face database. This work was supported by grants from the German Research Council (Deutsche Forschungsgemeinschaft, DFG, grant numbers Re 616/5-1 and Ka 2997/2-1).

## References

- Bentin, S., Deouell, L.Y., 2000. Structural encoding and identification in face processing: ERP evidence for separate processes. *Cognitive Neuropsychology* 17, 35–54.
- Boeschoten, M., Kemner, C., Kenemans, J., Engeland, H., 2005. The relationship between local and global processing and the processing of high and low spatial frequencies studied by event-related potentials and source modeling. *Brain Research Cognitive Brain Research* 24, 228–236.
- Burton, G., Moorhead, I., 1987. Color and spatial structure in natural scenes. *Applied Optics* 26, 157–170.
- Caharel, S., d'Arripe, O., Ramon, M., Jacques, C., Rossion, B., 2009. Early adaptation to repeated unfamiliar faces across viewpoint changes in the right hemisphere: evidence from the N170 ERP component. *Neuropsychologia* 47, 639–643.
- Collin, C., Liu, C., Troje, N., McMullen, P., Chaudhuri, A., 2004. Face recognition is affected by similarity in spatial frequency range to a greater degree than within-category object recognition. *Journal of Experimental Psychology: Human Perception and Performance* 30, 975–987.
- Costen, N., Parker, D., Craw, I., 1994. Spatial content and spatial quantisation effects in face recognition. *Perception* 23, 129–146.
- Costen, N., Parker, D., Craw, I., 1996. Effects of high-pass and low-pass spatial filtering on face identification. *Perception & Psychophysics* 58, 602–612.
- De Valois, R., De Valois, K., 1980. Spatial vision. *Annual Review Psychology* 31, 309–341.
- Doi, H., Sawada, R., Masataka, N., 2007. The effects of eye and face inversion on the early stages of gaze direction perception - an ERP study. *Brain Research* 1183, 83–90.
- Eger, E., Schyns, P., Kleinschmidt, A., 2004. Scale invariant adaptation in fusiform face-responsive regions. *Neuroimage* 22, 232–242.
- Eimer, M., 2000a. Event-related brain potentials distinguish processing stages involved in face perception and recognition. *Clinical Neurophysiology* 111, 694–705.
- Eimer, M., 2000b. The face-specific N170 component reflects late stages in the structural encoding of faces. *Neuroreport* 11, 2319–2324.
- Fiorentini, A., Maffei, L., Sandini, G., 1983. The role of high spatial frequencies in face perception. *Perception* 12, 195–201.
- Flevaris, A., Robertson, L., Bentin, S., 2008. Using spatial frequency scales for processing face features and face configuration: an ERP analysis. *Brain Research* 1194, 100–109.
- Ginsburg, A.P., 1978. Visual information processing based on spatial filters constrained by biological data. Doctoral Thesis, Air Force Aerospace Medical Research Laboratory.
- Goffaux, V., Hault, B., Michel, C., Vuong, Q., Rossion, B., 2005. The respective role of low and high spatial frequencies in supporting configural and featural processing of faces. *Perception* 34, 77–86.
- Goffaux, V., Jemel, B., Jacques, C., Rossion, B., Schyns, P.G., 2003. ERP evidence for task modulations on face perceptual processing at different spatial scales. *Cognitive Science* 27, 313–325.
- Grabowska, A., Nowicka, A., 1996. Visual-spatial-frequency model of cerebral asymmetry: a critical survey of behavioral and electrophysiological studies. *Psychological Bulletin* 120, 434–449.
- Graham, D., Field, D., 2007. Statistical regularities of art images and natural scenes: spectra, sparseness and nonlinearities. *Spatial Vision* 21, 149–164.
- Halit, H., de Haan, M., Johnson, M., 2000. Modulation of event-related potentials by prototypical and atypical faces. *Neuroreport* 11, 1871–1875.
- Halit, H., de Haan, M., Schyns, P., Johnson, M., 2006. Is high-spatial frequency information used in the early stages of face detection? *Brain Research* 1117, 154–161.
- Harmon, L., Julesz, B., 1973. Masking in visual recognition: effects of two-dimensional filtered noise. *Science* 180, 1194–1197.
- Herrmann, M., Ehlis, A., Ellgring, H., Fallgatter, A., 2005. Early stages (P100) of face perception in humans as measured with event-related potentials (ERPs). *Journal of Neural Transmission* 112, 1073–1081.
- Hoyer, P., Hyvärinen, A., 2002. A multi-layer sparse coding network learns contour coding from natural images. *Vision Research* 42, 1593–1605.
- Hsiao, F., Hsieh, J., Lin, Y., Chang, Y., 2005. The effects of face spatial frequencies on cortical processing revealed by magnetoencephalography. *Neuroscience Letters* 380, 54–59.

- Huynh, H., Feldt, L., 1976. Estimation of the Box correction for degrees of freedom from sample data in randomized block and split-plot designs. *Journal of Educational Statistics* 1, 69–82.
- Kanwisher, N., McDermott, J., Chun, M., 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience* 17, 4302–4311.
- Kaufmann, J.M., Schweinberger, S.R., 2008. Distortions in the brain? ERP effects of caricaturing familiar and unfamiliar faces. *Brain Research* 1228, 177–188.
- Kaufmann, J.M., Schweinberger, S.R., Burton, A.M., 2009. N250 ERP correlates of the acquisition of face representations across different images. *Journal of Cognitive Neuroscience* 21, 625–641.
- Keil, M., 2008. Does face image statistics predict a preferred spatial frequency for human face processing? *Proceedings of the Royal Society B: Biological Sciences* 275, 2095–2100.
- Livingstone, M., Hubel, D., 1988. Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science* 240, 740–749.
- Loftus, G., Harley, E., 2004. How different spatial-frequency components contribute to visual information acquisition. *Journal of Experimental Psychology: Human Perception and Performance* 30, 104–118.
- Martinez, A.M., Benavente, R., 1998. *The AR Face Database (vol. 24)*. CVC Technical Report.
- Morrison, D., Schyns, P., 2001. Usage of spatial scales for the categorization of faces, objects, and scenes. *Psychonomic Bulletin & Review* 8, 454–469.
- Nakashima, T., Kaneko, K., Goto, Y., Abe, T., Mitsudo, T., Ogata, K., et al., 2008. Early ERP components differentially extract facial features: evidence for spatial frequency- and-contrast detectors. *Neuroscience Research* 62, 225–235.
- Näsänen, R., 1999. Spatial frequency bandwidth used in the recognition of facial images. *Vision Research* 39, 3824–3833.
- Oldfield, R., 1971. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113.
- Olshausen, B., Field, D., 1996. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature* 381, 607–609.
- Paller, K., Gonsalves, B., Grabowecky, M., Bozic, V., Yamada, S., 2000. Electrophysiological correlates of recollecting faces of known and unknown individuals. *Neuroimage* 11, 98–110.
- Parker, D.M., Costen, N.P., 1999. One extreme or the other or perhaps the Golden Mean? Issues of spatial resolution in face processing. *Current Psychology: Developmental, Learning, Personality* 18, 118–127.
- Párraga, C., Troscianko, T., Tolhurst, D., 2000. The human visual system is optimised for processing the spatial information in natural visual images. *Current Biology* 10, 35–38.
- Redies, C., 2007. A universal model of esthetic perception based on the sensory coding of natural stimuli. *Spatial Vision* 21, 97–117.
- Redies, C., Hasenstein, J., Denzler, J., 2007a. Fractal-like image statistics in visual art: similarity to natural scenes. *Spatial Vision* 21, 137–148.
- Redies, C., Hänisch, J., Blickhan, M., Denzler, J., 2007b. Artists portray human faces with the Fourier statistics of complex natural scenes. *Network* 18, 235–248.
- Ruiz-Soler, M., Beltran, F., 2006. Face perception: an integrative review of the role of spatial frequencies. *Psychological Research* 70, 273–292.
- Schweinberger, S.R., Pickering, E., Burton, A.M., Kaufmann, J.M., 2002a. Human brain potential correlates of repetition priming in face and name recognition. *Neuropsychologia* 40, 2057–2073.
- Schweinberger, S.R., Pickering, E., Jentzsch, I., Burton, A.M., Kaufmann, J.M., 2002b. Event-related brain potential evidence for a response of inferior temporal cortex to familiar face repetitions. *Brain Research Cognitive Brain Research* 14, 398–409.
- Schyns, P., Bonnar, L., Gosselin, F., 2002. Show me the features! Understanding recognition from the use of visual information. *Psychological Science* 13, 402–409.
- Schyns, P., Oliva, A., 1999. Dr. Angry and Mr. Smile: when categorization flexibly modifies the perception of faces in rapid visual presentations. *Cognition* 69, 243–265.
- Scott, L.S., Tanaka, J.W., Sheinberg, D.L., Curran, T., 2006. A reevaluation of the electrophysiological correlates of expert object processing. *Journal of Cognitive Neuroscience* 18, 1453–1465.
- Sergent, J., 1982. The cerebral balance of power: confrontation or cooperation? *Journal of Experimental Psychology: Human Perception and Performance* 8, 253–272.
- Sergent, J., 1987. Failures to confirm the spatial-frequency hypothesis: fatal blow or healthy complication. *Canadian Journal of Psychology* 41, 412–428.
- Simoncelli, E., Olshausen, B., 2001. Natural image statistics and neural representation. *Annual Reviews of Neuroscience* 24, 1193–1216.
- Tanaka, J., Curran, T., Porterfield, A., Collins, D., 2006. Activation of preexisting and acquired face representations: the N250 event-related potential as an index of face familiarity. *Journal of Cognitive Neuroscience* 18, 1488–1497.
- Trenner, M.U., Schweinberger, S.R., Jentzsch, I., Sommer, W., 2004. Face repetition effects in direct and indirect tasks: an event-related brain potentials study. *Brain Research Cognitive Brain Research* 21, 388–400.
- Tolhurst, D., Tadmor, Y., Chao, T., 1992. Amplitude spectra of natural images. *Ophthalmic and Physiological Optics* 12, 229–232.
- Vinje, W., Gallant, J., 2000. Sparse coding and decorrelation in primary visual cortex during natural vision. *Science* 287, 1273–1276.
- Vinje, W., Gallant, J., 2002. Natural stimulation of the nonclassical receptive field increases information transmission efficiency in V1. *Journal of Neuroscience* 22, 2904–2915.
- Vuilleumier, P., Armony, J., Driver, J., Dolan, R., 2003. Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nature Neuroscience* 6, 624–631.
- Westheimer, G., 2001. The Fourier theory of vision. *Perception* 30, 531–541.
- Zeki, S., 1999. *An Exploration of Art and the Brain*. University Press, Oxford.

## 4 Diskussion

In der vorliegenden Arbeit wurden universelle Eigenschaften von ästhetischen Bildern anhand von künstlerischen Porträts untersucht und der Einfluss dieser Eigenschaften auf die Wahrnehmung von Gesichtern getestet. Dadurch konnten Erkenntnisse für die Neuroästhetik und die Gesichts-Wahrnehmungsforschung gewonnen werden.

Der erste Artikel dieser Arbeit (siehe Kapitel 3.1) zeigte, dass durch Künstler erstellte Porträts einen ähnlichen Anstieg im Fourierspektrum wie natürliche Szenen besitzen, obwohl Gesichtsfotos über einen anderen Anstieg im Fourierspektrum verfügen (Torralba und Oliva 2003, Bosworth et al. 2006). Bestehende Ergebnisse unserer (Redies et al. 2007) sowie einer amerikanischen Arbeitsgruppe (Graham und Field 2007) wurden dadurch bestätigt und ausgeweitet. Damit liefert die Studie einen weiteren Baustein für die Theorie, dass Künstler ästhetische Bilder an unser visuelles System anpassen und dabei basale Funktionsweisen des visuellen Systems nutzen (Redies 2007, Zeki 1999).

Im Anschluss daran manipulierten wir die Fourierspektren von Gesichtsfotos, um zu untersuchen, wie sich ein Anstieg desselben auf die Gesichterwahrnehmung mittels eines Gesichter-Lernexperiments mit Erfassung des EEGs auswirkt (siehe Kapitel 3.2). Eine Veränderung des Anstiegs des radiär gemittelten Fourierspektrums von Gesichtsfotos auf  $-2$  (ähnlich dem Befund bei natürlichen Szenen und ästhetischen Bildern) erleichterte das Erlernen und Wiederkennen unbekannter Gesichter. Ebenso scheint eine Veränderung des relativen Anteils an feinen und groben Strukturen zu Gunsten der feinen Strukturen (mehr HSF) das Gesichterlernen zu vereinfachen. Umgekehrt führten Verstärkung der groben Strukturen und ein Abschwächen der feinen Strukturen zu erschwertem Gesichterlernen. Diese Ergebnisse werden durch Literaturbefunde gestützt, die zeigen, dass neben natürlichen Szenen auch andere Stimuli mit  $1/f^2$ -Eigenschaften effizient codiert werden können (Párraga et al. 2000).



## **4.1 Kritische Betrachtung der Methodik**

In der vorliegenden Arbeit wurden  $1/f^2$ -Eigenschaften eindimensional erfasst, sodass keine Aussagen über die Isotropie des Fourierspektrums möglich sind (Koch et al. 2010). In Folgestudien könnte dies sowohl bei der Analyse als auch bei der Manipulation des Fourierspektrums berücksichtigt werden.

Des Weiteren wurden nur die Amplituden des Fourierspektrums betrachtet und die Phase vernachlässigt, obwohl darin ein Teil der Bildinformation steckt. Folgestudien könnten höhere statistische Verfahren anwenden, die auch die Phase in ihrem Modell berücksichtigen (Koch et al. 2010, Graham und Field 2007).

Rousselet et al. (2008) nutzten 2D-Gauß-Filter, um Randeffekte zwischen Gesicht und Hintergrund zu vermeiden. Wir manipulierten den Anstieg ohne diese Filter, und somit kann es zu Frequenzeffekten zwischen Gesichtern und Hintergrund gekommen sein. Deshalb wurde der Hintergrund bei unserer Methode nach der Manipulation des Fourierspektrums auf einen einheitlichen Grauwert gesetzt.

In unserem Lernexperiment wurden sowohl in der Lern- als auch in der Testphase die gleichen Fotos mit der gleichen Ansicht von Gesichtern verwendet, sodass hauptsächlich Aussagen über bildliches Lernen getroffen werden konnten. Des Weiteren haben alle Probanden dieselben Gesichter in der gleichen Frequenzmanipulation gesehen. Um weniger bildliches Lernen zu betrachten, sollte jeder Proband andere manipulierte Versionen der Gesichter gezeigt bekommen. Außerdem wurden die ungelerten Gesichter in der Testphase drei Mal wiederholt, weil uns nur eine begrenzte Anzahl an Stimuli zu Verfügung stand. Die Interpretation der Ergebnisse der ungelerten Gesichter ist dadurch eingeschränkt.

In den Verhaltensdaten zeigte sich ein Deckeneffekt für die Wiedererkennungsrate, weil mehrere Probanden 100% richtige Antworten für die verschiedenen Gesichtsm Manipulationen erreichten. Um signifikante Verhaltensdaten zu erhalten, sollte die Schwierigkeit der Aufgabe in zukünftigen Studien erhöht werden, um etwas geringere Wiedererkennungsdaten zu erhalten.

Die wichtigste der oben genannten Limitationen ist, dass unsere Ergebnisse auf bildliches Lernen zutreffen. Mit Hilfe der für diese Studie neu entwickelten Stimulus-Manipulation konnte der Einfluss des Anstiegs des Fourierspektrums auf die Gesichtswahrnehmung untersucht werden.

## **4.2 Fourierspektrum von Porträts und Gesichtern**

Für gemalte ästhetische Porträts wurden ähnliche  $1/f^2$ -Eigenschaften wie für natürliche Szenen gefunden. Künstler bilden Gesichter demzufolge nicht mit den gleichen statistischen Eigenschaften ab, die Fotos von Gesichtern besitzen. Außerdem weisen gemalte Porträts einen höheren relativen Anteil von feinen gegenüber groben Strukturen als Fotos von Gesichtern auf. In unserer Arbeit war der Anstieg des Fourierspektrums von Porträts unabhängig von Hintergrund, Kopfbedeckung, Geschlecht, Blickwinkel, Jahrhundert, Herkunftsland und Maltechnik. Damit konnte gezeigt werden, dass das Fourierspektrum nicht durch die Komplexität des Hintergrundes oder andere zuvor genannte Störfaktoren beeinflusst wurde.

Des Weiteren versuchten wir den Einfluss von Artefakten auf unsere Messung zu minimieren. Bei der Bestimmung des Anstiegs wurde nur der Bereich von 10 bis 256 cpi berücksichtigt, sodass Artefakte am Rande des Fourierspektrums nicht erfasst wurden. Beim Fotografieren, Scannen und Drucken werden Helligkeit und Farbe zu Pixelwerten transformiert und es kann zu nichtlinearen Effekten kommen, die als Gamma-Gradation bezeichnet werden. In einem Kontrollexperiment konnten wir zeigen, dass Änderungen der Gamma-Gradation einen sehr geringen Einfluss auf den Anstieg im Fourierspektrum besitzen. Ähnliche Ergebnisse wurden zuvor auch für Kontrast und mittleren Grauwert gezeigt (Ruderman 1997). Es ist demnach nicht anzunehmen, dass unsere Daten maßgeblich durch Reproduktionsartefakte beeinflusst wurden. Ursprünglich mit Farbe gemalte Porträts wurden in unserer Analyse zu Graustufen konvertiert, sodass die Farbinformation verloren ging. Dadurch enthielten diese gemalten Porträts weniger feine Strukturen und der Anstieg wurde steiler (Anstieg von -2.9).

Es konnte gezeigt werden, dass Künstler Gesichter mit Eigenschaften abbilden, die denen natürlicher Szenen ähneln. Da das visuelle System auf die Wahrnehmung natürlicher Szenen abgestimmt ist (Simoncelli und Olshausen 2001, Graham und Field 2009), deuten die Gemeinsamkeiten von ästhetischen Bildern und natürlichen Szenen darauf hin, dass Künstler ihre Werke an grundlegende Funktionen des visuellen Systems anpassen.

### **4.3 Fourierspektrum und Gesichterlernen**

Wir konnten zeigen, dass unbekannte STEEP SLOPE Gesichter mit einem steileren Anstieg (verstärkte grobe Strukturen und abgeschwächte feine Strukturen) schlechter gelernt werden können als unmodifizierte Gesichter. Dies wurde sowohl durch langsamere Reaktionszeiten als auch durch geringere Amplituden von N170, N250 und central positivity gezeigt (Erklärungen zu den EEG-Korrelaten siehe Kapitel 7.2). Die geringere Amplitude der N170 für STEEP SLOPE Gesichter kann dadurch erklärt werden, dass die N170 eher die Kodierung von feinen Strukturen der Gesichter repräsentiert (Nakashima et al. 2008). Die geringere N250 und central positivity deuten darauf hin, dass diese Gesichter zu einer weniger stabilen Repräsentation der Identität führen (Kaufmann et al. 2009, Tanaka et al. 2006).

Im Gegensatz zu den STEEP SLOPE Gesichtern gab es Hinweise darauf, dass SHALLOW SLOPE Gesichter mit einem Anstieg von  $-2$  ( $1/f^2$ -Eigenschaften, verstärkte feine Strukturen und abgeschwächte grobe Strukturen) besser gelernt werden konnten als unmodifizierte Gesichter. So fanden sich in den elektrophysiologischen Ergebnissen größere Amplituden der N170, N250 und central positivity, obwohl keine Unterschiede in den Verhaltensdaten nachgewiesen wurden. Die größere Amplitude der N170 spiegelt die Kodierung von feinen Strukturen der Gesichter wider (Nakashima et al. 2008). In anderen Studien wurden kleinere Amplituden für HSF- als für LSF-Gesichter gefunden (Goffaux et al. 2003, Halit et al. 2006). Dies ist vermutlich darauf zurückzuführen, dass diese Autoren HSF-Gesichter mit niedriger spektraler Energie genutzt haben. Die größere N250 und central positivity für SHALLOW SLOPE Gesichter in meiner Studie deuten darauf hin, dass diese Gesichter zu einer stabileren Repräsentation der Identität führen (Kaufmann et al. 2009). Allerdings war die N250 für SHALLOW SLOPE Gesichter nur in der Lernphase größer und nicht in der Testphase. Dies könnte zum einen an der unterschiedlichen Aufgabe in Lern und Testphase liegen (Gesichter betrachten und lernen vs. Gesichter aktiv wiedererkennen). Zum anderen könnten die feinen Strukturen der SHALLOW SLOPE Gesichter besonders nützlich sein, wenn Gesichter noch gänzlich unbekannt sind, wie das in der Lernphase der Fall ist. Nach den vier Wiederholungen riefen SHALLOW SLOPE und unmanipulierte Gesichter ähnliche N250 hervor und schnitten in den Verhaltensdaten ähnlich ab.

In einer Studie mit fMRT (Vuilleumier et al. 2003) wurde gezeigt, dass LSF-Gesichter im Gegensatz zu HSF-Gesichtern nicht zu Wiederholungseffekten führten. Feine Strukturen,

welche durch hohe Frequenzen im Amplitudenspektrum repräsentiert werden, spielen eine wichtige Rolle bei dem Erlernen unbekannter Gesichter (Fiorentini et al. 1983, Sergent 1986). So konnten Watier und Collin (2009) in einem Experiment, in dem Gesichter verglichen werden mussten, zeigen, dass für unbekannte Gesichter höhere Frequenzen als für bekannte Gesichter benötigt werden. Daraus schlussfolgerten sie, dass es Unterschiede bei der Frequenzwahl für bekannte und unbekannte Gesichter gibt.

Die SHALLOW SLOPE Gesichter unserer Studie, welche besser gelernt wurden, haben weniger spektrale Energie im niederfrequenten Bereich von 0-22 cpi als STEEP SLOPE Gesichter (siehe Abb. 13). Dies widerspricht bisherigen Studien, welche den Frequenzbereich von 8-16 cpi als wichtigsten Bereich für die Gesichtererkennung hervorgehoben haben (Costen et al. 1996, Näsänen 1999). Wenn sich der relative Anteil von groben zu feinen Strukturen verändert (also der Anstieg im Fourierspektrum), könnte die spektrale Energie in einem bestimmten Frequenzbereich (z.B. 8-16cpi) eine untergeordnete Rolle spielen.

Unsere Ergebnisse deuten darauf hin, dass Gesichter mit ähnlichem Fourierspektrum wie natürliche Szenen und ästhetische Bilder leichter gelernt werden. Dies könnte zum einen dadurch bedingt sein, dass für das Erlernen unbekannter Gesichter vor allem feine Strukturen von Bedeutung sind (Boeschoten et al. 2005, Goffaux et al. 2005, Schyns et al. 2002) und zum anderen dadurch, dass Stimuli mit  $1/f^2$ -Eigenschaften von unserem visuellen System besonders effizient codiert werden können (Párraga et al. 2000, Graham und Field 2009).

#### ***4.4 Bedeutung der Arbeit für die Gesichterforschung***

Soweit uns bekannt ist, handelt es sich hierbei um die erste Studie, welche den Einfluss des Anstiegs des Fourierspektrums auf die Gesichterwahrnehmung erforscht. Mit dieser Methode konnten konfundierende Störvariablen, wie sie bei Bandpass-gefilterten Stimuli auftraten, vermieden werden (z.B. unterschiedliche spektrale Energie der hohen und tiefen Frequenzen). Des Weiteren enthalten unsere Stimuli das ganze Fourierspektrum und wirken dadurch realistischer als Bandpass-gefilterte Gesichter. Objekte die wir im Alltag visuell wahrnehmen, variieren im Anstieg des Fourierspektrums und enthalten das gesamte Frequenzspektrum. Damit ist unsere Methode im Vergleich zu bisherigen Bandpassfilter-Methoden für die Gesichterforschung interessant und innovativ. Um den Einfluss vom Anstieg des Fourierspektrums auf die Gesichterwahrnehmung systematisch zu überprüfen,

könnten weitere Studien mit mehr als drei unterschiedlichen Anstiegen (-2; -3 und -3,5 in unserem Experiment) durchgeführt werden.

Die Aussagekraft unsere Ergebnisse ist begrenzt auf das bildhafte Lernen, da dieselben Gesichter in der Lern- und Testphase gezeigt wurden. Diese Beschränkung trifft auf einen Großteil der Studien zum Gesichterlernen zu und es gibt nur wenige Studien die verschiedene Gesichterversionen in Lern- und Testphase zeigen (Kaufmann et al. 2009). Durch das Verwenden gleicher Gesichter werden elektrophysiologische Effekte maximiert und wir halten deshalb dieses Vorgehen als angemessenen für eine erste Studie mit neuartiger Stimulusmanipulation. In weiterführenden Studien könnten Gesichter mit unterschiedlicher Ansicht gezeigt werden, um das Generalisationsvermögen über verschiedene Abbildungen einer Person zu erfassen. Des Weiteren verwendeten wir keine Kontrollstimuli, obwohl es eine intensive Diskussion darüber gibt, ob Effekte in der Gesichterforschung spezifisch für Gesichter sind oder auch für die Wahrnehmung anderer Objekte zutreffen (Bentin et al. 2007, Dering et al. 2009, Rossion und Jacques 2008, Thierry et al. 2007). Aussagen über die Wahrnehmung von Gesichtern können jedoch auch unabhängig von der Spezifitätsdebatte getroffen werden (Scott et al. 2006, Tanaka et al. 2006, Schweinberger 2011). Ob unsere Erkenntnisse auch für andere Objekte gelten, könnten weitere Studien unter Verwendung zum Beispiel von natürlichen Szenen und Häusern als Kontrollstimuli zeigen. Wenn  $1/f^2$ -Eigenschaften zu effizienterer Kodierung führen, ist anzunehmen, dass dies auch für andere Objekte als Gesichter gilt.

Unsere Ergebnisse lassen sich gut in das Modell des „diagnostic recognition approach“ einordnen (Ruiz-Soler und Beltran 2006), in dem Frequenzbestandteile des Fourierspektrums flexibel je nach Aufgabe und nach Frequenzzusammensetzung des Gesichtes genutzt werden können. Dadurch kann insbesondere erklärt werden, warum Ergebnisse über die Bedeutung des Fourierspektrums für die Gesichterwahrnehmung stark variieren und oft sehr unterschiedliche Frequenzbereiche hervorgehoben wurden. Unsere Ergebnisse gelten im Rahmen der Aufgabe unbekannte Gesichter zu erlernen, die durch einen steileren Anstieg im Fourierspektrum mit verstärkten groben und abgeschwächten feinen Strukturen erschwert wird. Um den Einfluss von  $1/f^p$ -Eigenschaften auf die Gesichterwahrnehmung unabhängig vom Gesichterlernen zu untersuchen, könnten in Folgeexperimenten Gesichter passiv wahrgenommen werden (Flevaris et al. 2008), sodass die Gesichter für die Aufgabe irrelevant sind.

Das Verstärken der feinen Strukturen von Gesichtern in unserer Studie könnte einen ähnlichen Effekt haben, wie er durch mit Linien gezeichnete Karikaturen hervorgerufen wird. In anderen Studien wurden solche Karikaturen besser gelernt als unmodifizierte Fotos von Gesichtern (Rhodes et al. 1997). In einer Studie mit fotorealistischen Karikaturen führten Karikaturen ebenfalls zu größerer N250 (Kaufmann und Schweinberger 2008, Kaufmann und Schweinberger 2012).

#### **4.5 Bedeutung der Arbeit für die Neuroästhetik**

In einer weiteren Studie unserer Arbeitsgruppe konnte gezeigt werden, dass außer ästhetischen Bildern auch Cartoons, Mangas und Comics  $1/f^2$ -Eigenschaften aufweisen (Koch et al. 2010). Somit gibt es auch andere von Menschenhand hergestellte Objekte, die für die ästhetische Wahrnehmung geschaffen wurden und ähnliche Fourierspektren wie natürliche Szenen besitzen. Sie unterscheiden sich zum Beispiel von wissenschaftlichen Illustrationen, die weniger starke ästhetische Wahrnehmung hervorrufen und ein flacheres Fourierspektrum haben. In der gleichen Studie konnte gezeigt werden, dass sich natürliche Szenen von ästhetischen Bildern durch eine größere Isotropie des 2D-Fourierspektrums unterscheiden (Koch et al. 2010). Ästhetische Bilder zeigen eine geringere Variabilität der Steigung des Fourierspektrums in den verschiedenen Orientierungen des Bildes mit der Folge, dass das Fourierspektrum in alle Richtungen gleichmäßig mit  $1/f^2$ -Eigenschaften abfällt. Da dies in unserer Manipulation vom Anstieg des Fourierspektrums nicht berücksichtigt wurde, könnte in einer Folgestudie versucht werden, Gesichter mit isotropen  $1/f^2$ -Eigenschaften zu erzeugen. Außerdem unterscheiden sich ästhetische Bilder von natürlichen Szenen durch ihre Leuchtdichte. Ästhetische Bilder haben einen wesentlich kleineren Dynamikumfang ihrer Leuchtdichte und eine weniger schiefe Leuchtdichtenverteilung (Graham 2008, Graham und Field 2008). Auch dies spricht dafür, dass Künstler in ihren Werken nicht die gleichen Eigenschaften der natürlichen Objekte widerspiegeln, wie sie in Fotografien abgebildet sind, sondern dass sie ihre Werke an Eigenschaften des visuellen Systems anpassen.

$1/f^2$ -Eigenschaften von ästhetischen Bildern werden zum einen möglicherweise von Künstlern genutzt, um die Perzeptibilität ihrer Werke durch das visuelle System zu erleichtern (Graham und Field 2008). Zum anderen nehmen die Autoren an, dass es auf Grund des Malprozesses sehr schwierig ist Bilder zu erzeugen, die keine  $1/f^2$ -Eigenschaften besitzen. So würde die menschliche Maltechnik, beispielsweise beim Malen eines Strichs oder eines

Farbkleckses, zu starker örtlicher Korrelation führen. Da Stimuli mit einem flachen Fourierspektrum ähnlich eines weißen Rauschens fast nur per Zufallsgenerator generierbar sind, ist nachvollziehbar, dass es für Künstler schwierig ist, Werke mit solch einem flachen Fourierspektrum zu erzeugen. Dies könnte durch weitere Experimente überprüft werden, in denen zum Beispiel gezielt versucht wird, Gesichter mit einem Anstieg zu malen, der von -2 abweicht. Doch warum erschaffen Künstler nicht Werke mit einem Anstieg von -3 oder -1,5? Grundsätzlich ist es für Künstler möglich, ästhetische Bilder zu erzeugen, welche keine  $1/f^2$ -Eigenschaften aufweisen (Redies et al. 2007). Zum Beispiel fallen die Fourierspektren von Bildern von geschriebenem Text nicht linear ab, sondern weisen zwei Abschnitte mit unterschiedlichen Steigungen auf, die beide von dem Wert -2 abweichen (T. Melmer et al.). Dennoch besitzen ästhetische Bilder im Durchschnitt  $1/f^2$ -Eigenschaften. Nach Graham und Redies liegt dies an der Anpassung der Werke durch den Künstler an das eigene visuelle System, um damit die Perzeptibilität zu gewährleisten bzw. zu erhöhen. Redies (2007) assoziiert  $1/f^2$ -Eigenschaften mit ästhetischem Empfinden, wobei diese Eigenschaften eher ein Beiprodukt der ästhetischen Produktion darstellen und nicht hinreichend für ästhetische Wahrnehmung sind. Redies und Graham interpretieren die  $1/f^2$ -Eigenschaften als Zeichen für universelle Gemeinsamkeiten bei ästhetischen Bildern. Neurobiologische Erkenntnisse der Neuroästhetik werden jedoch immer nur Teilaspekte der ästhetischen Wahrnehmung erklären können. Neben universellen Eigenschaften ästhetischer Bilder, welche grundlegende Funktionen des visuellen Systems ausnutzen, müssen auch andere Aspekte wie zum Beispiel soziokulturellen Faktoren beachtet werden (Goodman 1976, Danto 1981).

Das bessere Erlernen von Gesichtern mit  $1/f^2$ -Eigenschaften deutet darauf hin, dass außer natürlichen Szenen auch andere Stimuli mit solchen Eigenschaften effizient codiert werden können (Párraga et al. 2000). Durch die Änderung des Anstiegs kann die Wahrnehmung der Gesichter erleichtert werden. Wir nutzen Erkenntnisse der Neuroästhetik über die Manipulation der  $1/f^p$ -Eigenschaften von Gesichtern und konnten dadurch neue Erkenntnisse für das Erlernen unbekannter Gesichter gewinnen.

## 5 Schlussfolgerungen

In dieser Arbeit konnten Erkenntnisse der Neuroästhetik und elementare Funktionsweisen des visuellen Systems auf die Gesichtswahrnehmung übertragen werden.

Ästhetische gemalte Porträts zeigen  $1/f^2$ -Eigenschaften im Fourierspektrum, die denen natürlicher Szenen ähnlich sind. Sie unterscheiden sich in dieser Hinsicht stark von Gesichtsfotografien. Dies ist ein Hinweis darauf, dass Künstler ihre Abbildungen an Kodierungsmechanismen des visuellen Systems anpassen und nicht die Eigenschaften der Objekte abbilden, welche diese natürlicherweise besitzen. Da  $1/f^2$ -Eigenschaften nicht hinreichend für ästhetische Wahrnehmung sind, sollten weitere Studien auch andere statistische Eigenschaften ästhetischer Bilder untersuchen.

Wenn  $1/f^2$ -Eigenschaften auf Gesichterfotos übertragen werden, können diese Gesichter leichter gelernt werden. Gesichter mit steilerem Anstieg des Fourierspektrums wurden schlechter gelernt. Eine Veränderung des relativen Anteils von groben zu feinen Strukturen zu Gunsten der feinen Strukturen vereinfacht das Erlernen und Erkennen unbekannter Gesichter. Damit spielen feine Strukturen eine besonders wichtige Rolle bei der Erstellung neuer Repräsentationen von Gesichtern. Mit dieser Arbeit widerspreche ich bisherigen Studien, die den mittleren Frequenzbereich als besonders wichtig erachten. Unsere Ergebnisse beziehen sich zunächst nur auf bildliches Lernen, sodass in weiteren Studien der Zusammenhang von  $1/f^p$ -Eigenschaften und Gesichts- und Objektwahrnehmung systematisch erforscht werden könnte.

Fazit unserer Studien und Ergebnisse ist zum einen, dass verstärkte grobe und abgeschwächte feine Strukturen das Erlernen von Gesichtern erschweren. Zum anderen können  $1/f^2$  Eigenschaften, wie sie bei natürlichen Szenen auftreten, das Erlernen von Gesichtern erleichtern. Damit konnte gezeigt werden, dass eine Veränderung vom Anstieg des Fourierspektrums von Gesichtern, ähnlich wie sie durch Künstler durchgeführt wird, die Gesichterwahrnehmung erleichtert.



## 6 Literatur- und Quellenverzeichnis

- Adorno TW. 1970. *Ästhetische Theorie*. Frankfurt: Suhrkamp.
- Aks D, Sprott JC. 1996. Quantifying aesthetic preference for chaotic patterns. *Empirical Studies of the Arts*, 14 (1):1-16.
- Attwell D, Laughlin SB. 2001. An energy budget for signaling in the grey matter of the brain. *Journal of Cerebral Blood Flow and Metabolism*, 21 (10):1133-1145.
- Barlow HB. 1972. Single units and sensation: a neuron doctrine for perceptual psychology? *Perception*, 1 (4):371-394.
- Barlow HB, Olshausen BA. 2004. Convergent evidence for the visual analysis of optic flow through anisotropic attenuation of high spatial frequencies. *Journal of Vision*, 4 (6):415-426.
- Bear MF, Connors BW, Paradiso MA. 2007. *Neuroscience: exploring the brain*. Philadelphia: Lippincott Williams & Wilkins.
- Bentin S, Deouell LY. 2000. Structural encoding and identification in face processing: ERP evidence for separate processes. *Cognitive Neuropsychology*, 17 (1):35-54.
- Bentin S, Taylor MJ, Rousselet GA, Itier RJ, Caldara R, Schyns PG, Jacques C, Rossion B. 2007. Controlling interstimulus perceptual variance does not abolish N170 face sensitivity. *Nature Neuroscience*, 10 (7):801-802.
- Bodamer J. 1947. Die Prosop-Agnosie; die Agnosie des Physiognomieerkennens. *Archiv für Psychiatrie und Nervenkrankheiten, vereinigt mit Zeitschrift für die gesamte Neurologie und Psychiatrie*, 118 (1-2):6-53.
- Boeschoten MA, Kemner C, Kenemans JL, Engeland H. 2005. The relationship between local and global processing and the processing of high and low spatial frequencies studied by event-related potentials and source modeling. *Brain Research Cognitive Brain Research*, 24 (2):228-236.
- Bosworth RG, Bartlett MS, Dobkins KR. 2006. Image statistics of American Sign Language: comparison with faces and natural scenes. *Journal of the Optical Society of America A: Optics, Image Science and Vision*, 23 (9):2085-2096.
- Bouvier SE, Engel SA. 2006. Behavioral deficits and cortical damage loci in cerebral achromatopsia. *Cerebral Cortex*, 16 (2):183-191.

- Brady N, Field DJ. 2000. Local contrast in natural images: normalisation and coding efficiency. *Perception*, 29 (9):1041-1055.
- Bruce V, Young A. 1986. Understanding Face Recognition. *British Journal of Psychology*, 77:305-327.
- Burke E. 1757. A philosophical enquiry into the origin of our ideas of the sublime and beautiful. London: Dodsley.
- Burton AM, Bruce V, Johnston RA. 1990. Understanding face recognition with an interactive activation model. *British Journal of Psychology*, 81:361-380.
- Burton AM, Bruce V, Hancock PJB. 1999. From pixels to people: A model of familiar face recognition. *Cognitive Science*, 23 (1):1-31.
- Burton GJ, Moorhead IR. 1987. Color and spatial structure in natural scenes. *Applied Optics*, 26 (1):157-70.
- Busigny T, Rossion B. 2010. Acquired prosopagnosia abolishes the face inversion effect. *Cortex*, 46 (8):965-981.
- Busigny T, Graf M, Mayer En, Rossion B. 2010. Acquired prosopagnosia as a face-specific disorder: Ruling out the general visual similarity account. *Neuropsychologia*, 48 (7):2051-2067.
- Caharel S, d'Arripe O, Ramon M, Jacques C, Rossion B. 2009. Early adaptation to repeated unfamiliar faces across viewpoint changes in the right hemisphere: evidence from the N170 ERP component. *Neuropsychologia*, 47 (3):639-643.
- Campbell FW, Robson JG. 1968. Application of Fourier Analysis to Visibility of Gratings. *Journal of Physiology*, 197 (3):551-66.
- Carandini M, Demb JB, Mante V, Tolhurst DJ, Dan Y, Olshausen BA, Gallant JL, Rust NC. 2005. Do we know what the early visual system does? *Journal of Neuroscience*, 25 (46):10577-10597.
- Cavanagh P. 2005. The artist as neuroscientist. *Nature*, 434 (7031):301-307.
- Caywood MS, Willmore B, Tolhurst DJ. 2004. Independent components of color natural scenes resemble V1 neurons in their spatial and color tuning. *Journal of Neurophysiology*, 91 (6):2859-2873.

- Collin CA, Liu CH, Troje NF, McMullen PA, Chaudhuri A. 2004. Face recognition is affected by similarity in spatial frequency range to a greater degree than within-category object recognition. *Journal of Experimental Psychology Human Perception and Performance*, 30 (5):975-987.
- Cooper TJ, Harvey M, Lavidor M, Schweinberger SR. 2007. Hemispheric asymmetries in image-specific and abstractive priming of famous faces: Evidence from reaction times and event-related brain potentials. *Neuropsychologia*, 45 (13):2910-2921.
- Costen NP, Parker DM, Craw I. 1994. Spatial content and spatial quantisation effects in face recognition. *Perception*, 23 (2):129-146.
- Costen NP, Parker DM, Craw I. 1996. Effects of high-pass and low-pass spatial filtering on face identification. *Perception & Psychophysics*, 58 (4):602-612.
- Danto AC. 1981. *The transfiguration of the commonplace: a philosophy of art*. Cambridge: Harvard University Press.
- David SV, Vinje WE, Gallant JL. 2004. Natural stimulus statistics alter the receptive field structure of v1 neurons. *Journal of Neuroscience*, 24 (31):6991-7006.
- De Haan EHF. 1999. A Familial Factor in the Development of Face Recognition Deficits. *Journal of Clinical and Experimental Neuropsychology*, 21 (3):312-315.
- De Valois RL, De Valois KK. 1980. Spatial vision. *Annual Review of Psychology*, 31:309-341.
- Dering B, Martin CD, Thierry G. 2009. Is the N170 peak of visual event-related brain potentials car-selective? *Neuroreport*, 20 (10):902-906.
- Doi H, Sawada R, Masataka N. 2007. The effects of eye and face inversion on the early stages of gaze direction perception-an ERP study. *Brain Research*, 1183:83-90.
- Eger E, Schyns PG, Kleinschmidt A. 2004. Scale invariant adaptation in fusiform face-responsive regions. *Neuroimage*, 22 (1):232-242.
- Eimer M. 2000. Event-related brain potentials distinguish processing stages involved in face perception and recognition. *Clinical Neurophysiology*, 111 (4):694-705.
- Eimer M. 2000. The face-specific N170 component reflects late stages in the structural encoding of faces. *Neuroreport*, 11 (10):2319-2324.
- Fernandez D, Wilkins AJ. 2008. Uncomfortable images in art and nature. *Perception*, 37 (7):1098-1113.

- Field DJ. 1987. Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A*, 4 (12):2379-2394.
- Fiorentini A, Maffei L, Sandini G. 1983. The role of high spatial frequencies in face perception. *Perception*, 12 (2):195-201.
- Flevaris AV, Robertson LC, Bentin S. 2008. Using spatial frequency scales for processing face features and face configuration: an ERP analysis. *Brain Research*, 1194:100-109.
- Frazor RA, Geisler WS. 2006. Local luminance and contrast in natural images. *Vision Research*, 46 (10):1585-1598.
- Geisler WS. 2008. Visual perception and the statistical properties of natural scenes. *Annual Review of Psychology*, 59:167-192.
- Georghiades AS, Belhumeur PN, Kriegman DJ. 2001. From few to many: Illumination cone models for face recognition under variable lighting and pose. *Transactions on Pattern Analysis and Machine Intelligence*, 23 (6):643-660.
- Ginsburg AP. 1978. Visual information processing based on spatial filters constrained by biological data [Dissertation]. Ohio: Air force aerospace medical research lab wright-patterson.
- Gobbini MI, Haxby JV. 2007. Neural systems for recognition of familiar faces. *Neuropsychologia*, 45 (1):32-41.
- Goffaux V, Gauthier I, Rossion B. 2003. Spatial scale contribution to early visual differences between face and object processing. *Brain research. Cognitive brain research*, 16 (3):416-424.
- Goffaux V, Jemel B, Jacques C, Rossion B, Schyns PG. 2003. ERP evidence for task modulations on face perceptual processing at different spatial scales. *Cognitive Science*, 27 (2):313-325.
- Goffaux V, Hault B, Michel C, Vuong QC, Rossion B. 2005. The respective role of low and high spatial frequencies in supporting configural and featural processing of faces. *Perception*, 34 (1):77-86.
- Goguen JA, Myin E. 2004. *Art and the Brain III*. Thorverton: Imprint Academic.
- Goodman N. 1976. *Languages of art: An approach to a theory of symbols*. Indianapolis: Hackett.

- Grabowska A, Nowicka A. 1996. Visual-spatial-frequency model of cerebral asymmetry: A critical survey of behavioral and electrophysiological studies. *Psychological Bulletin*, 120 (3):434-449.
- Graham D. 2008. Global nonlinear compression of natural luminances in painted art. In: Stork DG, Coddington J. *Computer Image Analysis in the Study of Art*. SPIE, San Jose: 6810 (1):68100.
- Graham DJ, Field DJ. 2007. Statistical regularities of art images and natural scenes: Spectra, sparseness and nonlinearities. *Spatial Vision*, 21 (1-2):149-164.
- Graham DJ, Field DJ. 2008. Variations in intensity statistics for representational and abstract art, and for art from the Eastern and Western hemispheres. *Perception*, 37 (9):1341-1352.
- Graham DJ, Field DJ. 2009. Natural Images: Coding Efficiency. In: Larry RS. *Encyclopedia of Neuroscience*. Sechste Auflage. Amsterdam: Elsevier, 19-27.
- Graham DJ, Redies C. 2010. Statistical regularities in art: Relations with visual coding and perception. *Vision Research*, 50 (16):1503-1509.
- Graham NV. 2011. Beyond multiple pattern analyzers modeled as linear filters (as classical V1 simple cells): Useful additions of the last 25 years. *Vision Research*, 51 (13):1397-1430.
- Gross CG. 2002. Genealogy of the "grandmother cell". *Neuroscientist*, 8 (5):512-518.
- Grüter T, Grüter M, Carbon CC. 2008. Neural and genetic foundations of face recognition and prosopagnosia. *Journal of Neuropsychology*, 2:79-97.
- Guhr S. 2002. Visuell evozierte Flussgeschwindigkeitsänderungen in der A. cerebri posterior bei Normalprobanden und Patienten mit Leitungsverzögerungen im Sehbahnbereich: eine Untersuchung mit der funktionellen transkraniellen Dopplersonographie [Dissertation]. Berlin: Humboldt Universität.
- Hsagerhall CM, Purcell T, Taylor R. 2004. Fractal dimension of landscape silhouette outlines as a predictor of landscape preference. *Journal of Environmental Psychology*, 24 (2):247-255.
- Halit H, de Haan M, Johnson MH. 2000. Modulation of event-related potentials by prototypical and atypical faces. *Neuroreport*, 11 (9):1871-1875.
- Halit H, de Haan M, Schyns PG, Johnson MH. 2006. Is high-spatial frequency information used in the early stages of face detection? *Brain Research*, 1117 (1):154-161.

- Harmon LD, Julesz B. 1973. Masking in visual recognition: effects of two-dimensional filtered noise. *Science*, 180 (91):1194-1197.
- Haxby JV, Hoffman EA, Gobbini MI. 2000. The distributed human neural system for face perception. *Trends in Cognitive Sciences*, 4 (6):223-233.
- Hegel GWF. 1833. *Vorlesungen über die Ästhetik*. Frankfurt: Suhrkamp.
- Herrmann MJ, Ehlis AC, Ellgring H, Fallgatter AJ. 2005. Early stages (P100) of face perception in humans as measured with event-related potentials (ERPs). *Journal of Neural Transmission*, 112 (8):1073-1081.
- Hoyer PO, Hyvärinen A. 2000. Independent component analysis applied to feature extraction from colour and stereo images. *Network*, 11 (3):191-210.
- Hoyer PO, Hyvärinen A. 2002. A multi-layer sparse coding network learns contour coding from natural images. *Vision Research*, 42 (12):1593-1605.
- Hoyer PO, Hyvarinen A. 2002. Sparse coding of natural contours. *Neurocomputing*, 44:459-466.
- Hsiao FJ, Hsieh JC, Lin YY, Chang Y. 2005. The effects of face spatial frequencies on cortical processing revealed by magnetoencephalography. *Neuroscience Letters*, 380 (1-2):54-59.
- Hubel DH, Wiesel TN. 1962. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, 160:106-154.
- Huynh H, Feldt L. 1976. Estimation of the Box correction for degrees of freedom from sample data in randomized block and split-plot designs. *Journal of Educational Statistics*, 1:69-82.
- Hume D. 1826. Of the standard of taste. In: Green, Grose. *The philosophical works of David Hume*. London: Longman Green.
- Itier RJ, Taylor MJ. 2002. Inversion and contrast polarity reversal affect both encoding and recognition processes of unfamiliar faces: A repetition study using ERPs. *Neuroimage*, 15 (2):353-372.
- Juricevic I, Land L, Wilkins A, Webster MA. 2010. Visual discomfort and natural image statistics. *Perception*, 39 (7):884-899.
- Kandinsky W. 1912. *Über das geistige in der Kunst, insbesondere in der Malerei*. München: Piper.

- Kant I. 1992. Kritik der Urteilskraft. In: Weischedel W. Werkausgabe in zwölf Bänden. Frankfurt: Suhrkamp.
- Kanwisher N, Yovel G. 2006. The fusiform face area: a cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society-Biological Sciences*, 361 (1476):2109-2128.
- Kanwisher N, McDermott J, Chun MM. 1997. The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, 17 (11):4302-4311.
- Kaufmann JM, Schweinberger SR. 2008. Distortions in the brain? ERP effects of caricaturing familiar and unfamiliar faces. *Brain Research*, 1228:177-188.
- Kaufmann JM, Schweinberger SR. 2012. The faces you remember: Caricaturing shape facilitates brain processes reflecting the acquisition of new face representations. *Biological Psychology*, 89 (1):21-33.
- Kaufmann JM, Schweinberger SR, Burton AM. 2009. N250 ERP correlates of the acquisition of face representations across different images. *Journal of Cognitive Neuroscience*, 21 (4):625-641.
- Keil MS. 2008. Does face image statistics predict a preferred spatial frequency for human face processing? *Proceeding Biological Sciences*, 275 (1647):2095-2100.
- Kennerknecht I, Grueter T, Welling B, Wentzek S, Horst J, Edwards S, Grueter M. 2006. First report of prevalence of non-syndromic hereditary prosopagnosia (HPA). *American Journal of Medical Genetics Part A*, 140 (15):1617-1622.
- Koch M, Denzler J, Redies C. 2010.  $1/f^2$  Characteristics and isotropy in the fourier power spectra of visual art, cartoons, comics, mangas, and different categories of photographs. *PLoS One*, 5 (8):e12268.
- Koppe F. 2004. *Grundbegriffe der Ästhetik*. Paderborn: Mentis.
- Lee AB, Mumford D, Huang J. 2001. Occlusion Models for Natural Images: A Statistical Study of a Scale-Invariant Dead Leaves Model. *International Journal of Computer Vision*, 41 (1):35-59.
- Lee TW, Wachtler T, Sejnowski TJ. 2002. Color opponency is an efficient representation of spectral properties in natural scenes. *Vision Research*, 42 (17):2095-2103.
- Lennie P. 2003. The cost of cortical computation. *Current Biology*, 13 (6):493-497.
- Livingstone MS. 2002. *Vision and art: The biology of seeing*. New York: Harry N. Abrams.

- Loftus GR, Harley EM. 2004. How different spatial-frequency components contribute to visual information acquisition. *Journal of Experimental Psychology. Human Perception and Performance*, 30 (1):104-118.
- Maffei L, Fiorenti A. 1973. Visual - Cortex as a Spatial Frequency Analyzer. *Vision Research*, 13 (7):1255-1267.
- Martinez AM, Benavente R. 1998. The AR Face Database CVC Technical Report #24
- Melmer T, Koch M, Denzler J, Redies C. A comparison of statistical regularities in typographic text, handwriting, calligraphy, artistic ornaments and art. [Unveröffentliche Daten]
- Morrison DJ, Schyns PG. 2001. Usage of spatial scales for the categorization of faces, objects, and scenes. *Psychonomic Bulletin & Review*, 8 (3):454-469.
- Nakashima T, Kaneko K, Goto Y, Abe T, Mitsudo T, Ogata K, Makinouchi A, Tobimatsu S. 2008. Early ERP components differentially extract facial features: evidence for spatial frequency-and-contrast detectors. *Neuroscience Research*, 62 (4):225-235.
- Näsänen R. 1999. Spatial frequency bandwidth used in the recognition of facial images. *Vision Research*, 39 (23):3824-3833.
- O'Hare L, Hibbard PB. 2011. Spatial frequency and visual discomfort. *Vision Research*, 51 (15):1767-1777.
- Olshausen B, Field D. 1996. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381 (6583):607-609.
- Olshausen BA, Field DJ. 1996. Natural image statistics and efficient coding. *Network-Computation in Neural Systems*, 7 (2):333-339.
- Olshausen BA, Field DJ. 2000. Vision and the coding of natural images. *American Scientist*, 88 (3):238-245.
- Olshausen BA, Field DJ. 2004. Sparse coding of sensory inputs. *Current Opinion in Neurobiology*, 14 (4):481-487.
- Orban de Xivry J-J, Ramon M, Lefèvre P, Rossion B. 2008. Reduced fixation on the upper area of personally familiar faces following acquired prosopagnosia. *Journal of Neuropsychology*, 2 (1):245-268.
- Paller KA, Gonsalves B, Grabowecky M, Bozic VS, Yamada S. 2000. Electrophysiological correlates of recollecting faces of known and unknown individuals. *Neuroimage*, 11 (2):98-110.



- Palmer SE. 1999. Vision science: photons to phenomenology. Cambridge: MIT Press.
- Parker DM, Costen NP. 1999. One extreme or the other or perhaps the golden mean? Issues of spatial resolution in face processing. *Current Psychology*, 18(1): 118-127.
- Párraga CA, Troscianko T, Tolhurst DJ. 2000. The human visual system is optimised for processing the spatial information in natural visual images. *Current Biology*, 10 (1):35-38.
- Paul G. 1988. Philosophical theories of beauty and scientific research on the brain. In: Rentschler I, Herzberger B, Epstein D. *Beauty and the brain: biological aspects of aesthetics*. Basel: Birkhaeuser Verlag, 15-27.
- Paul G. 1998. *Abiturwissen philosophische Ästhetik: Grundbegriffe und Grundpositionen: das ästhetische in Kunst und Natur*. Stuttgart: E. Klett.
- Prasad S, Galetta SL. 2011. Anatomy and physiology of the afferent visual system. *Handbook of Clinical Neurology*, 102:3-19.
- Puce A, Allison T, Bentin S, Gore JC, McCarthy G. 1998. Temporal Cortex Activation in Humans Viewing Eye and Mouth Movements. *The Journal of Neuroscience*, 18 (6):2188-2199.
- Redies C. 2007. A universal model of esthetic perception based on the sensory coding of natural stimuli. *Spatial Vision*, 21 (1-2):97-117.
- Redies C, Hasenstein J, Denzler J. 2007. Fractal-like image statistics in visual art: similarity to natural scenes. *Spatial Vision*, 21 (1-2):137-148.
- Redies C, Hänisch J, Blickhan M, Denzler J. 2007. Artists portray human faces with the Fourier statistics of complex natural scenes. *Network*, 18 (3):235-248.
- Rentschler I, Caelli MML. 1988. Focusing in on art. In: Rentschler I, Herzberger B, Epstein D. *Beauty and the brain: biological aspects of aesthetics*. Basel: Birkhaeuser Verlag, 181-216.
- Rhodes G, Byatt G, Tremewan T, Kennedy A. 1997. Facial distinctiveness and the power of caricatures. *Perception*, 26 (2):207-223.
- Rhodes G, Byatt G, Michie PT, Puce A. 2004. Is the fusiform face area specialized for faces, individuation, or expert individuation? *Journal of Cognitive Neuroscience*, 16 (2):189-203.
- Riesenhuber M, Poggio T. 2000. Models of object recognition. *Nature Neuroscience*, 3 :1199-1204.

- Rossion B, Jacques C. 2008. Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170. *Neuroimage*, 39 (4):1959-1979.
- Rossion B, Caldara R, Seghier M, Schuller AM, Lazeyras F, Mayer E. 2003. A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. *Brain*, 126 (11):2381-2395.
- Rousselet GA, Husk JS, Bennett PJ, Sekuler AB. 2008. Time course and robustness of ERP object and face differences. *Journal of Vision*, 8 (12):1-18.
- Ruderman DL. 1997. Origins of scaling in natural images. *Vision Research*, 37 (23):3385-3398.
- Ruderman DL, Bialek W. 1994. Statistics of natural images: Scaling in the woods. *Physical Review Letters*, 73 (6):814-817.
- Ruiz-Soler M, Beltran FS. 2006. Face perception: an integrative review of the role of spatial frequencies. *Psychological Research*, 70 (4):273-292.
- Schelling FWJ. 1907. *Philosophie der Kunst*. In: Schröter M. Schellings Werke. München: C. H. Beck'sche Verlagsbuchhandlung.
- Schweinberger S, Pickering E, Burton A, Kaufmann J. 2002a. Human brain potential correlates of repetition priming in face and name recognition. *Neuropsychologia*, 40 (12):2057-2073.
- Schweinberger S, Pickering E, Jentsch I, Burton A, Kaufmann J. 2002b. Event-related brain potential evidence for a response of inferior temporal cortex to familiar face repetitions. *Brain Research. Cognitive Brain Research*, 14 (3):398-409.
- Schweinberger SR. 2011. Neurophysiological correlates of face recognition. In: Calder A, Rhodes G, Johnson M, Haxby J. *Oxford Handbook of Face Perception*. Oxford: Oxford University Press.
- Schweinberger SR, Burton AM. 2003. Covert Recognition and the Neural System for Face Processing. *Cortex*, 39 (1):9-30.
- Schyns PG, Oliva A. 1999. Dr. Angry and Mr. Smile: when categorization flexibly modifies the perception of faces in rapid visual presentations. *Cognition*, 69 (3):243-265.
- Schyns PG, Bonnar L, Gosselin F. 2002. Show me the features! Understanding recognition from the use of visual information. *Psychological Science*, 13 (5):402-409.

- Scott LS, Tanaka JW, Sheinberg DL, Curran T. 2006. A reevaluation of the electrophysiological correlates of expert object processing. *Journal of Cognitive Neuroscience*, 18 (9):1453-1465.
- Sergent J. 1982. The cerebral balance of power: confrontation or cooperation? *Journal of Experimental Psychology. Human Perception and Performance*, 8 (2):253-272.
- Sergent J. 1986. Microgenesis in face perception. In H. D. Ellis, M. A. Jeeves, F. Newcombe, & A. Young, *Aspects of face processing*. Dordrecht: Nijhoff.
- Sergent J. 1987. Failures to confirm the spatial-frequency hypothesis: fatal blow or healthy complication. *Canadian Journal of Psychology*, 41 (4):412-428.
- Simoncelli EP, Olshausen BA. 2001. Natural image statistics and neural representation. *Annual Review of Neurosciene*, 24:1193-1216.
- Spehar B, Clifford CWG, Newell BR, Taylor RP. 2003. Universal aesthetic of fractals. *Computers & Graphics*, 27 (5):813-820.
- Sprott JC. 1993. Automatic-generation of strange attractors. *Computers & Graphics*, 17 (3):325-332.
- Stollhoff R, Jost J, Elze T, Kennerknecht I. 2010. The early time course of compensatory face processing in congenital prosopagnosia. *PLoS ONE*, 5 (7):e11482.
- Tailor DR, Finkel LH, Buchsbaum G. 2000. Color-opponent receptive fields derived from independent component analysis of natural images. *Vision Research*, 40 (19):2671-2676.
- Tanaka J, Curran T, Porterfield A, Collins D. 2006. Activation of preexisting and acquired face representations: the N250 event-related potential as an index of face familiarity. *Journal of Cognitvie Neuroscience*, 18 (9):1488-1497.
- Taylor RP. 2002. Order in Pollack's chaos - Computer analysis is helping to explain the appeal of Jackson Pollock's paintings. *Scientific American*, 287 (6):116-121.
- Taylor RP, Micolich AP, Jonas D. 1999. Fractal analysis of Pollock's drip paintings. *Nature*, 399 (6735):422-422.
- Thierry G, Martin CD, Downing P, Pegna AJ. 2007. Controlling for interstimulus perceptual variance abolishes N170 face selectivity. *Nature Neuroscience*, 10 (4):505-511.
- Tolhurst DJ, Tadmor Y, Chao T. 1992. Amplitude spectra of natural images. *Ophthalmic & Physiological Optics*, 12 (2):229-232.

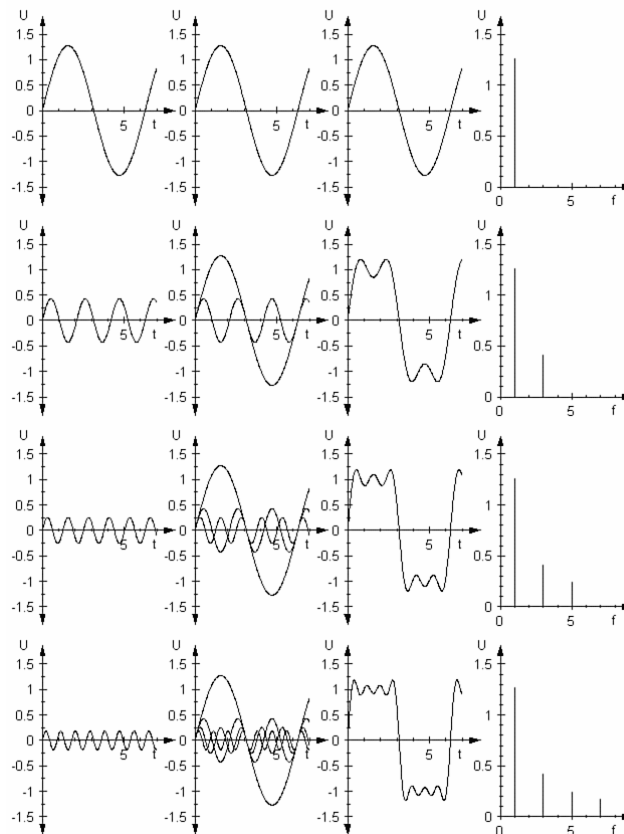
- Tolhurst DJ, Smyth D, Thompson ID. 2009. The sparseness of neuronal responses in ferret primary visual cortex. *Journal of Neuroscience*, 29 (8):2355-2370.
- Torralba A, Oliva A. 2003. Statistics of natural image categories. *Network-Computation in Neural Systems*, 14 (3):391-412.
- Trenner MU, Schweinberger SR, Jentzsch I, Sommer W. 2004. Face repetition effects in direct and indirect tasks: an event-related brain potentials study. *Brain Research. Cognitive Brain Research*, 21 (3):388-400.
- Valentine T, Brennen T, Brédart S. 1996. *The cognitive psychology of proper names: on the importance of being Ernest*. London: Routledge.
- van Hateren JH, Ruderman DL. 1998. Independent component analysis of natural image sequences yields spatio-temporal filters similar to simple cells in primary visual cortex. *Proceedings Biological Sciences*, 265 (1412):2315-2320.
- Vinje WE, Gallant JL. 2000. Sparse coding and decorrelation in primary visual cortex during natural vision. *Science*, 287 (5456):1273-1276.
- Vinje WE, Gallant JL. 2002. Natural stimulation of the nonclassical receptive field increases information transmission efficiency in V1. *Journal of Neuroscience*, 22 (7):2904-2915.
- Vuilleumier P, Armony JL, Driver J, Dolan RJ. 2003. Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nature Neuroscience*, 6 (6):624-631.
- Wachtler T, Albright TD, Sejnowski TJ. 2001. Nonlocal interactions in color perception: nonlinear processing of chromatic signals from remote inducers. *Vision Research*, 41 (12):1535-1546.
- Watier NN, Collin CA. 2009. Effects of familiarity on spatial frequency thresholds for face matching. *Perception*, 38 (10):1497-1507.
- Webster MA, Mollon JD. 1997. Adaptation and the color statistics of natural images. *Vision Research*, 37 (23):3283-3298.
- Werner JS, Ratliff F. 1999. Some origins of the lightness and darkness of colors in the visual arts and in the brain. *Techne*, 9-10:61-73.
- Westheimer G. 2001. The Fourier theory of vision. *Perception*, 30 (5):531-541.
- Willmore B, Tolhurst DJ. 2001. Characterizing the sparseness of neural codes. *Network*, 12 (3):255-270.

- Willmore BDB, Mazer JA, Gallant JL. 2011. Sparse coding in striate and extrastriate visual cortex. *Journal of Neurophysiology*, 105 (6):2907-2919.
- Zeki S. 1999. Art and the brain. *Journal of Consciousness Studies* 6-7, 76-96.

## 7 Anhang

### 7.1 Fourieranalyse

Mit der Fourieranalyse kann ein Signal in einzelne trigonometrische Komponenten verschiedener Frequenzen zerlegt werden. Umgekehrt kann zum Beispiel ein Rechtecksignal auch aus verschiedenen trigonometrischen Komponenten gebildet werden (siehe Abb. 15).

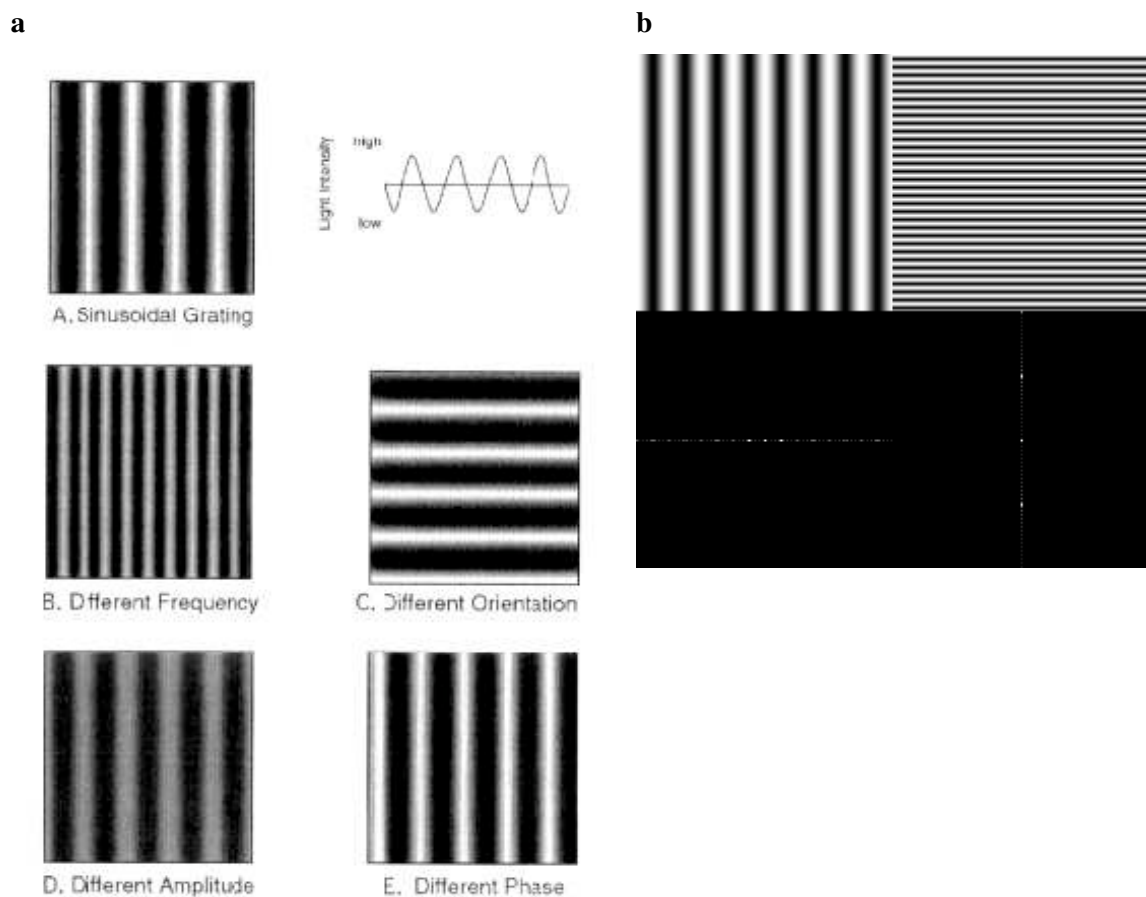


**Abb. 15** Fouriersynthese eines Rechtecksignals. In der rechten Spalte ist das Amplitudenspektrum abgebildet. In diesem Fall nimmt die Amplitude ab bei steigender Frequenz. Modifiziert nach factbook.org. GFDL-Lizenz

Die einzelnen Frequenzkomponenten können sich in Amplitude und Phase unterscheiden. Dabei zeigt die Amplitude, wie groß der Anteil einer bestimmten Frequenz ist und die Phase bestimmt, wo sich diese Frequenz befindet (siehe Abb. 16a).

Wird die Fourieranalyse für ein zweidimensionales Signal wie zum Beispiel ein Bild angewendet, erfolgt die Fourieranalyse zum einen in horizontaler und zum anderen in

vertikaler Richtung und man erhält ein zweidimensionales Fourierspektrum (siehe Abb. 16b). Bei komplexen Stimuli wie zum Beispiel Gesichtsfotografien erhält man ein komplexes Amplitudenspektrum (siehe Abb. 12). Wird dieses Fourierspektrum quadriert erhält man das Powerspektrum. Um nun den Anteil der verschiedenen Frequenzkomponenten zu erhalten, kann man die Amplitude radial mitteln und als Funktion der Frequenz logarithmisch darstellen und man erhält ein eindimensionales Amplitudenspektrum (siehe Abb. 8 und Abb. 13), mit dem der Anstieg des Fourierspektrums ermittelt wird. Dabei wird normalerweise der äußerste Rand des Frequenzspektrums nicht berücksichtigt, da dieser durch Artefakte beeinflusst werden kann. Für natürliche Szenen wurde ein Anstieg von  $-2$  ( $1/f^2$ -Eigenschaften) im Fourierspektrum gefunden (Burton und Moorhead 1987, Field 1987).



**Abb. 16 (a) Sinusmuster unterschiedlicher Frequenz, Amplitude, Phase und Orientierung. Aus Palmer (1999). (b) Sinusmuster und ihr Amplitudenspektrum. Tiefe Frequenzen finden sich im zentralen Bereich des Amplituden Spektrums (linker Teil des Bildes). Hohe Frequenzen zeigen sich im peripheren Bereich des Amplitudenspektrums (rechter Teil des Bildes). Aus J. M. Bayer [www.cs.unm.edu](http://www.cs.unm.edu)**

## 7.2 EEG- und fMRT-Korrelate der Gesichterwahrnehmung

Mittels evozierter Potentiale im EEG (siehe Abb. 17) können zeitliche Abläufe der visuellen Wahrnehmung von Gesichtern beschrieben werden (Schweinberger 2011).

Die okzipitale *P100* wird durch einfache Stimulusmerkmale beeinflusst wie zum Beispiel Kontrast und Leuchtdichte (Itier und Taylor 2002). Nach Herrmann et al. (2005) und Nakashima et al. (2008) beginnt Gesichterwahrnehmung bereits bei 100ms und beeinflusst die Amplitude der *P100*. So zeigten Nakashima et al. (2008), dass die *P100* bei der Gesichtererkennung die Kodierung von groben Strukturen (LSF) repräsentiert.

Die temporale *N170* ist mit früher struktureller Kodierung von Gesichtern assoziiert und gilt als unabhängig von Familiarität (Bentin und Deouell 2000, Eimer 2000) und Wiederholung (Eimer 2000, Cooper et al. 2007). Dennoch gibt es Hinweise, dass unter speziellen Bedingungen Wiederholung die *N170* beeinflussen kann (Caharel et al. 2009). Nach Nakashima et al. (2008) spiegelt die *N170* die Kodierung feiner Strukturen, welche durch HSF repräsentiert werden, wider. Damit zeigt sich in der *N170* die Erkennung eines Objektes als Gesicht, aber noch nicht die Erkennung ob dieses Gesicht bekannt oder unbekannt ist.

Die okzipitotemporale *N250* wird durch Wiederholung von Gesichtern erhöht und ist für bekannte Gesichter größer als für unbekannte (Kaufmann et al. 2009, Schweinberger et al. 2002b, Schweinberger et al. 2002a, Tanaka et al. 2006). Dieser Effekt vergrößert sich, wenn gleiche Bilder wiederholt werden, sodass die *N250* auch an der piktorialen Kodierung von Gesichtern beteiligt ist (Schweinberger et al. 2002b). Es zeigte sich, dass diese elektrophysiologische Komponente in der „fusiform face area“ entsteht, welche zuvor in fMRT Studien beschrieben wurde (siehe weiter unten).

Die *central positivity* zeigt größere old/new Effekte für bekannte als unbekannte Gesichter und gilt als Korrelat des Zugriffs auf semantische Informationen der Identität (Bentin und Deouell 2000, Eimer 2000, Paller et al. 2000).

Mittels fMRT können vor allem örtliche Abläufe der Wahrnehmung von Gesichtern beschrieben werden.

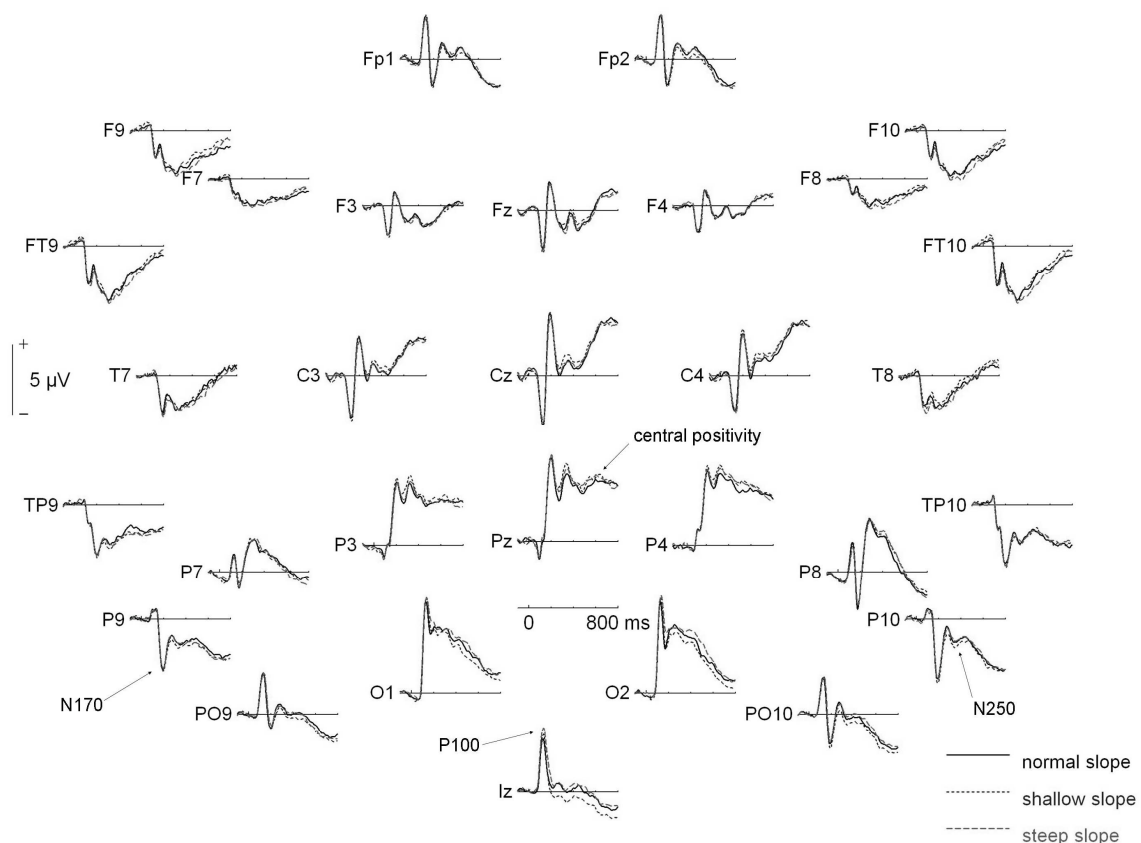
Die „*fusiform face area*“ ist im Temporallappen im Gyrus fusiformis lokalisiert und wird bei der Wahrnehmung von Gesichtern aktiviert (Kanwisher et al. 1997). Im Rahmen der



Spezifitätsdebatte wurde diskutiert, ob dieses Gehirnareal gesichterspezifisch ist oder auch zur Individuation von ähnlichen Objekten dient (Kanwisher und Yovel 2006). Letzteres wurde zum Beispiel durch Experimente geprüft, in denen Schmetterlinge passiv gesehen und anschließend gelernt und unterschieden werden mussten (Rhodes et al. 2004).

Die „*occipital face area*“ ist im Okzipitallappen lokalisiert und wird ebenfalls bei der Gesichterwahrnehmung aktiviert. Bei Schädigung dieses Bereiches kann es zu einer erworbenen Prosopagnosie kommen (Rossion et al. 2003, Bouvier und Engel 2006). Die „*occipitale face area*“ und die „*fusiform face area*“ verarbeiten unveränderliche bildliche Eigenschaften von Gesichtern (Grüter et al. 2008).

Der *Sulcus temporalis superior* ist an der Verarbeitung dynamischer Eigenschaften von Gesichtern wie zum Beispiel Blickwinkel und Gesichtsausdruck beteiligt (Haxby et al. 2000, Puce et al. 1998).



**Abb. 17 EEG-Korrelate aus unserem Lernexperiment. Die P100 findet sich okzipital. Die N170 und N250 zeigen sich temporal und die central positivity kann zentroparietal abgeleitet werden.**

## **7.3 Lebenslauf**

Marko Blickhan

geboren am 29.05.1985 in Boston, USA

### **Ausbildung**

- 2005-2012: Studium der Humanmedizin an der Friedrich-Schiller Universität Jena  
1. Abschnitt der Ärztlichen Prüfung: gut (2,5)  
2. Abschnitt der Ärztlichen Prüfung: sehr gut (1,5)
- 1997-2004: Carl-Zeiss Gymnasium, Spezialschule für Mathematik und Naturwissenschaften  
Abitur: Durchschnittsnote 1,2  
Seminarfacharbeit: Einfluss verschiedener Parameter auf die Trittfrequenz beim Radfahren
- 2001-2002 Austauschjahr an der South Eugene High School, USA

### **Stipendium**

- 2007-2011 Stipendium der Studienstiftung des deutschen Volkes

### **Weitere Qualifikationen**

- 2006-2012 wissenschaftliche Hilfskraft und Doktorand bei Prof. Christoph Redies in der AG Neuroästhetik
- 2008 fakultatives Seminar Neuroästhetik (Prof. Christoph Redies, Anatomie)
- 2006 fakultatives Seminar Humangenetik (Prof. Baniahmad, Humangenetik)

### **Arbeitserfahrung**

- 2006-2010 Tutor für Histologie
- 2005-2007 Tutor für Anatomie

### **Praktika im Krankenhaus**

- 04-07/2011 Praktisches Jahr am Spital Thusis, Schweiz. Innere Medizin  
12/10-03/11 Praktisches Jahr am Universitätsklinikum Lleida, Spanien. Chirurgie  
08-11/2010 Praktisches Jahr am Universitätsklinikum Jena. Neurologie  
03/2010 Famulatur am Krankenhaus Rosenheim. Neurologie  
03/2009 Famulatur am Krankenhaus Cazis, Schweiz. Psychiatrie  
09/2008 Famulatur am Krankenhaus Ntra. Sra. Pilar Barcelona, Spanien. Innere  
Medizin  
03/2008 Famulatur am Krankenhaus Landeck, Österreich. Chirurgie  
03/2006 Pflegepraktikum in Tamsweg, Österreich  
08/2006 Pflegepraktikum in Valdivia, Chile

### **Wehrdienst**

- 2004-2005 In Bad Reichenhall

### **Fremdsprachen**

Englisch, Spanisch, Griechisch

## **7.4 Danksagung**

Mein Dank gilt zunächst besonders meinem Doktorvater Christoph Redies. Er hat mich sehr gut und engmaschig betreut und mich von Anfang an in die Erstellung von wissenschaftlichen Publikationen eingebunden. Ich bin ihm besonders dankbar, dass er mich mit anderen Wissenschaftsbereichen sprich der Informatik und der Psychologie in Verbindung gebracht hat.

Des Weiteren gilt mein Dank Jürgen Kaufmann mit dem ich die psychologischen Experimente geplant und durchgeführt habe. Trotz meiner geringen Vorkenntnisse zur Durchführung von EEG-Studien und statistischen Auswertung hat er sich immer wieder viel Zeit für mich genommen. Jürgen Kaufmann hat ebenfalls wesentlich zu der Erstellung der Veröffentlichung beigetragen.

Joachim Denzler hat mir bei der Erarbeitung der Manipulation der Fourierspektren geholfen. Michael Koch hat mir ebenfalls bei der Stimulusmanipulation geholfen.

Stefan Schweinberger hat uns bei der Planung und Auswertung der Gesichter-Lernexperimente unterstützt. Insbesondere bei den zwei Korrekturen nach dem Reviewprozess war seine Hilfe unerlässlich.

Bettina Kamchen hat mir bei der Durchführung der EEG-Experimente tatkräftig zur Seite gestanden.

Außerdem gilt mein Dank meinem Vater Reinhard Blickhan, der meine Neugier an wissenschaftlichem Arbeiten und Denken geweckt hat.

## **7.5 Eigenanteil an den Publikationen**

Für die Publikation “Artists portray human faces with the Fourier statistics of complex natural scenes. Redies C, Hänisch J, Blickhan M, Denzler J. *Network*. 18(3):235-48. 2007“ habe ich Porträts eingescannt, die Kalibrierung des Scanners durchgeführt und eine Datenbank mit Porträts erstellt. Des Weiteren schrieb ich Computer-Programme zum Messen des Augenabstandes, zum Ausschneiden von Detailbildern und zur weiteren Auswertung der Daten des Fourierspektrums.

Für die Publikation „ $1/f(p)$  Characteristics of the Fourier power spectrum affects ERP correlates of face learning and recognition. Blickhan M, Kaufmann JM, Denzler J, Schweinberger SR, Redies C. *Biological Psychology*. 88(2-3):204-14. 2011“ entwickelte ich die Manipulation der Fourierspektren der Gesichter mit Unterstützung von Prof. Redies und Prof. Denzler. Die Idee, ein Gesicht-Lernexperiment mit den manipulierten Fotos durchzuführen, stammt von mir. Zusammen mit Dr. Kaufmann führte ich die EEG-Experimente durch und erhielt Unterstützung von ihm bei der Auswertung der Daten. Das Manuskript für diese Publikation wurde von mir geschrieben und anschließend von den Koautoren korrigiert.

## **7.6 Ehrenwörtliche Erklärung**

Hiermit erkläre ich,  
dass mir die Promotionsordnung der Medizinischen Fakultät der Friedrich-Schiller-Universität bekannt ist,

ich die Dissertation selbst angefertigt habe und alle von mir benutzten Hilfsmittel, persönlichen Mitteilungen und Quellen in meiner Arbeit angegeben sind,

mich folgende Personen bei der Auswahl und Auswertung des Materials sowie bei der Herstellung des Manuskripts unterstützt haben:

mein Doktorvater Prof. Dr. Dr. Christoph Redies,

meine Koautoren Dr. Jürgen M. Kaufmann (Ph.D.), Prof. Dr. S. Schweinberger, Prof. Dr. J. Denzler,

meine Lektoren Denise Blickhan, Dr. Janine Zimmermann, Dr. Rul v. Stülpnagel, Alexandra Fuchs und Nicole Luzar,

die Hilfe eines Promotionsberaters nicht in Anspruch genommen wurde und dass Dritte weder unmittelbar noch mittelbar geldwerte Leistungen von mir für Arbeiten erhalten haben, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen,

dass ich die Dissertation noch nicht als Prüfungsarbeit für eine staatliche oder andere wissenschaftliche Prüfung eingereicht habe und

dass ich die gleiche, eine in wesentlichen Teilen ähnliche oder eine andere Abhandlung nicht bei einer anderen Hochschule als Dissertation eingereicht habe.

Jena, 23.01.2012

Marko Blickhan