

Heinrich H. Bülhoff, Hanspeter A. Mallot,  
Rolf D. Ulrich, Felix A. Wichmann (Eds.)

# **8th Tübingen Perception Conference**

TWK 2005, Tübingen, Germany, 25<sup>th</sup>–27<sup>th</sup> February 2005  
Proceedings

Knirsch Verlag Kirchentellinsfurt

Prof. Dr. Heinrich H. Bülthoff  
Cognitive and Computational Psychophysics Department  
Max-Planck-Institut für biologische Kybernetik  
Spemannstr. 38  
72076 Tübingen, FRG

Prof. Dr. Hanspeter A. Mallot  
Eberhard-Karls-Universität Tübingen  
Institut für Zoologie  
Auf der Morgenstelle 28  
72076 Tübingen, FRG

Prof. Dr. Rolf D. Ulrich  
Eberhard-Karls-Universität Tübingen  
Psychologisches Institut  
Friedrichstr. 21  
72072 Tübingen, FRG

Dr. Felix A. Wichmann  
Department of Empirical Inference for  
Machine Learning and Perception  
Max-Planck-Institut für biologische Kybernetik  
Spemannstr. 38  
72076 Tübingen, FRG

1. Auflage 2005

© by the Max-Planck-Institut für biologische Kybernetik

Alle Rechte der Veröffentlichung, des teilweisen oder vollständigen Ab- oder Nachdrucks, der fotomechanischen Wiedergabe, der Verfilmung, zur Vervielfältigung und Verbreitung, durch Ton- und Bildträger und der Übertragung durch Rundfunkmedien im In- und Ausland sind vorbehalten.

Herstellung: Knirsch Verlag, D-72138 Kirchentellinsfurt  
ISBN 3-927091-70-7

## Preface

After last year's very diverse topical focus with symposia on the cerebellum, visual localization in space-time, and signal processing in the auditory system, this year's TWK focus is more firmly back on visual perception. However, the three symposia reflect the diversity of research approaches and methods in perceptual science.

The first symposium, organised by Andreas Schilling from the Computer Science Department of the University of Tübingen is on *Analysis by Synthesis Methods in Computer Vision and Perception*. The second symposium, jointly organised by Martin Giese, University of Tübingen, and Heiko Neumann, University of Ulm, is entitled *Mid- and High-level Motion Analysis—From Image Features to Behavior*, and our third and final symposium *Face Processing of Familiar Faces* is organised by Claus-Christian Carbon from the Psychology Department of the Free University of Berlin.

As in all the previous years one critically important aspect of the TWK is the extended poster session on Saturday. Altogether we have accepted 124 posters to be presented at the TWK 2005, more or less the same number we accepted last year. Given the size constraints imposed upon the TWK by the size of the foyer of the *Kupferbau* conference venue, we consider this to be just the right number of poster presentations.

The public lecture will be given by Professor Dirk Vorberg, head of the Cognitive Psychology section of the Psychology Department of the Technical University of Braunschweig. His talk, entitled *Unbewusste Wahrnehmung: Probleme, Befunde, Grenzen* will again be given in German as this is a lecture intended not only for scientists, but for the general public of Tübingen and beyond.

Ever since the 3<sup>rd</sup> TWK in 2000 we have awarded a prize for the best poster contributed by an undergraduate or graduate student. In 2004 the poster prize was jointly awarded to Stephan Berthoud of the Universität Zürich for his contribution *The Perception of Dichotic Shepard Tone Intervals—Fusion or Suppression?* and to Thilo Womelsdorf from the German Primate Center in Göttingen for his work on *Dynamic Spatial Shifts of Neuronal Receptive Fields Toward Attended Locations in Macaque Area MT*. A prize will again be awarded at this TWK. The prize is a cheque for 500 Euros, donated by the Förderverein für neurowissenschaftliche Forschung, e.V., whose support we gratefully acknowledge.

The Organizing Committee  
Tübingen, January 2005

## **Sponsoring Institutions**

Max-Planck-Institut für biologische Kybernetik, Tübingen, Germany  
Eberhard-Karls-Universität Tübingen, Germany  
Förderverein für neurowissenschaftliche Forschung, e.V.

## **Local Support Team**

Dagmar Maier, Heinrich Bülhoff's secretary, contacted all our speakers and arranged their accommodation, distributed our information material, reserved the Kupferbau, organised coffee, tea and snacks during the TWK, drafted-in students to help put up poster-boards, handles the on-site registration during the conference, and sorts out all the dozens of unexpected disasters. It is difficult to imagine how TWK 2005 could have happened without her.

Since 2003 the TWK website has been updated, refined and maintained by Jez Hill, post-doctoral fellow in Bernhard Schölkopf's Empirical Inference Department at the MPI, building on the system he had designed and programmed in the previous years. Abstracts are submitted directly into a SQL-database from which they can be posted to the net and, using his PHP-based  $\TeX$  generator, straight into the TWK Proceedings, typeset using  $\LaTeX$ .

Ralf Buckenmayer and Julia Liske of the administration of the Max-Planck-Institut für biologische Kybernetik were again dealt the burden to deal with the accounting side of the conference registration.

Reinhard Feiler, Walter Heinz, Werner Koch, Michael Renner and Uli Wandel, again helped in the background both during and before the conference—if you do not notice them, this is a sign of how well they work to ensure that the TWK runs smoothly. In addition and as in previous years, Walter Heinz produced the name-badges and the tickets for the dinner at the Museum and printed the banner outside the Kupferbau. Internet access during the TWK is possible thanks to Michael Renner. Reinhard Feiler helped organize technical support provided by *Bewegte Bilder Medien AG* and produced the programme flyers.

## **Friday 25<sup>th</sup> February 2005**

14:00 Registration and hanging of posters

14:55 Welcome

### **Symposium 1: Analysis by Synthesis Methods in Computer Vision and Perception (Andreas Schilling, Universität Tübingen)**

Symposium to be conducted in English

15:00 **Introduction to Analysis by Synthesis Methods**

Andreas Schilling (University of Tübingen)

15:10 **Active Appearance Models and Analysis by Synthesis**

Tim Cootes (University of Manchester)

15:40 **Analysis by Synthesis for Human Pose Tracking: Merging View-Based and Model-Based Approaches**

Cris Curio (MPI for Biological Cybernetics, Tübingen)

16:10 **Model-Based Interface Quality 3D Hand-Tracking**

Ferenc Kahlesz (University of Bonn)

16:40 Coffee break

17:10 **Modeling Issues in Visual Tracking and Servoing**

Danica Kragic (Royal Institute of Technology Stockholm)

17:40 **Analysis of Video Streams in Driver Assistance Systems—the Role of Prediction**

Uwe Franke (Daimler-Chrysler Research)

18:10 **Analysis by Synthesis in 3d-Reconstruction from Images**

Reinhard Koch (University of Kiel)

## Saturday 26<sup>th</sup> February 2005

### **Symposium 2: Mid- and High-level Motion Analysis—From Image Features to Behavior (Martin Giese and Heiko Neumann, Universität Tübingen, Universität Ulm)**

Symposium to be conducted in English

- 09:00 **Mid- and High-Level Motion Analysis—From Image Features to Behavior**  
Martin A. Giese (University Clinic Tübingen) & Heiko Neumann (University of Ulm)
- 09:10 **Features, Apertures and Occlusions: Determining Coherent Motion**  
Heiko Neumann (University of Ulm)
- 09:40 **Attention and Motion—Neural Mechanisms of Spatial and Feature-Based Influences**  
Stefan Treue, Julio Martinez-Trujillo & Thilo Womelsdorf (German Primate Center Göttingen)
- 10:10 Coffee break
- 10:40 **Contrast Effects Between Concurrently Perceiving and Producing Events**  
Marc Grosjean (MPI for Human Cognitive and Brain Sciences, München)
- 11:10 **Cognition in Motion: Do we Represent Change-Over-Time?**  
Ian M. Thornton (MPI for Biological Cybernetics, Tübingen)
- 11:40 **Learning as Basis for the Visual Recognition of Actions**  
Martin A. Giese (University Clinic Tübingen)
- 12:10 Lunch break

### **Poster sessions**

The posters will be divided into 4 groups. The authors should be present at their poster at least during the time allocated to their group as follows. The poster boards will be marked to indicate which group the poster is in. Note that we have extended the poster session to allow more time at the posters.

- 13:20 Group A
- 14:40 Group B
- 16:00 Group C
- 17:20 Group D
- 18:50 Poster prize to be awarded

### **Public evening lecture (to be given in German)**

- 19:00 **Unbewusste Wahrnehmung: Probleme, Befunde, Grenzen**  
Dirk Vorberg (Technische Universität Braunschweig)
- 20:15 Banquet in the Tübinger Kelter  
NB: places are limited. Tickets may be reserved during online registration, and any remaining tickets will be sold at the registration desk.

## **Sunday 27<sup>th</sup> February 2005**

### **Symposium 3: Face Processing of Familiar Faces (Claus-Christian Carbon, Freie Universität Berlin)**

Symposium to be conducted in English

- 09:00 **Face Processing of Familiar Faces**  
Claus-Christian Carbon (Free University of Berlin)
- 09:10 **Fame or Fake? Critical Information Needed for Decisions on Facial Veridicality**  
Claus-Christian Carbon (Free University of Berlin)
- 09:40 **Component, Configural and Temporal Routes in View-Based Face Recognition**  
Adrian Schwaninger (University of Zürich & MPI for Biological Cybernetics, Tübingen),  
Christian Wallraven (MPI for Biological Cybernetics, Tübingen) & Heinrich H. Bühlhoff  
(MPI for Biological Cybernetics, Tübingen)
- 10:10 **Neural Representations of Faces**  
Stefan Schweinberger (University of Glasgow)
- 10:40 Coffee break
- 11:10 **The Early Repetition Effect: Reliability and Methodological Influences**  
Grit Herzmann (Humboldt University Berlin)
- 11:40 **Prosopagnosia as an Impairment to Face-Specific Mechanisms: Elimination of the Testable Domain-General Explanations**  
Brad Duchaine (Harvard University)
- 12:10 **Development of Face Recognition: how Important is Social Information?**  
Cornelia Groß & Gudrun Schwarzer (University of Gießen)
- 12:40 Closing discussion, removal of posters, end of conference



# Contents

<b>Public evening lecture (to be given in German)</b>	<b>19</b>
D. Vorberg: <i>Unbewusste Wahrnehmung: Probleme, Befunde, Grenzen.</i> . . . . .	20
<b>Symposium 1: Analysis by Synthesis Methods in Computer Vision and Perception</b>	<b>21</b>
A. Schilling: <i>Introduction to Analysis by Synthesis Methods</i> . . . . .	22
T. Cootes: <i>Active Appearance Models and Analysis by Synthesis</i> . . . . .	24
C. Curio: <i>Analysis by Synthesis for Human Pose Tracking: Merging View-Based and Model-Based Approaches</i> . . . . .	25
F. Kahlesz: <i>Model-Based Interface Quality 3D Hand-Tracking</i> . . . . .	26
D. Kragic: <i>Modeling Issues in Visual Tracking and Servoing</i> . . . . .	27
U. Franke: <i>Analysis of Video Streams in Driver Assistance Systems—The Role of Prediction</i>	28
R. Koch: <i>Analysis by Synthesis in 3d-Reconstruction from Images</i> . . . . .	29
<b>Symposium 2: Mid- and High-level Motion Analysis—From Image Features to Behavior</b>	<b>31</b>
M. A. Giese & H. Neumann: <i>Mid- and High-Level Motion Analysis—From Image Features to Behavior</i> . . . . .	32
H. Neumann: <i>Features, Apertures and Occlusions: Determining Coherent Motion</i> . . . . .	33
S. Treue, J. Martinez-Trujillo & T. Womelsdorf: <i>Attention and motion—Neural Mechanisms of Spatial and Feature-Based Influences</i> . . . . .	34
M. Grosjean: <i>Contrast Effects Between Concurrently Perceiving and Producing Events</i> . . . . .	35
I. M. Thornton: <i>Cognition in Motion: Do we Represent Change-Over-Time?</i> . . . . .	36
M. A. Giese: <i>Learning as Basis for the Visual Recognition of Actions</i> . . . . .	37

<b>Symposium 3: Face Processing of Familiar Faces</b>	<b>39</b>
C.-C. Carbon:	
<i>Face Processing of Familiar Faces</i> . . . . .	40
C.-C. Carbon:	
<i>Fame or Fake? Critical Information Needed for Decisions on Facial Veridicality</i>	41
A. Schwaninger, C. Wallraven & H. H. Bühlhoff:	
<i>Component, Configural and Temporal Routes in View-Based Face Recognition</i> .	42
S. Schweinberger:	
<i>Neural Representations of Faces</i> . . . . .	43
G. Herzmann:	
<i>The Early Repetition Effect: Reliability and Methodological Influences</i> . . . . .	44
B. Duchaine:	
<i>Prosopagnosia as an Impairment to Face-Specific Mechanisms: Elimination of the Testable Domain-General Explanations</i> . . . . .	45
C. Groß & G. Schwarzer:	
<i>Development of Face Recognition: how Important is Social Information?</i> . . . .	46
<b>Attention and Awareness</b>	<b>47</b>
K. Bausenhardt, B. Rolke & R. Ulrich:	
<i>Locus of the Temporal Uncertainty Effect: Evidence from the Psychological Refractory Period</i> . . . . .	48
C. W. G. Clifford & J. A. Harris:	
<i>Contextual Modulation Outside of Awareness</i> . . . . .	49
A. Dinkelbach, B. Rolke & R. Ulrich:	
<i>A Benefit of Attentional Stimulus Processing? Temporal Discrimination Within a Spatial Cueing Paradigm</i> . . . . .	50
T. Eggert, S. Casana-Perez & A. Straube:	
<i>Deficits in the Control of Visual Attention in Patients with Cerebellar Lesions or with Parkinson's Disease</i> . . . . .	51
V. V. Gavrik:	
<i>Natural Semi-Logarithmic Decision Making and the Fechner's Paradox</i> . . . . .	52
E. Hein, B. Rolke & R. Ulrich:	
<i>Attention and Temporal Discrimination: Does the Type of Attentional Orienting Matter?</i> . . . . .	53
W. H. R. Miltner, S. Krieschel et al.:	
<i>Eye-Movements and Behavioral Responses to Threatening and Non-Threatening Stimuli During Visual Search in Phobic and Non-Phobic Subjects</i> . . . . .	54
B. Rolke, K. Bausenhardt et al.:	
<i>Temporal Discrimination Within the Attentional Blink</i> . . . . .	55
J. Saenger & E. Wascher:	
<i>Task Preparation in Task Switching—Insights from an ERP Study</i> . . . . .	56
A. Schankin & E. Wascher:	
<i>The N2pc as a Correlate of Recurrent Processes in Visual Areas</i> . . . . .	57
B. Spehar & R. Beecham:	
<i>Visual Selection of Perceptually Completed Objects: The Advantage of Modal Completion</i> . . . . .	58

Contents	11
<b>Auditory Perception</b>	<b>59</b>
S. Dietrich, A. Ischebeck et al.:	
<i>Prosody and Lexicality in Emotional Speech Perception</i> . . . . .	60
T. Fischer, M. Nusseck & H. Bruhn:	
<i>Timing Precision in Synchronization Tapping and Effects of Training</i> . . . . .	61
T. Mahler, P. Bayerl et al.:	
<i>Attention Driven Auditory Display</i> . . . . .	62
<b>Computer Vision and Robotics</b>	<b>63</b>
A. Birk:	
<i>Recognition of Occluded Objects through Evolutionary Learning</i> . . . . .	64
S. Fleck, S. Lanwer et al.:	
<i>The Smart Camera Paradigm and its Role in Tracking</i> . . . . .	65
S. Frintrop, G. Backer & E. Rome:	
<i>A Goal-Directed Visual Attention System</i> . . . . .	66
A. Tewes, C. von der Malsburg & R. P. Würtz:	
<i>A Flexible Object Model for Facial Gesture Recognition</i> . . . . .	67
Q. C. Vuong & M. Castrilln-Santana:	
<i>Improving Automatic Pedestrian Detection by Means of Human Perception</i> . . . . .	68
G. Westphal & R. P. Würtz:	
<i>Recognition of Pose and Illumination of Human Faces Through Combination of Feature-Based and Correspondence-Based Pattern Recognizers</i> . . . . .	69
<b>Multimodal Interactions: Sensory Integration</b>	<b>71</b>
H. Helbig & M. O. Ernst:	
<i>Cognitive Factors Facilitate Multimodal Integration.</i> . . . . .	72
C. Lange & M. O. Ernst:	
<i>Judging Size by Hand: No Benefit for Bimanual Estimates</i> . . . . .	73
B. E. Riecke, J. Schulte-Pelkum et al.:	
<i>Auditory Cues can Facilitate the Visually-Induced Self-Motion Illusion (Circular Vection) in Virtual Reality</i> . . . . .	74
C. Steckner:	
<i>Anticipation of Heaviness in Vision and Grasp</i> . . . . .	75
A. Wacker, A. Chatziastros & H. H. Bülthoff:	
<i>The Impact of Attention on Visual-Vestibular Interaction in Helicopter Control</i> . . . . .	76
<b>Multimodal Interactions: Sensory-Motor Integration</b>	<b>77</b>
D. Berger, C. Terzibas & H. H. Bülthoff:	
<i>Influence of Motion Cueing on Helicopter Stabilisation</i> . . . . .	78
D. Braun, L. Pracejus & K. R. Gegenfurtner:	
<i>No Correlation Between Smooth Pursuit and Perceived Motion in the Motion after Effect</i> . . . . .	79
K. Drewing & J. Trommershäuser:	
<i>Visually-Guided Pointing with Trajectory Perturbations: Tactile-Kinesthetic Detection and Movement Costs</i> . . . . .	80
S. Gulan, K. Seymour & U. J. Ilg:	
<i>Changes in Saccadic Latencies Due to Distractive Information : An Analysis Based on Latency Distributions</i> . . . . .	81

A. Horstmann & K.-P. Hoffmann:	
<i>Target Selection Characteristics in Eye-Hand Coordination</i> . . . . .	82
S. Klingenhoefer & F. Bremmer:	
<i>Perisaccadic Localization of Multisensory Targets</i> . . . . .	83
I. Liepelt, M. Trenner et al.:	
<i>The Relationship Between Action Pantomime, Semantic Object Knowledge and Multiple Object Use in Stroke Patients- a Multivariate Analysis</i> . . . . .	84
O. Masseck, B. Röhl & K.-P. Hoffmann:	
<i>The Visual System of Diurnal and Nocturnal Geckos (Gekkonidae, Reptilia): Optomotor Reaction and Associated Brain Structures</i> . . . . .	85
T. Schmidt, S. Niehaus & A. Nagel:	
<i>The Feedforward Dynamics of Response Priming by Color</i> . . . . .	86
<b>Natural Image Statistics</b>	<b>87</b>
J. Drewes, F. A. Wichmann & K. R. Gegenfurtner:	
<i>Classification of Natural Scenes using Global Image Statistics</i> . . . . .	88
R. Fleming:	
<i>On the Visual Perception of Translucent Materials</i> . . . . .	89
L. Gerstmayr & H. A. Mallot:	
<i>How does the Environment Influence the Shape of Receptive Fields in Cortical Simple Cells?</i> . . . . .	90
<b>Perception of Time</b>	<b>91</b>
E. Golovina:	
<i>Visual Durations Discrimination: Confidence Characteristics and Cognitive Styles</i> . . . . .	92
J. Nitschke & R. Ulrich:	
<i>Perceived Duration of Frequent and Infrequent Stimuli</i> . . . . .	93
G. Shenvi & E. A. Elliott:	
<i>Some Effects of Negative Delays Upon the Perception of Causal Relatedness</i> . . . . .	94
<b>Perceptual Development</b>	<b>95</b>
D. Poggel, C. Calmanti et al.:	
<i>The Toelz Temporal Topography Study: Mapping the Visual Field Across the Life Span</i> . . . . .	96
<b>Perceptual Neuroscience</b>	<b>97</b>
A. Gharabaghi, A. A. Sokolov et al.:	
<i>Biological Motion Processing in Neurosurgical Patients with Cerebellar Lesions</i> . . . . .	98
A. A. Karim, A. Schueler et al.:	
<i>Facilitating Effect of 15 Hz Repetitive Transcranial Magnetic Stimulation (rTMS) on Tactile Perceptual Learning</i> . . . . .	99
C. Mora-Ferrer & K. Behrend:	
<i>Cholinergic and GABAergic Effects on Temporal Transfer Properties in Goldfish Examined with the ERG and Behavioral Experiments</i> . . . . .	100
M. Pavlova, F. Marconato et al.:	
<i>Biological Motion Processing in Patients with Periventricular Lesions: MEG Findings</i> . . . . .	101

P. Ragert, C. Klaes et al.:	
<i>rTMS Induced Improvement of Human Visual Orientation Discrimination</i> . . . . .	102
T. Teichert, T. Wachtler et al.:	
<i>Relations Between Preferred Spatial Frequency and Size of Spatial Summation Field in V1 Multi-Unit Activity of Awake Macaque</i> . . . . .	103
<b>Research Methods and Techniques</b>	<b>105</b>
M. Kuss, F. Jäkel & F. A. Wichmann:	
<i>Bayesian Inference for Psychometric Functions</i> . . . . .	106
K. Spang, V. Diehl & M. Fahle:	
<i>Contrast Dependence of the BOLD Response in Visual Cortices</i> . . . . .	107
K. Swirszcz, M. Zvyagintsev et al.:	
<i>Amygdala Activation in Response to Fearful Faces: Lateralization Effects in fMRI Depending on Phase-Encoding Direction</i> . . . . .	108
T. G. Tanner, N. J. Hill et al.:	
<i>Efficient Adaptive Sampling of the Psychometric Function by Maximizing Information Gain</i> . . . . .	109
M. Wolf, J. P. Maurer & M. Bach:	
<i>Objective Estimation of Visual Acuity—High Correlation Across Normal and Diseased Vision with an Automated VEP-Based Algorithm</i> . . . . .	110
<b>Tactile Perception</b>	<b>111</b>
S. Franzkowiak, P. Ragert et al.:	
<i>End-Effect Suppression in a Multiple Choice Reaction Time Task Induced by Tactile Coactivation</i> . . . . .	112
<b>Visual Cognition: General</b>	<b>113</b>
V. Bochev, E.-J. Sim & M. Kiefer:	
<i>The Role of Visual and Action Features in the Representation of Natural and Artifactual Objects</i> . . . . .	114
C. Dahl & D. C. Kiper:	
<i>Canonicalness and Prototypicality of an Object Effects Remembering Performance</i> . . . . .	115
B. Garsoffky, S. Schwan & M. Huff:	
<i>Influencing Visual Recognition of Dynamic Scenes by the Use of Canonical Viewpoints</i> . . . . .	116
P. Jakob:	
<i>The Feeling of Privacy in Virtual Architecture</i> . . . . .	117
E.-J. Sim, S. Liebich et al.:	
<i>How does the Brain Acquire Object Categories? Diagnosticity of Visual and Action Features Drives Representation of Novel Objects</i> . . . . .	118
A. Stemme, G. Deco et al.:	
<i>Temporal Dynamics of the Interaction Between Working Memory and Attention: A Neuronal Model of a Wisconsin-DMS-Task</i> . . . . .	119
A. Thiel, R. Goertz & R. Sireteanu:	
<i>Curvature Versus Pointing in Visual Search</i> . . . . .	120

<b>Visual Cognition: Objects and Faces</b>	<b>121</b>
A. Altmann, W. Grodd et al.:	
<i>Similar Cortical Correlates Underlie Visual Object Identification and Orientation Judgment</i> . . . . .	122
S. W. Bock, P. W. Dicke & P. Thier:	
<i>Gaze Following in Humans: Experiments on its Precision</i> . . . . .	123
I. Bühlhoff:	
<i>Sensitivity to Changes in Identity, Caricature and Sex in Face Recognition</i> . . .	124
T. Cooke, C. Wallraven & H. H. Bühlhoff:	
<i>Representation of Shape and Texture Changes in Novel 3D Objects</i> . . . . .	125
M. Graf, H. Helbig & M. Kiefer:	
<i>Motor Representations in Visual Object Recognition</i> . . . . .	126
C. M. Hein, M. Bach & J. Kornmeier:	
<i>Old/Young Woman: ERP Correlates of the Disambiguation of an Ambiguous Face-Figure</i> . . . . .	127
S. P. Heinrich & K. Grill-Spector:	
<i>The Time Course of Object Repetition Effects in Human Visual Cortex</i> . . . . .	128
M. Herrmann, A. Muehlberger & A. Fallgatter:	
<i>Source Localization (LORETA) of Early Stages of Face Processing: An ERP Study</i> . . . . .	129
F. Hofer & A. Schwaninger:	
<i>Measuring the Effectiveness of Individually Adaptive Computer Based Training (CBT)</i> . . . . .	130
J. Jastorff, Z. Kourtzi & M. A. Giese:	
<i>Learning of Biological Motion: Combining fMRI and Theoretical Modeling</i> . .	131
H. Kirchner, Q. C. Vuong et al.:	
<i>Moving Objects in Ultra-Rapid Visual Categorisation Result in Better Accuracy, but Slower Reaction Times than Static Presentations</i> . . . . .	132
J. Kornmeier & M. Bach:	
<i>ERP Correlates to Perceptual Reversal of Ambiguous Figures: Effects of ISI Durations</i> . . . . .	133
W. A. Kreiner:	
<i>The Flower Pot Illusion</i> . . . . .	134
M. Nusseck, R. Fleming & H. H. Bühlhoff:	
<i>Abstraction of Physical Properties from Complex Object Interactions—The Case of Elasticity</i> . . . . .	135
K. Pilz, I. M. Thornton & H. H. Bühlhoff:	
<i>Motion Matters: Facial Motion Improves Delayed Visual Search Performance</i> .	136
A. A. Sokolov, A. Sokolov & M. Pavlova:	
<i>Perceived Dynamics of Static Images Enables Emotional Attribution</i> . . . . .	137
J. E. Weber, C.-C. Carbon et al.:	
<i>Characterization of Face Processing in Patients with Congenital Prosopagnosia by Simultaneous MEG/EEG Recordings</i> . . . . .	138
U. Weidenbacher, P. Bayerl et al.:	
<i>Mechanisms of Recovering Shape Properties from Perfectly Mirrored Objects</i> .	139

<b>Visual Cognition: Space Perception and Navigation</b>	<b>141</b>
S. D. Glaus & T. Jarchow: <i>The Influence of Body Tilt on Mental Rotation and Subjective Visual Vertical</i> . . .	142
M. Huff, B. Garsoffky & S. Schwan: <i>Can A-Priori Instructions Induce a Viewpoint Independent Mental Representation of a Dynamic Scene?</i> . . . . .	143
A. Weiss, S. Gillner & H. A. Mallot: <i>Contrast Dependence of Human Place Memory</i> . . . . .	144
J. M. Wiener, N. Rossmanith et al.: <i>Isovists Capture Locomotion-Relevant Properties of Space</i> . . . . .	145
<b>Visual Perception: Colour</b>	<b>147</b>
K. R. Gegenfurtner, S. Walter & T. Hansen: <i>Color Vision for Natural Objects: Color Constancy</i> . . . . .	148
T. Hansen & K. R. Gegenfurtner: <i>Color Vision for Natural Objects: Color Discrimination</i> . . . . .	149
K. Heier & S. Schultz: <i>Color Constancy: Increment-Decrement-Asymmetries Depend on Illuminant Color</i> . . . . .	150
S. Schultz: <i>Color Constancy and its Relation to Models of Dichromacy</i> . . . . .	151
T. Wachtler & R. Hertel: <i>Contribution of On- and Off-Pathways to Color Categories of Dichromats</i> . . .	152
S. Walter: <i>Can Differences in Colour Categorization be Caused by Lens Brunescence?</i> . . .	153
<b>Visual Perception: Depth</b>	<b>155</b>
M. Hofbauer & T. Eggert: <i>Larger Perceived Distance Yields Shorter Responses in Time-To-Collision Task</i>	156
S. J. Maier, A. E. Welchman & H. H. Bühlhoff: <i>The Role of Extra-Retinal Cues in Velocity Constancy</i> . . . . .	157
<b>Visual Perception: Eye Movements</b>	<b>159</b>
K. Beykirch, C. von Laßberg & J. Krug: <i>Are there Correlations Between Vertical VOR and Multiaxial Spatial Orientation?</i>	160
A.-M. Brouwer, V. Franz et al.: <i>Looking During Grasping</i> . . . . .	161
M. Dorr, M. Böhme et al.: <i>Variability of Eye Movements on High-Resolution Natural Videos</i> . . . . .	162
D. de Grave, V. Franz & K. R. Gegenfurtner: <i>The Coding of Combined Pointing Movements and Saccades in a Length Illusion</i>	163
A. Kaminiarz, M. Rohe & F. Bremmer: <i>Localization of Visual Targets During Optokinetic Eye Movements</i> . . . . .	164
C. S. Konen, R. Kleiser et al.: <i>The Cortical Interaction of Attention and Intention</i> . . . . .	165
C. S. Konen, R. Kleiser et al.: <i>I Know where You'll Look: An fMRI Study of Oculomotor Intention and a Change of Motor Plan</i> . . . . .	166

C. von Laßberg, K. Beykirch & J. Krug: <i>Visual Pursuit in Gymnasts</i> . . . . .	167
B. M. Schäfer: <i>The Contribution of Torsional Oculomotor Dynamics to Spatial Interactions of Vertebrates</i> . . . . .	168
M. Spering, D. Kerzel & K. R. Gegenfurtner: <i>The Effect of Distractors on Smooth Pursuit Eye Movements</i> . . . . .	169
M. Stritzke & J. Trommershäuser: <i>Eye Movements During a Video-Game-Like Pointing Task: What Attracts the Eye while the Hand Moves?</i> . . . . .	170
B. White, K. R. Gegenfurtner & D. Kerzel: <i>Structured Random Noise as a Visual Distractor for Saccadic Eye Movements</i> . . . . .	171
N. E. Ziegler & D. Kerzel: <i>Exogenous and Endogenous Attention Shifts During Smooth Pursuit Eye Movements</i> . . . . .	172
<b>Visual Perception: Motion</b>	<b>173</b>
D. Barteit, P. Bayerl & H. Neumann: <i>Illusory Contours and the Amodal Grouping of Visual Motion</i> . . . . .	174
P. Bayerl & H. Neumann: <i>The Role of Attention in Figure-Ground Separation of Transparent Visual Motion—A Modeling Approach</i> . . . . .	175
B. Friedrich, F. Caniard et al.: <i>Motion-Induced Shift and Navigation in Virtual Reality</i> . . . . .	176
M. Gehres & C. Neumeyer: <i>Perception of Object-Movement in Goldfish (<i>Carassius Auratus</i>)</i> . . . . .	177
S. Getzmann: <i>Shifting the Onset of a Moving Sound Source: An Auditory Fröhlich Effect?</i> . . . . .	178
N. Hagemann & A. Ziegler: <i>The Advantage of Left-Handers in Sports Due to a Perceptual Frequency Effect</i> . . . . .	179
P. Haselmayer, C. Neumeyer & C. von Campenhausen: <i>Spectral Sensitivity of Apparent Motion does not Follow <math>V(\lambda)</math> but <math>[V(\lambda)]^2</math> !</i> . . . . .	180
D. Hupfeld & K.-P. Hoffmann: <i>Global Motion Thresholds in Albinotic and Pigmented Ferrets (<i>Mustela Putorius Furo</i>)</i> . . . . .	181
F. I. Kandil & M. Lappe: <i>Mechanisms of Spatio-Temporal Interpolation</i> . . . . .	182
S. Kuhlmann & M. Lappe: <i>Recognition of Biological Motion from Blurred Natural Scenes</i> . . . . .	183
S. McDonald & Z. Kourtzi: <i>Perception of Motion Induction for Naturalistic Images in the Human Visual Cortex</i> . . . . .	184
R. Philipp, C. Distler & K.-P. Hoffmann: <i>A Motion Sensitive Area in Ferret Extrastriate Visual Cortex: An Analysis in Pigmented and Albino Animals</i> . . . . .	185
P. Pretto & A. Chatziastros: <i>Apparent Speed in the Footstep Illusion is not Totally Contrast Dependent</i> . . . . .	186

C. L. Roether & M. A. Giese:  
*Recognizing Emotions Expressed in Human Walking: Integration of Synergies* . 187

P. Sarkheil, J. Jastorff et al.:  
*Categorization of Complex Dynamic Patterns in the Human Brain* . . . . . 188

R.-B. Schmidt-Hoffmann & C. Mora-Ferrer:  
*Contrast-Sensitivity of Goldfish, Measured with the Optomotor Response, is Independent from Glycine- and NMDA-Receptor Mediated Actions* . . . . . 189

A. E. Welchman:  
*Does the Paradoxical Perception of Motion in Depth Reflect an “Optimal” Solution?* . . . . . 190

M. Zvyagintsev, H. Menning et al.:  
*Audio-Visual Perception of Self-Induced Apparent Motion* . . . . . 191

**Visual Perception: Spatial Vision** **193**

K. Georg, F. H. Hamker & M. Lappe:  
*Saccadic Compression with near Threshold Stimuli in Light and Dark Adapted Conditions* . . . . . 194

N. Koechy:  
*Rapid Natural Scene Discrimination: Orientation and Context Effects.* . . . . 195

F. Scharnowski & M. Herzog:  
*Time Will Tell: The Temporal Order of Events Determines Feature Integration* . 196

H. Strasburger:  
*Character Recognition and Ricco’s Law* . . . . . 197

M. Zirnsak, F. H. Hamker et al.:  
*A Neurocomputational Model of Perisaccadic Compression and Receptive Field Dynamics Based on Oculomotor Feedback* . . . . . 198

**Index of Authors** **199**



**Public evening lecture (to be given in German)**

Prof. Dr. Dirk Vorberg  
Technische Universität Braunschweig

Saturday 26<sup>th</sup> February 2005

**Unbewusste Wahrnehmung: Probleme, Befunde, Grenzen.**

Dirk Vorberg

Institut für Psychologie der TU Braunschweig

d.vorberg@tu-bs.de

Ob Reize unser Verhalten auch dann beeinflussen können, wenn wir sie nicht bewusst registrieren ('unterschwellige' oder 'subliminale Wahrnehmung'), ist eine alte Streitfrage, und in den populären Medien finden sich regelmäßig Berichte über den Einsatz solcher Techniken und ihre Gefahren, z.B. in der Werbung. Wenn man Erwartungs-Effekte ausschließt, zeigen sich nur vernachlässigbare Wirkungen unbewusster Reize. Das scheint die Meinung von Skeptikern zu stützen, dass Belege für unbewusste Wahrnehmung überwiegend auf methodische Artefakte zurückgehen. Diese Ansicht lässt sich angesichts einer Fülle neuer Befunde nicht mehr aufrechterhalten. Sie zeigen zweifelsfrei, dass sich vorbereitete Handlungen durch maskierte Reize bahnen ("Priming") oder hemmen lassen, obwohl sie nicht bewusst wahrgenommen werden können. Die aktuelle Forschung zweifelt nicht mehr an der Existenz unbewusster Wahrnehmung, sondern fragt nach den Bedingungen, unter denen sie möglich ist. Für Handlungs-Priming sind sie weitgehend bekannt; seine Wirkungen sind überaus kurzlebig, was ausschließt, es für unkontrollierbare Manipulationen einzusetzen. Kontrovers bleibt, worin sich semantisches, evaluatives und affektives Priming von Handlungs-Priming unterscheiden, welche qualitativen Unterschiede zwischen unbewusster und bewusster Wahrnehmung bestehen, und vor allem, welches ihre neuronalen Korrelate sind.

## **Symposium 1**

### **Analysis by Synthesis Methods in Computer Vision and Perception**

Andreas Schilling  
Universität Tübingen

## Introduction to Analysis by Synthesis Methods

Andreas Schilling

University of Tübingen

andreas@wsi-gris.uni-tuebingen.de

“Analysis-by-Synthesis” (AbS) methods have recently gained much interest in the area of processing visual images. In AbS techniques, the knowledge (which is always needed for recognition) is represented in the form of some parameterized model of reality that can be used to create or render 2d-images or views. Another possibility is to represent the knowledge directly in the form of images—a kind of representation that is supposed to be mainly used in the brain. It is still an open question to which degree humans use their ability to synthesize new views from known objects for recognition and thus also employ some sort of model based recognition. However, in the area of technical computer vision, AbS approaches are commonly used, though known by other names like model based vision or tracking.

The basic principle of AbS is very simple: the parameters of the model are adjusted in a feedback loop until its rendered representation matches with the measured sensor image. This principle can be generalized to include arbitrary sensors like radar, lidar, ultrasound etc.

There are however, several important research questions that are addressed in the following talks of this more technically oriented symposium:

**Representation of the model:** There exists a broad spectrum of possibilities ranging from simple wireframe models to images combined with warping rules to geometric meshes, combined with texture and transformation rules. An important point is how the large number of parameters (often exceeding tens of thousands) can be reduced to a small but meaningful set of parameters using statistical tools. T. Cootes, who is one of the pioneers in this area and laid the foundations for the work on morphable face models carried out here in Tübingen by Blanz and Vetter, will give the first talk about Active Appearance Models and AbS. Later D. Kragic, who used model based vision for a robot system, that can operate in a domestic environment with changing lighting and background conditions will address the issue of models in robot vision.

**Initialization of the model** One of the key problems with AbS is to find initial parameters for the optimization loop. C. Curio will show us, how to combine advanced algorithms from machine learning with AbS in order to not only initialize the model, but also support robust dynamic tracking. His application is the tracking of human poses.

**Determination of the level of abstraction,** on which the prediction by the model is compared to the measurement—Choice of difference measures and optimization algorithms AbS does not necessarily mean, that photorealistic computer generated images are compared to camera input. It is one of the open research questions to decide, which abstraction level should be used for the comparison with reality and, which similarity measures and optimization algorithms to use for the AbS loop. Which level of preprocessing or filtering should be applied to the sensor data? F. Kahlesz will show us by the example of hand tracking, how a practical implementation for real-time tracking can look like and how computer graphics algorithms as well as computer graphics hardware can be used to support the vision task. U. Franke will give us insights into algorithms used in modern driver assistance systems for real time tracking, where the comparison takes place on a rather abstract level, and the relations between AbS and Kalman prediction become obvious.

A key reason for the massively increasing interest in AbS is the fact, that today we are able to generate nearly photorealistic images from 3d models faster and better than ever before. The connection between computer graphics and computer vision opens up new possibilities and a new area of research. R. Koch was one of the first researchers who published about AbS in computer vision already in 1993. He will show us some of the mentioned possibilities introducing his work about scene analysis for the purpose of rendering novel views from a given set of real camera views of a 3D scene.

**Active Appearance Models and Analysis by Synthesis**

Tim Cootes

University of Manchester

tim.cootes@man.ac.uk

Statistical models of shape and appearance have been shown to be powerful tools for image interpretation, as they can explicitly deal with the natural variation in objects of interest. Such models can be built from suitably labelled training sets. Given a model of appearance we can match it to a new image using efficient optimisation algorithms, which seek to minimise the difference between a synthesized model image and the target image (Analysis by Synthesis). The Active Appearance Model algorithm exploits the linear nature of the model so as to be able to accurately estimate the optimal model parameters very quickly. This talk will describe the appearance models and matching algorithms and demonstrate their application to a number of image interpretation problem domains.

## References:

T.F. Cootes, D. Cooper, C.J. Taylor and J. Graham, "Active Shape Models—Their Training and Application." *Computer Vision and Image Understanding*. Vol. 61, No. 1, Jan. 1995, pp. 38–59.

A.Lanitis, C.J.Taylor, T.F.Cootes, "Automatic Interpretation and Coding of Face Images Using Flexible Models. *IEEE PAMI* Vol.19 No.7 pp.743–756, July 1997.

T.F.Cootes, G.J. Edwards and C.J.Taylor. "Active Appearance Models", *IEEE PAMI*, Vol.23, No.6, pp.681–685, 2001

## Analysis by Synthesis for Human Pose Tracking: Merging View-Based and Model-Based Approaches

Cris Curio

MPI for Biological Cybernetics, Tübingen

`cristobal.curio@tuebingen.mpg.de`

Action recognition by the brain is thought to combine the recognition of body configurations with some form of feature tracking over time (for review see [1]). So far, in the field of computer vision these two mechanisms are typically addressed as separate problems. Proposed solutions have either been based on independent recognition of learned body configurations [2,3], or on high-level model-based stochastic tracking mechanisms [4–6] in both 2D and 3D space. A new system will be presented that integrates dynamic view-based pose configuration estimation and model-based articulated tracking, resulting in increased robustness and improved generalization properties for a small set of training data. A combination of kernel-based nonlinear regression analysis and competitive particle filtering maps together image features robustly onto points of a smooth manifold of body postures (action space). This analysis-by-synthesis based posture estimation is used as a prior for the automatic initialization and support of the tracking of a flexible articulated 2D model. It will be discussed how this computational approach represents perceptual ambiguities and robustly solves the correspondence problem for the reconstruction of human body pose under strong self-occlusions.

- [1] Giese, M., and T. Poggio (2003). Neural mechanisms for the recognition of biological movements. *Nature Review Neuroscience*, 4(3):179–192.
- [2] Brand, M. (1999). Shadow puppetry. In *International Conference on Computer Vision*, 1237–1244.
- [3] Rosales, R., and S. Sclaroff (2001). Learning body pose via specialized maps. In *Neural Information Processing Systems NIPS-14*.
- [4] Rehg, J., and T. Kanade (1995). Model-based tracking of self-occluding articulated objects. In *International Conference on Computer Vision*, pages 612–617.
- [5] Sidenbladh, H., M. Black, and D. Fleet (2000). Stochastic tracking of 3d human figures using 2d image motion. In *European Conference on Computer Vision (2)*, pages 702–718.
- [6] Wachter, S., and H.-H. Nagel (1999). Tracking of persons in monocular image sequences. *Computer Vision and Image Understanding*, 74:174–192.

## Model-Based Interface Quality 3D Hand-Tracking

Ferenc Kahlesz

University of Bonn

fecu@cs.uni-bonn.de

The feedback subsystem of 3D Human-Computer Interaction has seen tremendous advancement in recent years, as Computer Graphics become increasingly able to render stunningly immersive 3D scenes. This holds not only for expensive VR environments, but also for commodity desktop PCs, thanks to the advancement of GPUs. A natural demand from the user would be to use her own hand(s) for 3D interaction with these virtual environments. Unfortunately, the feedforward part, i.e. subsystems providing 3D user input, was not able to advance on par with visualization. The problems, among others, are that, on the one hand, hardware devices used for true 3D input are expensive and intrusive (e.g. datagloves, electromagnetic tracking systems) and, on the other hand, less intrusive systems utilizing vision-based methods do not yet allow true 3D input. As a consequence, we still have to use intrusive and cumbersome 3D input devices in virtual environments (and 3D desktop applications), which diminishes user performance and experience.

Our goal is to provide the user with a “virtual dataglove”: a complete, nonintrusive, vision-based replacement of the traditional dataglove. Our “virtual dataglove” measures the 3D pose of the hand (global motion) and the movement of the fingers (local motion). We call the global 3D parameters and the jointangles of the hand together the state of the hand. The user’s hand is then observed by one or more cameras and for every camera frame, the state of the hand is determined. This problem belongs to the area called ‘articulated’ or ‘deformable object tracking’. Basically, there are two types of approaches to deal with articulated tracking: image- (or contour-) and model-based methods.

We pursue a model-based approach. First, camera images are segmented via skin-color segmentation into binary hand-contour images. Then, a  $21+6 = 27$  degree-of-freedom 3D hand model is fitted into these contour images. Using calibrated cameras, our OpenGL hand-model can be rendered into the images with the same (interior and exterior) camera parameters as the real cameras have. The error of the fit is the amount of pixels in the XOR image of the contour and the rendered image. This error function can be effectively computed on up-to-date graphics hardware, which enables a large number of error function evaluations in every camera frame. The fitting error is minimized by an iterative Downhill-Simplex optimization algorithm. Two kinds of constraints of the possible handstates are exploited to narrow the search space of the optimization: the joint angles of the fingers are biometrically constrained, and the dimensionality of the 21 DOF state vector can be reduced to about 11 DOFs using PCA. In addition, we exploit temporal coherence by providing a good initial-value for the Downhill-Simplex optimization using the result of the optimization from the last frame and first-order motion prediction.

## Modeling Issues in Visual Tracking and Servoing

Danica Kragic

Royal Institute of Technology Stockholm

danik@nada.kth.se

One of the key research areas in robotics today is mobile manipulation. For performing realistic tasks in dynamic, domestic environments one of the biggest requirements is robust and flexible visual system. The use of visual information is important for place recognition in terms of navigation and SLAM, real-time feedback for visual servoing control, scene segmentation and object recognition, estimation of objects' pose (position and orientation) for grasping purposes.

For humans, everyday activities such as pointing, grasping, reaching, catching, various tool manipulation are strongly dependent on rich coordination between the eye and the hand. Each of these actions require attention to different attributes in the environment—while pointing requires only an approximate location of the object in the visual field, a reaching or grasping movement require more exact information about the object's pose. An extensive study of human visually guided grasps has shown that the human visuomotor system takes into account the three dimensional geometric features rather than the two dimensional projected image of the target objects to plan and control the required movements. Compared to most of the current robotic visual servoing systems, which are image based and based on 2D feature tracking, the information used by humans is much more complex and permits humans to operate in large range of environments. In terms of robotic manipulation, it is usually required to accurately estimate the pose of the object to, for example, allow the alignment of the robot arm with the object or to generate a feasible grasp and grasp the object.

Using prior knowledge about the object properties such as size, texture or shape, a special representation can further increase the robustness of the tracking system. Realistic environments (tables, shelves) and natural objects (such as food packages, cups, etc.) offer very little place for assumptions such as, for example, uniform color or simple texture attributes. In terms of pose estimation, a number of model-based tracking systems have been proposed with a common feature of using wire-frame models to estimate current pose/velocity of the object. Most of them deal with tracking of particular targets (e.g. cars), usually uniform in color with a moderately varying backgrounds.

This talk will address the issue of models related to a vision system used in a robotic mobile manipulation framework. Combination of foveal and peripheral stereo vision system will be presented related to scene segmentation, object recognition and model based tracking. Position and image based visual servoing control will be discussed in terms of their requirements and robustness.

**Analysis of Video Streams in Driver Assistance Systems—The Role of Prediction**

Uwe Franke

Daimler-Chrysler Research

uwe.franke@daimlerchrysler.com

Cars of the near future will be equipped with sensors that perceive their environment in order to recognize dangerous situations as early as possible. Besides Radar and Lidar, computer vision will play a prominent role due to its excellent resolution in space and time. Vision based safety systems such as Lane Departure Warning and Lane Departure Protection are already on the market. Current research concentrates on more complex tasks like understanding inner city traffic, especially at intersections.

In the 80s, mobile robots were programmed to analyse the situation using state-of-the-art computer vision methods, drove some meters, stopped and started thinking again. Then, the new situation was linked to the previously extracted information. Equipped with the most powerful computers available at that time, those vehicles “drove” extremely slowly.

In the mid 80s, Dickmanns introduced the principles of control theory, namely observer theory, to the analysis of image sequences. His idea was to estimate the relevant parameters of a dynamic system by local measurements and to use the prediction inherent in Kalman Filters to guide the vision process efficiently. 1986, his vehicle was able to drive at about 90 km/h on a German highway with only a couple of 8086 processors. Thanks to the approach used, even states that are not directly measurable can be observed.

Today, most image sequence analysis systems used in robotics and computer vision for driver assistance are based on prediction and correction. This approach allows incorporating data from different sensors (sensor fusion) and the exploitation of continuity constraints in an optimal manner and delivers estimates with minimal noise.

Well known examples are the tracking of cars in front with estimation of the relative motion or the lane recognition. In these and many other cases, the knowledge obtained in the past is used to predict the position and the motion of relevant objects. This information is used to direct the attention to relevant image parts and to tune the image processing algorithms to the expected structures in order to increase robustness in real life scenes.

One problem of these tracking approaches is that the systems tend to see what they expect and run the risk to overlook transient events. For example, a lane tracker may follow an exit instead of preferring the correct lane. This problem can be overcome by the recently introduced multi-filter systems that are able to run different hypotheses in parallel. This is close to the principle of “Analysis by Synthesis”. Here, different hypotheses are set up and allowed to evolve over time. An appropriate judging based on the innovation errors decides which guess explains the situation best.

Modern Condensation Tracker and Particle Filter schemes can be looked upon as a further generalization of this prediction-verification scheme. In PF-estimation, each particle represents one hypothesis that is predicted in the Kalman sense, but the measurement and correction step is much more general. One advantage of these filters is that they can cope with multi-modal solutions.

## Analysis by Synthesis in 3d-Reconstruction from Images

Reinhard Koch

University of Kiel

rk@mip.informatik.uni-kiel.de

In this contribution we will exploit the interdependency between computer graphics and computer vision for realistic 3D surface reconstruction. Image-based rendering and 3D scene reconstruction from real image data are important research topics in computer vision and computer graphics. One goal of these methods is to render novel views of a scene from a set of given real camera views of a 3D scene. Depending on the method, different intermediate representations, like light fields or surface models, are produced to achieve realistic rendering of the novel views.

In image-based 3D reconstruction, a highly realistic 3D surface representation of the observed scene is built with the help of computer vision techniques. From this representation, novel views can be rendered using computer graphics methods. Hence, there is an interdependency between graphics and vision. The visual reconstruction of a 3D model from real images is a complicated inverse problem, while computer graphics will render novel images from the given 3D model in a straight forward way.

We can exploit the inverse properties between vision and graphics in an analysis by synthesis feedback loop. Given some coarse prior knowledge of the 3D object to be modeled, we can reconstruct a simplified model of the object using computer vision techniques. Then, the reconstructed model can be synthesised back into the original view by graphics rendering and compared to the real image data. The observed image differences will be fed back into the vision component to improve the object reconstruction.

References: Koch, R.: Dynamic 3D Scene Analysis through Synthesis Feedback Control. *IEEE Trans. Patt. Anal. Mach. Intell.*, Special issue on analysis and synthesis, Vol. 15 (6), pp. 556–568, June 1993.

Koch, R.: 3-D Modeling of Objects from Image Sequences. in: B. Jaehne, H. Haussecker, P. Geissler(eds.): *Handbook of Computer Vision and Applications*. Academic Press, 1999.

Evers-Senne, J.F., Koch, R.: Image Based Interactive Rendering with View Dependent Geometry. *Eurographics 2003*, Granada, Spain, September 2003.



## **Symposium 2**

### **Mid- and High-level Motion Analysis—From Image Features to Behavior**

Martin Giese and Heiko Neumann  
Universität Tübingen, Universität Ulm

**Mid- and High-Level Motion Analysis—From Image Features to Behavior**Martin A. Giese<sup>1</sup> and Heiko Neumann<sup>2</sup><sup>1</sup>University Clinic Tübingen, <sup>2</sup>University of Ulm

martin.giese@tuebingen.mpg.de

Much of the previous research on motion perception has focused on simple mechanisms for motion detection and integration, such as explaining motion coherence or the solution of the aperture problem. Recently, a vivid interest in the study of higher functions in motion perception has emerged and research has started to address questions like: How do attention and top-down effects influence the processing of motion information? How does, potentially learned, information about movements and shapes interact with the perception of motion? What are the processes that underlie the perception of motion of complex articulated objects and biological organisms? Another interesting question, which has attracted much attention, is whether the recognition of complex body movements is linked to the motor planning of movements, potentially exploiting a common representation for perception and action. The symposium tries to give an overview of several recent lines of research in mid- and high-level motion vision, presenting results from psychophysics, electrophysiology, imaging, and computational modelling.

The talk by Heiko Neumann will illustrate how a local integration of motion and form information can be accomplished by exploiting relatively simple physiologically plausible dynamic neural networks. By the appropriate combination of feed-forward and feedback connections, and appropriate connection between different cortical areas an accurate estimation of object motion in presence of occlusions and discontinuities of the optic flow field can be accomplished. Stefan Treue will present physiological data from monkey cortex that sheds some light in the neural basis of the modulation of motion processing by attention. Neural correlates of spatial and feature-based attentions are discussed. The contribution of Marc Grosjean will discuss psychophysical evidence demonstrating interference between execution and perception of different motor actions. Such evidence supports the idea of a common substrate for the encoding of events in perception and action. Ian Thornton will discuss how internal representations of static and dynamic stimuli interact in high-level perception, addressing problems like the perception of body and facial movements. Particularly, he will present psychophysical evidence for an important role of top-down influences in the processing of complex movement and form stimuli. The contribution of Martin Giese discusses experimental evidence from psychophysics and functional imaging that supports the hypothesis that the recognition of biological movements and actions is based on the learning of prototypical patterns. These results support a computational theory that assumes that action patterns are neurally encoded in terms of spatio-temporal templates for sequences form and optic flow patterns.

## **Features, Apertures and Occlusions: Determining Coherent Motion**

Heiko Neumann

University of Ulm

heiko.neumann@informatik.uni-ulm.de

Motion detection is locally ambiguous and prone to error since the measurements are constrained by, e.g., the aperture problem for extended luminance boundaries. Yet localized intrinsically two-dimensional structures often contribute robust features that can be robustly tracked over time. However, when several objects move they often partially occlude each other. In such cases, erroneous feature motions occur at junctions that are defined by intrinsic shape boundaries as well as some that are extrinsic to the shape.

We developed a neural model of motion detection and integration [1] that consists of two model areas (V1 and MT) each of which is subdivided into two layers of model neurons with gradual activation dynamics. At the first stage input activations are integrated over a neighborhood in the space-velocity map, while the second stage realizes a shunting center-surround inhibition to normalize localized activities. Feedback is signaled along recurrent modulatory MT-V1 connections to generate space-time reentry patterns delivered by coarse scale representations in area MT. Together with the normalization stage localized features (at corners) will be tracked over time. Also, by way of modulatory feedback, directions will be given competitive advantages that cohere with the disambiguated motion at the higher processing stage. Recurrent interaction leads to filling-in of disambiguated motion signals along extended boundaries, thus, solving the aperture problem [2]. It will be shown, in addition, how the resonant feed-forward and feedback interaction produces an inertial effect of motion capturing and how feedback from even higher cortical stages bias responses to generate attention effects.

When an object moves in front of another surface, however, the occlusion generates localized features which often signal wrong motion directions. How does the visual system segregate valuable from invaluable features while integrating ambiguous measures into a representation of coherent shape motions? For example, in barber pole stimuli that were flanked by monocular occluders the intrinsic terminator signals strongly bias the perceived motion and override opposing extrinsic terminator signals [3]. We suggest an excitatory V2/V3-MT mechanism that enhances motion responses in a directionally unspecific way in the neighborhood of an occlusion (T-junction configuration). Subsequent competition suppresses such erroneous feature motions. Consequently, MT feedback signals that remain strong enough will enhance more localized V1 motion responses to segregate around occlusions. This demonstrates that segmentation of motion signals at occlusions can be steered by interactions between the form and motion pathways to generate representations of independently moving forms.

[1] Bayerl and Neumann, *Neural Comp.* 16, 2004.

[2] Pack and Born, *Nature* 409, 2001.

[3] Pack et al., *J. of Neurosci.* 24, 2004.

**Attention and motion—Neural Mechanisms of Spatial and Feature-Based Influences**

Stefan Treue, Julio Martinez-Trujillo and Thilo Womelsdorf

German Primate Center Göttingen

treue@gwdg.de

The neural signature of visual attention is the selective modulation of cortical responses to visual stimuli based on their behavioral relevance. The last decade has greatly advanced our understanding of where in the visual cortex such modulation can be observed and what rules it follows. The talk will review single cell recording studies from the visual cortex of macaque monkeys trained to perform attentional tasks with moving stimuli to demonstrate how and when neuronal responses are influenced by spatial and feature-based attention.

The animals were trained to attend to one (the target) of several moving stimuli on a computer monitor and to respond to changes in the target, ignoring changes in the other stimuli (the distractors). At the beginning of each trial the animal was instructed as to which stimulus would be the target in the upcoming trial. Sensory stimulation within a given cell's receptive field was the same in all trials but which stimulus was designated the target varied. This allowed assessing the effect of different attentional conditions.

The results demonstrate a profound influence of attention on neuronal firing rates and demonstrate that attentional modulation is multiplicative, preserving the selectivity of individual neurons, but increasing the influence of attended and reducing the influence of unattended stimuli. They show that attentional modulation can be based on the attended location as well as the attended stimulus feature (i.e. motion direction). The data also show how non-multiplicative effects, such as an increased selectivity of population responses to attended stimuli and shifts in receptive field positions can be created from multiplicative modulations of neuronal responses.

These results demonstrate that attention has a powerful and systematic influence and that even at rather early stages of visual information processing neural signals reflect the behavioral relevance of stimuli.

## **Contrast Effects Between Concurrently Perceiving and Producing Events**

Marc Grosjean

MPI for Human Cognitive and Brain Sciences, München

`grosjean@psy.mpg.de`

The simultaneous perception and production of events has recently been shown to lead to a bi-directional suppression (as opposed to an assimilation) between what is perceived and what is produced. For example, producing a medium-amplitude movement trajectory while concurrently watching a large-amplitude motion trajectory leads to a reduction in size of the produced movement, whereas the size of the large-amplitude motion is perceived as larger than it really is [1]. These perception-action interactions were further investigated here by asking participants to produce movements on a graphics tablet while concurrently perceiving a functionally unrelated motion on a screen. Both the produced movements and the motions presented on the screen could vary in direction. As was expected, the results revealed that the perception of motion direction and the direction of movement production mutually shifted away from each other along the orientation dimension. However, in terms of production, the CE was actually preceded within a given action by an assimilation effect and was only obtained when the overlapping dimension between the perceived and produced events was task relevant. In terms of perception, the CE did not depend on the angular distance between the produced and perceived direction, but did depend on the categorization of the produced movement (e.g., right vs. left). Finally, when both types of CEs were obtained in the same task, their sizes were uncorrelated across participants. Taken together, these results suggest that the CE on production might be, at least in part, strategic in origin and that the source of the two types of CEs may actually differ. The implications of these findings for modeling within the framework of common coding and that of internal forward models will be discussed.

[1] Schubö, Aschersleben, & Prinz, 2001

**Cognition in Motion: Do we Represent Change-Over-Time?**

Ian M. Thornton

MPI for Biological Cybernetics, Tübingen

ian.thornton@tuebingen.mpg.de

In this talk I want to suggest that motion—or more generally change-over-time—are aspects of vision that have wide ranging implications for our understanding of brain function. An observer from another field of study might find such a claim surprising if they were to randomly sample a selection of recent vision research papers. While the perception of motion is a well established domain of research—as the existence of this symposium illustrates—it is nonetheless a specialist area, taught and studied in isolation from other topics in vision. While such specialisation occurs for many other topics as well, motion processing has another barrier to overcome. The problem is that vision science continues to be dominated by what I will call the “pictorial brain” assumption. Briefly stated this assumption is that visual processing begins at the level of some static retinal “image” and ends with representations that attempt to capture the static, spatial structure of the input. Central to this assumption is that processing operates on a series of “snapshots”. What is missing from such a view of visual processing is any consideration of time. I will argue that time is not a separable dimension of vision that can be added, post-hoc, to a collection of extracted features. Our concepts of visual features and in particular visual representations need to be modified to capture the temporal continuity that is so much a part of the physical environment in which we have evolved. Perhaps this caricature of “pictorial” processing described above is very far away from the way you think about vision. But it is worth reflecting for a moment on how you do think visual processing proceeds. What are the basic building blocks of vision? What is a feature? What is the nature of the input? Is it discrete or continuous? Does time, temporal continuity, change-over-time, play any role in your basic concepts? Part of the problem here is that pictorial metaphors are so embedded in the way we work, in the language we adopt and the tools we use that they are very rarely questioned. There can be no denying that the vast majority of experimental paradigms continue to present static stimuli, whether in the context of human, animal and machine observers. As the input is bound to partly determine the output, the use of such static stimuli can be seen as problematic. More dangerous, in my opinion, is the influence that such pictorial metaphors have on our hypotheses and models, how they constrain our thinking about vision and visual representations. The issues that I will raise in this talk are not new ones. Many of these concerns have been raised more than once during the relatively brief history of vision science. As yet, however, there has been no acceptable solution that has allowed time to be more generally integrated into our current models of visual processing. In this talk, I will also not be able to provide such a general solution. My goal here is rather to bring this issue back into the spotlight and to emphasise the importance of considering both space and time in vision. I will describe recent experiments from our group and others illustrating that the visual system often responds quite differently to dynamic rather than static input. In some domains at least, such differences often manifest themselves in terms of measurable performance advantages for dynamic stimuli.

## **Learning as Basis for the Visual Recognition of Actions**

Martin A. Giese

University Clinic Tübingen

`martin.giese@tuebingen.mpg.de`

Humans and animals can recognize biological movements with astonishing robustness and accuracy. The neural and computational mechanisms that underlie the recognition of such complex non-rigid motion patterns remain largely unclear. Motivated by a learning-based neural theory for the recognition of biological movements and actions, we conducted a series of experimental studies in order to provide direct evidence for a critical role of learning in biological motion recognition.

The underlying computational model [2] assumes that the visual recognition of biological movements and actions is based on learned neurally encoded templates for sequences of form and optic flow patterns. The model has been shown to be consistent with a variety of previous experimental results on biological motion perception. The model predicts the existence of an efficient visual learning process that encodes complex movement patterns independent of their biological relevance. In addition, the model predicts that learning of novel biological motion patterns should be associated with specific activity changes, in particular in motion-selective structures in the visual pathway.

Learning of biological motion was studied in a series of psychophysical experiments that required subjects to discriminate between motion morphs, which were generated by motion morphing by linearly combining prototypical movement trajectories (cf. [1]). It is demonstrated that subjects can learn novel biological motion patterns with a relatively small number of stimulus presentations. Consistent with the predictions from the learning-based neural theory, the learned representations are view-dependent, and the speed of the learning process does not depend on the biological relevance of the learned motion patterns.

Based on the same discrimination task, an fMRI study was designed that exploits an event-related adaptation paradigm for demonstrating correlates of the visual learning process along the visual processing stream. We found systematic changes of the adaptation of the BOLD signal that seem consistent with the predictions from the neural theory: For artificial patterns biological motion-selective areas (e.g. the STS) show an emerging selectivity for the novel patterns after discrimination training. For novel movement stimuli that are similar to real human movements we find a gradual selectivity before training, which gets enhanced after discrimination training.

These direct demonstrations of learning-induced changes in the recognition of biological motion together with the fact that a learning-based computational theory explains many classical results suggest that learning might be a fundamental clue for the understanding of motion recognition.

- [1] Giese M A and Lappe M (2002): Measurement of generalization fields for the recognition of biological motion. *Vision Research*, 38, 1847–1858.
- [2] Giese M A and Poggio T (2003): Neural mechanisms for the recognition of biological movements and action. *Nature Reviews Neuroscience* 4, 179–192.

Supported by the Deutsche Volkswagenstiftung, the DFG (SFB 550), HFSP, and the Hermann Lilly Schilling Foundation.



## **Symposium 3**

### **Face Processing of Familiar Faces**

Claus-Christian Carbon  
Freie Universität Berlin

**Face Processing of Familiar Faces**

Claus-Christian Carbon

Free University of Berlin

ccc@experimental-psychology.de

The ability to recognize and identify familiar faces is an essential skill for everyday life. Six talks from special fields of face research investigate which specific mechanisms are needed to successfully process familiar faces. Carbon tests the critical facial information which is needed to recognize familiar faces using veridicality decision tasks. Schwaninger demonstrates the specific importance of configural and componential information for identifying three-dimensional faces. The approaches of Schweinberger and Herzmann test a specific EEG component called N250r which is attributed to familiarity and repetitions effects of faces. Duchaine investigates face processing under extreme conditions. With the data of a person who suffers from prosopagnosia, which is a specific cognitive impairment of processing faces, Duchaine tests several hypotheses about the processing of faces. The last contribution demonstrates that face processing develops over a life span. Congenitally existing performances to recognize faces are fine-tuned and optimized over a long period of time.

**Fame or Fake? Critical Information Needed for Decisions on Facial Veridicality**

Claus-Christian Carbon

Free University of Berlin

ccc@experimental-psychology.de

The identification of a face is a key competence in everyday life. Although all faces possess the same first-order relationship between cardinal features (eyes, mouth, nose) and are therefore highly homogenous, humans are capable of perceiving rather subtle configural changes in faces. In the present study it was investigated which critical facial information is needed to decide whether a facial image is veridical or configurally altered. Participants had to evaluate stimuli for which the available facial information was systematically reduced. Based on images with the full facial information (full), additional versions with only the cardinal features (parts), only the parts plus the contour of the hairline (contour), the parts plus a standard oval indicating a kind of average facial outline (oval), and with the parts plus the texture of the inner section of the face (texture) were constructed. In the veridicality decision task, based on matching these versions with an original face version, only full and texture versions lead to accurate decisions, whereas contour faces showed medium accuracy. For parts and oval information, the performance was near chance level. Thus, critical configural information for faces is retrieved not only from the simple spatial relationship between the cardinal features (parts), but from the whole inner configuration including the texture and the parts.

**Component, Configural and Temporal Routes in View-Based Face Recognition**Adrian Schwaninger,<sup>1</sup> Christian Wallraven<sup>2</sup> and Heinrich H. Bülthoff<sup>2</sup><sup>1</sup>University of Zürich & MPI for Biological Cybernetics, Tübingen,<sup>2</sup>MPI for Biological Cybernetics, Tübingen

aschwan@allgpsy.unizh.ch

Face recognition relies on detecting subtle differences between facial parts (components) and their spatial relations, so-called configural information (for a recent review see [2]). Psychophysical experiments are reported, in which a new method was used to test component and configural processing separately [3]. An important role for both types of information was revealed in unfamiliar as well as familiar face recognition and only quantitative but no qualitative differences were found. The results could be modeled using a computational framework based on key-frames in which the component and configural route were explicitly implemented [4]. In a new series of experiments, the Inter-Extra-Ortho paradigm by Bülthoff and Edelman [1] was applied to investigate face recognition across viewpoint. Systematic effects of viewpoint were found that are consistent with local view interpolation schemes and the proposed framework based on key-frames [5]. In a recent study the role of component and configural information in view-based face recognition was investigated by combining the Inter-Extra-Ortho paradigm with the psychophysical method for selective testing of processing component and configural information. Interestingly, processing of parts was more dependent on viewpoint than processing of configural information. The same result was found for the computational model mentioned above, in which the component and configural processing route have been implemented. The psychophysical and computational results are also discussed in respect to a third route to recognition in which different views of a face are temporally associated [6].

- [1] Bülthoff, H.H., & Edelman, S. (1992). Psychophysical support for a two-dimensional view interpolation theory of object recognition. *Proceedings of the National Academy of Sciences of the United States of America*, 89, 60–64.
- [2] Schwaninger, A., Carbon, C.C., & Leder, H. (2003). Expert face processing: Specialization and constraints. In G. Schwarzer & H. Leder, *Development of face processing* (pp. 81–97), Göttingen: Hogrefe.
- [3] Schwaninger, A., Lobmaier, J., Collishaw, S.M. (2002). Component and configural information in face recognition. *Lectures Notes in Computer Science*, 2525, 643–650.
- [4] Wallraven, C., Schwaninger, A., & Bülthoff, H.H. (2004). Learning from humans: computational modeling of face recognition. *Early cognitive vision workshop, ECVW 2004*.
- [5] Wallraven, C., Schwaninger, A., Schuhmacher, S., & Bülthoff, H.H. (2002). View-Based Recognition of Faces in Man and Machine: Re-visiting Inter-Extra-Ortho. *Lectures Notes in Computer Science*, 2525, 651–660.
- [6] Wallis, G.M., & Bülthoff, H.H. (2001). Effects of temporal association on recognition memory. *PNAS* 98, 4800–4804.

**Neural Representations of Faces**

Stefan Schweinberger

University of Glasgow

s.schweinberger@psy.gla.ac.uk

Cognitive models of face recognition make a distinction between several levels of representation (e.g., pictorial, abstractive, semantic). Recordings of event-related brain potentials (ERPs) may specify neural correlates of these representations in the spatiotemporal domain. We study repetition priming and associative priming. Repetition effects for faces act on multiple levels of representation and yield a modulation of an N250r(repetition) response (200–350 ms) over inferior temporal areas predominantly over the right hemisphere, which presumably is generated in fusiform cortex, and an N400 response (300–600 ms) over central-parietal areas. In contrast, the N170 does not exhibit repetition effects. While the N250r response is absent in associative priming, an N400 effect is preserved. Moreover, the N250r seems to be a selective brain response for faces when compared with other visually homogeneous categories (ape faces, inverted faces, cars) - although all those stimuli elicit a clear N170. Based on these and other results, I will argue that the relationship between different ERP components and different representational levels in face recognition is beginning to emerge. Specifically, pictorial/structural processing is reflected in the N170, visual representations for recognition relate to the N250r, and the N400 reflects semantic representations for people.

**The Early Repetition Effect: Reliability and Methodological Influences**

Grit Herzmann

Humboldt University Berlin

`grit.herzmann@cms.hu-berlin.de`

When a familiar face is seen twice in a repetition priming paradigm an increased negativity at temporal-parietal and a increased positivity at frontal electrodes around 250 ms after stimulus onset is seen in event-related potentials when compared to unrepeatd familiar faces. This effect is called the early repetition effect, N250r or ERE. It increases in amplitude from unfamiliar, to publicly familiar and is largest for personally familiar faces. The aim of the present study was to assess whether newly learned previously unfamiliar faces would elicit an ERE and to compare it to repetition effects to newly learned names under comparable conditions. 40 unfamiliar faces and unfamiliar person names were learned separately to criterion without any other context information. One week later recognition of these stimuli was tested. The obtained ERE for faces showed the same characteristics as the ERE previously observed to personally and publicly familiar faces. Some differences emerged between the EREs for faces and names, indicating domain specificities of the underlying processes. Preliminary analyses indicate promising retest reliabilities at least for the ERE.

**Prosopagnosia as an Impairment to Face-Specific Mechanisms: Elimination of the Testable Domain-General Explanations**

Brad Duchaine

Harvard University

brad@wjh.harvard.edu

For more than 35 years, researchers have debated whether face recognition involves face-specific mechanisms or domain-general mechanisms that are used for faces and other types of objects. Evidence has come from many sources, but studies of prosopagnosic subjects have provided some of the most powerful evidence in favor of face-specific mechanisms. However, there are many explanations for prosopagnosia other than the face-specific explanation, and no studies of prosopagnosics have eliminated all of the alternative explanations in one individual so the results do not conclusively support a particular explanation. Here we address all of the existing testable explanations of prosopagnosia in a developmental prosopagnosic. Edward has had lifelong difficulties with face perception and clear impairments with a variety of face processing abilities. Edward performed normally with nearly every non-face test he was assessed with, and his results eliminated all of the testable explanations. As a result, we argue that Edward's prosopagnosia is accounted for by a failure to develop normal face-specific mechanisms. Because his only demonstrated impairments are with faces, his case also indicates that the development of face perception mechanisms involves developmental mechanisms different from those used to build other visual recognition mechanisms.

**Development of Face Recognition: how Important is Social Information?**

Cornelia Groß and Gudrun Schwarzer

University of Gießen

cornelia.gross@psychol.uni-giessen.de

Viewing a face implicates two fundamental psychological processes. The first process concerns the recognition of the identity of a face and the second one implies the perception of information that facilitates social communication. For identification, the invariant structure of a face has to be extracted from fast changing social cues (orientation, emotional expression, facial speech) and has to be compared to existing representations. For the recognition of social cues those easily changeable features like the orientation of a face or emotional expression have to be interpreted meaningfully. Until now, the development of face recognition and the development of the perception of social cues have been investigated in different lines of research. Therefore, the goal of the present contribution is to investigate to what extent both processes interact with each other. The first series of experiments investigated whether 7 month old infants succeed in extracting the structure of a face from different orientations and whether this ability is influenced by the emotional expression of the face. The results showed a significant influence of emotional expression such that infants were not able to extract the structure of a face from different orientations when the faces showed neutral facial expressions. In contrast, infants were well able to do so when the faces showed a positive or a negative facial expression. The second series of experiments examined how structural and social information in a face interact with each other in older children. Five ten year old children and adults were asked to classify faces according to facial identity, emotional expression and facial speech. Reaction times were measured and analyzed. The results revealed an independent processing of facial identity whereas processing of social information like emotional expression or facial speech was influenced by faces' identity. The results of both series of experiments will be discussed with regard to different neural systems activated by structural and social information of a face.

## **Attention and Awareness**

## **Locus of the Temporal Uncertainty Effect: Evidence from the Psychological Refractory Period**

Karin Bausenhart, Bettina Rolke and Rolf Ulrich

Universität Tübingen

karinbausenhart@web.de

Reaction time (RT) is especially short if a warning signal precedes the imperative stimulus, because the warning signal reduces the temporal uncertainty about the stimulus appearance. This temporal uncertainty effect depends on the time interval (i.e. the foreperiod duration) between warning signal and imperative stimulus. In addition, the distribution of the foreperiod duration modulates this effect. Many studies have located the temporal uncertainty effect within late stages of RT, though there is also some evidence that earlier stages might be involved. We employed a psychological refractory period paradigm to investigate further the locus of this effect. Within this dual task paradigm, RT2 to a second stimulus is prolonged, in case both stimuli are separated by a short stimulus onset asynchrony (SOA). Theoretical explanations of this increase in RT2 proceed from a central bottleneck process, which can operate only on one stimulus at a given time. According to this bottleneck model, a temporal uncertainty effect for RT1 to the first stimulus should carry over to RT2 at short SOAs, if the foreperiod manipulation affects early processing stages. If this manipulation affects, however, late stages, temporal uncertainty should not influence RT2. In our experiment, subjects performed two temporally overlapping RT tasks to visual stimuli (SOA: 50, 400, 1000 ms). A visual warning signal preceded the first stimulus and the foreperiod interval was 400, 800, or 1600 ms. Foreperiod durations were either varied randomly across the trials in a single block (variable condition) or kept constant within a single block of trials (constant condition). Foreperiod duration produced a reliable effect on RT1 in the variable condition. A carry-over of this effect on RT2 was found at the shortest SOA and this effect diminished at longer SOAs. This interaction suggests that temporal uncertainty operates at early stages within the processing chain between stimulus and response.

**Contextual Modulation Outside of Awareness**

Colin W. G. Clifford and Justin A. Harris

The University of Sydney, Australia

colinc@psych.usyd.edu.au

Contextual effects are ubiquitous in vision and reveal fundamental principles of sensory coding. Here, we demonstrate that an oriented surround grating can affect the perceived orientation of a central test grating even when backward masking of the surround prevents its orientation from being consciously perceived. The effect survives introduction of a gap between test and surround of over a degree even under masking, suggesting either that contextual information can effectively propagate across early visual cortex in the absence of awareness of the signaled context or that it can proceed undetected to higher-levels of processing at which such horizontal propagation may not be necessary. The effect under masking also shows partial interocular transfer, demonstrating processing of orientation by binocular neurons in visual cortex in the absence of conscious orientation perception. This pattern of results is consistent with the suggestion that simultaneous orientation contrast is mediated at multiple levels of the visual processing hierarchy and supports the view that propagation of signals to, and possibly back from, higher visual areas is necessary for conscious perception.

This research was supported by an Australian Research Council Discovery Project and Queen Elizabeth II Fellowship to CC.

## **A Benefit of Attentional Stimulus Processing? Temporal Discrimination Within a Spatial Cueing Paradigm**

Angela Dinkelbach, Bettina Rolke and Rolf Ulrich

Universität Tübingen

angela.dinkelbach@student.uni-tuebingen.de

It is well documented that visual spatial attention enhances performance in various tasks. Surprisingly, however, Yeshurun and Levy [1] reported a negative effect of attention on a temporal discrimination task. In a spatial cueing paradigm, discrimination performance of a temporal gap within a dot was impaired if attention was oriented towards the spatial position of the dot (valid condition) compared to a neutral condition, in which attention was not oriented to a specific location (neutral condition).

Because this surprising effect is important to constrain theories on visual attention, we further investigated this effect. Specifically, we assessed whether this negative attentional effect might reflect a confounding rather than a genuine attentional effect. Yeshurun and Levy [1] employed a rather unusual neutral cue, which was a long horizontal bar extending over the whole screen. By contrast, their valid cue was short and appeared above the stimulus location. Because both cues differed in their spatial extension, it might be possible that the negative effect reported by Yeshurun and Levy is due to a stronger visual activation elicited by the neutral cue. This stronger activation, however, might enhance stimulus processing and thus have contributed to the negative effect reported by these authors. In order to test this alternative account, we used a standard cueing paradigm with valid and invalid trials. In contrast to Yeshurun and Levy, however, the cue in both conditions was physically identical. The invalid cue (25% of all trials) oriented attention away from the stimulus location, whereas the valid cue (75% of all trials) directed attention to the location.

Temporal discrimination performance was better in the invalid than in the valid condition. Thus, our results strengthen the results of Yeshurun and Levy [1] and rule out an alternative interpretation of their results. In conclusion, attention seems to impair temporal discrimination and this surprising attentional effect appears to be a robust phenomenon.

- [1] Yeshurun, Y. & Levy, L. (2003). Transient spatial attention degrades temporal resolution. *Psychological Science*, 14(3), 225–231.

## Deficits in the Control of Visual Attention in Patients with Cerebellar Lesions or with Parkinson's Disease

Thomas Eggert,<sup>1</sup> Susana Casana-Perez<sup>1</sup> and Andreas Straube<sup>2</sup>

<sup>1</sup>Dept. of Neurology, Ludwig-Maximilians-University, Munich,

<sup>2</sup>Dept. of Neurology, Ludwig-Maximilians-University, Munich

eggert@nefo.med.uni-muenchen.de

Previous studies showed that immediately before saccade execution, the character discrimination in the visual periphery is improved at the location of a saccade target. This effect can be used to measure control of visual attention. Moreover, healthy controls are not able to focus their visual attention at one location while simultaneously preparing a saccade to another location [1]. These results suggest a very close coupling between saccade preparation and control of attention.

This study investigates whether attention mechanisms are involved in oculomotor disorders of patients with cerebellar lesions (9 subjects) or with parkinson's disease (9 subjects). Patients were compared with a healthy control group (10 subjects). Eye movements were recorded with a differential infrared reflexion device (IRIS). Both patient groups showed impaired saccadic accuracy and increased saccade latencies. Saccade dynamics, quantified by the peak velocity were almost normal in both patient groups.

Control of attention was assessed by measuring the cue effect on the performance in a two alternative character discrimination task performed in the near visual periphery (4 deg eccentricity). In a 2 x 2 design two motor tasks (fixation/ saccade to the cued location) were combined with two cue types, a central symbolic cue or a peripheral transient cue. The control group showed higher percentage of correct responses when the discrimination character was presented at the cued location compared to the non-cued location. This difference was significant for all conditions except for the fixation task with the peripheral transient cue. In contrast, patients with cerebellar lesions did not show significant cue effects in any of the conditions although the average percentage of correct responses across all conditions (66%) did not differ significantly from that of the controls (71%). Thus, even with saccades executed to the cued target location, cue effects on discrimination performance did not reach significance for cerebellar patients.

The number of correct responses of Parkinson patients was very close to chance level (56%). These patients showed a cue effect only under the saccade condition with the peripheral transient cue.

These data suggest that control of visual attention is impaired in cerebellar and in parkinson patients. This deficit did not only depend on whether a saccade towards the target was under preparation or not. This supports the idea that, in contrast to most findings in healthy subjects, patients with central oculomotor deficits are able to generate saccades that are not obligatory linked to a concomitant shift of spatial visual attention.

[1] Deubel, H. and Schneider, W.X. (1996): *Vis. Res.* 36:12, 1827–1837

This study has been supported by the DFG- Graduiertenkolleg (DFG-GRK-267)

## Natural Semi-Logarithmic Decision Making and the Fechner's Paradox

Vitaly V. Gavrik

BRYG, Köln

`gavrik@arcor.de`

The semi-logarithmic Fechner's law faces the apparent paradox of infinite negative reaction while its stimulus is tending to zero [1]. The Stevens' power law [2] that is commonly used to avoid the difficulty is approximate. The semi-logarithmic functions were accurately measured [3] in the experiments on the single-stimulus decision making or their entire set was exhaustively calculated from the over-error spread of multistimulus decision data [4] in the psychophysics of image naturalness for all the variety of scene images and optical properties considered [5,6]. A feedback process to compare a varying input with its memorized range should result namely in the kind of decision-making function observed [7]. Nevertheless, the latter continued seemingly to face the Fechner's paradox.

A threshold break has been observed on the experimental curves along with another break to the constant reaction values when the stimulus lay within its high-reliability range memorized by the life experience. Our tests have shown every stimulus to be perceived in the relation to its threshold. A threshold stimulus was not perceptible yet and the logarithm of its relative value was zero. The reaction was absent and the function was zero also below the threshold. The roles played by the plateaus and the semi-logarithmic portions differed fundamentally. The current perception appeared mainly to occur within the high-reliability plateau. The comparison examined whether the input lay within its tolerances for various objects. A set of high-reliability inputs should produce a familiar mental image or simply a judgment that they had come from a known object and can draw no further attention. An input within the zero range excludes the object according to the possibility multiplication rule. The fast comparison should be performed in many neural channels simultaneously. It has take as little as dozens to hundreds of milliseconds.

The plateau length has both reflected the stimulus statistics and played the detector role for the attention dominants including unknown objects and unusual variations in the properties of familiar objects. The dominants have short plateaus whose boundaries are readily passed on to the semi-logarithmic portions when the simple comparison above becomes impossible. The inputs draw active attention and undergo the comparison for a judgment to be made on fuzzy-logic possibility that the input could come from a definite object. The slower process has taken as much as dozens to hundreds of seconds. It should underlie the further plateau changes and the memorization of new objects.

- [1] Fechner, G.T. (1890): *Z.f.Psychol.* 1, 29–46, 109–120.
- [2] Stevens, S.S. (1961): *Science* 133, 80–86.
- [3] Gavrik, V.V. and Sokolova, I.M. (1990): *J. Optical Technol.* 57, 496–498; Gavrik, V.V. (1991): *Ibid.* 58, 688–692; Gavrik, V.V. and Kirillova, L.N. (1997): *Ibid.* 64, 123–126.
- [4] Gavrik, V.V. (1999): *Proc. Intl. Joint Congr. on Neural Networks*, Washington, DC, IEEE, 2715–2718.
- [5] Gavrik, V.V. (1995): *J. Optical Technol.* 62, 232–237.
- [6] Gavrik, V.V. (1998): *Proc. SPIE* 3409, 52–63.
- [7] Gavrik, V.V. (2001): *Beitr. 4. Tuebinger Wahrnehmungskonf., Buelthof, H.H. et al., eds., MPI for Biological Cybernetics*, 171–172.

## Attention and Temporal Discrimination: Does the Type of Attentional Orienting Matter?

Elisabeth Hein, Bettina Rolke and Rolf Ulrich

Universität Tübingen

`elisabeth.hein@uni-tuebingen.de`

Yeshurun and Levy [1] have claimed that visual spatial attention impairs temporal discrimination. They employed an exogenous cue to direct attention to the stimulus location or provided a neutral cue in a control condition, which provided no advance information about the stimulus location. Subjects had to detect a single temporal gap within a stimulus. Surprisingly, discrimination performance was impaired when attention was directed to the target location compared to the neutral condition. In two experiments we investigated whether this surprising finding generalizes to another perceptual task and to another form of attentional orienting. In both experiments, subjects judged the temporal order of two adjacent dots. In the first experiment, attention was directed automatically by an uninformative exogenous cue. In the second experiment, attention was oriented voluntarily with an endogenous arrow-cue. We compared discrimination performance in valid trials (i.e. when attention was shifted toward the target location) and in invalid trials (i.e. when attention was shifted toward the non-target location). In agreement with the negative cueing effect reported by Yeshurun and Levy [1], discrimination performance of temporal order was impaired in the valid condition when attention was shifted with an exogenous cue. By contrast, a positive cueing effect was obtained when attention was oriented by an endogenous cue. In conclusion then, the type of attentional orienting matters: Automatically oriented attention impairs temporal order discrimination, whereas voluntary oriented attention enhances discrimination performance of temporal order judgment.

[1] Yeshurun, Y. and Levy, L. (2003): *Psychol Sci.* 14, 225–31

Graduiertenkolleg Kognitive Neurobiologie

## **Eye-Movements and Behavioral Responses to Threatening and Non-Threatening Stimuli During Visual Search in Phobic and Non-Phobic Subjects**

Wolfgang H. R. Miltner, Silke Krieschel, Holger Hecht, Ralf Trippe and Thomas Weiss

Department of Biological and Clinical Psychology, Friedrich Schiller University

wolfgang.miltner@uni-jena.de

Many recent studies have shown that the attention of anxious and phobic subjects is biased by threat. However, it still remains unsolved whether this bias is present during conditions of voluntary and involuntary attention. The present studies were aimed to address this question by means of two visual target-search experiments where spider-phobic and non-phobic subjects searched for a feared/fear-relevant or neutral target (mushroom) presented in visual matrices of neutral objects (flowers)[1]. For phobic subjects spiders served as feared stimuli, while for non-phobic subjects spiders were defined as fear-relevant stimuli. During half of the matrices depicting a neutral target a spider was added as distractor while during half of the matrices presenting a spider target a neutral mushroom was used as distractor. Manual responses to targets were recorded, along with eye-movements while subjects were scanning the matrices. Key-press responses of all subjects were faster to neutral than to feared or fear-relevant targets. However, when neutral targets were presented together with a feared distractor, phobic subjects showed significant delays of responses not observed in control subjects. The eye-movement data suggest that the eyes of spider-phobic subjects were automatically pulled to these feared distractors before fixating the neutral target. This effect was absent in control subjects. In a third experiment, subgroups of these subjects were exposed to matrices depicting only flowers (homogenous matrices) or to matrices presenting a feared/fear-relevant (spider) or fear-irrelevant picture (mushroom) against a background of neutral pictures (flowers) (inhomogeneous matrices). Subjects were requested to indicate by pressing a key as fast as possible whether matrices were composed of homogenous or inhomogeneous stimuli. Again, fear-irrelevant, neutral discrepant stimuli were found more quickly than feared ones with no significant differences in search times or eye-movements between phobic subjects and controls. Results indicate that threat seems to capture the attention of phobic subjects only when it is part of the background to which subjects are not explicitly requested to attend.

[1] Miltner, W.H.R. et al. (2004). *Emotion*, 4, 4, 323–339

Research was supported by the Deutsche Forschungsgemeinschaft (Mi-165/6–1/2)

**Temporal Discrimination Within the Attentional Blink**

Bettina Rolke, Karin Bausenhardt, Elisabeth Hein and Rolf Ulrich

Universität Tübingen

bettina.rolke@uni-tuebingen.de

Recent studies suggest a dissociating effect of visual spatial attention on the discrimination performance of figural and temporal tasks. Within these studies, attention was oriented by exogenous spatial cues. Discrimination performance for a figural task (detecting a gap within a Landolt-square) was enhanced in the cued compared to a neutral condition. In contrast, temporal discrimination (detecting a temporal gap within a dot) was impaired in the cued condition. The aim of the present study was to investigate whether this dissociating effect is limited to visual spatial attention or generalizes to another type of attentional manipulation. We investigated the effect of attention within the rapid serial visual presentation task (RSVPT). If subjects process a target stimulus in this task, other stimuli which appear shortly after this target are lost. Theoretically, this loss of stimuli has been attributed to the limited attentional resources of our information processing system: Within the time window of attentive processing of one stimulus, no resources are left for the attentive processing of other stimuli. Consequently, these other stimuli are lost in this time window (attentional blink, AB). In two experiments, subjects performed either a figural discrimination or a temporal discrimination task on stimuli presented within a variable stimulus onset asynchrony (SOA) to a first target. We expect an AB for the figural task. If the impairment of temporal discrimination is a general phenomenon of visual attention, temporal discrimination performance should be enhanced during the AB time window. Contrary to this prediction, the results of our experiments revealed an AB for both tasks. Thus, the negative effect of attention on temporal discrimination, which is present in visual spatial attention tasks cannot be generalized to the AB. This finding suggests that the attentional mechanism underlying visual spatial attention differs from the one underlying the AB.

## Task Preparation in Task Switching—Insights from an ERP Study

Jessica Saenger and Edmund Wascher

MPI for Human Cognitive and Brain Sciences

`jessica.saenger@cbs.mpg.de`

The task switching paradigm is widely used to focus on the role of executive processes, examining costs and benefits in performance as subjects alternate between tasks or repeat the same task from trial to trial. Typically, reaction time is faster and accuracy is higher when subjects repeat the same task than when they alternate between tasks. This raises the question of how these additional “switch costs” are generated. Foreperiods that can be used for task preparation result in a decrease of the switch costs [1,2]. However, additional time for response preparation is available “residual switch costs” still remain. ERP and neuroimaging studies have shown that both, parietal and frontal cortical regions are involved in the processes taking place in task switching but the underlying neurophysiological mechanisms are still unclear. In two experiments we used a cuing paradigm, where subjects had to perform two numerical judgment tasks. They had to decide if a given target (a digit ranging from 1 to 9, excluding 5) is more or less than 5 in one task or if a given target is odd or even in the other task. The current task was indicated by a cue (square or diamond), which were presented either 100 or 1000 ms before the target. Participants had to respond by a left or right keypress. By means of event-related potentials (ERP), measured from 60 scalp electrodes, we investigated the relation between “switch costs” and “residual switch costs”. The present ERP data showed, that for short preparation times “switch costs” are clearly reflected in a frontocentral N2 effect, where a more efficient allocation of attention takes place and a high effort of control is necessary. Whereas, “residual switch costs” were reflected in a late centroparietal positive component, resembling an additional component in the P3 complex. Furthermore this P3 component can be divided into two functional subcomponents. An early task switching effect, which is followed by a slow but sustained task repetition effect. Thus, the present data indicate that we searched out qualitatively different neurophysiological mechanisms (N2 and P3) for the different task preparation times (100 and 1000 ms) at different brain areas (frontal and parietal) that are involved in task switching. Due to this we assume that the observed “residual switch costs” are an effect of task-repetition benefit and that persisting activation of response rules might play a major role task switching.

- [1] Meiran, N. (1996). *Journal of Experimental Psychology: Learning, Memory & Cognition*, 22: 1423–1442.
- [2] Meiran, N., Chorev, Z. & Sapir, A. (2000). *Cognitive Psychology*, 41: 211–253.

## The N2pc as a Correlate of Recurrent Processes in Visual Areas

Andrea Schankin and Edmund Wascher

MPI for Human Cognitive and Brain Sciences

`schankin@cbs.mpg.de`

Changes between two successively presented pictures are hard to detect when their presentation is interrupted by a short blank. This phenomenon is called change blindness. It is assumed that changes are detected only when attention is directed to the location of the change. The N2pc component was identified as a correlate of attention in visual search tasks. It is characterised by a stronger contralateral than ipsilateral activation in relation to the position of a stimulus. Until now, the discussion holds on what attentional process is reflected by the N2pc: an attentional filter process, the selection of a target stimulus or the shift of attention to a task-relevant location. Because the generator of the N2pc is assumed to be located in the ventral occipito-temporal cortex, it is also probable that the N2pc component reflects recurrent processes from higher to lower visual areas. In several experiments, a matrix of 9 by 9 different coloured dots were presented. In each case, blue dots formed the capital letter “M” or “W”, respectively, that only differed in four dots in the centre of the matrix. In order to control the allocation of attention, the first task of the subjects was to discriminate between both letters. The presentation of the matrix was interrupted by a short blank. During the blank one of the dots could change either its colour (experiment 1) or its luminance (experiments 2 and 3). The second task of the subjects was answering the question whether (experiments 1 and 2) or where (experiment 3) they had seen any change. The event-related potential (ERP) was conducted at 64 electrodes including 4 electrodes for controlling eye movements. In all experiments, an N2pc component could be found for detected changes. Additionally, a smaller but clear N2pc component occurred only for undetected luminance changes. Probably, the colour information is processed only by the parvocellular visual pathway, whereas the luminance information is processed by both the magnocellular and parvocellular stream that results in a small N2pc component. No N2pc component could be observed for reported but not presented changes. Thus, the N2pc probably reflects recurrent processes from higher to earlier visual areas in both visual pathways. However, only recurrent processes in the ventral visual pathway results in awareness.

## Visual Selection of Perceptually Completed Objects: The Advantage of Modal Completion

Branka Spehar and Rowena Beecham

The University of New South Wales, Sydney, Australia

b.spehar@unsw.edu.au

When attention is directed only to a part of an object, other parts of that object enjoy an attentional advantage compared to equally distant image locations that do not belong to that object [1]. Although it has been shown that this object-based attentional advantage extends also to objects that require perceptual completion [2], it is still unclear whether different types of perceptual completion are equally effective in triggering object-based attentional selection. Direct comparisons are difficult as various types of completion are typically specified by different local image cues, making it hard to attribute the observed differences to either the initial processing of different local image cues or to the differential dynamics of perceptual interpolation. This study compared the object-specific attentional advantage between modally- and amodally-completed objects specified by image cues equated in local geometry and luminance contrast magnitude. Under these conditions, displays consistent with modal completion, but not those consistent with amodal completion, displayed the same-object advantage, highlighting the advantage of modal compared to amodal completion in the process of attentional selection.

- [1] Egly, R., Driver, J., & Rafal, R.D. (1994). *Journal of Experimental Psychology: General*, 123, 161–177.
- [2] Moore, C.M., Yantis, S., & Vaughan, B. (1998). *Psychological Science*, 9, 104–110.

## **Auditory Perception**

**Prosody and Lexicality in Emotional Speech Perception**

Susanne Dietrich,<sup>1</sup> Anja Ischebeck,<sup>1</sup> Kai Alter<sup>2</sup> and Hermann Ackermann<sup>3</sup>

<sup>1</sup>MPI for Human Cognitive and Brain Sciences Leipzig,

<sup>2</sup>School of Neurology, Neurobiology & Psychiatry, Newcastle, UK,

<sup>3</sup>Department of Neurology, University of Tübingen

dietrich@cbs.mpg.de

It is still under debate how the brain disentangle lexical and prosodic information. We therefore investigated verbal interjections with an affective and neutral intonation in an functional magnetic resonance imaging study. Verbal interjections differed with regard to the complexity of their segmental content (lexicality). For instance, a lexical interjection such as ‘hurra’ can be understood as a happy interjection even if it is neutrally spoken. In the case of an interjection which consists of a vowel only, such as ‘a’, however, only prosody provides the affective meaning. For interjections with clear affective content, provided either by its segmental or its suprasegmental (prosodic) structure, activations were observed bilaterally in the anterior part of the superior temporal gyrus. On the other hand, neutrally spoken interjections with a high degree of lexicality (segmental information) led to activations in more posterior parts of the superior temporal region (STR), and anterior STR close to Heschl’s gyrus. Activations in the anterior part of the superior temporal region have been associated with intelligibility or the integration of given information into a context. The stronger activations in the posterior parts of the superior temporal gyrus for the segmentally complex stimuli indicate that this region could be specialized to process complex speech signals such as sentences and words rather than single vowels.

## Timing Precision in Synchronization Tapping and Effects of Training

Timo Fischer,<sup>1</sup> Manfred Nusseck<sup>2</sup> and Herbert Bruhn<sup>3</sup>

<sup>1</sup>Universität Hamburg, <sup>2</sup>MPI for Biological Cybernetics, Tübingen, <sup>3</sup>Universität Flensburg

timo@fischart.de

The present experiment adds another study to the experiments on negative synchronization errors: Asked to follow a leading signal by tapping, subjects tend to tap 20 to 80 ms before the physical onset of the leading signal. Little is known about the effects of training over long tapping series, although the relevance for musical performance is obvious: In concerts musical experts (e.g. drummers) don't show any synchronization error.

Two experiments were conducted with music students as subjects, who had to follow a bongo sound with an isochronous inter-stimulus-interval of  $ISI = 500$  ms by their tapping. Each experiment consisted of four trials, lasting 1, 3, 5 and 10 minutes (about 2280 taps in each session). To investigate the effect of teaching, 8 of 13 subjects got a detailed verbal and visual feedback after each trial. Compared with the control group subjects reduced the negative synchronization error immediately. The overall effect was significant, although the subjects began all trials after the feedback again with a negative error (back-lash). The negativity was every time reduced immediately after a few taps.

In the second experiment four subjects without feedback during the four trials were asked to repeat all trials in a second session one week later. They got a feedback only after having finished the 10-minute-trial within the first week. Also these subjects showed a reduced negative synchronization error as one week before. The back-lash effect could be observed, but it was not as impressing as in the first experiment.

The results raise the question, whether the negative synchronization error might be endogenously generated reactions which disappear by being overlearnt. Further research will be conducted.

## Attention Driven Auditory Display

Thorsten Mahler, Pierre Bayerl, Michael Weber and Heiko Neumann

Universität Ulm

thorsten.mahler@informatik.uni-ulm.de

The interdisciplinary field of image sonification aims at the transformation of images to auditory signals. The basic question is how information from the image domain can be transformed to the sound domain to produce an intuitively comprehensive audio signal. Three classes of sonification approaches have previously been proposed [1,2]: First, parameter mapping, where image data (e.g. position, luminance) is directly mapped to the parameters of the sound signal (e.g. amplitude, frequency, duration); second, model-based sonification, where virtual sound objects (e.g. instruments) are controlled by the visual input; third, auditory scene generation, where the input data is utilized to generate an auditory scene. The starting point for our work was the sonification system “vOICe” introduced by Meijer [3]. This system is a variation of the parameter mapping approach, where image luminance steers the sound amplitude, vertical image location the sound frequency and horizontal location time and stereo. A drawback of this approach is that the entire data contained in the image is sonified, regardless of the relevance of the information. Other systems are designed for a very special purpose (e.g. [2]). Unlike previous approaches we aim to sonify images of any kind. We propose that models of visual attention [4] and visual grouping [5] can be utilized to dynamically select relevant visual information to be sonified. For the auditory synthesis we employ an approach, which takes advantage of the sparseness of the selected input data. Horizontal image locations are directly mapped into the sound signal using auditory stereo. Vertical information is encoded by time. Additional audio parameters, such as frequency, can be controlled by local image features, such as orientation. Furthermore, we introduce a sequential playback mode where image features at different locations are played in a successive manner. Instead of being played simultaneously, this enhances the perception of relative differences in the stereo signal. In conclusion, the presented approach proposes a combination of data sonification approaches, such as auditory scene generation, and models of human visual perception. It extends previous pixel-based transformation algorithms by incorporating mid-level vision coding and high-level control. The mapping utilizes elaborated sound parameters that allow non-trivial orientation and positioning in 3D space.

- [1] Hermann, Ritter (2004): Proceedings of the IEEE, Special Issue Engineering and Music, 92
- [2] Hermann, Nattkemper, Ritter, Schubert (2000): Proceedings of the Mathematical and Engineering Techniques in Medical and Biological Sciences, CSREA Press
- [3] Meijer (1992): IEEE Transactions on Biomedical Engineering, 39
- [4] Itty, Koch, Niebur (1998): IEEE Transactions on Pattern Analysis and Machine Intelligence, 20
- [5] Fischer, Bayerl, Neumann, Christobal, Redondo (2004): European Conference on Computer Vision 2004, LNCS 3023

## **Computer Vision and Robotics**

## Recognition of Occluded Objects through Evolutionary Learning

Andreas Birk

International University Bremen

a.birk@iu-bremen.de

The recognition of partially occluded objects through vision is still a largely open problem, both from the application oriented viewpoint of machine vision and robotics as well as from the basic research perspective of cognitive science and related fields.

A machine learning approach is presented here that is capable of recognizing geometric figures even in scenes where they are partially occluded. The approach has several special properties. First and foremost, it is based on the idea that perception is an active process where representations are generated that try to match the data coming from the sensors. Concretely, the underlying representations consist in the experiments presented here of programs that generate data, namely images based on a simple turtle graphics language.

The programs representing a particular figure by producing the related image are generated in an evolutionary learning process. This process operates on a population of programs in an online fashion, i.e., it tries to generate a representation of the visual input in real-time. In doing so, evolutionary operators that mimic mutation and crossover as well as hill-climbing on constants are used as optimization techniques. The evolutionary learning process is driven by a fitness function that is based on two aspects. First, it measures the similarity of an input image with the images produced by the programs from the current population. This special similarity function operates on pixel level and it can be very efficiently computed. Second, the length of a program representing a figure is taken into account. In the spirit of Kolmogoroff-complexity, a shorter program producing the same data is preferred.

Given the input images of colored geometric figures like triangles, rectangles or more complex polygons, the evolutionary process learns programs that generate the according images, i.e., programs that generate data that matches the sensor input. The interesting aspect is what happens when scenes composed of several figures are presented to the system. When for example a scene with a red triangle and a blue rectangle is used, the result is a program that consists of two according parts. Even under occlusions, the “right” representations are learned when the separated figures have been perceived shortly before. When for example the blue rectangle is partially hidden behind the red triangle, it is nevertheless recognized by the system when it has been perceived shortly before and it is hence still part of the population in the learning algorithm.

## The Smart Camera Paradigm and its Role in Tracking

Sven Fleck, Sven Lanwer, Peter Biber and Wolfgang Straßer

University of Tübingen

fleck@gris.uni-tuebingen.de

Today's computer vision systems typically see cameras only as simple sensors. The processing is performed after transmitting the complete raw sensor stream via a costly and often distance-limited connection to a centralized processing unit (PC). We think it is more natural to also physically embody the processing in the camera itself: what algorithmically belongs to the camera is also physically performed in the camera. The idea is to compute the information where it becomes available—directly at the sensor—and transmit only results that are on a higher level of abstraction. This follows the emerging trend of self contained and networking capable Smart Cameras. We present a Smart Camera prototype for probabilistic tracking. Tracking plays a central role for many applications including robotics (visual servoing, RoboCup), surveillance (person tracking) and also human-machine interface, motion capture, augmented reality and 3DTV. Our Smart Camera consists of a CCD sensor, a FPGA, a PowerPC processor and an Ethernet connection. It implements a particle filter using color histograms. Particle filters have become a major way of tracking objects. They propagate (approximated) probability density functions (pdfs) over time and can handle multiple hypotheses and nonlinear systems. This Smart camera approach offers several benefits, especially: 1. Low bandwidth requirements: The raw image is processed directly in the camera. Hence, only the approximated pdf of the target's state has to be transmitted using relatively few parameters. This allows the use of standard networks (i.e. Ethernet) with virtually unlimited range. 2. No additional computing outside the camera has to be performed: the output could e.g. be directly connected to a robot control unit for visual servoing. 3. Higher resolution: As the raw video stream does not need to comply with the bandwidth any more, sensors with higher spatial or temporal resolutions can be used. 4. As a consequence, this approach offers optimal scaling for multi-camera systems to work together in a decentralized way (e.g. as basis for airport surveillance). To increase the tracking robustness, the camera automatically adapts to slow appearance (e.g. illumination) changes during runtime. This is performed by blending the appearance at the most likely position with the target reference appearance in histogram space. The rate of adaption is variable and depends on the filter's tracking confidence to circumvent overlearning. We finally present first concepts how the Smart Camera paradigm gets part of our Analysis by Synthesis framework (which is also particle filter enabled).

## A Goal-Directed Visual Attention System

Simone Frintrop,<sup>1</sup> Gerriet Backer<sup>2</sup> and Erich Rome<sup>1</sup>

<sup>1</sup>Fraunhofer Institut AIS, <sup>2</sup>Krauß Software GmbH

`simone.frintrop@ais.fraunhofer.de`

Suppose you are looking for your key. You know it to be somewhere on your desk but it still takes you several fixations until your roaming view hits the key. If you have a salient key fob contrasting with the desk, you will detect the key faster. This is according to the separation of visual processing into two subtasks: First, a fast parallel attention system detects object candidates and, second, complex recognition restricted to these regions verifies the hypothesis. The focus of this work is on the first task, i.e., on the pre-selection of object candidates with a computational attention system.

In human perception, the focus of attention is guided by two factors: bottom-up attention detects regions standing out from the rest of the scene, for example a black sheep among white ones, and top-down attention guides the view according to knowledge, expectations and goals. Existing computer models of visual attention focus mainly on the bottom-up aspect. Here, we present a robust computational attention system that enables goal-directed search.

A standard bottom-up architecture [1] based on the Feature Integration Theory [2] is extended by a top-down component, enabling the weighting of features depending on previously learned weights. In learning mode, the region of the target is provided by the user in a user-friendly way by drawing a rectangle around the target with the mouse. The system determines autonomously which region inside the rectangle is most salient and used this region for learning. The learned weights consider not only the properties of the target (excitation) but also of the background (inhibition). In search mode, the system uses the learned weights to excite or inhibit the features in the scene and directs the focus to the region which most likely contains the target.

Detailed performance results are presented on artificial images as well as on a wide variety of real-world images. The target is typically among the first 3 focused regions making the system a robust and time-saving front-end for object recognition.

[1] Itti, L. et al. (1998): PAMI, 20(11), 1254–1259

[2] Treisman, A.M. and Gelade, G. (1980): Cogn. Psych. 12, 97–136

## A Flexible Object Model for Facial Gesture Recognition

Andreas Tewes, Christoph von der Malsburg and Rolf P. Würtz

Neuroinformatik, Universität Bochum

`andreas.tewes@neuroinformatik.rub.de`

Elastic Graphs, labeled with Gabor jets, have proven to be a powerful concept for object detection and recognition. They are robust against small variations in the shape and texture of objects, but it is not possible to model these variations explicitly. Since this is necessary for describing flexible objects like the human face with emphasis on facial gestures, we have developed a Flexible Object Model (FOM) as a derivative of the concept of Elastic Graphs. The FOM, still using the Elastic Graph for representing the current object state, uses available matching techniques and can therefore immediately be deployed wherever Elastic Graphs have been used so far. While an Elastic Graph can be seen as a snapshot of one object state concerning shape and texture, the FOM is capable of describing the variations present in a whole learning set. The FOM consists of an implicit description of shape variations as well as mappings which synthesize the Gabor jets as functions of the current shape. Both, the deformations as well as the mappings, are learned from samples without any user assistance. The samples were collected from video sequences of subjects performing selected facial gestures. Each sequence starts and ends in a neutral expression. The landmarks were manually labeled with nodes in the first frame and afterwards tracked throughout the sequence using a method based on the phases of Gabor wavelets. Finally, each frame leads to a sample consisting of shape (arrangement of nodes) and texture (Gabor jets attached to each node). Using a Neural Gas for clustering leads to a sparse representation of possible object shapes while the correlation between shape and texture is modeled by a linear mapping between the former and the latter.

Instead of finding landmarks in a given image using predefined graph moves together with constant texture, the FOM makes it possible to exploit the previously learned “gesture moves” together with texture which is always adapted and therefore consistent to the current shape.

Finding landmarks in faces of different gestures not only leads to gesture recognition but is also a necessary preprocessing step for improving the performance of face recognition under gesture variations. It is possible to estimate the neutral facial expression of a subject by replacing those landmarks which differ from their neutral position in shape and texture by their neutral correspondents of the FOM while keeping the remaining landmarks unchanged.

Funding by the European Commission in the Research and Training Network MUHCI (HPRN-CT-2000-00111) is gratefully acknowledged.

## Improving Automatic Pedestrian Detection by Means of Human Perception

Quoc C. Vuong<sup>1</sup> and Modesto Castrilln-Santana<sup>2</sup>

<sup>1</sup>MPI for Biological Cybernetics, Tübingen,

<sup>2</sup>Universidad de Las Palmas de Gran Canaria, Spain

quoc.vuong@tuebingen.mpg.de

The present study investigates the detection of pedestrians by humans and by computer vision systems. This simple task is accomplished easily and quickly by human observers but still poses a challenge for current systems. In the computer vision community, different automatic detection systems have been designed using simple features to detect faces, heads and shoulders, full bodies and leg regions. With these regions, these systems perform fairly well but still have high miss rates. However, we found a small correlation between the performance of these systems and human observers. This finding motivated us to systematically analyze human performance on a pedestrian detection task that tests whether these regions are the most semantically useful and whether other regions can also provide useful information. For that purpose, we used a psychophysical “bubbles” technique[1] to isolate those regions used by humans for detection. In this technique, images containing aligned pedestrian and non-pedestrian scenes are revealed through a mask of small randomly distributed gaussian windows (“bubbles”). Across observers, masks leading to correct “present” responses are summed and normalized to reveal image regions that were useful for detection. Our results indicate that observers relied predominantly on head and leg regions, and to a lesser extent on arm regions. These results confirm some of the regions already considered by automatic pedestrian detectors. An important question is: Among those regions that are particularly discriminable are some more critical than others for certain pattern matching problems? To address this question, the regions selected by human observers were applied to a general object detection framework designed by Viola and Jones[2]. This framework has already been successfully applied to different objects categories, and it is well known in the face detection community because it provides real-time performance. In sum, we believe that this perceptually-based approach can be useful for assigning different weights to regions and points not only based on their discriminability but also on their perceptual significance for the problem considered.

[1] Gosselin, F. and Schyns, P.G. (2001): *Vis. Res.* 41, 2261–2271.

[2] Viola, P. and Jones, M. (2001): In *IEEE Conf. on Comp. Vis. and Pat. Recog.*

## Recognition of Pose and Illumination of Human Faces Through Combination of Feature-Based and Correspondence-Based Pattern Recognizers

Günter Westphal and Rolf P. Würtz

Neuroinformatik, Universität Bochum

guenter.westphal@neuroinformatik.rub.de

We present a method to recognize pose and illumination of still images of human faces through combination of feature-based and correspondence-based pattern recognizers.

In the feature-based part nearest neighbor classifiers, which decide whether a feature is present in an image, are learned from training images. The discriminative power of the single classifiers is modelled to be equivalent to the information of the respective features with respect to the learning set. For processing faces, the features are small localized grid graphs, whose nodes are attributed with Gabor amplitudes on five scales and in eight orientations. Combining the classifiers weighted by the information measures into a linear discriminant yields a preselection of model candidates. This weighted voting is a successful extension of existing multiple feature subset methods.

If the preselection yields several model candidates, each of them is verified using a variant of elastic graph matching, a standard correspondence-based technique for face recognition. To further differentiate between model candidates with similar features it asserts that the features be in similar spatial arrangements for the model to be selected. Model graphs are constructed efficiently by assembling the model features already extracted into larger graphs. If model ambiguities still remain after this step, denser correspondence maps can be created and evaluated by, e.g., pyramid matching, until a unique candidate can be identified.

We report the results of experiments on a database of computer-generated human faces with variation of two pose angles and illuminations composed of a diffuse and a point light source at varied angles. The method achieved high ( $> 95\%$ ) recognition rates of pose and illumination with good generalization over identity. First experiments on databases of real images are also encouraging. The results are remarkable, because our method requires neither an internal 3D model nor the evaluation of correspondence maps with a large number of models. Consequently, the evaluation is very fast, allowing for about five decisions per second on average.



## **Multimodal Interactions: Sensory Integration**

**Cognitive Factors Facilitate Multimodal Integration.**

Hannah Helbig and Marc O. Ernst

MPI for Biological Cybernetics, Tübingen

helbig@tuebingen.mpg.de

Ernst & Banks (2002) showed that humans integrate visual and haptic signals in a statistically optimal way if they are derived from the same spatial location. Integration seems to be broken if there is a spatial discrepancy between the signals (Gepshtein et al., in press). Can cognitive factors facilitate integration even when the signals are presented at two spatial locations? We conducted two experiments, one in which visual and haptic information was presented at the same location. In the second experiment, subject looked at the object through a mirror while touching it. This way there was a spatial offset between the two information sources. If cognitive factors are sufficient for integration to occur, i.e. knowledge that the object seen in the mirror is the same as the one touched, we expect no difference between the two experimental results. If integration breaks due to the spatial discrepancy, we expect subjects' percept to be less biased by multimodal information. To study integration, participants looked at an object through a distortion lens. This way, for both the "mirrored" and "direct vision" conditions, there was a slight shape conflict between the visual and haptic modalities. After looking at and feeling the object simultaneously participants reported the perceived shape by either visually or haptically matching it to a reference object. Both experiments revealed that the shape percept was in-between the haptically and visually specified shapes. Importantly, there was no significant difference between the two experimental results regardless of whether subjects matched the shape visually or haptically. However, we found a significant difference between matching by touch and matching by vision. Haptic judgments are biased towards the haptic input and vice versa. In conclusion, multimodal signals seem to be combined if observers have high-level cognitive knowledge about the signals belonging to the same object, even when there is a spatial discrepancy.

[1] Ernst, M.O. and Banks, M. (2002). *Nature*, 415, 429–433.

[2] Gepshtein, S., Burge, J., Banks, M.O. and Ernst, M. (in press). *Journal of Vision*

This work was supported by the Max Planck Society and was conducted within the EU-Project TOUCH-HapSys (IST-2001–38040)

## Judging Size by Hand: No Benefit for Bimanual Estimates

Christoph Lange and Marc O. Ernst

MPI for Biological Cybernetics, Tübingen

`christoph.lange@tuebingen.mpg.de`

When redundant sources of sensory information are available, integrating these sources is beneficial for a system in order to decrease the perceptual noise and so to increase the accuracy of the overall percept [1]. However, a potential cost may come from incorrect binding of the different sources, which will generally evoke perceptual illusions [2].

Here we ask whether humans take advantage of a bimanual size estimate originating from an object of constant size. The stimulus was a cylindrical object that subjects felt with their left and right hand simultaneously. To display the bimanual haptic stimulus we used two PHANToM force-feedback devices and measured size discrimination performance using a 2IFC paradigm. Subjects' task was to decide which of two bimanual stimuli was bigger. From the resulting psychometric functions we determined the JND for bimanual vs unimanual size discrimination and the point of subjective equality (PSE).

There was no difference in the bimanual vs the unimanual JNDs which indicates that subjects did not benefit from having available two size estimates in the bimanual situation. We therefore conclude that there is no integration of size estimates between hands when an object is touched bimanually. We may speculate that this failure of integration to occur results from the fact that naturally the object is touched at slightly different spatial locations and so information is not truly redundant even though subjects were told to touch an object of constant size. Not being able to integrate information that comes from two different spatial locations may help to prevent misbinding of different sensory sources.

[1] Ernst, M.O. and M.S. Banks (2002): Nature 415, 429–433.

[2] Da, W., Kanai R., Shimojo S. (2004): Nature 429, 262.

This research was supported by the 5th FP European grant TOUCH-HapSys (IST-2001–38040).

## **Auditory Cues can Facilitate the Visually-Induced Self-Motion Illusion (Circular Vection) in Virtual Reality**

Bernhard E. Riecke, Jörg Schulte-Pelkum, Franck Caniard and Heinrich H. Bühlhoff

MPI for Biological Cybernetics, Tübingen

bernhard.riecke@tuebingen.mpg.de

There is a long tradition of investigating the self-motion illusion induced by rotating visual stimuli (“circular vection”). Recently, Larsson et al. (2004)[1] showed that up to 50% of participants could also get some vection from rotating sound sources while blindfolded, replicating findings from Lackner (1977)[2]. Compared to the compelling visual illusion, though, auditory vection is rather weak and much less convincing.

Here, we tested whether adding an acoustic landmark to a rotating visual photorealistic stimulus of a natural scene can improve vection. Twenty observers viewed rotating stimuli that were projected onto a curved projection screen (FOV: 54°x40.5°). The visual scene rotated around the earth-vertical axis at 30°/s. Three conditions were randomized in a repeated measures within-subject design: No-sound, mono-sound, and 3D-sound using a generic head-related transfer function (HRTF).

Adding mono-sound showed only minimal tendencies towards increased vection and did not affect presence-ratings at all, as assessed using the Schubert et al. (2001) presence questionnaire [3]. Vection was, however, slightly but significantly improved by adding a rotating 3D-sound source that moved in accordance with the visual scene: Convincingness ratings increased from 60.2% (mono-sound) to 69.6% (3D-sound) ( $t(19)=-2.84$ ,  $p=.01$ ), and vection buildup-times decreased from 12.5s (mono-sound) to 11.1s (3D-sound) ( $t(19)=2.69$ ,  $p=.015$ ). Furthermore, overall presence ratings were increased slightly but significantly. Note that vection onset times were not significantly affected (9.6s vs. 9.9s,  $p>.05$ ).

We conclude that adding spatialized 3D-sound that moves concordantly with a visual self-motion simulation does not only increase overall presence, but also improves the self-motion sensation itself. The effect size for the vection measures was, however, rather small (about 15%), which might be explained by a ceiling effect, as visually induced vection was already quite strong without the 3D-sound (9.9s vection onset time). Merely adding non-spatialized (mono) sound did not show any clear effects. These results have important implications for the understanding of multi-modal cue integration in general and self-motion simulations in Virtual Reality in particular.

[1] Larsson, P., Västfjäll, D., & Kleiner, M. (2004): Proceedings of Presence 2004, 252–258

[2] Lackner, J.R. (1977): *Aviat. Space and Env. Med.* 48(2), 129–131

[3] Friedmann, F., & Regenbrecht, H. (2001): *Presence—Teleop. and Virt. Env.*, 10(3), 266–281

SUPPORT: EU grant POEMS-IST-2001–39223 (see [www.poems-project.info](http://www.poems-project.info)) and Max Planck Society.

## Anticipation of Heaviness in Vision and Grasp

Cornelius Steckner

focas

csteckner@hotmail.com

There is a vision-based physical object mapping in grasp to guide the motor control of the hand. This object mapping drives the common visual and haptic Müller-Lyer illusion [1, 3].

According to its time-line, the mapping process is an anticipation of object qualities [4]. But not only shape is mapped in advance, but also the shape related physical qualities. The heaviness of the physical object is among these anticipated qualities [2].

But how to have direct experimental access to these anticipated qualities such as heaviness?

The anticipatory hand movements just show the forthcoming access of shape and might be filmed or measured from the distance of the finger tips. But these distances reveal the anticipated shape and not the shape-related physical qualities. To succeed, the guiding map in vision and grasp itself has to be detected. Consequently, we have to look for common instances in vision and grasp.

M.E. Cheveul in his “De la loi du contraste simultane des couleurs” (1839; Engl. ed. 1854) already had presented a generalisation of the class of Müller-Lyer illusions, as a “contrast considered with respect to the size of two contiguous objects of unequal size”. In the view of the related experiment, all these illusions were the effects of simultaneous contrast enhancements. These occur whenever specific pairs of qualities are simultaneously related, which might be pairs of colours, or pairs of arrows or even lines.

This idea of contrast enhancement leads to the conclusion, to use also simultaneously paired qualities in anticipatory experiments: The related effect of contrast should enhance the associated qualities and allow their experimental access. Accordingly, to have access to the anticipation of heaviness, we use both hands in the experimental design to enhance the subjective contrast of the focused quality.

- [1] Gentaz, E., Camos, V., Hatwell, Y., Jacquet, M.Y. (2004). The visual and haptic Müller-Lyer illusion : correlation study. *Current Psychology Letter*, 13, vol. 2. [<http://cpl.revues.org/document431.html>]
- [2] Kawai, S. (2002), Heaviness perception. II. Contributions of object weight, haptic size, and density to the accurate perception of heaviness or lightness, *Experimental Brain Research*, 147(1), 23–28.
- [3] Rudel, R., Teuber, H.-L. (1963). Decrement of visual and haptic Müller-Lyer illusion on repeated trials: A study of cross-modal transfer. *Quarterly Journal of Experimental Psychology*, 15, 125–131.
- [4] Sarabon, N., Dimitrijevic, M.M., Zidar, J., Dimitrijevic, M.R. (2004). Remarks on neurocontrol of the hand and significance of afferent input, *Zdrav Vestn* 73, Suppl. II, 11–17.

## The Impact of Attention on Visual-Vestibular Interaction in Helicopter Control

Andreas Wacker, Astros Chatziastros and Heinrich H. Bühlhoff

MPI for Biological Cybernetics, Tübingen

`andreas.wacker@tuebingen.mpg.de`

Ego-motion perception is a crucial ability of humans, particularly with regard to control of posture. Therefore the role of multisensory integration, which handles the information from multiple senses, seems to be a basic necessity. Currently, most experimental work on visual-vestibular interaction is based on modification of one or more sensory signals. Cognitive variables like previous knowledge, expectancy or attention are largely disregarded in this research area. A few recent research, however, suggests that especially attention has a significant impact on the way information from different senses is combined [1,2].

To investigate the influence of attention in an ego-motion perception task, we employed a realistic, physics-based helicopter simulation on a Stewart motion platform. Participants performed a ‘hover task’, in which they were required to obtain and maintain a pre-defined orientation around the earth vertical by simultaneously controlling four (yaw, pitch, roll, height) of the six helicopter’s degrees-of-freedom. In a between-group design, participants were instructed to focus either on the visual input (projection screen) or on the vestibular input (motion platform). During the hover task, subjects were distracted by task-irrelevant disturbances (visual lateral offset, vestibular lateral offset, or combined lateral offset). Given some inherent ambiguity between translational and rotational movements, we expected visually attending subjects to respond strongest to the task-irrelevant disturbances during a visual offset. Similarly, vestibular instruction was expected to lead to a stronger deflection under a mere platform offset. 12 naive subjects and 11 experienced helicopter pilots participated in this experiment.

The results did not reveal any systematic influence of the mode of instruction on the control behavior. Further analysis showed that pilots and non-pilots exhibited different patterns of responses under the different conditions. Non-pilots were distracted only by visual disturbances. In contrast the pilots responded only to the combined offset (visual together with vestibular). Furthermore, female subjects benefited from the vestibular whereas the male subjects profited by the visual focus of attention, resulting in a better hover performance.

The findings indicate that lateral disturbances are in part erroneously interpreted as rotational movements during helicopter flying. The specific response patterns suggest that non-pilots were using a rather visually based control strategy, whereas pilots relied on the integrated, visual and vestibular, information—the later finding possibly reflecting the pilots’ previous knowledge of helicopter dynamics. The apparent failure to influence the control behavior by instructionally enforced focus of attention will be discussed in terms of different concepts of attention and in terms of potential differences in the task demands of open-loop compared to close-loop studies.

[1] Furman, J., Müller, M., Refern, M. and Jennings, J. (2003): *Exp. Br. Res.* 152, 383–392

[2] Yardley, L., Gardner, M., Lavie, N. and Gresty, M. (1999): *Neuropsychologica* 37, 1293–1301

Thanks to the Max Planck Society for funding this research and the pilots from the HTC Helicopter Training and Charter GmbH for participating in this experiment.

## **Multimodal Interactions: Sensory-Motor Integration**

## **Influence of Motion Cueing on Helicopter Stabilisation**

Daniel Berger, Cengiz Terzibas and Heinrich H. Bühlhoff

MPI for Biological Cybernetics, Tübingen

daniel.berger@tuebingen.mpg.de

Hovering a helicopter is a difficult task. Helicopter dynamics are unstable, comparable to an inverse pendulum. Helicopter pilots need to actively control the position and orientation of the helicopter at every moment in time. It is assumed that they use not only visual information, but also body cues (vestibular/somatosensory) for this task.

To test this hypothesis, we measured human stabilisation performance in a helicopter simulator (similar to a Robinson R-22 helicopter) using a Stewart platform. We compared four body motion cueing conditions: platform off, platform translations, platform rotations, and both translations and rotations. Washout-filters were used for platform translations whereas pitch and roll rotations were directly transferred to the platform. A computer-generated scene was projected onto a screen (70x54 degrees visual field) 1.15m in front of the participants. The stimulus consisted of a textured ground plane with a target ball fixed to the ground and a second ball as position marker (fixed to the helicopter 15 meters in front). The visual scene was always shown from an observer in the simulated helicopter. The task was to control the helicopter so that the two balls remained as close as possible to each other, while minimizing heading drift. Each trial lasted for two minutes. Trials with different motion cueing were interleaved.

Participants controlled three of the four axes of the helicopter: pitch (forward-backward acceleration), roll (sideways acceleration), and yaw (heading orientation). Height above ground was automatically controlled by the simulator. Only subjects who after training managed to stabilize the helicopter with a mean distance of less than 5 meters were included in the study. For every trial, we measured the mean distance between target and helicopter position marker, mean velocity of the marker, and mean pitch and roll angles of the helicopter. We then analyzed the influence of the motion cueing conditions on these measured variables.

We found that participants could better stabilize the helicopter if body motion cues were available. Body rotations had a larger beneficial effect on stabilization performance than body translations. The mean distances from the target were smaller in the sideways direction than in the forward-backward direction for all four conditions.

We conclude that participants can use body motion cues to increase their stabilisation performance when controlling a simulated helicopter. Possibly humans are not very sensitive to small visual acceleration, and body motion can help to overcome this limitation. As the simulated helicopter always accelerates in the direction in which it is tilted, body tilt can provide immediate feedback to the pilot in which direction he is accelerating, even before he can identify an acceleration visually. These findings have important implications for the design of helicopter simulation platforms.

## **No Correlation Between Smooth Pursuit and Perceived Motion in the Motion after Effect**

Doris Braun, Lars Pracejus and Karl R. Gegenfurtner

Universität Giessen

doris.braun@psychol.uni-giessen.de

Smooth pursuit eye movements can be observed in response to a stationary sinewave grating after prolonged adaptation to a moving grating [1]. The magnitude of the perceptual motion after effect was equal to the after effect on smooth pursuit eye movements, indicating that the same mechanisms might underly performance in both tasks. If this is indeed the case, then there should be a trial-by-trial correlation between eye movement speed and perceptual judgments.

We used a nulling paradigm to measure the magnitude of the MAE. A moving sinewave grating of 0.25 cycle per degree, 8 degree per second, and 40 per cent contrast was initially presented for 30 seconds, and then again for 5 seconds at the beginning of each trial. The subjects were instructed to keep their eyes on a central fixation spot. The adapting grating was followed by a test grating that was either stationary or would slowly move in the adapting or the opposite directions. At the same time eye movements were monitored during the adapting and the testing phase. At the end of the trial, the subject had to give a rating whether the grating was stationary, moving in the adapting direction, or in the opposite direction.

Even though the magnitude of the effects of adaptation was similar for pursuit and perception, there was no significant trial-by-trial correlation ( $\rho < 0.1$ ) at any of the test speeds. Furthermore, when we used a squarewave grating as the test stimulus, we still observed a perceptual motion after effect, but there was no motion after effect on smooth pursuit eye movements.

We conclude that both pursuit and perception carry out similar motion computations, but that they are implemented in different neural circuits.

[1] Pracejus et al, TWK 2004.

**Visually-Guided Pointing with Trajectory Perturbations: Tactile-Kinesthetic Detection and Movement Costs**

Knut Drewing and Julia Trommershäuser

Universität Gießen

knut.drewing@psychol.uni-giessen.de

Pointing movements towards a visual goal call for a visually-guided specification of the finger trajectory. When this trajectory is physically perturbed by a force acting on the finger, detection and compensation for the perturbations are additionally required for successful movement execution. Here we asked if participants are able to detect external trajectory perturbations using tactile-kinesthetic cues, how the executed movement is affected by the perturbation and, whether the perturbation interferes with the goal of the task. Participants were instructed to rapidly hit a visual target, which was presented within a three-dimensional visuo-haptic virtual environment. Late responses and failures to hit the target were penalized. During the pointing movement the finger was not visible. During the initial movement phase, participants were presented with a force pulse, which was applied to their right finger tip. Force perturbations were given orthogonally to the movement direction. We determined detection thresholds for perturbations from six different directions (up, down, upper right/left; lower right/left) using a two-interval forced choice paradigm. Five participants completed the experiment. Surprisingly, detection thresholds for the applied perturbations (threshold about 0.10 N) were just slightly higher than tactile detection of forces in a single-task context (about 0.05N [1]). Detection performance did not depend on the direction of the perturbation, but was better for short perturbations (30 ms presentation time) compared to longer perturbations (50 ms presentation time). Shorter perturbations differed from longer perturbations by a steeper increase in force amplitude (10% of the duration). Locally, perturbations ( $>$  about 0.07 N) affected the movement kinematics significantly as compared to trajectories without perturbation. However, the distribution of movement end points at the location of the visual target did not correlate with the direction of the perturbation. These results are a first hint that the brain is able to detect trajectory perturbations during pointing movements using tactile-kinesthetic cues and can compensate for these perturbations without extra costs.

[1] Lederman, S.J. and Klatzky, R.L. (1999): *Presence* 8, 86–103.

Deutsche Forschungsgemeinschaft (Emmy-Noether-Programm, Grant TR 528/1–2)

**Changes in Saccadic Latencies Due to Distractive Information : An Analysis Based on Latency Distributions**

Stefan Gulan, Kiley Seymour and Uwe J. Ilg

Universität Tübingen, Hertie-Institut

`stefan.gulan@student.uni-tuebingen.de`

It was previously shown that saccadic latency in humans can be increased by sudden changes in luminance or orientation of a structured background. This result was led back to the limited capacity of channels in the visuomotor pathway. We expand this finding in a series of studies from a rather qualitative to a more quantitative result : The intervall in which a distraction is possible is refined to lie below 14 ms SOA. Next we seek to incorporate a measure of the strength of a distracting stimulus the increase in saccadic latency is functionally dependant on the amount of change in background luminance. The findings from the experiments will also be examined from the viewpoint of the LATER-model which allows a distinction between different possible cases of latency changes based on latency distributions. In our case it points to a difference in available visual information due to different background changes. As eye and hand movement commands partly share a common pathway, i.e. the superior colliculus, we also repeat the amentioned experiment with combined eye and hand movements. The result that the latency of hand movements underlies the same functional dependance on sudden background changes further strengthens the relation between goal-oriented behaviour by these two kinds of actions. These results indicate that information from luminance-changes is an important factor for the building of saliency maps, since it cannot be deliberately suppressed at critical times which demand fast and top-down modulated responses.

Supported by DFG, SFB 550, A3

**Target Selection Characteristics in Eye-Hand Coordination**

Annette Horstmann and Klaus-Peter Hoffmann

Ruhr-Universität Bochum

horstmann@neurobiologie.rub.de

During a goal-directed movement of the hand to a visual target the controlling nervous system depends on information provided by the visual system. Thus a coupling between these two systems is reasonable. In a choice condition with two or more equivalent objects present at the same time two questions arise: 1. How is one particular target chosen? 2. If a coordinated movement of eyes and arm is planned: Do we reach for an object we have selected by looking at it or do we look to the object we selected to grasp? To answer those questions we examined the preference of human subjects to select the left or the right target depending on the action to be performed (eye, arm or coordinated movement) and the horizontal position of the target in peripersonal space. Two targets were presented at the same distance to the left and right of a fixation point and the stimulus onset asynchrony (SOA) was adjusted until both targets were selected equally often. This balanced SOA is then a quantitative measure of selection preference. We compared these preferences at three horizontal positions for the different movement types (eye, arm, both). The subjects showed at the peripheral positions a preference for the target that was more central with respect to the head, regardless of movement type. At the central position they showed on average a bias to select the right target. Comparing the preferences for the different movement types the preferences for 'arm' and 'coordinated' movement types were correlated strongest. The requirements of the arm movement seem to outweigh the requirements of the coordinated movement's saccadic part. These findings provide evidence that the control of gaze is in this case a means to an end, namely to conduct the arm movement properly and that a common spatial framework for eye-hand coordination exists.

## Perisaccadic Localization of Multisensory Targets

Steffen Klingenhoefer and Frank Bremmer

Philipps-University Marburg

steffen.klingenhoefer@physik.uni-marburg.de

Recent psychophysical studies have demonstrated visual mislocalization effects at the time of saccades: In case spatial visual references are available immediately after the saccade, perceived target locations are compressed around the endpoint of the saccade. In total darkness targets appear to be shifted in the direction of the saccade. In both cases mislocalization starts about 100 ms before a saccade and reaches a peak at the time of saccade onset. The goal of our present study was to investigate whether also auditory spatial perception is influenced by saccadic eye movements and, if this was the case, whether it would show the same spatial and temporal properties as in the visual domain.

Human subjects were seated in a light tight and sound attenuated chamber performing blocks of either visual saccades or control tasks with steady fixation at either 7.5° left (F1) or 7.5° right (F2) from straight ahead. In half of the trials with fixation at F1 the second fixation point F2 was presented during the trial acting as a visual distractor. In all tasks target stimuli were presented at one of 4 possible positions symmetrically located around F2. Target stimuli in auditory localization tasks were bursts of white noise (duration: 5ms) presented via a moveable loudspeaker. In visual control trials bars (0.5 x 20 degree) were flashed for 10 ms. Saccades were directed from F1 to F2 with the target stimuli being presented within +/-200ms relative to saccade onset.

In visual control trials all perceived target positions were shifted perisaccadically in the direction of the saccade, i.e. due to missing visual references we did not observe a compression effect in the visual domain. In auditory localization tasks subjects reliably discriminated different speaker locations in all tasks, yet during fixation trials localization was strongly influenced by the current eye position. In saccade and distractor trials perceived locations were shifted (compressed) towards F2. In saccade trials this effect was significantly more pronounced (Wilcoxon Rank test,  $p < 0.05$ ). Perisaccadic mislocalization was stationary, i.e. it did not reveal the dynamics observed in the visual domain.

As these crossmodal results differ from the pure visual control data we conclude that the effect of perisaccadic mislocalization of auditory targets is an independent phenomenon. In other words: the spatial and temporal properties of perisaccadic localization of visual targets do not generalize across senses.

**The Relationship Between Action Pantomime, Semantic Object Knowledge and Multiple Object Use in Stroke Patients- a Multivariate Analysis**

Inga Liepelt,<sup>1</sup> Maja Trenner,<sup>2</sup> Stefanie Freund<sup>3</sup> and Thomas Platz<sup>3</sup>

<sup>1</sup>Neurologische Universitätsklinik, Tübingen,

<sup>2</sup>Abteilung für Neurologie, Charite Universitätsmedizin Berlin,

<sup>3</sup>Abteilung für neurologische Rehabilitation, Charite Universitätsmedizin Berlin

inga.liepelt@uni-tuebingen.de

The execution of activities of daily living may depend on different cognitive processes. Traditionally the term ideational apraxia describes an impaired manipulation of single or multiple objects. Disturbances in semantical object knowledge or object function knowledge are often also defined as ideational or conceptual apraxia. Both tasks are referred to a semantic level of action planning (ideational component). Therefore it is assumed that disturbances in semantical object knowledge and object use are related to each other. Nevertheless the impairment of other processes, for example disturbances in action production (ideomotor component) can possibly also lead to deficits in object use. In a group of 58 stroke patients the relationship between task performance in action production, object-object association knowledge, tool function knowledge (subtest results from the Berlin Apraxia Test) and multiple object use (Multiple Object Test) are investigated. The results of confirmatory factor analysis indicate that both, ideational and ideomotoric components, are related to difficulties in multiple object use. Correlation analysis showed that deficits in action sequencing and omissions of action elements were moderately related to task performances of action pantomime (ideomotor component). In contrast, content errors in object use were moderately associated with deficits of semantical object knowledge (ideational component).

This work was supported by the DFG within the GK 'Clinical and Cognitive Neuroscience' in Berlin.

**The Visual System of Diurnal and Nocturnal Geckos (Gekkonidae, Reptilia): Optomotor Reaction and Associated Brain Structures**

Olivia Masseck, Beate Röhl and Klaus-Peter Hoffmann

Ruhr-Universität Bochum

masseck@neurobiologie.rub.de

In all vertebrates the horizontal OKR (hOKR) stabilizes the image on the retina during horizontal movements. The eyes /heads move with the same velocity and direction as the surrounding. If the horizontal movement holds on, the pursuit eye movements (which keep the image on the retina) are followed by quick phases in the opposite direction. The occurrence of the monocular hOKN varies in different vertebrates. A symmetrical monocular hOKN means that both stimulus directions (temporo-nasal, nasal-temporal) are of the same gain. Asymmetric monocular hOKN means that one direction is preferred. There are a lot of hypotheses about the symmetry of monocular hOKN. The most prominent is the so called “fovea-theory” developed by Tauber and Atkin [2]: Only foveate vertebrates are able to perform a symmetrical hOKN. Another one is the decussation-theory devised by Fukuda and Tokita which maintains that the degree of optic nerve decussation is most important for a symmetrical hOKN: the more fibers stay ipsilateral the more symmetrical the monocular hOKN will get. To proof these theories for lizards behavioural and neuroanatomical studies on foveate geckos: *Lygodactylus spec.* and on afoveate geckos: *Lepidodactylus lugubris* were done. For visual stimulation an optokinetic drum with a Julesz-stimulation pattern was used. Two stimulus velocities 20 deg/s and 50 deg/s were tested in a clockwise (CW) and counter clockwise direction (CCW). The results of the optokinetic measurements were surprising. The foveate geckos did not perform a symmetrical monocular hOKN. A monocular hOKN in naso-temporal direction could not be elicited at all. Thus for foveate geckos a fovea is not the main substrate for driving a symmetrical monocular hOKN. Neuroanatomical studies were done in vitro, with DiI a fluorescent antero- and retero-grad tracer. Crystal DiI was applied on one optic nerve or in one optic tract. The tracer studies showed ipsilateral retinofugal projections to the main optic areas. Thus the results disprove Fukuda’s and Tokita’s [1] decussation-theory for lizards as well.

[1] T.Fukuda and T.Tokita (1957) *Acta Oto-Laryngol*[2] E.S.Tauber and A.Atkin (1969) *Science*

**The Feedforward Dynamics of Response Priming by Color**

Thomas Schmidt, Silja Niehaus and Annabel Nagel

Universität Göttingen

thomas.schmidt@psych.uni-goettingen.de

Visual stimuli elicit a wave of early brain activation that reaches most cortical areas within 120 ms [1]. Because typical cells can fire only once in the interval before activity is passed on to the next area, it has been proposed that this feedforward sweep must be largely free from intracortical feedback, whereas conscious perception is supposed to be possible only with recurrent processing. However, it is not clear yet whether feedforward waves can proceed to the level of overt behaviour uncompromised by forthcoming waves. By measuring pointing responses to color targets preceded by color stimuli priming, either the same or opposite response as the targets [2], we found that early pointing kinematics depended mainly on properties of the primes and were largely independent of motor and perceptual effects of the actual targets, provided that the prime-target interval was long enough (about 70 ms). With shorter intervals, even the earliest phases of the pointing response were influenced by properties of the mask. We propose that the pointing movements are under continuous control of successive feedforward waves of prime and target; these waves have an effective duration well exceeding the first few spikes of the wavefront and have substantial overlap for short prime-target intervals. We argue that speeded pointing actions are exclusively controlled by stimulus input, independent of visual awareness, provided that cognitive control mechanisms have established a unique mapping of stimuli to responses.

[1] Lamme, V.A.F., & Roelfsema, P.R. (2000). *Trends Neurosci.* 39, 729–736.

[2] Schmidt, T. (2000). *Psychol. Sci.* 13, 112–118.

Supported by the German Science Foundation, Schm1671/1

## **Natural Image Statistics**

## Classification of Natural Scenes using Global Image Statistics

Jan Drewes,<sup>1</sup> Felix A. Wichmann<sup>2</sup> and Karl R. Gegenfurtner<sup>1</sup>

<sup>1</sup>University of Giessen, Giessen, Germany,

<sup>2</sup>MPI for Biological Cybernetics, Tübingen, Germany

jan.drewes@psychol.uni-giessen.de

The algorithmic classification of complex, natural scenes is generally considered a difficult task due to the large amount of information conveyed by natural images. Work by Simon Thorpe and colleagues showed that humans are capable of detecting animals within novel natural scenes with remarkable speed and accuracy. This suggests that the relevant information for classification can be extracted at comparatively limited computational cost. One hypothesis is that global image statistics such as the amplitude spectrum could underly fast image classification (Johnson & Olshausen, *Journal of Vision*, 2003; Torralba & Oliva, *Network: Comput. Neural Syst.*, 2003).

We used linear discriminant analysis to classify a set of 11.000 images into animal and non-animal images. After applying a DFT to the image, we put the Fourier spectrum of each image into 48 bins (8 orientations with 6 frequency bands). Using all of these bins, classification performance on the Fourier spectrum reached 70%. In an iterative procedure, we then removed the bins whose absence caused the smallest damage to the classification performance (one bin per iteration). Notably, performance stayed at about 70% until less than 6 bins were left. A detailed analysis of the classification weights showed that a comparatively high level of performance (67%) could also be obtained when only 2 bins were used, namely the vertical orientations at the highest spatial frequency band. When using only a single frequency band (8 bins) we found that 67% classification performance could be reached when only the high spatial frequency information was used, which decreased steadily at lower spatial frequencies, reaching a minimum (50%) for the low spatial frequency information. Similar results were obtained when all bins were used on spatially pre-filtered images.

Our results show that in the absence of sophisticated machine learning techniques, animal detection in natural scenes is limited to rather modest levels of performance, far below those of human observers. If limiting oneself to global image statistics such as the DFT then mostly information at the highest spatial frequencies is useful for the task. This is analogous to the results obtained with human observers on filtered images (Kirchner et al, VSS 2004).

[1] Kirchner et al, VSS 2004

[2] Johnson & Olshausen, *Journal of Vision*, 2003

[3] Torralba & Oliva, *Network: Comput. Neural Syst.*, 2003

## On the Visual Perception of Translucent Materials

Roland Fleming

MPI for Biological Cybernetics, Tübingen

roland.fleming@tuebingen.mpg.de

Many commonly occurring substances are somewhat translucent (e.g. wax, jade, fruit-flesh, and cheese). When light strikes a translucent material, it passes through the surface and scatters a number of times within the body of the object before re-emerging. This causes light to ‘bleed’ through translucent objects, giving them a distinctive visual softness and glow. What image cues are responsible for this characteristic appearance? How do we distinguish between translucent and opaque materials? Here we use a combination of image statistics and psychophysics to study the perception of translucent materials.

There has been a large amount of previous work on the perception of materials that transmit light. Almost all of this work is based on simple physical models of transparency, (e.g. the episcotister model of Metelli, 1974), in which the object of interest is a thin filter or screen. However, recent advances in computer graphics (Jensen et al. 2001; Jensen and Buhler, 2002) make it possible to simulate the complex physics of solid translucent objects. We have used this model to study how a wide range of factors influence the perception of translucency, including highlights, colour, contrast, brightness, blurriness, and conditions of illumination.

Our main findings are as follows: (1) We have found that it is possible to enjoy a vivid impression of translucency even when many of the cues that were traditionally believed to be important for the perception of transparency (e.g. X-Junctions, contrast conditions) are absent from the display. (2) We argue that sub-surface light scattering is too complex for the visual system to infer translucency by inverse optics. Accordingly we suggest that the visual system tracks low-level image statistics that reliably correlate with changes in translucency. (3) We find that perceived translucency varies dramatically with conditions of illumination. We compare how well a number of candidate cues can predict these variations.

In conclusion, there is a wide range of materials that have not been studied before, and which we are only just beginning to understand. Many intuitions that we have about which cues are crucial for recovering opacity turn out to be at best incomplete.

[1] Metelli (1974). *Scientific American*, 230(4).

[2] Jensen, Marschner, Levoy & Hanrahan (2001), SIGGRAPH 2001.

[3] Jensen & Buhler (2002). *ACM Trans. on Graphics*, 21(3).

RWF was supported by the Max Planck Society

## How does the Environment Influence the Shape of Receptive Fields in Cortical Simple Cells?

Lorenz Gerstmayr and Hanspeter A. Mallot

Universität Tübingen

lorenz.gerstmayr@uni-tuebingen.de

Perceptual systems have evolved to optimally process the kind of stimuli occurring most frequently and to transform the incoming stimuli to an efficient representation reducing statistical redundancy. It has been shown [1] that Independent Component Analysis (ICA) is an appropriate statistical method to estimate this transformation for visual stimuli. Under the assumption that images are a linear combination of certain base images (Linear Image Synthesis Model), ICA can estimate these base images from a set of training images. It has also been shown that the base images resemble the receptive fields of Simple Cells in the visual cortex and can therefore be described by Gabor functions.

In this work we investigated how statistical properties of the environment influence the shape of the receptive fields of simple cells and how these properties are coded in the receptive fields. Therefore we sorted images into three different categories, namely natural images (showing landscapes, flowers and animals), man-made images (showing mainly buildings), and aerial images of urban areas, with each category containing approximately 2500 grayscale images. The three classes were characterized by shorter contours of various orientations, shorter contours in mainly horizontal and vertical orientation and long contours of various orientations respectively. For all categories the independent components were computed, Gabor functions were fitted to the resulting base images and the parameters of the Gabor functions were statistically analyzed.

For all three categories, the resulting Gabor functions are mostly antisymmetric. The distributions of the spatial frequencies differ for the classes, but all have a maximum for the same spatial frequency. The results also show that the simple cells are much more elongated for aerial images than for natural and man-made images. For man-made images horizontal and vertical orientations are dominating while for aerial images the orientation is almost equally distributed. For natural images we found that horizontal, vertical and diagonal oriented fields are most frequent. We conclude that the statistical properties of natural, man-made and aerial images can be found again in the shape of the receptive fields of simple cells.

For future work these findings about the receptive fields will be used together with a model of edge detection and contour grouping in the human visual system for a point of interest detection.

[1] E.P. Simoncelli and B.A. Olshausen, *Annu. Rev. Neurosci.*, 2001

## **Perception of Time**

## Visual Durations Discrimination: Confidence Characteristics and Cognitive Styles

Elena Golovina

Institute of Psychology Russian Academy of Sciences, Moscow; student of Neurocognitive program, LMU, München

golovina@psy.uni-muenchen.de

Confidence and its relationships with individual cognitive styles were investigated in a visual duration discrimination task in which pairs of visual flashes either the same or different in duration were presented. The difference between durations was chosen on an individual basis to achieve 70–80% of correct responses [5]. Observers estimated the flash durations in each pair to be either the “same” or “different” and then marked their responses as “confident” or “unconfident” and instruction stressed the accuracy of responses. The following indices were calculated: proportion of correct responses: PC; mean weighed confidence category used:  $Con = 1/N \times (0.5 \times N1 + 1 \times N2)$ , where N1 was a number of “unconfident” responses, N2—a number of “confident” responses, N—total number of responses; Bias=Con—PC. A positive bias-value meant “overconfidence”, i.e. an observer believed s/he performed better than s/he actually did. A negative bias-value meant “underconfidence”, i.e. an observer believed s/he performed worse than s/he actually did [1,4,5]. Cognitive styles are interpreted as individual preferences in perception, action and performance operations [3]. The style “Cognitive control of differentiation” means a subject’s intention to perceive persons and objects to be grouped into a few classes included many items (a wide control) or into many classes included a few items (narrow control) [2]. The style “Field dependence-independence” means a subject’s intention to perceive figures to be dependent of a surrounding or not [6]. The following results were obtained. 1. Overconfidence in visual duration discrimination was obtained in Russian sample. 2. Subjects characterized by a wide control of differentiation appeared to be less confident in their discrimination responses but estimated the accuracy of their answers adequately (i.e. showed no over-, neither underconfidence) as compared with those of a narrow control. 3. Subjects characterized by a middle level of field-(in)dependancy were more confident in visual durations discrimination than field-independent and field-dependent ones.

- [1] Baranski, J.V., Petrusic, W.M. (1999): Realism of confidence in sensory discrimination. *Perception and Psychophysics*, 61,1369–1383.
- [2] Gardner, H. (1968): *Perceptual and Motor Skills*, 26, 311–330.
- [3] Nosal, C.S. (1990): *Psychologisne modele umyslu*. Warsaw: Patswowe Wydawnictwo Naukowe.
- [4] Olsson, H., Winnman, A. (1996): Underconfidence in sensory discrimination/ *Perception and Psychophysics*, 58, 374–382.
- [5] Skotnikova, I.G., Rammsayer, T., Brandler, S.(2001): Cross-cultural differences in confidence judgements/ *Fechner Day’s 2001*, Germany, 608–613.
- [6] Witkin, H.A., Goodenough, D.R., (1981): *Cognitive styles: essence and origins*. Intern. Univ. Press, N.Y.

Supported by Russian Foundation for Humanities(RFH) (grant No 03–06–00076a).

## Perceived Duration of Frequent and Infrequent Stimuli

Judith Nitschke and Rolf Ulrich

Universität Tübingen

`judith.nitschke@uni-tuebingen.de`

Two experiments were conducted to determine whether perceived stimulus duration depends on whether subjects have to process an expected or an unexpected visual stimulus. Subjects compared the duration of a constant standard stimulus with a variable comparison stimulus. An adaptive psychophysical procedure was employed to assess the perceived duration of the comparison. Changes in expectancy were induced by presenting one type of comparison more frequently than another type. Thus, in 70% of all trials the frequent and hence expected comparison of constant shape and color was presented, whereas in the remaining 30% of the trials the infrequent and hence unexpected comparison differing in color and shape was shown. Experiment 1 used standard durations of 100 and 400 ms, whereas Experiment 2 employed durations of 400 and 800 ms. Stimulus frequency did not affect perceived duration in Experiment 1. In Experiment 2, however, the constant error was larger for frequent than for infrequent comparisons. This pattern of results is consistent with the notion that subjects tend to perceive expected stimuli as shorter than unexpected stimuli, if the duration of those stimuli does not fall below a critical value. Our findings do also rule out a potential confound of previous research, which holds that directed spatial attention increases perceived stimulus duration.

### Some Effects of Negative Delays Upon the Perception of Causal Relatedness

Goutami Shenvi<sup>1</sup> and Mark A. Elliott<sup>2</sup>

<sup>1</sup>Univesität München, <sup>2</sup>Universität München

`goutami.shenvi@psy.uni-muenchen.de`

We examined the effects of negative delays on the perception of causality using a variation of the paradigm originated by Michotte [1] and as an extension to similar work conducted by Kanizsa and Vicario [2]. In our design, on some trials a second Object B started to move prior to collision with the launching Object A. Given that contact A-B was made we expected reports akin to 'A launched B' following some negative delays. However, rather than obtaining unequivocal measures related to launching, with variations in reportage over the range of negative delays, Experiment 1 revealed a tendency for subjects to adjust their pattern of responses following positive delays as a function of their reportage following negative delays. In fact, observers tended to equilibrate their pattern of causality and no causality responses such that the different proportions of responses were symmetrical across negative and positive delays. In Experiment 2, which introduced a further report alternative aiming describe better way causality reportage by means of metaphor, report equilibration was found to equilibrate across the different classes of causality reports and one alternative class of no causality report. The two experiments described here strongly suggest that causality reportage governed by a tendency for subjects to try to balance the number of reports they make such that each response alternative equally represented.

- [1] Michotte, A. (1954). *La Perception de la Causalité*. (Études de Psychologie), Louvain: Studia Psychologica (2<sup>o</sup> Édition).
- [2] Kanizsa, G., & Vicario, G. (1968). *Ricerca Sperimentale Sulla Percezione*. Università degli Studi di Trieste.

## **Perceptual Development**

## The Toelz Temporal Topography Study: Mapping the Visual Field Across the Life Span

Dorothe Poggel,<sup>1</sup> Claudia Calmanti,<sup>2</sup> Bernhard Treutwein<sup>3</sup> and Hans Strasburger<sup>4</sup>

<sup>1</sup>Center f. Innovative Visual Rehabilitation, Boston,

<sup>2</sup>Generation Research Program, Bad Toelz,

<sup>3</sup>IuK, Ludwig-Maximilians Universitaet, München, <sup>4</sup>Universitaet Goettingen

dapoggel@bu.edu

Like many other cerebral functions, the capacity for processing visual information is generally believed to deteriorate with age, especially in the temporal domain. However, the course, patterns, and mechanisms of this decline over the life span are largely unknown. Since visual field maps are useful for elucidating processing mechanisms in the visual pathway, our goal was to describe and compare maps of different temporal and non-temporal visual functions and their change over the life span.

We compared the visual field topography of 95 healthy subjects (10–90 years) with respect to luminance thresholds (static perimetry), temporal resolution (double-pulse resolution, DPR), simple reaction time (RT), and contrast thresholds ( $R_{\text{contrast}}$ ) as well as the correlations of these variables with several non-topographical attentional and visual functions.

DPR thresholds increased with eccentricity and age. The decline of performance with age was more pronounced in the periphery. RT increased only slightly and uniformly across the visual field in older subjects. Perimetric luminance thresholds rose in a pattern similar to the DPR maps. Contrast thresholds, alertness, divided and spatial attention, and saccadic exploration showed age-related changes and complex correlation patterns with the main outcome variables. Performance was almost constant up to the age of sixty; a marked decrease was observed only late in life (80–90 years).

The age-related decrease of visual performance was confirmed, but the topographical patterns of functional loss differ between outcome measures, even between those within the domain of temporal information processing. Since interindividual variability is high, especially in old subjects, age is a weak predictor of functionality. Age-related deterioration cannot be attributed exclusively to optic media and the retina, but visual and cognitive brain processes also affect the characteristics of visual field maps, particularly attentional functions. Our data provide a useful normative basis for psychophysical and neuropsychological studies.

[1] Poggel, D.A. & Strasburger, H. (2004). *Acta Neurobiologiae Experimentalis*, 64: 427–437.

[2] Treutwein B, Rentschler I (1992) *Clinical Vision Science* 7: 421–434.

This project was funded by a grant from the Deutsche Forschungsgemeinschaft to HS (Str 354/3–1). We thank Luise Schulze, Andrea Geyer, and Inge & Patrick Das Gupta for help with data acquisition and analysis.

## **Perceptual Neuroscience**

**Biological Motion Processing in Neurosurgical Patients with Cerebellar Lesions**

Alireza Gharabaghi, Arseny A. Sokolov, Marcos Tatagiba and Marina Pavlova

University Hospital of Tübingen

alireza.gharabaghi@uni-tuebingen.de

Veridical visual processing of biological movement is of immense value for a variety of daily-life situation (e.g., safe car driving) and social communication. Although neurobiological mechanisms underlying biological motion perception have been explored extensively by neuroimaging techniques [1], the role of subcortical structures is not well understood. Recent PET and fMRI data indicate that a gradient of activation during visual perception of biological motion is located in the medial and/or lateral cerebellum [2–4]. Here we report the results of pre- and postoperative psychophysical examinations of neurosurgical patients with cerebellar lesions. We ask (i) whether biological motion processing is compromised in these patients; and (ii) how, if at all, biological motion processing is altered after neurosurgery. Our preliminary findings reveal that these patients exhibit a lower sensitivity to a point-light walker camouflaged by an additional set of moving dots in comparison to healthy controls. In spite of changes in the patients' general health status, postoperative examinations did not reveal any further deterioration in biological motion processing. Combination of psychophysical and brain imaging techniques is required to determine an exact localisation of the cerebellar regions involved in visual processing of biological motion. These findings may help to assess the risk/benefit ratio in relation to the neurosurgical strategy.

- [1] Pavlova, M. et al. (2005): *Cereb. Cortex* 15, in press.
- [2] Grossman, E. D. et al. (2000): *J. Cogn. Neurosci.* 12, 711–720.
- [3] Vaina, L. M. et al. (2001): *Proc. Natl. Acad. Sci. USA* 98, 11656–11661.
- [4] Ptito, M. et al. (2003): *Neuroimage* 19, 246–252.

## **Facilitating Effect of 15 Hz Repetitive Transcranial Magnetic Stimulation (rTMS) on Tactile Perceptual Learning**

Ahmed A. Karim, Anne Schueler, Eva Friedel, Yiwen Li Hegner and Ben Godde

Institute of Med. Psychology and Behavioral Neurobiology, University of Tübingen

ahmed.karim@uni-tuebingen.de

Recent neuroimaging studies have revealed that tactile perceptual learning can lead to substantial reorganizational changes of the brain [1, 2]. We report here for the first time that combining high-frequency (15 Hz) rTMS over the primary sensory cortex (SI) with tactile discrimination training is capable to facilitate tactile perceptual learning. Most notably, facilitating the excitability of SI by 15-Hz rTMS improved tactile perceptual learning in spatial but not in temporal discrimination tasks. These findings give causal support to recent correlative data from our research group obtained by fMRI studies indicating a differential role of SI in spatial and temporal discrimination learning [3]. To our knowledge this is the first study to show that operant learning can be facilitated by rTMS. The introduced combination of rTMS and tactile discrimination training may provide new therapeutical potentials in facilitating neuropsychological rehabilitation of functional deficits after lesions of the somatosensory cortex [4].

- [1] Pascuel-Leone, A. and Torres, F. (1993): *Brain* 116, 39–52.
- [2] Recanzone, G.H., Mezernich, M.M., Jenkins, W.M., Grajaski, K., Dinse, H.R. (1992): *J Neurophysiology* 67, 1031–1056.
- [3] Hodzic, A., Karim, A.A., Veit, R., Erb, M., Godde, B. (2004): *J Neuroscience* 24, 442–446.
- [4] Karim, A.A., Schueler, A., Hegner, L.Y., Friedel, E., Godde, B. (submitted).

Supported by a grant of the Volkswagen-Stiftung (I 73035) to B.G and by a grant of the medical faculty in Tuebingen, Germany (AKF 13800) to A.K.

## Cholinergic and GABAergic Effects on Temporal Transfer Properties in Goldfish Examined with the ERG and Behavioral Experiments

Carlos Mora-Ferrer<sup>1</sup> and Konstantin Behrend<sup>2</sup>

<sup>1</sup>J. Gutenberg Universität Mainz, <sup>2</sup>J. Gutenberg Universität Mainz

carlos.mora-ferrer@uni-mainz.de

**Purpose:** Cholinergic and GABAergic mechanisms have been demonstrated to be involved in retinal motion coding. Whether and how blockade of GABAergic and cholinergic receptors affect temporal transfer properties (TTP) was investigated under comparable experimental conditions in both the ERG and behavioral experiments. **Methods:** Temporal transfer properties were determined from vitreal ERG recordings under photopic conditions as well as from behavioral, two-choice forced procedure, experiments before and after injection of curare (all concentrations given are estimated intra-ocular concentrations, 10 microM, nicotinic ACh-R antagonist), atropine (100 microM, muscarinic ACh-R antagonist), picrotoxine (30 microM, GABA<sub>A</sub>/c-R antagonist, ERG only), or bicuculline (30 microM and 67 microM, GABA<sub>A</sub>-R antagonist). **Results:** The gain-frequency plots determined from ERG recordings and from behavioral experiments are very similar in shape [1]. The ERG data can be described as a 3rd order low-pass filter with resonance having an upper limit frequency of about 30 Hz. Atropine shifts the gain-frequency characteristic to a lower limit frequency without changing its shape. Curare changes the limit frequency and the characteristic of the gain-frequency plot to that of a pure low-pass, i.e. the resonance is removed. Bicuculline removes the resonance and in addition lowers the order of the low-pass filter and the limit frequency. Post-injection of bicuculline, the gain-frequency plot generated with behavioral data has a comparable limit frequency to the ERG-data but falls off rather steeply. **Conclusions:** Nicotinic and GABAergic receptor mediated mechanisms appear to be involved in the generation of the resonance property of the TTP while muscarinic receptors do not seem involved. The generation of the resonance property of the TTP occurs most probably in the inner retina since actions of nicotinic ACh-receptors have been reported in the inner retina. The site of action must be a cell type visible in the ERG, i.e. not the ganglion cells. Changes of the resonance part of the TTP could occur due to either changes in inhibitory strength or an phase- shift of the inhibition. This is in line with the GABAergic control of bipolar cell output as deduced from modulatory action of dopamine [1]. So far, all data indicate that the well balanced timing of glutamatergic excitation and GABAergic feedback inhibition at the bipolar cell axon terminals is critical for the high frequency end of the TTP and thus the temporal resolution ability of goldfish.

[1] Mora-Ferrer, C. and Behrend, K., Dopaminergic modulation of photopic temporal transfer properties in goldfish retina investigated with the ERG, *Vision Res*, 44 (2004) 2067–81.

Forschungsfond der Universität Mainz (CMF)

**Biological Motion Processing in Patients with Periventricular Lesions: MEG Findings**

Marina Pavlova,<sup>1</sup> Fabio Marconato,<sup>2</sup> Alexander Sokolov,<sup>3</sup> Christoph Braun<sup>1</sup> and  
Ingeborg Krägeloh-Mann<sup>4</sup>

<sup>1</sup>University of Tübingen, <sup>2</sup>University of Padova, <sup>3</sup>University of Ulm,  
<sup>4</sup>University Hospital of Tübingen

marina.pavlova@uni-tuebingen.de

Although neurobiological mechanisms underlying biological motion perception are being explored extensively by neuroimaging techniques, the role of subcortical structures is not well understood [1]. Here we ask whether and, if so, how cortical activity in response to point-light biological motion is modulated by early damage to periventricular regions. To this end, we assessed changes in the MEG response to point-light configurations recorded in adolescents who were born premature with signs of bilateral periventricular leukomalacia (PVL) at MRI scan and in healthy term-born controls. Participants performed a one-back repetition task with a set of stimuli consisting of a point-light walker and a scrambled configuration. Patients produced a greater number of misses in response to the walker, whereas no difference was found between the patients and controls in response to the scrambled configuration. Two bilateral spatio-temporal dipoles in the left and the right occipito-parieto-temporal junction, and one additional dipole in the superior parietal cortex were fitted to the group means of neuromagnetic response. In controls, the strength of the parietal dipole at latencies of 360–430 ms post-stimulus was significantly weaker to the scrambled configuration than to the canonical walking figure. Irrespective of stimulus type, the strength of the left dipole in the occipito-parieto-temporal junction was significantly greater in patients than in controls. The findings suggest that parieto-occipital periventricular lesions lead to non-stimulus-specific disinhibition of cortical neuromagnetic activity resulting in compromised processing of point-light configurations. One of the possible mechanisms of this influence is that periventricular lesions break the reciprocal interrelations between the parietal cortex and thalamus impinging on posterior thalamocortical fibers.

[1] Pavlova, M. et al. (2005) *Cereb. Cortex* 15: in press.

Supported by the DFG (KR 1316/5-P)

## rTMS Induced Improvement of Human Visual Orientation Discrimination

Patrick Ragert,<sup>1</sup> Christian Klaes,<sup>1</sup> Martin Tegenthoff<sup>2</sup> and Hubert R. Dinse<sup>1</sup>

<sup>1</sup>Ruhr Univ Bochum, <sup>2</sup>Ruhr Univ Bochum, BG-Kliniken Bergmannsheil

patrick.ragert@rub.de

At a cellular level, plastic changes of synaptic transmission can be evoked solely by patterned electrical stimulation lacking meaning, attention or task constraints. In an attempt to explore the efficacy of patterned, passive stimulation for perceptual learning in humans, we have recently demonstrated that 5 Hz repetitive transcranial magnetic stimulation (rTMS) applied over the hand representation of somatosensory cortex leads to improvement of tactile 2-point discrimination of the fingers. Here we demonstrate that 5 Hz rTMS applied over the foveal visual cortex representation is similarly effective in the visual domain by evoking persistent improvement of orientation discrimination performance.

Orientation discrimination was tested using the protocol introduced by Schoups et al.1995. The stimulus was a circular 2.5 deg diameter noise field positioned centrally, consisting of light and dark bars (0.075 to 0.31 deg, relative phases randomized). Intertrial interval was 1 s, stimulus exposure 300 ms. Subjects had to respond within 600 ms after stimulus onset. The reference orientation, which was never shown, was right oblique (45 deg); for control of generalization, a reference of -45 deg was also tested. The subjects had to decide whether the noise field was tilted clockwise or counterclockwise to the reference, auditory feedback was given for errors. An up-down staircase procedure was used converging on an just noticeable orientation differences (JND) corresponding to an 84 % criterion. Starting value was an orientation difference of  $\pm 7$  deg, step size was 20 %. Daily sessions consisted of 10 blocks of 100 trials each. To position the coil over the foveal representation of primary visual cortex, we induced phosphenes within the central visual field using single pulse TMS. For rTMS, two sessions separated by 30 min through the tangentially oriented coil grip backwards were applied at an intensity 10 % below the phosphene induction threshold resulting in a total of 2500 pulses. First, orientation discrimination was tested over 10 sessions. In the 11th session rTMS was applied. About 15 min after termination, orientation discrimination was re-tested to study changes in discrimination performance. To assess the time course of recovery, additional tests were performed 24 hours and 1 week later. In the control group, subjects were tested with an identical schedule, but without rTMS application. For each subject we found perceptual changes occurring over the first 5 to 6 sessions that were highly specific for the learned orientation with little transfer to an orientation 90 deg away. After reaching a stable plateau of discrimination in the remaining sessions, rTMS improved performance by about 10 % for both the learned and the non-learned orientation. Most notable, the improvement continued to develop resulting in an even higher gain of performance 24 h after rTMS. After 1 week, the improvement recovered, yet not reaching baseline performance. None of these effects could be observed in the control group. The results provide further evidence that stimulation protocols resembling those used in LTP studies applied from outside directly to selected brain regions can induce persistent improvement of discrimination performance.

## Relations Between Preferred Spatial Frequency and Size of Spatial Summation Field in V1 Multi-Unit Activity of Awake Macaque

Tobias Teichert,<sup>1</sup> Thomas Wachtler,<sup>1</sup> Alexander Gail,<sup>2</sup> Markus Wittenberg,<sup>1</sup> Frank Michler<sup>1</sup> and Reinhard Eckhorn<sup>1</sup>

<sup>1</sup>Universität Marburg, <sup>2</sup>California Institute of Technology

tobias.teichert@physik.uni-marburg.de

**Introduction.** Many neurons in striate cortex are selective for orientation and spatial frequency (SF). Models of early visual processing often use sets of localized SF-filters, based on oriented gabor-wavelets, that are scaled in size inversely proportional to their optimal SF. This results in identical relative SF-tuning widths. In striate cortex, preferred SFs of neurons at a given eccentricity cover a range of 3 to 4 octaves [1]. If these “neural SF-filters” represent a gabor- or other wavelet basis, the sizes of their receptive fields (RF) are expected to increase inversely proportional to their preferred SF.

**Methods.** We recorded multi-unit activity in the upper layers of awake macaque primary visual cortex. Preferred SF and spatial-summation-field (SSF) size were determined by observing responses to gabor-patches of six different sizes (0.1–5.0 deg) and seven spatial frequencies (0.7–8.0 cycles/deg) centered on the classical RF. Preferred SF and SSF-size were estimated using a bootstrap method and interpolating the response strength on the 6x7 patchsize/spatial frequency grid.

**Results.** SF preference covered the expected range, with average preferred SF at 3.3 +/- 3.8 cpd. At eccentricities between 3 and 5 deg, average SSF size, was .7 +/- .1 deg. We found a significant negative correlation ( $p < .001$ ) between preferred SF and SSF-size. On the other hand, we could show that, contrary to predictions from Wavelet analysis, there was a significant increase of relative size of the SSF with increasing preferred SF. As could be expected from linear filter theory, relative size of the SSF showed a negative correlation with the SF tuning width as estimated from the bootstrap method.

**Conclusions.** These results suggest that multi-unit complex-type cell activity in the upper layers of V1 shows similarities to localized linear frequency analysis. On the other hand, the relationship between SSF-size and preferred SF corresponds neither to Wavelet- nor Windowed Fourier- analysis.

[1] De Valois RL, Albrecht DG, Thorell LG *Vision Research* 22:545–559

Supported by DFG (to RE) and DFG GK NeuroAct (to TT)



## **Research Methods and Techniques**

## **Bayesian Inference for Psychometric Functions**

Malte Kuss, Frank Jäkel and Felix A. Wichmann

MPI for Biological Cybernetics, Tübingen

`kuss@tuebingen.mpg.de`

In psychophysical studies of perception the psychometric function is used to model the relation between the physical stimulus intensity and the observer's ability to detect or discriminate between stimuli of different intensities. We propose the use of Bayesian inference to extract the information contained in experimental data to learn about the parameters of psychometric functions. Since Bayesian inference cannot be performed analytically we use a Markov chain Monte Carlo method to generate samples from the posterior distribution over parameters. These samples can be used to estimate Bayesian confidence intervals and other characteristics of the posterior distribution. We compare our approach with traditional methods based on maximum-likelihood parameter estimation combined with parametric bootstrap techniques for confidence interval estimation. Experiments indicate that Bayesian inference methods are superior to bootstrap-based methods and are thus the method of choice for estimating the psychometric function and its confidence-intervals.

MK was supported by the German Research Council (DFG) through grant RA 1030/1.

**Contrast Dependence of the BOLD Response in Visual Cortices**Karoline Spang,<sup>1</sup> Volker Diehl<sup>2</sup> and Manfred Fahle<sup>3</sup><sup>1</sup>Dept. Human Neurobiology, Univ. Bremen, <sup>2</sup>Dept. Radiology, Bremen,<sup>3</sup>Human Neurobiology, Univ. Bremen

karospang@hotmail.com

Visual resolution strongly increases with stimulus contrast, and there is consensus that the BOLD response in early visual cortices increases with contrast, too. But the exact shape of the transducer function relating the BOLD response to stimulus contrast is still under debate and may differ for different cortical areas, indicating different properties of the corresponding neuronal networks. We tested the amplitude of the BOLD response in five topographically identified cortical areas (V1, V2v, V2d, V3v & V3d) as well as in area MT as a function of stimulus contrast in 8 normal observers. A checkerboard subtending approximately 20 degrees (deg) of visual angle was presented in the bore of a Siemens 1.5 tesla "Vision" scanner at a contrast reversal frequency of 4 Hz, and with a check size of about 3 deg. The BOLD response during 30 sec blocks of checkerboard presentations was compared with both the response during 30 sec blocks of "full field" flicker of the same temporal frequency and with the response to a homogeneous stationary grey field. The contrast of the checkerboard (and between the two states of the homogeneous field) was set to either 100%, 50%, 25%, 10%, 5%, 2.5%, or 1%. The amplitudes of the BOLD response for the representations of both the central and the peripheral visual field were calculated by means of Brainvoyager Software for small regions of interest. The results show a monotonic increase of the BOLD response amplitude with stimulus contrast for both the checkerboard and the homogeneous field stimuli. Amplitudes increase rather linearly with the log of stimulus contrast, at least for the representation of the central visual field in V1. The peripheral representation, on the other hand, shows a strong depression, rather than an increase, for high stimulus contrasts especially of the checkerboard in the retinotopic visual areas. Thus, the exact form of the transducer function relating stimulus contrast to cortical BOLD response depends strongly on the position in the visual field, on the type of baseline used for comparison, and on the cortical area investigated. Our study shows that cortical oxygen levels change as a result of visual stimulation, that this change depends on stimulus contrast, and that the influence of stimulus contrast differs between cortical areas and between visual field positions. The data confirm long-range interactions between the representations of the central versus peripheral visual field in early cortical representations and indicate rather complex interactions between different visual field representations.

## Amygdala Activation in Response to Fearful Faces: Lateralization Effects in fMRI Depending on Phase-Encoding Direction

Krystyna Swirszcz,<sup>1</sup> Mikhail Zvyagintsev,<sup>1</sup> Hermann Ackermann<sup>1</sup> and Klaus Mathiak<sup>2</sup>

<sup>1</sup>Universität Tübingen, <sup>2</sup>Universität Aachen

krystyna.swirszcz@med.uni-tuebingen.de

Most of functional neuroimaging studies of emotion processing report lateralized activation of the amygdala, more often to the left than to the right [1]. Several hypotheses have been suggested to explain this difference in lateralization. One of the hypotheses ascribes asymmetries to susceptibility artifacts from adjacent sinuses [2]. Echo-planar imaging (EPI) is widely used for functional MRI but suffers from dephasing and strong asymmetric distortions, most prominent along the phase-encoding (PE) direction. Asymmetries of amygdala activation can be expected to depend upon those artifacts. Specialized multi-echo EPI can acquire opposite phase-encoding directions in a single shot. An implementation of this imaging procedure should be used to investigate the effects of PE on amygdala lateralization.

Six healthy male subjects participated in the present study. For the stimulation, masked and unmasked presentation of neutral and fearful faces [3] was applied. All examinations were conducted on a 3 T whole-body MR scanner (Magnetom Trio, Siemens, Erlangen). For functional measurements, a multi-echo EPI sequence was applied to collect 16 coronal slices (matrix size 64x48, voxel size 3x3x4 mm, TR = 2 s, TEs = 23/47/71 ms, FA = 70°). Preprocessing encompassed undistortion with contrast-to-noise optimization [4], normalization to standard space and spatial smoothing with 12 mm Gaussian kernel. Statistical analysis was performed using single-subject analyses based on the general linear model of a block design (SPM2 software package). The obtained amygdala activations were compared between the single echo methodology—comparable to standard EPI—and the undistorted and optimized images.

Amygdala activations could be obtained in 5 out of the 6 subjects. Responses to the unmasked fearful faces adapted in about 8 sec. No reliable amygdala responses to the masked fearful faces were observed on the single subject level. PE direction and polarity determined the distortions at the level of the amygdala. In the preliminary data exploration, PE polarity and image contrast affected the activation patterns. The optimized image combination exhibited the most consistent activation patterns.

Our study supports the hypothesis that asymmetries in amygdala activation depend on phase encoding direction. In addition to improving functional sensitivity, single-shot multi-echo EPI with alternating PE gradients might ameliorate artifactual asymmetries. Performing fMRI studies on lateralization at the level of the amygdala involves a number of pitfalls. The control of PE direction and multi-echo EPI can help to reduce some of these limitations.

[1] Baas D. et al. (2004): *Brain Res. Rev.* 45, 96–103

[2] LaBar, K.S. et al. (2001): *NeuroReport* 1616, 3461–3464

[3] Ekman, P. and Friesen, W.V. (1976): *Pictures of Facial Affect*, Consulting Psychologists Press, Palo Alto, NM

[4] Weiskopf et al. (2005): *NeuroImage*, in press

## Efficient Adaptive Sampling of the Psychometric Function by Maximizing Information Gain

Thomas G. Tanner, N. Jeremy Hill, Carl E. Rasmussen and Felix A. Wichmann

MPI for Biological Cybernetics, Tübingen

tanner@tuebingen.mpg.de

A psychometric function can be described by its shape and four parameters: position or threshold, slope or width, false alarm rate or chance level, and miss or lapse rate. Depending on the parameters of interest some points on the psychometric function may be more informative than others. Adaptive methods attempt to place trials on the most informative points based on the data collected in previous trials. We introduce a new adaptive bayesian psychometric method which collects data for any set of parameters with high efficiency. It places trials by minimizing the expected entropy [1] of the posterior pdf over a set of possible stimuli. In contrast to most other adaptive methods it is neither limited to threshold measurement nor to forced-choice designs. Nuisance parameters can be included in the estimation and lead to less biased estimates. The method supports block designs which do not harm the performance when a sufficient number of trials are performed. Block designs are useful for control of response bias and short term performance shifts such as adaptation. We present the results of evaluations of the method by computer simulations and experiments with human observers. In the simulations we investigated the role of parametric assumptions, the quality of different point estimates, the effect of dynamic termination criteria and many other settings.

- [1] Kontsevich, L.L. and Tyler, C.W. (1999): Bayesian adaptive estimation of psychometric slope and threshold. *Vis. Res.* 39 (16), 2729–2737.

## Objective Estimation of Visual Acuity—High Correlation Across Normal and Diseased Vision with an Automated VEP-Based Algorithm

Marc Wolf, J. Peter Maurer and Michael Bach

Univ.-Augenlinik Freiburg

wolfmarc@aug.ukl.uni-freiburg.de

**Purpose:** To assess the visual acuity with fully objective evaluation based on visual evoked potentials.

**Methods:** 40 normal subjects, and 24 patients (10 cataract, 3 corneal clouding, 7 epiretinal gliosis, 3 AMD, 1 chorioretinitis; decimal visual acuity (VA) range 0.14–1.1) participated in the study. Checkerboard stimuli with 6 check sizes covering 0.05–0.4° (or 0.09–0.8° for visual acuity below 0.35) were presented in brief-on-off mode (40 ms on, 93 ms off) at 7.5 Hz. In normal subjects, the stimuli were optically degraded by various degrees of dioptrical defocus (n=35) or Bangerter occluders (n=45) resulting in a decimal VA range of 0.128–2.6. Steady-state VEPs were recorded with a Laplacian montage ((2Oz-(RO+LO)). Fourier analysis yielded the magnitude at the stimulus frequency  $A_s$  and—as noise estimate—the average of the two neighboring frequencies ( $N$ ).  $A_s$  and  $N$  determine the significance level  $p$  of the response, and from their ratio the non-noise-contaminated response  $A^*$  can be calculated. Tuning curves were obtained by plotting  $A^*$  vs. the dominant spatial frequency of the corresponding checkerboard. A relatively simple fully automated algorithm used the significance level ( $p < 5\%$ ) and  $A^*$  to automatically select an appropriate region in the high spatial-frequency range on which a linear regression was performed. Comparing of several classes of algorithms, the best result was obtained by a stepwise heuristic. “VEP acuity” obtained as the inverse of the zero-extrapolated spatial frequency. Subjective VA was obtained with the automated “Freiburg Acuity Test”.

**Results:** The brief-on presentation evoked high VEP amplitudes. However, many tuning curves displayed the well-known “notch”. The analysis algorithm successfully ignored the notch, if present. The resulting VEP acuities, when compared to subjective VA, correlated by  $r=0.90$  and coincided within  $\pm 1$  octave in 95% in all normals, including the visually degraded runs. In patients, the VEP acuity and subjective VA coincided within one octave in 87.5% of the cases

**Conclusions:** The fully automated analysis avoided subjective problems in peak-trough assessment. The slightly lower agreement of VEP acuity and subjective VA in patients indicates that optical degradation does not model all pathologies well, but it was also more difficult for aged patients to operate the automatic subjective acuity test. The results provide quantitative limits to assess patients where subjective testing is impossible or problematic, e.g. in possible malingering.

## **Tactile Perception**

## End-Effect Suppression in a Multiple Choice Reaction Time Task Induced by Tactile Coactivation

Stephanie Franzkowiak, Patrick Ragert, Christine Heinisch, Tobias Kalisch,  
Claudia Wilimzig, Gregor Schöner and Hubert R. Dinse

Ruhr-Universität Bochum

patrick.ragert@rub.de

Utilizing the knowledge about brain plasticity accumulated over the last years, we suggested to design specific stimulation protocols through which it is feasible to change purposefully brain organization and thus perception and behavior. We have therefore developed a so-called coactivation protocol that applies the idea of patterned “in-vitro” stimulation protocols to human subjects. Here we report that coactivation leads to persistent changes of selection processes as revealed by multiple choice reaction time (RT) measurements during a finger selection task (adopted according to [1]). During RT measurements, an image of each hand was displayed on a monitor and a finger was selected by a visual marker. Subjects had to press the key corresponding to the selected finger on a hand-shaped 10-button keyboard. To apply coactivation, a small device consisting of a small solenoid was taped to the tip of the right middle finger (right MF). Coactivation-stimuli were drawn from a Poisson process at interstimulus-intervals between 100 to 3000 ms (average stimulation frequency one Hz, pulse duration 10 ms). Duration of coactivation is three hours. First, RT measurements were performed to obtain a stable baseline (pre). Then, the right MF was coactivated for a period of three hours to induce learning processes. Next, RT measurements were repeated (post). To assess stability and recovery, additional tests were performed 24 hours and one week later. For control, subjects were tested with an identical schedule, but without coactivation. Our results revealed that RTs were longest for the MF (d3) of each hand, but shortest for d1 and d5 of each hand, indicative for a strong end-effect. After coactivation was applied to the right MF, we observed a slight speeding up of the RT of all fingers, most probably due to practice. However, RTs of the right MF showed a highly significant shorting of RT thereby almost eliminating the end-effect for the right, but not for the left hand. Analysis of the time course of effects revealed that the effects of end-effect elimination persisted even after one week after coactivation. Our results demonstrate that learning processes induced by passive stimulation modify processing in early stages of somatosensory cortical areas that control selection processes. They further show that the end-effect is not restricted to higher cognitive processing, but emerges from lower level processing.

[1] Alegria J; Bertelson P (1970) *Acta Psychologica* 33: 36

## **Visual Cognition: General**

## The Role of Visual and Action Features in the Representation of Natural and Artifactual Objects

Viktor Bochev, Eun-Jin Sim and Markus Kiefer

Universität Ulm

`viktor.bochev@medizin.uni-ulm.de`

In an event-related potential (ERP) study, we tested the assumption of a modality-specific organization of semantic memory. Neuropsychological [1] neuroimaging [2] and electrophysiological [3] studies have shown that neuroanatomically distinct cortical regions are differentially involved in representing objects from natural categories like animals (visual association cortex) and artifactual categories like tools (parietal and frontal motor areas). Based on the neuroanatomical locations of these category-related effects, it is assumed that these regions code visual and action-related knowledge. These findings have been taken as evidence that semantic knowledge is essentially modality-specific in the sense that it is functionally and neuroanatomically grounded in the sensory and motor systems. However, the nature of the representations underlying category-related brain activation is still a matter of debate. In our study, we investigated whether category-related brain activation would depend on the feature probed in the task (visual, action-related). In a feature verification task, participants ( $n = 20$ ) had to decide whether a given verbal feature description (referring to shape or action) matches with the name of a natural or artifactual object while ERPs were recorded from 64 channels. Reaction time analysis showed that visual features were verified faster for natural objects and action features were responded to faster for artifactual objects. Congruent with the behavioral results, we found between 150 and 200 ms after onset of the target name a stronger activation over occipital areas for natural categories in the visual task. In contrast, greater fronto-central activation was obtained for artifactual categories in the action task. The present results are in line with the suggestion that visual and action-related features are differentially relevant for representing objects from natural and artifactual categories, respectively. The observed interaction between task and category supports the notion of modality-specific semantic systems.

[1] Humphreys, G. W. and Forde, E. M. E. (2001): *Behav Brain Sci*, 24, 453–509.

[2] Martin, A., & Chao, L. L. (2001): *Curr Op Neurobiol*, 11, 194–201.

[3] Kiefer, M. (2005). *J Cognit Neurosci*, 2, 1–13.

Supported by a grant from the German Research Community to MK (DFG KI 804/1–1)

**Canonicalness and Prototypicality of an Object Effects Remembering Performance**Christoph Dahl<sup>1</sup> and Daniel C. Kiper<sup>2</sup><sup>1</sup>Institute of Neuroinformatics, <sup>2</sup>Institute of Neuroinformatics, Zürich

christoph.dahl@tuebingen.mpg.de

In apparent motion displays, the remembered final position of a translating object is often displaced in the direction consistent with its implied path of motion. In literature, this forward displacement is known as representational momentum, suggesting an underlying dynamical process for perceptual representations. There is only little evidence for representational momentum in transforming targets. In a series of experiments, we investigated whether representational momentum can be generalized to form transformations such as rotation and morphing. Observers saw a series of three sequentially presented images (SOA of 250ms each), implying either depth rotation or morphing, followed by a test image, which they had to adjust to match the last presented (third) image by pressing to keys. Stimuli were rendered from 3D-generated models of either a simple cuboid model or a complex face model. Responses were measured as displacement from the position of the last presented image. For the rotation transformation we report a significant backward displacement using the complex object. There was no significant effect using the simple object. For the morphing transformation a significant backward displacement was found using the complex object and a significant forward displacement using the simple object. In addition, we examined how the observer's performance relates to the canonical perspective (for rotation transformations) and to the prototypical exemplar (for morphing transformations), which was determined by the observer before the experiment. For the rotation sequences, we calculated the distance from the last presented image to the canonical perspective for each trial and correlated it with the corresponding memory distortion. We found a negative correlation for the complex object, but no significant correlation for the simple object. Accordingly, we calculated the distance from the last presented image to the prototypical exemplar of the morphing sequences and determined the correlation with the corresponding memory distortion. Again, the complex object showed a negative correlation, whereas the simple object revealed a positive correlation. Representational momentum, as demonstrated for translating objects, has not been found for transforming objects in general. Representational momentum was only detected when simple objects were morphed. It did not occur for complex objects. Remembering the final image of rotating and morphing objects seem not related to remembering the final position of a translating target. Canonicalness and prototypicality seem to influence remembering the final shape or perspective of an object. With increasing distance from the canonical perspective or the prototypical exemplar memory distortion increases.

## Influencing Visual Recognition of Dynamic Scenes by the Use of Canonical Viewpoints

Bärbel Garsoffky, Stephan Schwan and Markus Huff

IWM—Knowledge Media Research Center, Tübingen

b.garsoffky@iwm-kmrc.de

In two experiments, the role of canonical viewpoints for the mental representation of dynamic scenes was examined. The experiments were based on the following previous findings: (a) Visual recognition of dynamic scenes is viewpoint dependent, i.e. recognition is best if the viewpoint does not change between learning and test phase and recognition becomes worse, if the viewpoints in learning and test phase differ [1]. (b) Participants show high agreement when asked which viewpoints are the best to present a visual dynamic event [2]. The latter finding suggests that the notion of canonical viewpoints holds not only for static objects, but also for dynamic scenes [3,4]. By combining these two findings, it was hypothesized that the viewpoint-dependency would be less pronounced for canonical, thereby cognitively more robust and informative viewpoints. Both experiments were based on a recognition paradigm: Participants saw a dynamic event (learning phase) and then had to decide if video stills presented the same or another event (test phase); the viewpoint on the event was either the same or differed between learning and test phase. In addition the two experiments examined if the usage of canonical viewpoints in either the learning phase (experiment 1) or the test phase (experiment 2) weaken the viewpoint deviation effect in visual recognition according to their information richness. Experiment 1 found a viewpoint deviation effect both for canonical and less canonical viewpoints in the learning phase, even if it was weaker for the canonical views. In experiment 2 the use of canonical viewpoints in the test phase had no influence on the viewpoint deviation effect. This means that the influence of canonical viewpoints on the viewpoint deviation effect is more important during storage than during retrieval.

[1] Garsoffky, B., Schwan, S. & Hesse, F. W. (2002): *JEP:LMC* 28, 1035–1050

[2] Garsoffky, B., Schwan, S. & Hesse, F. W. (2004): *CogSci2004*

[3] Blanz, V., Tarr, M. J. & Bühlhoff, H. H. (1999): *Perception* 28, 575–599

[4] Palmer, S., Rosch, E. & Chase, P. (1981): In Long, J. & Baddeley, A. (eds.), *Attention and Performance IX*, 135–151

## The Feeling of Privacy in Virtual Architecture

Patrick Jakob

Bauhaus-Universität Weimar

pajakob@gmx.de

The following hypothesis forms the base for the experiment: A room entered by an opening transfer a stronger feeling of privacy than the same room open to one side. The aim is an approach to the question of public and private spaces in virtual architecture. In an experiment it was examined if a wall in virtual architecture can transfer the same feeling of privacy and if this can be aimed by entering an opening. In real architecture a room usually consists of a base plate and a ceiling plate surrounded by a wall. With an opening the possibility to enter this room was offered. The feeling of privacy was transferred to the user by knowing the material character of the wall [1]. In contrast to real architecture the wall in virtual architecture exists. A room surrounded by a wall does not transfer the same feeling of privacy than a room in real architecture [2].

Two virtual environments were designed. The first environment consisted of a square room with a diameter of forty meters. Each side of the room became a niche which varies from a small to a big one. The second environment was based on the same room with the four niches. Three of them were closed by a wall and provided with an opening. The smallest niche was copied from the first environment without any changes. The participants were divided into two groups and assigned to one of the environments. Their task was to find a place in the environment for a private talk with their best friend. It was logged which place the participant choose. Additionally a questionnaire was developed which measured the feeling of privacy in the virtual architecture.

As a result of statistical examination the participants in the environment with the closed niches and an opening had a stronger feeling of privacy. A coherence between the character of a room and the feeling of privacy in virtual architecture was demonstrated. The hypothesis was corroborated that a room entered by an opening offered the feeling of privacy. It has also showed that the material character of the wall as a protection from visual and personal contacts was transferred in virtual architecture.

[1] van der Laan, Dom (1992): Der architektonische Raum. Fünfzehn Lektionen über die Disposition der menschlichen Behausung, Leiden

[2] Mörtenböck, Peter (2001): Die virtuelle Dimension, Architektur, Subjektivität und Cyberspace, Wien

## How does the Brain Acquire Object Categories? Diagnosticity of Visual and Action Features Drives Representation of Novel Objects

Eun-Jin Sim,<sup>1</sup> Sarah Liebich,<sup>1</sup> James Tanaka<sup>2</sup> and Markus Kiefer<sup>1</sup>

<sup>1</sup>Universität Ulm, <sup>2</sup>University of Victoria

eun-jin.sim@medizin.uni-ulm.de

Theories of semantic memory organization suggest that semantic knowledge is modality-specific in the sense that it is functionally and neuroanatomically grounded in the sensory and motor systems. Supporting evidence for this view comes from neurophysiological studies showing that neuroanatomically distinct cortical regions are differentially involved in representing objects from natural categories like animal (visual association cortex) and artifactual categories like tools (parietal and frontal motor areas). It has been suggested that the relevance of a given feature for category membership is the driving force for the emergence of such category-related effects. For that reason, we developed a set of 64 novel visual 3D objects ('nobjects') which were given nonsense names. These nobjects could be distinguished according to their shapes and actions afforded by a detail feature. Shapes and afforded actions were also given nonsense names. Feature relevance was varied by forming two classes of categories: In one class of categories, the shape of the nobject was diagnostic for category membership (visual categories), and in the other class, the afforded action was diagnostic (action categories). Across 16 training sessions, 12 participants acquired knowledge on these nobjects. In order to enhance the acquisition of action knowledge, participants had to pantomime the afforded action. After training, participants performed a feature verification task while ERPs were recorded from 64 channels. They had to decide whether a given feature name (referring to shape or action) matches with the nobject name. Reaction time analysis showed that shapes were verified faster for nobjects from visual categories while action features were responded to faster for nobjects from action categories. In the ERP analysis, we found that as early as 100 ms after name onset action categories elicited a more negative potential than visual categories over parietal regions in the action task only. Starting at 200 ms shape verification elicited a more negative potential over occipital areas, whereas action verification was associated with a more negative potential over frontal areas. Visual categories elicited a strong ERP response over occipital cortex independent of the task. Action categories evoked strong occipital activity only in the shape task. Over frontal regions, action categories were associated with a more negative potential particularly in the action task. The observed interactions between task and category suggest that feature diagnosticity drives the formation of object representation. The topography of these effects support the view that object knowledge is represented in modality-specific memory systems.

Supported by a grant from the German Research Community to MK (DFG KI 804/1-1)

## Temporal Dynamics of the Interaction Between Working Memory and Attention: A Neuronal Model of a Wisconsin-DMS-Task

Anja Stemme,<sup>1</sup> Gustavo Deco,<sup>2</sup> Astrid Busch,<sup>1</sup> Werner X. Schneider,<sup>1</sup> Joe Krummenacher<sup>1</sup>  
and Josef Zihl<sup>1</sup>

<sup>1</sup>LMU München, <sup>2</sup>Universitat Pompeu Fabra, Barcelona

stemme@psy.uni-muenchen.de

The poster presents the work conducted in the context of a DFG-project whose aim is to construct an explicit mechanistic model of the cognitive processes involved in visual information processing within the Prefrontal Cortex (PFC). For this purpose, we developed a Wisconsin-like “What-Where-Delayed-Matching-to-Sample” - task to be accomplished by human subjects and a neurodynamical model.

In this task a pair of visual stimuli, separated by a fixed delay, is presented to the subjects who have to indicate whether these stimuli match with respect to one of two possible feature dimensions: Same object (“What” Rule) or same position on the screen (“Where” Rule). The valid rule changes without notice (Wisconsin-like paradigm) at random intervals. The analysis of the response times of 12 healthy subjects revealed an influence of the irrelevant stimulus dimension; hence, answers were fastest in cases where both stimulus dimensions (object and position) matched (426 ms). In trials with a match in the relevant stimulus dimension (i.e., “object-rule with object-match” or “space-rule with space-match”), reaction times (RTs) were slower (477 ms) and still increased for trials with a match in the irrelevant stimulus dimension (i.e., “object-rule with space-match” or “space-rule with object-match”; 534 ms).

The neuronal model is based on the framework of Brunel and Wang ([1]) and constitutes a biophysically detailed model. The model consists of various pools of neurons: Two rule pools, responsible for the maintenance of the active rule, and two times two memory pools reflecting the stimulus dimensions “object” and “space”. The identification of pools is based on neurophysiological studies recording single cell data in behaving monkeys. This model was used to simulate the tasks described above. The simulations allowed the replication of the RT distributions as generated by the subjects. Thus, the model allows to gain detailed insights into the basic processes involved in visual information processing and will later be used to predict patients behaviour in the named task.

The poster presents the experimental setup used and depicts the results obtained by the human subjects. Further on, it shows how the neurodynamical model is setup to do the same task as the human subjects. Finally, the simulation results in terms of spiking rates of the participating neuronal pools as well as response times achieved by the model are outlined.

[1] Nicolas Brunel and Xiao-Jing Wang, *Journal of Computational Neuroscience*, 11(1):63–85, 2001

DFG- Project: FOR 480/1–1

## Curvature Versus Pointing in Visual Search

Aylin Thiel, Ralf Goertz and Ruxandra Sireteanu

Universität Frankfurt

aylin.thiel@gmx.de

In a visual search tasks, subjects search for a target item amidst a number of distracting items. If the time needed to complete the search is independent of the number of distractors, the search is said to be parallel. If the time increases roughly linearly with the number of distractors, the search is called serial. Item for which search is parallel are identified as “features”. However, it is not always clear what a feature is and under which conditions a certain stimulus property can become a feature. In extensive experiments we wondered whether items which initially do not elicit parallel search can become “features” after extended practice. Furthermore, we were interested to know whether “featurization” transfers between targets differing from distractors in shape, but sharing the direction of pointing, or differing in the pointing direction, but similar in shape.

We used two different kinds of stimuli: wedges and semicircles, both pointing either upwards or downwards. A target was defined as a single item differing in shape or pointing in the opposite direction from the distractors. In a first experiment, we constantly used wedge symbols as distractors and varied the target items in three different ways: the target item could either be a wedge symbol pointing in the opposite direction from the distractors, or a semicircle pointing either in the same or in the opposite direction. This experiment was carried out with 21 subjects.

We found that the wedge targets require longer reaction times than both semicircle targets. Furthermore, reaction times did not differ between the two different semicircle targets. In all tasks, responses were significantly faster for 8 items than for 16 items, with a higher increase in non-target trials. Presence of semicircle targets was detected significantly faster than their absence, whereas there was no target effect for wedges targets.

These results show that information about “curvature” is much more important than the “pointing” direction. Differences in the shape of the target appear to be more relevant for the learning process than differences in luminance distribution.

## **Visual Cognition: Objects and Faces**

## Similar Cortical Correlates Underlie Visual Object Identification and Orientation Judgment

Christian Altmann,<sup>1</sup> Wolfgang Grodd,<sup>2</sup> Zoe Kourtzi,<sup>3</sup> Heinrich H. Bühlhoff<sup>3</sup> and Hans-Otto Karnath<sup>1</sup>

<sup>1</sup>Center of Neurology, Hertie Institute for Clinical Brain Research, University of Tübingen, Tübingen, Germany,

<sup>2</sup>Section for Experimental MR of the CNS, Department of Neuroradiology, University of Tübingen, Tübingen, Germany, <sup>3</sup>MPI for Biological Cybernetics, Tübingen, Germany

christian.altmann@gmx.net

Early neuropsychological observations described a patient that presented with impaired knowledge of object orientation but spared object recognition skills [1]. Similar cases exhibiting this peculiar dissociation have been reported since Best's discovery. Interestingly, a recent study observed a double dissociation between object identification and orientation [2]. These findings were interpreted as evidence that separate cortical centers underlie visual object recognition and processing of spatial features. Accordingly, visual object perception has been suggested to follow two different routes in the human brain: a ventral, view-invariant occipital-temporal route processes object identity, whereas a dorsal, view-dependent occipital-parietal route processes spatial properties of an object. Using fMRI, we addressed the question whether these routes are exclusively involved in either object recognition or representation of object orientation. To this end, we presented subjects with images of natural objects and involved them either in an object identification or object orientation judgment task. For both tasks we observed activation in ventro-temporal as well as parietal areas bilaterally, with significantly stronger responses for the orientation judgment in ventro-temporal areas. Our findings suggest that object identification and orientation judgment do not follow strictly separable cortical pathways, but rather involve both the dorsal and the ventral stream.

[1] Best, F. (1917): *v.Graefes Archiv für Ophthalmologie*, 93, 49–150.

[2] Turnbull, O.H. (1997): *Neuropsychologia*, 35, 153–163.

This work was supported by grants from the Deutsche Forschungsgemeinschaft (Ka 1258/2–3, SFB 550-A4) awarded to H-O K and by the Max Planck Society. We would like to thank Hans-Joerg Mast for technical assistance with the data acquisition.

## Gaze Following in Humans: Experiments on its Precision

Simon W. Bock, Peter W. Dicke and Peter Thier

Hertie-Institute for Clinical Brain Research, Department of Cognitive Neurology, Tübingen

sbock@uni-tuebingen.de

Joint attention can be thought of as the basis of understanding other individuals as intentional beings, whose attention to outside objects may be shared, followed into, and directed in various ways. Gaze following, a prerequisite of joint visual attention, may be paraphrased as “looking where someone else is looking”. Many studies have tested this ability in human and nonhuman primates, but as yet no quantitative measure of gaze following judgements has been obtained. However, knowledge of the precision of gaze following is indispensable for the understanding of the underlying neural circuitry. In the present study we therefore investigated the capability of adult human ‘receivers’ to direct their gaze to the position of one out of many objects, which was defined by the gaze of an adult human ‘sender’. Sender and receiver faced each other at 100 cm distance, looking at the opponent through a ring of 90 pinhead objects (object size  $0.44^\circ$ , object spacing  $1.03^\circ$  at  $15^\circ$  visual angle eccentricity, 50cm distance). The sender looked at object positions following a given pseudo-randomized order and the receiver reported the perceived target of the sender’s gaze. Receivers’ performance was tested under different conditions (binocular vs. monocular vision, both sender’s eyes visible vs. one eye visible only, receiver following sender’s saccade to object location vs. static gaze of sender to object location, real sender vs. computer presented image of sender). In order to assess gaze following precision, we considered pooled responses of nine receivers. Their deviations from the sender’s target are normally distributed and show surprisingly high precision (standard deviation  $3.16^\circ$  visual angle over all conditions). Experiments comparing binocular vision (SD  $3.03^\circ$ ) with monocular vision (SD  $2.97^\circ$ ) of the receiver did not reveal statistically significant differences. If only one eye of the sender was visible, receivers tended to be slightly less accurate (SD  $3.65^\circ$ ). On the other hand, receivers appeared to be more accurate if they were allowed to see the sender’s saccade to the target location as compared to a condition in which the sender’s gaze was made accessible only after the saccade had been carried out (SD  $2.72^\circ$  vs. SD  $3.03^\circ$ ). Precision of receivers was similar when the real human sender was replaced by a computer presented image of the sender (SD  $3.16^\circ$ ). In summary, human gaze following is not only very precise but, moreover, also surprisingly robust to manipulations of the sender cues available for guiding the receiver’s eyes.

The authors would like to thank Dr. C. Falch for valuable suggestions.

## Sensitivity to Changes in Identity, Caricature and Sex in Face Recognition

Isabelle Bühlhoff

MPI for Biological Cybernetics, Tübingen

`isabelle.buelthoff@tuebingen.mpg.de`

It is known that we are quite accurate at judging the sex of unfamiliar faces [1]. Furthermore sex categorization is performed more rapidly, on average, than familiarity or identity decisions [2]. In one of our recent studies on face perception, with unfamiliar faces [3] we were surprised to find that discrimination performance was much lower for faces differing in sex quality than when the facial features were morphed between two identities. Here, we investigated if this observation holds also for familiar faces. The motivation for this series of experiments was to find out if memory of familiar faces was showing similar differences; participants being more inaccurate when they had to remember the specific femininity or masculinity of a well known face than when identity-related changes of facial features were involved. Participants had to identify the veridical faces of familiar work colleagues among ten distractor faces that were morphing variations of the original faces. Distractor faces varied either in identity, caricature or sex. In the identity face sets, distractor faces were morphs between the original face and unfamiliar faces mixed in different proportions. In the caricature face sets, distractors were different caricatures of the original face. Finally, in the sex face sets, distractor faces were different feminized and masculinized versions of the veridical face. Participants performed best when the original face was presented among identity distractors. They had a tendency to choose feature enhancing caricatures over the original faces in caricature sets. Participants were very poor at finding the original faces in the sex sets. Generally our findings with unfamiliar faces show that sex-related changes in facial features are less obvious to the observers than identity-related changes. Furthermore our study on familiar faces suggests that we do not retain sex-related facial information in memory as accurately as identity-related information. These results have important implications for models of face recognition and how facial features are represented in the brain.

- [1] Bruce V., Burton A.M., Hanna E., Healey P., Mason O., Coombes A., Fright R., Linney A. (1993), *Perception*, 22, 131–52
- [2] Bruyer, R., Galvez, C. & Prairial, C. (1993), *Brit. J. Psychol.*, 84, 433–441
- [3] Bühlhoff, I. and Newell, F.N. (2004), *Vis. Cogn.*, 11, 823–855

## Representation of Shape and Texture Changes in Novel 3D Objects

Theresa Cooke, Christian Wallraven and Heinrich H. Bülthoff

MPI for Biological Cybernetics, Tübingen

`theresa.cooke@tuebingen.mpg.de`

Object recognition and categorization researchers have proposed that similarity may offer an organizational principle for the representation of objects in the brain and have demonstrated that such a principle is capable of supporting the distinct tasks of object recognition and categorization. But what is it that makes two objects perceptually similar? And what role does sensory modality play in shaping perceptual similarity?

As a first step towards addressing this question, we used computer graphics modeling techniques to create novel, three-dimensional objects that varied parametrically in shape and texture. Five levels of ‘texture blurring’ and five levels of ‘shape blurring’ were selected as the creation parameters for 25 objects. In a psychophysical experiment, fifteen human subjects viewed pairs of these objects and rated the similarity between them on a seven-point scale. Similarity ratings of the objects were also generated from a set of standard computer vision techniques based on pixel-wise differences, cross-correlations, edge images, and Gabor jets. Multidimensional scaling (MDS) was then performed on the similarity data. MDS on the perceptual similarities led to an optimal two-dimensional ‘perceptual map’ of the stimuli whose dimensions corresponded to texture and shape blur. Furthermore, the configuration of the stimuli in this perceptual space closely resembled the configuration of the stimuli in the space defined by the creation parameters. Interestingly, the perceptual distance between stimuli along the shape dimension decreased as the amount of texturing increased, revealing an interaction between shape and texture in the visual similarity judgments. The five texture levels were quite evenly spaced in the perceptual map, whereas the spacing between shape levels hinted at the formation of two distinct groupings based on shape. Comparisons against computational measures revealed that simple pixel-based and correlation-based similarities were closest to the perceptual similarities, whereas similarities based on edge detection and Gabor jets overemphasized the importance of the texture dimension compared to perceptual similarities.

This study establishes an experimental paradigm for generating information-rich maps of a set of stimuli and comparing maps derived from different similarity measures. Perceptual similarity measures can be compared against computational similarity measures generated using various object features, providing an opportunity to evaluate the perceptual validity of those features. In future work, this approach will be used to compare perceptual maps based on visual similarities against maps based on haptic similarities and in studies of multimodal categorization.

## Motor Representations in Visual Object Recognition

Markus Graf,<sup>1</sup> Hannah Helbig<sup>2</sup> and Markus Kiefer<sup>3</sup>

<sup>1</sup>MPI für Kognitions- und Neurowissenschaften, Arbeitsbereich Psychologie, München,

<sup>2</sup>MPI for Biological Cybernetics, Tübingen, <sup>3</sup>Universität Ulm, Abteilung Psychiatrie III

markus.graf@cbs.mpg.de

It has been proposed recently that object recognition relies on coordinate transformations, i.e. on similar processes as visuomotor control [1]. Thus, the two visual streams involved in object recognition and object-directed action may rely on common computational principles, which provides the possibility for interactions between the two streams. Existing behavioral and neurophysiological findings suggest that viewing manipulable objects automatically potentiates possible actions [e.g., 2,3]. We investigated whether action knowledge has a functional role in visual object recognition. More specifically, we used a priming paradigm to test whether objects are recognized better when viewed after another object which affords congruent as compared to incongruent motor interactions. Two grey-scale pictures of artifactual manipulable objects were presented sequentially (tools, kitchen utensils, musical instruments). Subjects were required to name the objects. The stimuli were briefly presented and masked. The presentation time of the second object was adjusted individually in an adjustment phase so that naming accuracy approached 80%. In the congruent condition both objects afforded a similar motor interaction, and dissimilar motor interactions in the incongruent condition. Stimulus pairs in both conditions were matched for baseline naming accuracy, word frequency, word length, as well as visual and semantic similarity. We found that naming accuracy was higher in the congruent than in the incongruent condition (Experiment 1 and 2). This action congruency effect indicates that object naming is facilitated by a previous activation of an appropriate action representation. In two further experiments we investigated the nature of the representations underlying the action congruency effect. The effect was reduced or absent when the prime stimulus was inverted (Experiment 3), and when the prime was presented as a word (Experiment 4). This suggests that the action representations underlying this congruency effect are closer to specific (parameterised) motor representations than to abstract semantic representations. Overall, the findings suggest that the recognition of manipulable objects involves not only visual but also action representations. These are not abstract semantic representations, but are relatively close to motor representations.

[1] Graf, M., Kaping, D. & Bühlhoff, H.H. (in press): *Psychol Science*.

[2] Tucker, M., & Ellis, R. (1998): *JEP:HPP* 24, 830–846.

[3] Chao, L.L., & Martin, A. (2000): *NeuroImage* 12, 478–484.

This work was supported by the Max Planck Society, the European Commission (IST 2000–29375 COGVIS), and by the German Research Community (DFG Ki 804/1–1).

## Old/Young Woman: ERP Correlates of the Disambiguation of an Ambiguous Face-Figure

Christine M. Hein, Michael Bach and Jürgen Kornmeier

Uni-Augenklinik Freiburg

post-fuer-die-tine@t-online.de

**Background.** Normally we perceive the world as stable. With ambiguous figures, however, a sudden unavoidable perceptual reversal occurs, although the figure itself stays unchanged. We recently found a very early (120 ms) occipital ERP component to endogenous orientation reversals of a Necker lattice. It was followed by a second ERP component at 300 ms at the central electrode position. With exogenously induced reversals of unambiguous lattice-variants a similar, but earlier (50 ms), centrally distributed ERP signal was found; no early occipital signal could be detected, however. In the present experiment we were interested if perceptual reversals of a “semantic” type of ambiguous figures, namely the old/young-woman-figure, provides a similar ERP pattern as reversals of the lattice-stimuli. **Methods.** In separate experiments an ambiguous old/young woman figure and two unambiguous variants were presented repeatedly for 800 ms after blank periods of 40 ms. The unambiguous stimuli changed randomly in 50% of the cases. Our subjects compared successive stimuli and indicated in different blocks whether they perceived a “reversal” or “stability” (same percept). EEG was recorded from 13 channels for 13 subjects; stimulus-onset was used as time reference for averaging. The ERP-differences (reversal minus stability) were analyzed. **Results.** Unambiguous stimuli. A sequence of ERP correlates of the exogenous reversal can be detected, beginning with a first correlate at temporal and central electrodes at 160 ms. It is followed by a negativity at occipital locations at around 220 ms and a negativity at the parietal/temporal locations at 280 ms. Ambiguous old/young woman. Only one broad negativity could be detected. It is most prominent at temporal locations at around 300 ms and includes central, frontal and parietal locations. No counterpart to the early occipital component, found with the Necker cube, exists with the old/young woman. **Discussion.** In case of unambiguous stimuli, the sequence of reversal correlates may reflect recurrent activity of the neural network during a reversal of object representation, including face-specific areas at temporal locations. In case of ambiguous stimuli the temporal jitter of endogenous reversal instances may have caused the distinct ERP components to merge into one broad negativity. The prominent temporal/central reversal correlate at 300 ms for both stimulus types indicates similar perceptual processing at this step. Disambiguation of the ambiguous input must have already been completed. The missing latency-shift from unambiguous to ambiguous stimuli here compared to the lattice-stimuli may suggest that processing of faces is more efficient compared to 3D-space-objects.

[1] Kornmeier J, Bach M (2004) *Psychophysiology* 41:1–8.

[2] Kornmeier J, Bach M (2004) *Vision Research*, in press

Support by the Deutsche Forschungsgemeinschaft (BA 977–11) is gratefully acknowledged

**The Time Course of Object Repetition Effects in Human Visual Cortex**

Sven P. Heinrich and Kalanit Grill-Spector

Stanford University

sven.heinrich@stanford.edu

Repetition of the same object typically yields a reduction of the fMR signal (adaptation) over object-selective cortex. EEG studies report repetition effects, but results vary from decreased cortical activity starting at around 200 ms (or later) to increased activity around 170 ms. Since fMRI does not yield information about the time course and the EEG effects are variable it is difficult to judge whether these repetition-effects are driven by bottom-up or top-down processing. Here we measured EEG and fMRI activation in the same subjects and for the same stimuli consisting of repeated and non-repeated animals.

Six subjects participated in 61-channel EEG experiments yielding time course information. For two of these subjects we estimated the location of the neural activity from fMRI data and EEG dipoles. Subjects viewed repeated images of the same dog (18 repetitions), the same bird (18 repetitions), or non-repeated images of dogs and birds. Images were presented in an event-related design for 150 ms followed by a blank interval of 1850 ms. Conditions appeared in pseudo-randomized order and were intermixed with scrambled images and fixation trials. Subjects participated in 10 blocks of 108 trials and were asked to categorize images to “bird”, “dog” or “other” while fixating. In each block, a different set of images was used. All subjects showed an increase for repeated images peaking around 150 ms, that is well explained by bilateral occipito-temporal dipoles. In later time intervals, there was a high inter-subject variability. Consistent with earlier studies, but in contrast to the EEG, fMRI activations in object-selective cortex were reduced for repeated images. One possible reason for the difference between EEG and fMRI may be that the early increase seen in the EEG is lost due to temporal integration in fMRI. Based on the EEG time course we conclude that at least some repetition effects occur during bottom-up processing.

Supported by DFG (HE 3504/2–1) and NSF (345920)

**Source Localization (LORETA) of Early Stages of Face Processing: An ERP Study**

Martin Herrmann, Andreas Muehlberger and Andreas Fallgatter

Universität Würzburg

martin.herrmann@mail.uni-wuerzburg.de

Face processing can be measured with the event-related potential (ERP) N170, a bilateral negative occipital-temporal component peaking around 170 ms (N170) with increased amplitudes to faces. Recent studies [1,2] reported that even at 100 ms a face specific brain activity could be found. While the origin of N170 was repeatedly described in the fusiform gyrus, the neural origin and the functional meaning of the P100 effect remains unclear. This study investigated 72 subjects while viewing faces and control stimuli. Both components, the P100 and the N170, showed significantly higher amplitudes to faces compared to the control stimuli. LORETA source localization revealed significantly stronger brain activity for faces than for control stimuli in the gyrus fusiformis for the peaks of both components. For the N170, but not for the P100, we additionally found extensive activation in the occipital, temporal and parietal cortex. These results underscore the role of the gyrus fusiformis in the processing of face and contradict the assumption that the P100 effect for face processing was generated in the occipital cortex.

- [1] Herrmann M.J., Ehlis A.-C., Ellgring H. and Fallgatter A.J, Early stages (P100) of face perception in humans as measured with ERPs, *J Neural Transm* (in press)
- [2] Itier R.J. and Taylor M.J., N170 or N1? Spatiotemporal differences between object and face processing using ERPs, *Cereb Cortex* 14 (2004) 132–42.

**Measuring the Effectiveness of Individually Adaptive Computer Based Training (CBT)**

Franziska Hofer and Adrian Schwaninger

University of Zurich, General Psychology (Cognition)

fhofer@allgpsy.unizh.ch

The role of human factors in airport security has become more and more important in recent years. Although airport security technology has evolved remarkably in the last two decades, the human operator is still the last decision instance when judging a bag as harmless or dangerous. Thus, the effectiveness of this technology is limited to the personnel who operate it. The task of an airport security screener is to recognize all kinds of prohibited items in X-ray images of passenger bags. This is a challenging object recognition task in which the screener has to know what prohibited objects look like from different viewpoints, in different bag complexities and with different degrees of superimposition by other objects in the bag. In this study we evaluated a computer based training system (X-Ray Tutor) for improvised explosive devices (IEDs). A Latin Square counterbalanced design was used to measure CBT effectiveness in a longitudinal study of six months. 72 airport security screeners participated in this study. Four tests with IEDs that had not been seen before were applied. We found large improvements of detection performance suggesting a strong generalization of learned to new IED threat items. Moreover, substantial reduction in response time was found, which shows that CBT is a very effective tool for increasing the effectiveness as well as the efficiency in aviation security screening.

This research was financially supported by Zurich Airport Unique.

## Learning of Biological Motion: Combining fMRI and Theoretical Modeling

Jan Jastorff,<sup>1</sup> Zoe Kourtzi<sup>2</sup> and Martin A. Giese<sup>1</sup>

<sup>1</sup>ARL, Dept. of Cognitive Neurology, University Clinic Tübingen,

<sup>2</sup>MPI for Biological Cybernetics, Tübingen

jan.jastorff@uni-tuebingen.de

Learning has been proposed to contribute to the recognition of biological movements. We have investigated the neural correlates of such learning processes using event-related fMRI adaptation. This paradigm entails repeated presentation of a stimulus resulting in a decrease of the fMRI response, compared to stronger responses after a change in a stimulus dimension. This stronger response indicates sensitivity of the measured neural populations to this changed dimension. Novel biological movements were generated by linear combination of triples of prototypical trajectories of human movements. Subjects had to discriminate between identical, very similar, moderately similar and completely different point-light stimulus pairs. The difficulty of the discrimination task could be precisely controlled by choosing appropriate weight vectors of the prototypes in the linear combinations. Subjects were able to learn the discrimination between these novel biological motion stimuli. The fMRI results indicate that several visual areas are involved in this learning process. More specifically, lower-level motion-related areas (hMT+/V5 and KO/V3B) show an emerging sensitivity for the differences between the discriminated stimuli, and higher-level areas (STS and FFA) show an increase of sensitivity after training. In addition, we find an overall reduction of the BOLD activity after training. Our present work focuses on modeling these BOLD signal changes during discrimination learning using a hierarchical physiologically-inspired neural model for biological motion recognition [1]. We show that learning of novel templates for complex movement patterns, encoded by sequences of body shapes and optic flow patterns, can be implemented by hebbian learning. Our model combines competitive and time-dependent hebbian plasticity in order to establish new spatio-temporal templates exploiting physiologically plausible local learning rules. Our results demonstrate that these mechanisms can account for the emerging sensitivity for novel movement patterns observed in fMRI. We conclude that our model provides a first step to formulate and test quantitative hypotheses about the neuronal plasticity mechanisms that underlie the learning of complex biological and non-biological movement patterns.

[1] Giese, M.A. and Poggio, T. (2003): *Nat. Rev. Neurosci.* 4, 179–192

DFG, Deutsche Volkswagenstiftung, Herman-lilly-Schilling Foundation and Max Planck Society

**Moving Objects in Ultra-Rapid Visual Categorisation Result in Better Accuracy, but Slower Reaction Times than Static Presentations**

Holle Kirchner,<sup>1</sup> Quoc C. Vuong,<sup>2</sup> Simon J. Thorpe<sup>1</sup> and Ian M. Thornton<sup>2</sup>

<sup>1</sup>CerCo, CNRS Toulouse, <sup>2</sup>MPI for Biological Cybernetics, Tübingen

`holle.kirchner@cerco.ups-tlse.fr`

Ultra-rapid categorisation studies have analysed human responses to briefly flashed, static natural scenes in order to determine the time needed to process different kinds of visual objects. Recently, Kirchner and Thorpe reported that reaction times can be extremely fast if subjects are asked to move their eyes to the side where an animal had appeared. Accuracy was remarkably good with the fastest reliable saccades occurring in only 130 ms after stimulus onset. Vuong and colleagues in a 2AFC task with apparent motion displays and manual responses further showed that humans can be detected more easily than machines. In the present study we combined the two approaches in order to determine the processing speed of static vs. dynamic displays. In blocked conditions, human subjects were asked to detect either an animal or a machine which in half of the trials were presented either static or in apparent motion. On each trial, an animal and a machine were presented simultaneously on the left and right of fixation, and the subjects were asked to make a saccade or to press a button at the target side. Manual responses and saccadic eye movements both resulted in good accuracy, while reaction times to animals were significantly faster than to machines. Only saccadic eye movements showed a clear advantage of dynamic over static trials in accuracy, but the analysis of mean reaction times pointed to a speed-accuracy trade-off. This might be explained by different response modes as seen in the latency distributions. We conclude that form processing can be improved by stimulus motion, but the speed of this process can be observed much more directly in eye movement latencies as compared to manual responses.

This work was supported by the European Research Training Network “Perception for Recognition and Action”.

**ERP Correlates to Perceptual Reversal of Ambiguous Figures: Effects of ISI Durations**

Jürgen Kornmeier and Michael Bach

Uni-Augenklinik Freiburg

juergen.kornmeier@uni-freiburg.de

**Background.** Ambiguous figures lead to unstable percepts, which reverse spontaneously during prolonged inspection. The EEG can trace the time course of the neural mechanisms. To extract event related potentials (ERPs) from the EEG via averaging, a precise time reference is necessary but difficult to acquire in case of endogenous events. In previous experiments with the Necker lattice, discontinuous stimulus presentation (800 ms on-time, 400 ms inter stimulus interval, “ISI”) synchronized reversal instances to stimulus onset. Stimulus onset provided a better time reference for EEG-averaging and thus an improved temporal resolution of the underlying neural processes. To better approximate our presentation mode to continuous observation, we here reduced ISIs from 400 ms to 40 ms (less than a blink). How do the ERPs to perceptual reversals depend on the ISI? **Methods.** An ambiguous “Necker lattice” appeared repeatedly for 800 ms after breaks of 40 ms. In addition, unambiguous lattices (with depth cues) were presented, with same or reversed orientation. Our subjects compared successive stimuli and indicated whether they perceived a “reversal” or “stability” (two successive stimuli in the same orientation) at stimulus-onset. EEG was recorded from 13 channels in 9 subjects; the difference traces between reported “reversal” and “stability” were analyzed. **Results.** 40-ms-ISI: Highly significant ( $p < 0.001$ ) central ERP negativities were seen for both, ambiguous and unambiguous lattices. In case of the unambiguous lattice, this negativity occurs about 50 ms earlier compared to the Necker lattice. In case of the Necker lattice an additional early occipital positivity occurred at 120 ms ( $p < 0.01$ ). Similarities across ISIs: In case of the Necker lattice the early occipital positivity occurs with both ISIs. Differences between ISIs: For both stimulus types, an occipital/parietal negativity at 250 ms is found with 400-ms-ISI whereas a central negativity at 300 ms with higher amplitudes (factor of 4) is found for the 40-ms-ISI. **Discussion:** We interpret the earliest occipital signal as a direct correlate of the perceptual instability. This instability is resolved during early visual processing, leading to a delay of all subsequent perceptual processes. This is mirrored in the increased latencies of the ERP correlates. The negativity, subsequent to the early positivity, may indicate “normal” perceptual processing after disambiguation has been resolved. In case of shorter ISIs the increased latency and the central distribution of this negativity may reflect additional preparation of the immediate response in contrast to the longer ISI, where the response was executed after stimulus-offset.

Support by the Deutsche Forschungsgemeinschaft (BA 977–11) is gratefully acknowledged.

## The Flower Pot Illusion

Welf A. Kreiner

Universität Ulm

welf.kreiner@chemie.uni-ulm.de

A visual illusion is reported which in its simplest form may be described in the following way: On a white surface an area, e.g., an upright rectangle, is surrounded by a black border line. At its bottom, the interior of the rectangle is black, while on moving towards the upper end, the black gradually turns into grey and finally fades away into white. When the width of the contour line is chosen such that it approaches the resolution limit of the eye, the rectangle seems to change its shape, giving the impression of being wider at the top and narrower at the bottom. Different illusion effects are achieved when the shading changes in a symmetric fashion, i.e., when the black (or the white) region is half way up the rectangle. Wave-like distortion of the contour is observed when the black and white pattern is oriented parallel to the diagonal. Various effects can be observed when arranging several rectangles adjacent to each other, one of them reminiscent of the Münsterberg / café wall illusion. The illusion disappears as soon as the contrast of the border line is such that its visibility is not influenced by the hue of the adjacent area, or it exhibits a too low contrast. The illusions observed are traced back to two main effects:

1. An area exhibiting fine structure appears to be larger than an unstructured area [1].
2. The irradiation illusion [2]

Both are explained from the fact that the capacity  $C$  of any data processing system is limited [3]. Assuming the visual system to be such a data processing device, one can draw the conclusion that the attentional angle has to be narrowed down in order to increase resolution [4], e.g., when narrow lines have to be resolved. In case of the irradiation illusion, it is assumed that the size of the attentional angle corresponds more or less to the size of the white area, from which information is expected. Further, it is assumed that the conspicuity range (solid angle, from which information is obtained) is projected onto an internal visual memory screen of constant size. Then, so the explanation, in the case of a small angle, the perceived size of the object is somewhat increased and vice versa.

Both effects are demonstrated on rectangles as described above, but seen on backgrounds shaded in different ways.

- [1] Gouye, P. (1703): *Histoire de l'Acad. Royale des Sciences*. Année 1700, 10–12
- [2] von Helmholtz, H. (1867): *Handbuch der Physiologischen Optik*. Voss (Leipzig)
- [3] Shannon, C.E. (1948): *The Bell System Technical Journal* 27, 379–423
- [4] Kreiner, W.A. (2004): *Z. Phys. Chem.* 218, 1041–1061

## **Abstraction of Physical Properties from Complex Object Interactions—The Case of Elasticity**

Manfred Nusseck, Roland Fleming and Heinrich H. Bülthoff

MPI for Biological Cybernetics, Tübingen

`manfred.nusseck@tuebingen.mpg.de`

For humans it is useful to be able to visually infer an object's physical properties (e.g. weight, hardness or elasticity). One potentially important source of information is the way that an object moves and interacts with other objects in the environment. For example, the way that a ball bounces could inform us about its elasticity.

There have been several explorations of what are the necessary and typical visual cues in a bouncing event. However, in most previous work the stimuli consisted of a ball bouncing repeatedly on a simple horizontal plane. Warren, Kim and Husney (1987) showed that under these circumstances, there are at least three heuristic cues to elasticity: relative height, relative period and relative velocity of the bounces. We wanted to test whether the visual system can interpret more complex bouncing events in which these simple cues are not present in the display. Can subjects abstract something more sophisticated from the trajectories of bouncing objects, or must they rely on these simple heuristics?

To test this, we used the *Virtools Physics Pack* to simulate a ball falling through an array of horizontal cylindrical pegs housed in a vertical box. The ball fell from a random location above the box, bounced a number of times and finally fell out of a hole at the bottom. This stimulus allows us to completely randomize the duration that the ball needs to get through this box, the number of collisions, the velocity of the ball and the height of the rebounds, because the angle of the collisions with the pegs is always different.

Subjects performed an elasticity matching task. Subjects were presented with two peg-boxes simultaneously. The left-hand box was the Test ball, whose elasticity was chosen at random by the computer. The right-hand box contained the Match ball, whose elasticity could be adjusted by the subject. The subject's task was to adjust this elasticity of the Match ball until it appeared to have the same behavior as the Test ball.

The results show that subjects generally performed poorly in this task, despite dramatic variations in the elasticity of the ball. However, we found large individual differences, in which some subjects were able to perform the task above chance levels. Our results suggest that subjects normally rely on simple heuristics to estimate elasticity (e.g. bounce height), which, by design, were eliminated from our stimuli. Further research is needed to investigate which additional, complex cues were used in some cases to abstract the behaviour of the ball.

**Motion Matters: Facial Motion Improves Delayed Visual Search Performance**

Karin Pilz, Ian M. Thornton and Heinrich H. Bühlhoff

MPI for Biological Cybernetics, Tübingen

`karin.pilz@tuebingen.mpg.de`

In the last few years there has been growing interest in the role that motion might play in the perception and representation of facial identity. In separate experiments, we explored how learning is affected by two different types of movement: the non-rigid motion of the face, typically associated with expression and communication, and the equally familiar rigid motion that occurs whenever a person approaches you in depth. Traditional old/new recognition tasks have previously yielded mixed results in the context of facial motion. Therefore, we decided to use a delayed visual search paradigm. Observers were familiarised with two target individuals, one seen in motion, the other via static snapshots. All images were non-degraded and consisted of video sequences in Experiment 1 and 3D heads on a walking avatar body in Experiment 2. After a delay of several minutes, observers were asked to search for their targets in static search arrays. Crucially, during this test phase, the static search arrays were identical regardless of how the face was first seen during learning. Nevertheless, in both experiments, faces that were learned in motion were found more quickly and more accurately than faces that had been learned from snapshots. These findings provide further evidence that facial motion can affect identity decisions, both in the presence of intact form cues and across extended periods of time.

**Perceived Dynamics of Static Images Enables Emotional Attribution**Arseny A. Sokolov,<sup>1</sup> Alexander Sokolov<sup>2</sup> and Marina Pavlova<sup>1</sup><sup>1</sup>University Hospital of Tübingen, <sup>2</sup>University Hospital of Ulm

arseny.sokolov@student.uni-tuebingen.de

Perception of intentions and dispositions of others is an essential ingredient of adaptive daily-life social behaviour. Dynamics of moving images allows for veridical perception of social attributes (e.g., [1,2]). Anecdotal observations in art, science, and popular culture indicate that dynamic imbalance can be revealed in static images [3]. Here we ask whether perceived dynamics of abstract figures is related to emotional attribution. Participants first estimated instability of geometric figures rotated in 15-deg steps in the image plane, and then ranked the intensity of basic emotions that can be ascribed to the figures. No substantial link occurred between the deviation of the figures from the vertical orientation and perceived instability. Irrespective of shape, the strong positive correlation was found between negative emotions and perceived instability. By contrast, the positive emotions were inversely linked with the figure deviation from vertical orientation. The work demonstrates for the first time that dynamics conveyed by static images enables specific emotional attribution, and agree well with the assumption that neural networks for production of movements and understanding the dispositions of others are intimately tied [4]. The findings are also of importance for exploring the ability to reveal social properties through dynamics in abnormal development, for example, in patients with early brain injury [5] or autistic spectrum disorders [6].

[1] Heider, F. and Simmel, M. (1944): *Am. J. Psychol.* 57, 243–259.

[2] Runeson, S., and Frykholm, G. (1983): *J. Exp. Psychol. General* 112, 585–615.

[3] Cutting, J. (2002): *Perception* 31, 1165–1193.

[4] Gallese, V. et al. (2004): *Trends Cogn. Sci.* 8, 396–403.

[5] Pavlova, M. et al. (2005): *Cereb. Cortex* 15, in press.

[6] Blake, R. et al. (2003): *Psychol. Sci.* 14, 151–157.

### Characterization of Face Processing in Patients with Congenital Prosopagnosia by Simultaneous MEG/EEG Recordings

Joachim E. Weber,<sup>1</sup> Claus-Christian Carbon,<sup>2</sup> Martina Grueter,<sup>3</sup> Tilmann H. Sander,<sup>4</sup>  
Thomas Grueter,<sup>3</sup> Gabriel Curio,<sup>1</sup> Lutz Trahms<sup>4</sup> and Andreas Lueschow<sup>1</sup>

<sup>1</sup>Charité—University Medicine Berlin, <sup>2</sup>Institute of Psychology, Freie Universität Berlin,  
<sup>3</sup>University of Münster, <sup>4</sup>Physikalisch-Technische Bundesanstalt, Berlin

j.e.weber@t-online.de

Congenital prosopagnosia (cPA) is assumed to be a severe and selective impairment in face processing from early childhood. 13 individuals with cPA were compared to a group of matched controls by means of simultaneous measurements of MEG and EEG. Evoked responses were recorded in a viewing task with a sequence of faces and houses. The EEG/MEG evoked components N170 and its analogon the M170 are considered to be linked to structural encoding of faces rather than to later processes responsible for individual identification. It is well known for normal subjects that the M/N170 elicited by faces is larger than for other visual objects indicating intercategory discrimination at this stage of processing. Our measurements clearly show that the M/N170 elicited by faces and houses is decreased for the prosopagnosics compared to the controls but in contrast to previous single case studies a significant amplitude difference between the two categories is preserved. This preserved intercategory amplitude difference for the M/N170 in subjects with cPA can explain their remaining ability to discriminate faces from houses and this finding adds to existing evidence that M/N170 label a stage of intercategory discrimination. Another interesting result concerns the latency of the M170 for faces, which peaks about 20 ms later in subjects with cPA compared to normals.

I would especially like to thank the Sonnenfeld foundation for the financial support given to this study.

## Mechanisms of Recovering Shape Properties from Perfectly Mirrored Objects

Ulrich Weidenbacher,<sup>1</sup> Pierre Bayerl,<sup>1</sup> Roland Fleming<sup>2</sup> and Heiko Neumann<sup>1</sup>

<sup>1</sup>Department of Neural Information Processing, University of Ulm,

<sup>2</sup>MPI for Biological Cybernetics, Tübingen

uw1@informatik.uni-ulm.de

When we look at a perfectly mirrored object, such as a polished kettle, we generally have a remarkably strong impression of its 3D shape. This leads to the question of whether there is a mechanism to completely recover the shape of a mirrored object from a single static image (e.g. a photograph). Without explicit knowledge of the surrounding scene, this is theoretically impossible because many possible combinations of illumination from the surrounding scene and surface properties can generate the same image (i.e. it is an ill-posed problem). Therefore, the only way to extract information about object shape is to constrain the possible combinations of object shape and illumination. If we assume that the reflected scene contains isotropic contrast information, then there is a close relation between the surface curvature of an object (specifically the second derivatives of the surface function) and the distortions of the reflected scenery [1]. In this contribution we present two different computational methods for analysing images of mirrored objects to recover certain properties of 3D shape. Our first method is a statistical approach, based on principal components of the image gradient computed in a local neighborhood, known as the structure tensor. In this context, the eigenvectors of the tensor tell us the orientation of curvature and the eigenvalues of the tensor give us information about the anisotropy of curvature (ratio of maximal and minimal curvature). Our second method is a biologically motivated approach, based on Gabor filters and grouping. We apply an iterative refinement in a simple model of cortical feedforward/feedback processing [2]. Context information is collected by cells with long-range lateral connections. This information is fed back to enhance regions where local information matches the top-down reentry pattern provided by the larger context. Our approach shows that under the assumption mentioned above, it is possible to recover two characteristic curvature properties of mirrored objects: (i) the direction of maximal and minimal curvature and (ii) the anisotropy of curvature. Our simulations demonstrate that both methods (the statistical and biological motivated approach) lead to comparable results and that the models perform well even if the assumption of isotropic contrasts in the scenery is violated to a certain degree.

[1] Fleming, Torralba and Adelson. *Journal of Vision*, 4, 2004

[2] Neumann and Sepp, *Biol. Cybern.*, 81, 1999



## **Visual Cognition: Space Perception and Navigation**

## The Influence of Body Tilt on Mental Rotation and Subjective Visual Vertical

Sarah D. Glaus<sup>1</sup> and Thomas Jarchow<sup>2</sup>

<sup>1</sup>Universität Zürich, <sup>2</sup>Man Vehicle Laboratory, MIT, Cambridge MA, USA

sglaus@allgpsy.unizh.ch

According to earlier spatial-orientation studies, an A-effect occurs when an illuminated bar is adjusted to a subjectively vertical position (SVV) in the dark. The A-effect increases quasi linear up to 135° body tilt and collapses in body tilts greater than 135°. In mental rotation studies the question remains whether subjects respond faster to objects that are in line with physical or with retinal frame of references, e.g. when subjects are tilted and the retinal and the gravitational reference frame are misaligned. The present study combines SVV and mental rotation with the aim to (1) replicate the classical SVV pattern, and to (2) investigate whether a similar pattern can be found in reaction times of mental rotation task. In this experiment, subjects were placed in the positions of 0° (=upright position), 45°, 90°, 135°, and 180° (=upside down) on a tilt board in a dark room. In each position, subjects performed two separate tasks: One consisted of adjustments of an illuminated bar to the subjective vertical position (SVV). In the other, subjects were asked to decide whether the writing of meaningless consonant triplets was normal or mirrored. The triplets were rotated in the plane of the writing and reaction time of subjects was recorded. Results on the SVV task are in accordance with the findings of earlier studies. Further results show that in the mental rotation task subjects responded fastest to triplets that corresponded to the respective SVV, i.e. an “A-effect” in mental rotation is found. This effect occurs up to body tilts of 90°. At 135°, the fastest reaction time is found for triplets oriented between SVV and retinal orientation, which could be interpreted as transition from a SVV-coordinate system to a retinal one. At 180°, fastest reaction time is found for retinal orientation, in contrast to SVV-results. The authors conclude that two separate systems are involved in mental rotation, one in line with SVV and one with retinal coordinates, the former applying to smaller, the latter to larger tilt domains.

## Can A-Priori Instructions Induce a Viewpoint Independent Mental Representation of a Dynamic Scene?

Markus Huff, Bärbel Garsoffky and Stephan Schwan

Knowledge Media Research Center (IWM-KMRC), Tübingen, Germany

m.huff@iwm-kmrc.de

The main objective of this study is to investigate whether a-priori instructions help observers to develop a viewpoint-independent representation of dynamic scenes. Although some previous studies showed that the recognition of dynamic scenes is viewpoint-dependent [1], others revealed that learners can intentionally develop an abstract representation of a given content by means of cognitive elaborations [2]. Viewpoint-independent representations are regarded as a specific type of abstraction. Thus it seems possible that increasing the number of elaborations may induce the development of a viewpoint-independent representation. The participants were exposed to a video showing ball-races that comprised four computer animated balls of different colors and velocities. The subjects were given different types of a-priori instructions: (1) no instruction (no elaboration), (2) paying attention to the colors of the balls (irrelevant elaboration) and (3) focusing on the overtaking processes of the balls (relevant elaboration). After having watched the video, subjects had to decide in a recognition task whether displayed video stills (at different viewpoints and different points of time) were part of the previous scene or not. Additionally, participants of condition 2 and 3 had to answer one question after the recognition test regarding the a-priori instruction. We expected a viewpoint-dependency effect in the no-elaboration condition. In the relevant-elaboration condition, however, no viewpoint-dependency effect should turn up, because the participants should have developed a more elaborated mental representation of the scene. The irrelevant-elaboration condition should have no advantage over the no-instruction condition, because the distractor video stills in the recognition test showed the same ball colors like in the displayed video scene. The analysis of the sensitivity rate showed a significant viewpoint-dependency effect and a marginal significant effect for the instruction condition, but not the predicted interaction of viewpoint deviation and instruction condition. Separate analyses of hits and false alarms revealed the following results: A significant viewpoint-dependency effect appeared in both measures, whereas a significant effect for the instruction condition could be found only in the false alarm rate. The participants produced more false-alarms in the irrelevant-elaboration condition than in the relevant-elaboration condition, in the no-elaboration condition the lowest false-alarm rate was observed. Overall the findings suggest that instead of a viewpoint-independent representation in the relevant-elaboration condition, participants showed a representation which abstracts from the particular temporal characteristics and remained viewpoint dependent.

- [1] Garsoffky, B., Schwan, S., & Hesse F. W. (2002). Viewpoint dependency in the recognition of dynamic scenes. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 28(6), 1035–1050.
- [2] Salomon, G., & Perkins, D. N. (1989). Rocky roads to transfer: Rethinking mechanisms of a neglected phenomenon. *Educational Psychologist*, 24, 113–142.

## Contrast Dependence of Human Place Memory

Anja Weiss, Sabine Gillner and Hanspeter A. Mallot

Universität Tübingen

anja\_mw@web.de

We investigated the accuracy of humans to find back to a certain location in a small scale environment, comparable to the task rats had to solve in the Morris water maze [1]. For bees it was proposed, that they locate certain points in space with a snapshot memory [2], that is they remember images of the surrounding rather than objects. In our experiment the task of human subjects was to return to a previous visited point in a cylindrical room ( $r = 2.25$  m) where no objects were placed in. The walls of the room were covered with a rainbow-like colour gradient in horizontal direction. In vertical direction the colour did not change. For this experiment we used the method of virtual reality. Subjects could walk freely, wearing video goggles where the simulated room was shown. The movement of the subject was tracked and combined with the simulation in that way that a movement was simulated according to the behaviour of the subject. The dependant variable was the contrast of the colour gradient. A reduction in contrast should disturb the performance based on a snapshot memory, but not of an object memory. 6 subjects participated in the experiment (within subject design), each were tested in 6 different contrast conditions (10 % - 100 %). We found a contrast dependant homing accuracy for points in the periphery of the room. For points in the centre of the room there was no such contrast dependence, indicating that for points in the periphery and in the middle of the room different homing strategies have been used by the subjects. Theoretical considerations predict [3] that the homing accuracy should drop with the distance to objects in space. An additional source of information might be the depth information served by the edge between the ground plane and the wall of the cylinder, which should not be influenced by a variation in contrast. In summary we could show that humans are able to use a snapshot memory to find a certain place in space and that they combine image information with geometrical cues of the environment.

[1] Morris R. (1984) *J Neurosci Methods* 11(1): 47–60

[2] Cartwright BA & Collet T (1983) *J Comp Physiology* 151: 521–543

[3] St'urzl W (2004) PhD thesis, University of T'ubingen

## Isovists Capture Locomotion-Relevant Properties of Space

Jan M. Wiener,<sup>1</sup> Nicole Rossmannith,<sup>2</sup> Andreas Reichelt<sup>2</sup> and Gerald Franz<sup>2</sup>

<sup>1</sup>MPI for Biological Cybernetics, Tübingen and University of Tübingen,

<sup>2</sup>MPI for Biological Cybernetics, Tübingen

jan.wiener@tuebingen.mpg.de

The systematic investigation of spatial behavior and cognition requires generic description systems for spatial form and configuration that capture behaviorally relevant properties of environments. Isovists (i.e. viewshed polygons describing the the visual space from single observation points, cf. [1]) have always been conceived as closely related to spatial behavior and, indeed, case studies found correlations between isovist measurands and large-scale statistical pedestrian dispersal [2] as well as navigation performance [3]. The relations between visuo-spatial properties and small-scale locomotion behavior, however, are widely unknown. The study presented here quantitatively explored correlations between isovist derivatives and behavior in architectural environments at the level of trajectories.

In a navigation experiment, 16 participants were asked to find the place offering the best overview in each of 16 virtual indoor spaces. Subjects were advised to approach this position as accurately, directly, and quickly as possible. During this, their positional and orientational trajectories were recorded. Participants experienced the virtual environments from an egocentric perspective on a standard CRT offering a simulated field of view of 90x63 degrees; they interacted with the simulation using a customary joystick.

The analysis tested for subjects' performance with respect for finding the best overview place. Mean characteristic values derived from the individual trajectories were compared to task performance as well as to global isovist measurands of the corresponding virtual indoor scenes. Several strong correlations were found. For example, subjects' performance was significantly correlated with the global isovist measurand 'jaggedness', describing aspects of the shape complexity of spatial situations ( $r=-.66$ ,  $p<.01$ ). On the level of trajectories, jaggedness was negatively correlated with average angular velocity during locomotion, i.e. subjects tended to turn more slowly in more complex environments ( $r=-.72$ ,  $p<.01$ ). Furthermore, in scenes featuring a high variance in visual stability, subjects showed a decreased tendency to stop and to reorient when solving the navigation task ( $r=-.55$ ,  $p<.05$ ). Correspondingly, a similar negative correlation with overall performance was found ( $r=-.56$ ,  $p<.05$ ).

Altogether, the results suggest that isovists and their derivatives have predictive power not only for overall performance in the task tested, but also for spatial behavior at the level of trajectories, and therefore promise novel insights into the perceptual basis of locomotion.

[1] Benedikt, 1979, *Environment and Planning B*, 6, 47–65

[2] Turner & Penn, 1999, *Space Syntax II Proceedings*

[3] Wiener & Franz, 2004. *Spatial Cognition IV*, Springer, Berlin, in press



## **Visual Perception: Colour**

**Color Vision for Natural Objects: Color Constancy**

Karl R. Gegenfurtner, Sebastian Walter and Thorsten Hansen

Universität Giessen

gegenfurtner@uni-giessen.de

We investigated whether memory colours—the typical colours of objects—do contribute to colour constancy. We presented subjects with digitised photographs of naturally occurring fruit and vegetable objects. The subjects had to adjust the photographs (by changing the mean colour in a two-dimensional colour-opponent space) until the isolated objects appeared neutral gray or in the memory color. The background illumination was neutral gray or shifted towards red, green, yellow, or blue. Subjects also had to make adjustments to uniformly coloured discs, uniformly coloured shapes corresponding to the objects, or to a version of photographs where individual pixels were randomly swapped. The objects were typically perceived in their memory colour when the average pixel colour was gray. For the objects to appear gray, subjects adjusted the average colour by about 10%-20% in the colour direction opposite to the memory colour. For example, the banana was adjusted to a bluish average colour until it appeared gray. No such effect was found for the uniformly coloured discs or shapes or for the photographs with the randomly swapped pixels. These results show a reliable high-level memory effect on colour constancy.

## Color Vision for Natural Objects: Color Discrimination

Thorsten Hansen and Karl R. Gegenfurtner

Justus-Liebig-Universität Gießen

thorsten.hansen@psychol.uni-giessen.de

The ability to distinguish different colors is probably the most basic competence of human color vision. Traditional approaches to study chromatic discrimination, including the seminal study of MacAdam [1], have used patches of homogeneous color to study human chromatic discrimination. Natural objects such as a yellow banana, however, are not composed of a single color. Instead, they are characterized by a distribution of different colors. Here we study chromatic discrimination using digital fruit images.

In a 4AFC paradigm, four stimuli were briefly presented on a CRT monitor in a 2x2 arrangement. Three of the stimuli were identical (test stimuli) and the fourth one (comparison stimulus) differed. The stimuli were either homogeneous patches of light, or digital photographs of fruit objects (e.g., banana). The stimuli were displayed on a homogeneous background whose chromaticity was also systematically varied. The mean color of the comparison stimulus was varied along eight different directions in color space from the test stimulus. Discrimination thresholds were measured along these eight directions and ellipses were fitted to the resulting threshold contours.

In agreement with earlier studies, we found that discriminability was best when the test stimuli had the same average color as the adapting background. However, when fruit objects were used as stimuli, thresholds were elevated and threshold contours were elongated in a way that reflected the distribution of hues in the stimulus. For test stimuli that had an average color different from the background, threshold contours for fruit objects and homogeneous patches were identical.

We conclude that the distribution of colors within natural objects can have a profound effect on color discrimination and needs to be taken into account when predicting discriminability.

[1] MacAdam, D. L. (1942): JOSA 32, 246–274

**Color Constancy: Increment-Decrement-Asymmetries Depend on Illuminant Color**

Katrin Heier and Sven Schultz

Martin-Luther-Universität Halle-Wittenberg

heikatti@web.de

For models of color contrast, distinction between increments and decrements has proven useful [1]. In a study of color constancy, Bäuml [2] found evidence for asymmetries in the processing of increments and decrements even in complex scenes. In particular, color constancy performance turned out to be better for decremental than for incremental lights. The question arises which role in visual processing increment-decrement asymmetries may play. We investigated if the strength of increment-decrement asymmetries varies systematically with the illuminant. In the experiment, we presented simulations of selected surfaces and illuminants on a CRT-monitor. The stimulus was a 5x5 arrangement of 25 mattly colored squares situated on a homogeneous background. The task for the three subjects was to find an achromatic adjustment for the central square, i.e. to set it so that it looked neither reddish nor greenish and neither bluish nor yellowish. We used these settings as measure for the subjects' color constancy performance. Subjects carried out adjustments for five incremental and five decremental stimuli, with luminances between 5 and 40 cd/m. Increments and decrements in this context were defined in relation to the average light signal reaching the eye from the remaining surfaces. Each stimulus was rendered with one of ten homogeneous simulated daylight illuminants covering a color temperatur range from 4000 to 10000K. In all subjects, we found asymmetries in the processing of increments and decrements. The magnitude of increment-decrement asymmetries varied systematically with illuminant color. The lower the color temperature of the illuminant the stronger were the increment-decrement asymmetries. In case of substantial increment-decrement asymmetries, better color constancy performance was observed for decrements than for increments. The systematic variation of increment-decrement asymmetries with color temperature could lead to the assumption that the visual system uses information from the strength of the asymmetry to estimate the current illumination.

[1] Mausfeld, R. & Niederée, R. (1993): *Perception* 22, 427–462

[2] Bäuml, K.-H. (2001): *JOSA A* 18, 2419–2429

## Color Constancy and its Relation to Models of Dichromacy

Sven Schultz

Martin-Luther-Universität Halle-Wittenberg

s.schultz@psych.uni-halle.de

Dichromats are persons who lack a certain type of cone photoreceptor. In the standard model of dichromacy the dichromatic color space is usually described as a two-dimensional subspace of the trichromatic color space. In the present study, a test of this standard dichromatic model was conducted which relies on the analysis of color constancy performance of dichromats and on characteristics of asymmetries in the processing of increments and decrements. I investigated the color constancy of dichromats with an achromatic matching task and compared the results to the predictions of the dichromatic model of color vision. Furthermore, I analyzed the appearance and strength of increment-decrement-asymmetries in complex scenes as a second test of the model. In the experiment, simulations of surfaces and illuminants were presented on a CRT monitor. The stimuli consisted of 5x5 arrays of matte surfaces, which were rendered under a spatially uniform illumination. Five different illuminants were used. Three of them were simulations of CIE daylight (D65, D40, D10). The other two were artificial illuminants that appeared red and green to normal trichromats. Subjects were three dichromats and—as controls—two normal trichromats. The subjects made achromatic settings for the central surface of the patch (test surface), i. e. they had to adjust the test surface so that it looked neither reddish nor greenish and neither bluish nor yellowish. The settings were made for eight incremental and eight decremental test surfaces. Increments and decrements were defined in relation to the mean light signal reaching the eye from the whole stimulus [1]. For the daylight illuminants, color constancy of the dichromats was comparable to that of the trichromatic subjects as predicted by the dichromatic model. For the red and the green illuminants, color constancy and the variability of the settings of the dichromats were much better than predicted by the dichromatic model and even comparable to trichromatic performance. Furthermore, in the daylight conditions all subjects showed strong increment-decrement-asymmetries. For the red and the green illumination, increment-decrement-asymmetries only occurred in the S-cone system of dichromatic subjects but not in trichromatic subjects. Both findings for dichromatic subjects in the red/green illuminant conditions were not consistent with the dichromatic model of color vision. The observed differences between dichromats and trichromats in this study could reflect differences in the respective S-cone systems [2]. These may partially compensate for the missing cone type.

[1] Bäuml, K. H. (2001): JOSA A 18, 2419–2429

[2] Kurtenbach, A., Meierkord, S. and Kremers, J. (1999): JOSA A, 16, 1541–1548

## Contribution of On- and Off-Pathways to Color Categories of Dichromats

Thomas Wachtler<sup>1</sup> and Rainer Hertel<sup>2</sup>

<sup>1</sup>Philipps-Universität, Marburg, <sup>2</sup>Albert-Ludwigs-Universität, Freiburg

thomas.wachtler@physik.uni-marburg.de

Protanopes and deuteranopes, despite lacking a chromatic dimension at the receptor level, use the color terms “red” and “green”, together with “blue” and “yellow”, to describe their color percepts elicited by monochromatic stimuli of different wavelengths [1]. A plausible multi-stage model of color processing could explain these findings [2]. To further investigate the color categories of dichromats we performed hue scaling experiments with stimuli defined in cone-opponent color space [3,4]. Subjects were dichromats with only a single X-chromosomal opsin gene, as determined by molecular genetic analysis, and color-normal trichromats. Stimuli were 0.5 or 1 degree square patches with chromaticities regularly spaced along the azimuth of cone-opponent color space and with luminance contrast ranging between +/-50% with respect to the neutral background. In each trial, the stimulus was presented for 500ms, after which the subject described the perceived color by relative proportions of basic colors. For stimuli brighter than the background, “blue” and “yellow” dominated the dichromats’ responses in different azimuth ranges, with only low percentages of “red” and “green”. However, at the transitions between the “blue” and “yellow” regions, “green” was reported with proportions up to 60%, consistent with the results for monochromatic stimuli. Around isoluminance, hue proportions showed an elevated trial-by-trial variability, and color category regions were less distinct. For stimuli darker than the background, the category regions were different than for bright stimuli. In particular, regions dominated by “red”, with hue proportion up to 90%, were found. In contrast to the results of dichromats, hue scaling by color-normal trichromats showed only weak dependence on luminance. We extended our earlier model to account for luminance increments and decrements by explicitly considering parallel On- and Off- pathways, originating from the two types of cones of dichromats. We assume multiple, neurophysiologically plausible stages, including antagonistic center-surround mechanisms and mutual inhibition between pathways. The model explains the color categories of dichromats as determined in our hue scaling experiments, including the dependence on luminance. The results suggest that color categories may be traced back to low-level mechanisms of color processing. Even when one cone type is missing, as in dichromats, the visual system is able to map the signals from the reduced receptor color space onto color categories comparable to those of trichromats.

[1] Boynton, R.M. and Scheibner, H. (1967): *Acta Chromatica* 1, 205–220

[2] Wachtler, T. et al. (2004): *Vis. Res.* 44, 2843–2855

[3] MacLeod, D.I.A. and Boynton, R.M. (1979): *JOSA* 69, 1183–1186

[4] Derrington, A.M. et al. (1984): *J. Physiol.* 357, 241–266

## Can Differences in Colour Categorization be Caused by Lens Brunescence?

Sebastian Walter

Justus-Liebig-Universität Giessen

sebastian.walter@psychol.uni-giessen.de

Many languages make no distinction between “green” and “blue” in contrast to most Indo-European languages such as German, English or French. They only have one term for these colours, generally called “grue”. Usually they integrate short-wavelength stimuli into the “green” category. For some languages the focus of the grue category is located in the “blue” area. Lindsay and Brown suggested that the grue category is caused by altered perception through lens brunescence [1]. Chronic exposure to a high amount of UV-B radiation, as in the sunlight of the tropics, causes increased aging of the human eye, which means a yellowing of the lens. As a result of this yellowing, the bluish colours should appear greenish, or greyish if tending to purple. I investigated the colour categorization in German with and without simulation of aged lenses. On a computer monitor (8° x 11°) 417 coloured patches (2°) were presented per condition in random order for 500 ms each. These patches were distributed regular across the available colour gamut at a constant luminance. Subjects had to assign each patch to one of the colour categories “gelb” (yellow), “orange”, “rot” (red), “violett” (purple), “blau” (blue), “türkis” (turquoise), “grün” (green), and “grau” (grey). In one experimental condition subjects looked through Kodak filters simulating a strongly aged lens (filter condition). In the other condition they were looking onto the computer monitor without filters (normal condition). The results show that for all subjects, when looking through filters simulating aged lenses, colour categorization differed very much from the normal condition. The green category expanded deep into the now strongly reduced blue category for the majority of test persons. Partially previously blue colours were categorized as grey. Mixing green, blue and turquoise was common in the filter condition. A minority of subjects reduced the normal-condition green category in favour of blue. This experiment shows that different colour perception caused by lens brunescence could be an explanation for the existence of a grue colour category in areas with high UV radiation.

[1] Lindsay, D. T. & Brown, A. M (2002). *Psychological Science*, Vol. 13, No. 6, 506–512



## **Visual Perception: Depth**

## Larger Perceived Distance Yields Shorter Responses in Time-To-Collision Task

Markus Hofbauer and Thomas Eggert

Dept. of Neurology, Ludwig-Maximilians-University, Munich

mhofbauer@nefo.med.uni-muenchen.de

The two main controversial theories for time-to-collision (ttc) estimates are based either on the ratio between retinal size and retinal expansion (tau-strategy) or on the ratio between perceived distance and perceived speed (indirect strategy). This study investigates how changes in perceived distance induced by monocular depth cues affect ttc estimates. In the first experiment, we measured ttc responses. In the second experiment, we measured perceived distance responses using the same stimuli. The stimuli were presented on a projection screen (1.6x1.3 m<sup>2</sup>) in a dark room. The observers distance to the screen was 3.9 m.

In the first experiment, we presented two different sequences in a time-to-collision task. In one sequence, only the moving object was presented on a black background. In the other sequence, the moving object was embedded in a rich surrounding. In this sequence, we presented the moving object in a perspectively correct represented room with a textured floor and together with three human silhouettes distributed along the depth of the room. I.e. we added a variety of additional depth cues in this presentation. In both sequences, the trajectory of the object was passing the observer at the left hand side. Response times were measured in each sequence. We observed a shorter response time for the sequence with additional depth cues.

In the second experiment, we measured the effect of the visual environment used in the first experiment on the perceived depth of a stationary object. In a 2AFC experiment observers decided in which presentation the object appeared closer to him or her. We measured the point of subjective equality. Embedded in the visual environment, the object appeared further away from the observers than in darkness.

Our results are consistent with the hypothesis that perceived distance is also used in a ttc estimation if depth information is available. At first sight the results from the two experiments seems contradictory, because in comparably conditions response time is shorter while perceived object distance is longer. This result suggests that perceived distance modifies the tau-strategy in a specific way.

We conclude that the ratio between known object size and perceived distance can replace retinal size in a tau-strategy. This explains how the two theories (tau- and indirect strategy) can be combined for ttc judgements.

This study was supported by Bayerische Forschungsstiftung (FORBIAS).

### The Role of Extra-Retinal Cues in Velocity Constancy

Simon J. Maier, Andrew E. Welchman and Heinrich H. Bülthoff

MPI for Biological Cybernetics, Tübingen

`simon.maier@tuebingen.mpg.de`

To estimate the real world speed of an object the velocity of the retinal projection must be scaled by the perceived distance. If observers perceive objects travelling with the same speed at different distances from the eye as equally fast, they are said to exhibit Velocity Constancy. However, not all studies examining Velocity Constancy support the idea that observers can scale speeds for the viewing distance. In fact they suggest that subjects perceive angular rather than objective velocities [1]. The degree to which velocity constancy is observed depends on the information provided by the stimulus and its surround [2,3,4]. So far, studies on velocity constancy and distance have not considered the separate contribution of vergence as a cue to distance. Here, we investigate whether eye vergence (as an extra-retinal cue to distance) and horizontal disparity (as an retinal distance cue) contribute to Velocity Constancy. Subjects viewed two sequentially-presented dots or spheres moving horizontally in the frontoparallel plane. They were required to report whether or not the speed of the second sphere exceeded the objective velocity of the first one. The eye vergence was manipulated using a haploscope. Varying the disparity of the second sphere with respect to the background plane, the constancy of the subject's velocity judgments at different disparities was investigated. Vergence defined distances did not influence the subject's velocity percept, whereas the distances defined by disparity did.

- [1] Epstein W. (1978) Two factors in the perception of velocity in distance. *Perception & Psychophysics* 24:105–114.
- [2] McKee S. P., Welch L. (1989) Is there a constancy for velocity? *Vision Res.* 29(5):553–561.
- [3] Wallach H. (1939) On the constancy of visual speed. *Psychophysical Review* 46:541–552.
- [4] Zohary E., Sittig A.C. (1993) Mechanisms of velocity constancy. *Vision Res.* 33(17):2467–78.



## **Visual Perception: Eye Movements**

### Are there Correlations Between Vertical VOR and Multiaxial Spatial Orientation?

Karl Beykirch,<sup>1</sup> Christoph von Laßberg<sup>2</sup> and Jürgen Krug<sup>3</sup>

<sup>1</sup>Universität Tübingen; MPI for Biological Cybernetics, Tübingen,

<sup>2</sup>Universität Leipzig; Universität Tübingen, <sup>3</sup>Universität Leipzig

kab@tuebingen.mpg.de

Stangl et al. [1] describe differences of vestibulo-ocular reflex (VOR) gain in groups of athletes accustomed to rotational movements. The authors suggest that the physiological purpose of these differences is to enable better spatial orientation during whole body rotations. Is the reverse possible: to use the VOR gain to estimate an individual's per-rotatory spatial orientation ability? To answer this question, we looked for differences of the VOR gains in a group of high-performance gymnasts (gym) versus a group of non-athletes (control). The results from the gymnasts were also examined for correlations with a rank of their individual per-rotatory spatial orientation (PSO) abilities, performed by their coaches. The subjects' (gym:  $n = 9$ , age: 10–13 years, control:  $n = 10$ , age: 11–12 years) eye movements were recorded using a video nystagmography system (SMI). They were seated with head fixed in a software controlled multi-axial whole body rotator. The test included 4 horizontal and 4 vertical sinusoidal whole body rotations with various frequencies and peak velocities (0.4, 0.2 and 0.1 Hz at 25°/s; 0.1 Hz at 50°/s). VOR gain was calculated as the ratio of the amplitude of the best-fit sine wave for the slow-component eye velocity to the amplitude of the stimulus velocity. The PSO ranking list was based on the independent estimation of three coaches, who work with the group of gymnasts daily. The results showed clear gain differences between the 4 tests (the higher the speed or frequency, the higher the gains). It was not possible to show group differences between groups for either horizontal or during vertical rotations. But there were clear inter-individual differences in both groups, characterized through high significant individual stability (individual correlations between the tests mostly with  $p < 0.01$  [Pearson]). All the individual coach rankings showed also significant correlations [ $p < 0.05$ , Spearman]). The analysis of correlations between vertical gains and PSO showed a clear and significant correlation (0.4Hz:  $r_S = 0.647$ ,  $p = 0.109$ ; 0.2Hz:  $r_S = 0.786$ ,  $p = 0.036$ ; 0.1Hz@25°/s:  $r_S = 0.619$ ,  $p = 0.102$ ; 0.1Hz@25°/s:  $r_S = 0.821$ ,  $p = 0.023$ ). Horizontal motion did not show any correlations with PSO. Differences in VOR gain in subjects accustomed to rotational movements could not be confirmed. Correlations between the various tests for each subject show they are stable indicators of VOR response. Correlations between vertical VOR-gains and PSO of the athletes could indicate a link between VOR parameters and individual per-rotatory multi-axial orientation abilities.

[1] Stangl, W. et al. (1998) *Leistungssport* 30 (1), 58–62

SFB 550/B8

### Looking During Grasping

Anne-Marie Brouwer,<sup>1</sup> Volker Franz,<sup>2</sup> Dirk Kerzel<sup>3</sup> and Karl R. Gegenfurtner<sup>2</sup>

<sup>1</sup>MPI for Biological Cybernetics, Tübingen, <sup>2</sup>Justus-Liebig-Universität Giessen,

<sup>3</sup>University of Geneva, Switzerland

anne-marie.brouwer@tuebingen.mpg.de

In a grasping task, Johansson et al. [1] found that subjects look at the position to which the finger tips are guided. However, in their experiment, only the contact position of the thumb was visible. We investigated what happens if the contact positions of both finger and thumb are visible. We recorded eye and finger movements. In a first experiment, subjects always grasped with the index finger at the top and the thumb at the bottom of a flat shape that was mounted on a horizontal bar. In order to see whether a salient feature of a shape would affect the fixation positions, we presented an (asymmetric) cross in 4 orientations (with the crossing of the bars representing the salient feature). In order to see whether gaze is attracted to the position where the finger has to be guided relatively precisely, we presented a triangle in two orientations that subjects had to contact at the base and at the pointed top (i.e., a higher required precision to contact the top than the base). We found that the crossing of the bars attracted the gaze whereas the top of a triangle did not. A prominent result was that subjects fixated above the center of the shape. In order to distinguish between subjects fixating the upper part of the shape versus being attracted by the index finger, we mounted a square and a triangle in two orientations on a vertical bar. We asked subjects to grasp first with one hand and then with the other so that the shape remained constant but the contact positions of the index finger and thumb were reversed. Subjects still looked above the center. In addition, for the square the gaze was now attracted to the thumb so that the average fixation location was to the left or to the right half of the square, depending on the used hand. The latter result indicates that gaze not only depends on the shape but also on features of the grasp. This is also supported by the finding that the variability in fixation locations was more in the horizontal direction for the second experiment than for the first, i.e. the variability tends to be along the grasping axis. We conclude that both features of the shape and the grasp affect gaze location.

[1] Johansson, R.S. et al. (2001): *J Neurosci* 21(17), 6917–6932

Supported by the European Commission (Grant HPRN-CT-2002–00226) and the Deutsche Forschungsgemeinschaft (DFG grant FR2100/1–1).

## Variability of Eye Movements on High-Resolution Natural Videos

Michael Dorr,<sup>1</sup> Martin Böhme,<sup>1</sup> Jan Drewes,<sup>2</sup> Karl R. Gegenfurtner<sup>2</sup> and Erhardt Barth<sup>1</sup>

<sup>1</sup>Universität zu Lübeck, <sup>2</sup>Justus-Liebig-Universität Gießen

dorr@inb.uni-luebeck.de

We collected and analysed a large data set of eye movements with video clips of natural scenes. 59 subjects watched 18 different videos of approximately 20s duration each. Videos had a high spatial resolution (HDTV standard, 1280x720 pixels) and 30 Hz temporal resolution and showed a variety of outdoor scenes in and around Luebeck. They were displayed at a visual angle of 48 x 27 degrees. Overall, we could analyse 950 trials with about 45,000 saccades.

Two different algorithms, K-means clustering and a superposition of Gaussians, were used to derive fixation maps. From these maps, we extracted  $n=1, \dots, 20$  regions that were fixated most frequently. For each video frame, we then counted how many fixations fell within these extracted regions with varying size. Clearly, these counts vary for different frames even of a single video as there exist some “hot spots” where attention is inevitably drawn towards only a small number of locations while in other frames bottom-up saliency is very low. Nevertheless, we also averaged these counts over all 600 frames to give an overall variability score. Our results indicate that, depending on the video, a single ( $n=1$ ) circular region per frame with a diameter of about 2 degrees can predict 8% to 31% of all fixations. With more regions, these numbers increase to up to 65–85% ( $n=20$ ). For some single frames, even a single region can predict up to 60% of fixations ( $n=1$ ). Clustering fixations by the K-means algorithm yielded significantly lower predictability (28–52% for  $n=20$  regions, averaged over the video).

We have investigated the variability of eye-movement patterns on a particular image sequence, i.e. in how many different ways a movie can be “seen”. This variability was quantified in terms of how well the most frequently fixated regions can predict the individual fixations. Attempts to reduce this variability for better visual communication and interaction are made in [1].

[1] <http://www.inb.uni-luebeck.de/Itap>

Research is supported by the German Ministry of Education and Research (BMBF) under grant number 01IBC01B.

## The Coding of Combined Pointing Movements and Saccades in a Length Illusion

Denise de Grave, Volker Franz and Karl R. Gegenfurtner

Justus-Liebig-Universität Gießen

denise.d.grave@psychol.uni-giessen.de

When making an eye or hand movement to a visual target, different sources of information can be used. Either visual judgments of direction and distance (or length) of the required displacement can be used (vector coding), or the final position (position coding), or a combination of both. In an earlier study [1] it was shown that pointing movements mainly use vector coding. To find out which source of information is used for combined eye and hand movements, we used the Brentano illusion, a version of the Müller-Lyer illusion. This illusion primarily influences judgements of length, but not the position information. Thus, a task will only be influenced by this illusion if the task requires a visual estimate of length. In this study we investigated the influence of the Brentano illusion on pointing and saccadic eye movements when these movements are performed in the same trial. Ten subjects fixated in the middle of the touch-screen and also started the pointing movement with their index finger at this position. After the stimulus was presented for 200 milliseconds they made saccades and pointing movements in four directions (up, down, left, right). Movements were always from an outer vertex of the Brentano illusion toward the middle vertex. Pointing movements as well as saccades showed an effect of the illusion, indicating that length information was used. A trial-by-trial correlation for pointing and saccades ( $r = 0.05 \pm 0.03$ ) was non-significant. This might either be due to small between trial variability of the illusion relative to the noise in each response or to an independent, parallel generation of the illusion effects for the two responses. In both tasks an effect of the illusion is found (pointing:  $26.19\% \pm 2.72\%$ ; saccades:  $20.14\% \pm 2.51\%$ ), indicating that length is used (vector coded). There is no significant difference in illusion effect between the pointing task and the saccadic eye movement task. We conclude that the results favour the interpretation of using the same information in eye and hand movements.

- [1] de Grave, Brenner & Smeets (2004). Illusions as a tool to study the coding of pointing movements. *Exp Br Res*, 155, 56–62.

## Localization of Visual Targets During Optokinetic Eye Movements

Andre Kaminiarz, Marc Rohe and Frank Bremmer

Philipps-Universität Marburg

andre.kaminiarz@physik.uni-marburg.de

Although the visual world appears stable during eye movements previous psychophysical studies have demonstrated a perceptual distortion of space during smooth pursuit and saccadic eye movements. Both types of eye movements have in common that they are voluntarily controlled. To our best knowledge, localization has yet not been tested for reflexive eye movements. In our present study we therefore asked how accurately human subjects can localize briefly presented visual targets during optokinetic eye movements, which can be considered an alternating pattern of smooth and fast, saccade-like resetting eye movements. Seven human subjects with normal or corrected to normal vision participated in the experiments, which were performed in a light tight chamber. The subjects' head was supported by a chin rest. Eye movements were sampled at 500 Hz with an EyeLink2 eye tracker (SR Research Inc.). Visual stimuli were presented within a circular aperture of 25° diameter on a computer monitor placed 57 cm in front of the subjects' head. In blocks of trials subjects performed either control or eye movement tasks. In control trials subjects freely viewed a homogeneous gray monitor for 4000 ms. After 3500 ms a visual target (white circle, 0.5 degree in diameter) was flashed for 10 ms at one of five possible locations. In the eye movement condition a random dot pattern was moving horizontally for 6000 ms in pseudorandomized order either to the left or to the right eliciting reliably an optokinetic nystagmus (OKN) in all subjects. After 5500 ms the same visual target as in the control condition was presented with the same spatial and temporal properties. After the end of a trial (control or eye movement condition) subjects had to indicate the perceived horizontal location of the target with respect to a ruler. Localization during free viewing (Base line condition) was not veridical but rather biased towards a location centered on the vertical meridian in head centered space. Localization during OKN, corrected for the bias obtained in the base line condition, was shifted into the direction of the slow eye movement. Our results therefore clearly indicate that localization of visual targets is influenced not only during voluntarily controlled but rather also during reflexive eye movements.

Supported by Deutsche Forschungsgemeinschaft (Br 2282/1)

### The Cortical Interaction of Attention and Intention

Christina S. Konen,<sup>1</sup> Raimund Kleiser,<sup>2</sup> Rüdiger J. Seitz<sup>2</sup> and Frank Bremmer<sup>1</sup>

<sup>1</sup>Universität Marburg, <sup>2</sup>Universität Düsseldorf

`christina.konen@physik.uni-marburg.de`

Everyday life often necessitates a dissociation between our directed attention and the intention to direct our gaze. Accordingly, the differential role of visuomotor related areas in the one or the other process is an issue of an ongoing debate. We used an event-related fMRI study with concurrent eye tracking to elaborate a differentiation between attention and intention in these areas. Thirteen subjects were asked to fixate a central target, while they directed attention to a colored cue in the left or right visual field. Regardless of its location, the color of the cue provided information about the direction of the upcoming saccade. The attention to the peripheral cue and the intention to perform the saccade were thus either directed to the same side or to opposite sides. The offset of the cue signaled the start of the memory saccade. Data analysis was performed with BrainVoyager QX and SPM2. The results showed a variety of activated areas demonstrating the cortical stages of visuomotor processing. We observed stronger activation of the early visual areas V1/V2 and V5 during the presentation of a contralateral compared to an ipsilateral stimulus. The ‘meaning’ of the cue entailed an increased fMRI signal of the color sensitive area V4. A different response pattern emerged from the supplementary eye field (SEF) and the posterior parietal cortex (PPC) where the maximum of the fMRI signal change was found for the conditions when attention and intention were directed to the same side compared to attention and intention being directed to opposite sides. In contrast, the BOLD amplitude within the frontal eye field (FEF) was strongest when attention and intention were spatially divergent. Taken together, the visuomotor related areas split up regarding their involvement in attentional and intentional processes. Activation of the early visual cortex and area V5 is mainly the result of visual and / or attentional processes. Area V4, however, seems to be the first stage along the visual pathways representing the encoding of the ‘meaning’ of the cue. The PPC as well as the SEF are highly influenced by spatially congruent attention and intention. The FEF reveals the contrary time course. These modulations stress the prominent roles of the PPC as well as the SEF in pro- and of the FEF in anti-saccades and their differential roles for attentional and intentional processes.

## **I Know where You'll Look: An fMRI Study of Oculomotor Intention and a Change of Motor Plan**

Christina S. Konen,<sup>1</sup> Raimund Kleiser,<sup>2</sup> Rüdiger J. Seitz<sup>2</sup> and Frank Bremmer<sup>1</sup>

<sup>1</sup>Universität Marburg, <sup>2</sup>Universität Düsseldorf

christina.konen@physik.uni-marburg.de

Electrophysiological studies in monkeys showed that the intention to perform a saccade is reflected in the neural activation of the posterior parietal cortex (PPC) [1]. In such case, even the covert change in motor planning is indicated at the neuronal level. We asked whether such covert intentional oculomotor processes are detectable in humans as well. We addressed this issue with an event related fMRI study and concurrent eye tracking. Thirteen subjects were instructed to fixate a central target, which changed its color in order to indicate the direction of the subsequent saccade. Without the advanced knowledge of the subjects, the color changed again in half of the trials to instruct a spatially opposite saccade. This allowed us to investigate the effect of covertly changing the motor plan without changing the direction of gaze and the spatial focus of attention. The offset of the fixation target signaled the start of the memory saccade. Data analysis was performed with Brain Voyager QX and SPM2. The results revealed a widespread activation of areas rendering the sensory-motor transformation. Area V5 was labeled by an enhanced and prolonged signal in case two cues were successively presented compared to a single cue. The same was true for the PPC. The BOLD amplitude of the supplementary eye field (SEF) showed no significant differences across conditions, except for the delayed peaks when two cues were presented. In contrast, the signal of the frontal eye field (FEF) was higher modulated by an individual cue compared to the double-cue. Even more interesting, the analysis on a single subject level exposed that the BOLD signal change of the posterior segment of the intraparietal sulcus (IPS) correlated with the change in motor plan, i.e. activation strongly decreased when the cue instructed an ipsiversive saccade while it strongly increased when it instructed a contraversive saccade. Our data show that intentional processes affect the brain in different ways. The double-cue induced a synergistic effect of the BOLD signal in area V5 and the PPC, whereas the reverse was true for the FEF. Noteworthy, the IPS within the PPC seems to be of particular importance for the encoding of intention. The fMRI signal was correlated with the specific motor command mediated by the cue.

[1] Bracewell R.M. et al. (1996): J. Neurophysiol. 76, 1457–1464

### Visual Pursuit in Gymnasts

Christoph von Laßberg,<sup>1</sup> Karl Beykirch<sup>2</sup> and Jürgen Krug<sup>3</sup>

<sup>1</sup>Universität Leipzig; Universität Tübingen,

<sup>2</sup>Universität Tübingen; MPI for Biological Cybernetics, Tübingen, <sup>3</sup>Universität Leipzig

christoph.lassberg @med.uni-tuebingen.de

In comparison with non-athletes, there is little doubt that gymnasts have better spatial orientation during complex sport-specific movements, like double or triple “twisting somersaults” or similar exercises. An understanding of the specific processes of spatial training and their function in a higher level of multi-axial spatial orientation remains unclear. One must consider the role that motor learning in any sensorimotor system contributing to orientation might play, e.g. smooth pursuit. Are there measurable differences in the dynamics of these behaviors between gymnasts and non-athletes due to their respective levels of training? We sought to investigate the relationship between smooth pursuit performance and the spatial orientation needed during fast multi-axial whole body rotations. The subjects’ (gym:  $n = 9$ , age: 10–12 years, control:  $n = 10$ , age: 10–12 years) smooth pursuit eye movements were recorded using a monocular video nystagmography system (SMI). They were seated in the dark with head and body fixed. The stimulus was a laser target moving horizontally with a sinusoidal velocity profile. Maximum stimulus velocities of  $60^\circ/\text{s}$ ,  $120^\circ/\text{s}$ ,  $140^\circ/\text{s}$ ,  $160^\circ/\text{s}$ , were used, with a short break between tests. The gymnast group was tested again after a three-week break in their gymnastic training. Pursuit gain was calculated as the ratio of the amplitude of the best-fit sine wave for the slow-component eye velocity to the amplitude of the stimulus velocity. Any additional training effect on pursuit velocity, e.g. video games, could be excluded. Although in both groups the gains were reduced with higher maximum stimulus velocity, the results show significantly higher gains for the gymnasts during the  $120^\circ/\text{s}$  test (gain  $\pm$  s.d., gym.:  $0.23 \pm 0.10$ ,  $n=9$ , control:  $0.14 \pm 0.08$ ,  $n=10$ ;  $p=0.022$ ,  $Z=-2.289$ , Mann-Whitney-U) and  $140^\circ/\text{s}$  (gain  $\pm$  s.d., gym.:  $0.19 \pm 0.12$ ,  $n=9$ , control:  $0.08 \pm 0.05$ ,  $n=10$ ;  $p=0.010$ ,  $Z=-2.536$ , Mann-Whitney-U). For  $60^\circ/\text{s}$  and  $160^\circ/\text{s}$  there was only a tendency toward higher gains for the gymnasts. After the break in training, the gymnasts gains were significantly reduced at  $120^\circ/\text{s}$  and  $140^\circ/\text{s}$  ( $120^\circ/\text{s}$ :  $0.20 \pm 0.09$ ,  $n=9$ ;  $p=0.028$ ,  $Z=-2.198$ ;  $140^\circ/\text{s}$ :  $0.12 \pm 0.07$ ,  $n=9$ ;  $p=0.033$ ,  $Z=-2.136$ , Wilcoxon). There was also a clear reduction at  $60^\circ/\text{s}$  and  $160^\circ/\text{s}$ , but not significant. Smooth pursuit in healthy humans shows a saturation (gain  $> 0.8$ , binocular) of approximately  $100^\circ/\text{s}$  [1]. So tests with peak velocities of  $120^\circ/\text{s}$  and  $140^\circ/\text{s}$  are clearly above this saturation level. The results demonstrate that high-performance gymnasts show a training-dependent modification of their maximum velocity for smooth pursuit.

[1] Meyer, C.H. et al. (1985): Vis. Res. 25, 561–563

SFB 550/B8

## **The Contribution of Torsional Oculomotor Dynamics to Spatial Interactions of Vertebrates**

Bernhard M. Schäfer

UniNet GmbH Heidelberg

bernhard\_m.schaefer@t-online.de

A conspicuous character in the evolution of vertebrate vision systems is the occurrence of 3D eye movements. The trochlearis complex (muscle, cranial nerve and nucleus) on the one hand appears in all species with external eye muscles, on the other hand shows a great variability over species. This fact may be taken as evidence for a key role in evolution and development of the visual systems in vertebrates. We present an evolutionary scenario for the roles of the three muscle pairs (axes) in spatial interactions i.e. control of self motion, scene perception, object localisation etc. In the primitive condition large field optic flow detectors and the eye muscles act in a closed loop to produce nystagmus like patterns, which are then used to control self motion. Differentiation into smaller local flow detectors for different spatial frequencies lead to the possibility to separate the visual field into depth layers through motion parallax. The responses of the local flow detectors are pooled corresponding to the movements in the main axes. The pooled responses of the local flow detectors at a contour (which are shift invariant) to oscillatory movements in all three axes determine the head centric position of the contour. This leads to a position code, that is invariant under eye movements, which do not alter the torsion axis.

## The Effect of Distractors on Smooth Pursuit Eye Movements

Miriam Spering,<sup>1</sup> Dirk Kerzel<sup>2</sup> and Karl R. Gegenfurtner<sup>1</sup>

<sup>1</sup>Universität Gießen, <sup>2</sup>Université de Genève

miriam.spering@psychol.uni-giessen.de

If two potential targets for smooth pursuit move simultaneously, the eyes move according to the vector average of the position and velocity of the two stimuli before one stimulus is selected as a pursuit target [1]. In previous studies, the target and distractor stimulus often had identical appearances and no previous information was given that allowed to decide which stimulus would become the target. We investigated the case when the target direction is known to the observer, and the simultaneously moving distractor stimulus is irrelevant to the oculomotor task. Smooth pursuit eye movements were elicited by horizontally moving Gaussian dots. Observers (N = 12, 5 highly trained and 7 untrained and naive to the experiment's purpose) had to make an initial saccade to a stimulus that appeared to the left or right of a central fixation spot. The stimulus then started to move centripetally along the horizontal meridian at a constant velocity of 10 deg/s and split into two dots at a random point in time. The dot that continued to move horizontally was the target and the observer was instructed to smoothly pursue this stimulus. The distractor moved into one of eight directions, diagonally or straight up or down, and either in the same direction as the target or in the opposite direction. Results show that the trajectory of the smooth pursuit eye movement deviated contralateral to the motion direction of the distractor: When the distractor moved upwards, the eyes tended to curve downwards and vice versa. The distractor had a significant effect on horizontal eye velocity: The eyes slowed down substantially, especially when the distractor moved into the opposite direction of the target. Vertical eye velocity was higher in the direction contralateral to the distractor. These findings are surprising, considering the evidence for vector averaging for the concurrent processing of two stimuli. In contrast to the prediction of vector averaging, eye position and velocity go in the direction contralateral to the distractor stimulus. Previous research has shown that trajectories of saccadic eye movements curve away from distractor stimuli that need to be ignored [2]. Saccadic deviation is interpreted as a consequence of the necessity to suppress the extant oculomotor program. These studies propose a strong link between spatial attention and oculomotor programming. Further experiments need to clarify whether the effect on smooth pursuit is attention-based or purely sensory.

[1] Lisberger, S. G. and Ferrera, V. P. (1997): *Jn of Neurosc*, 17, 7490–7502.

[2] Sheliga, B. M., Riggio, L. and Rizzolatti, G. (1995): *Exp Brain Res*, 105, 261–275.

Miriam Spering is supported by the DFG graduate program “Brain and Behavior” (DFG 885/1).

## **Eye Movements During a Video-Game-Like Pointing Task: What Attracts the Eye while the Hand Moves?**

Martin Stritzke and Julia Trommershäuser

Justus-Liebig-University Giessen

`martin.stritzke@psychol.uni-giessen.de`

Vision is needed for guiding the hand towards a visually specified target. In everyday situations, however, hitting or missing the target with the hand leads to different consequences than hitting or missing it with the eye. Here we researched to what extent the visual properties of the stimulus and the hand contribute to the characteristics of eye movements in a task in which movements of the hand, but not the eye lead to monetary consequences for the movement planner. In a series of four experiments, we simultaneously measured eye movements and touch point of the finger during a video-game-like pointing task [1]. In the first three experiments, subjects were instructed to rapidly touch a target region (green) on a screen while trying not to hit a nearby penalty region (red). Each target hit yielded a gain of points; each penalty hit incurred a loss of points. Late responses were penalized. At the end of each trial, subjects received feedback about whether they hit the target or not. In the first experiment, the penalty was a filled red disk and the target a hollow green circle, the background was grey. In the second experiment, the penalty was hollow and the target was filled. In the third experiment, the stimulus configuration was the same as in the first experiment, but disappeared as soon as the pointing movement was initiated. Penalty value, overlap of the circles and stimulus locations were varied. In a control experiment (experiment 4), subjects performed a visual judgement indicating whether the target was on the right side or on the left side of the penalty. Four subjects completed all experiments. In experiments 1, 2 and 4, subjects made two saccades on average. This number was slightly lower in the third experiment. In all experiments, the majority of landing points of the first saccade was within the region of the filled circle, i.e. the more salient stimulus. The landing point of the second saccade shifted closer towards the touch point of the finger. In most trials, the second saccade was concluded before the finger hit the screen. We conclude that eye movements during pointing tasks are guided by both the visual properties of the stimulus and the hand.

[1] Trommershäuser, J., Maloney, L. T., and Landy, M. S. (2003): *JOSA A*, 20, 1419–1433.

Deutsche Forschungsgemeinschaft (Emmy-Noether-Programm, Grant TR 528/1–2)

**Structured Random Noise as a Visual Distractor for Saccadic Eye Movements**Brian White,<sup>1</sup> Karl R. Gegenfurtner<sup>1</sup> and Dirk Kerzel<sup>2</sup><sup>1</sup>Universität Giessen, <sup>2</sup>Universität Genf

brian.white@psychol.uni-giessen.de

It has been suggested that the remote distractor effect is the direct result of visual stimulation of a collicular fixation zone near the time of target onset. The distributed network of the cells responsible for this effect is believed to extend over a large area, responding to distractors up to 10 deg in the periphery. Several studies also implicate the superior colliculus as the substrate behind an inhibited saccadic response due to a display change. We examined this further by presenting a spatially continuous form of non-target stimulation (a patch of random noise) of various sizes while observers made a saccade to a simultaneously presented peripheral Gabor-target (4 c/deg, 4, 7 or 10 deg, left or right of fixation). A small patch (2.3 x 2.3 deg) of centrally displayed noise produced a significant increase in saccadic latency consistent with the remote distractor effect. In contrast, a large patch (36 x 36 deg, essentially a background change) did not increase latency despite the fact that it also provided raw stimulation to the region representing the fixation zone. This finding is inconsistent with several studies reporting an inhibited saccadic response following a large transient onset. We also found a steady decrease in saccade latency as patch size increased from 1.6 x 1.6 deg to 4.5 x 4.5 deg. We conclude that non-target stimulation of the region representing the fixation zone is not in itself sufficient to produce the increase in latency typically found with remote distractors. The results are consistent with the idea that only a spatially confined object leads to a discharge of collicular fixation neurons.

**Exogenous and Endogenous Attention Shifts During Smooth Pursuit Eye Movements**Nathalie E. Ziegler<sup>1</sup> and Dirk Kerzel<sup>2</sup><sup>1</sup>Justus-Liebig-University Giessen, <sup>2</sup>Université de Genève

nathalie.e.ziegler@psychol.uni-giessen.de

Studies on the interaction of visual attention and saccadic eye movements argue in favor of an obligatory coupling of saccade goal and focus of visual attention (e.g., [1]). In a similar fashion, Khurana and Kowler [2] proposed a coupling of visual attention and the target of smooth pursuit eye movements. The current study investigated whether exogenous and endogenous attention shifts [3] are possible during smooth pursuit eye movements. Subjects foveated a stationary fixation cross (fixation conditions) or a moving cross (smooth pursuit conditions). An exogenous or endogenous cue indicated the location of the upcoming discrimination target (symbols “E” and mirror “E”) with a probability of 80%. The difference in discrimination performance between valid and invalid cue conditions was taken as a measure of visual attention. The discrimination target was either stationary or moved at the same velocity as the pursuit target. Stationary discrimination targets in space were stationary on the retina with eye fixation, but not with smooth pursuit and vice versa for moving targets. Effects of exogenous and endogenous cueing were about the same size for fixation and pursuit conditions. Further, it did not matter whether the discrimination target was stationary or moving. However, retinal motion of the discrimination target influenced smooth pursuit gain. Pursuit gain decreased when the discrimination target was stationary in space and moved across the retina, but not when it moved with the pursuit target and was approximately stationary on the retina. The reduction of pursuit gain with the onset of the discrimination target was present for exogenous and endogenous cues, but it occurred earlier in the exogenous than in the endogenous condition (~150 ms vs. ~350 ms after discrimination target onset). The results suggest that the coupling of visual attention to the target of a smooth pursuit eye movement is not as strong as has been proposed.

[1] Deubel, H. and Schneider, W. X. (1996): *Vis. Res.* 36(12), 1827–1837

[2] Khurana, B. and Kowler, E. (1987): *Vis. Res.* 27(9), 1603–1618

[3] Posner, M. I. (1980): *QJEP* 32(1), 3–25

Supported by the Deutsche Forschungsgemeinschaft (DFG KE 825/3–1 and 825/4–1,2)

## **Visual Perception: Motion**

## Illusory Contours and the Amodal Grouping of Visual Motion

Dennis Barteit, Pierre Bayerl and Heiko Neumann

Universität Ulm

dennis.barteit@informatik.uni-ulm.de

The human visual system is able to group spatially separated visual motion information and thus interpret the visual motion of objects, even if only parts of it are visible. In this work, we investigate the ability of perceiving coherent motion from partly occluded objects. Our experiments examine the perception of a circulating outlined square occluded by opaque squares which are defined by illusory contours, such as Kaniza squares or Varin figures (also squares). Thus, the occluding objects have the same luminance as the background and do not have a closed contour. As a consequence, the moving elements in the presented stimuli consist of moving line segments (fragmented parts of the moving square) for all experiments. We employ a test paradigm where observers have to rate the perceived coherence in visual motion on a rate scale (compare McDermott, *Perception*, 30, 2001). The first two experiments investigate the influence of occluding squares defined by (A) filled Pacmans of different sizes and (B) Pacmans defined by groups of circle segments (line-end Pacmans). Our results for (A,B) indicate that the perceived motion coherence is roughly proportional to the perceived strength of the illusory contour (Shipley, Kellman, *Percept. Psychophys.*, 52(1), 1992; Leshner, Mingolla, *Vis. Res.*, 33(16), 1993). In a third experiment (C) we changed the length of the moving line-elements, such that the line ends are inconsistent with the illusory contour. Experiments where we shortened the line segments revealed a gradual decrease of the strength of perceived coherence, while the percept of coherent motion abruptly collapses in experiments where line ends were prolonged. This may be explained by the fact that in the latter case the line segments overlap with illusory contours which indicates that the moving bars are seen in the front of the illusory contours. Hence, the bars are not occluded by any objects and thus are not likely to be grouped together. In conclusion, we demonstrate how illusory contours generate surface properties which influence the perception of coherent motion similar as real surfaces. We show that the effect of such illusory contours on the amodal grouping of moving fragments is roughly proportional to the strength of the illusory contour. We further demonstrate that conflicting cues, where the presented bars occlude the illusory contours, instead of being occluded by the illusory surface, abruptly suppress the amodal grouping of visual motion.

## The Role of Attention in Figure-Ground Separation of Transparent Visual Motion—A Modeling Approach

Pierre Bayerl and Heiko Neumann

Universität Ulm

`pierre@neuro.informatik.uni-ulm.de`

The role of attention in motion integration has been widely investigated, such as for the perception of transparent motion patterns (e.g., [1]). Yet, the underlying neural mechanisms involved in the interpretation of such transparent cues remain unclear. In terms of a limited capacity for the representation of visual motion, transparency presents a difficult problem for motion segmentation because multiple velocities have to be represented at each spatial location. We analyze a neural model of feedforward/feedback processing in area V1 and MT that is capable to solving the motion aperture problem [2] and propose an extension of it that explains how transparent motion is integrated and segregated in early parts of the dorsal pathway. We restrict our investigations on dense motion, such as projections of moving textured surfaces which may be transparent and tilted in space or random dot kinematograms. First, we analyze how an unspecific attention signal from higher areas (e.g., to select any rightward motion in the entire visual field) biases the disambiguation process realized by feedback between model V1 and MT. Results from model simulations demonstrate that in the presence of transparent motion cues, attentional modulation selects objects within a range of certain velocities. If attention is changed the model is able to switch to other objects with other velocities (depending on the attention signal). Second, we employ a pair of figure/ground layers of motion sensitive cells to represent transparent motion. We demonstrate that even without explicit transparency detection the motion of the transparent figure is segregated from ground motion in the presence of transparency, while no such separation occurs for opaque motion. Our simulations substantiate that attention is necessary for the separation process, even if it only coarsely selects a moving pattern in velocity and spatial domain. In conclusion, the presented contributions demonstrate the role attention for the integration and segregation of transparent motion. We show that attentional modulation allows selecting desired motion information in the presence of transparency. For a layered representation of image motion, model simulations reveal that attention is crucial for the separation of transparent motion into different layers. Most important, the model separates motion signals only in the presence of transparency and does not rely on an explicit detection of transparent regions.

[1] Felisberti and Zanker. *Perception Supp.*, 33, 2004

[2] Bayerl, Neumann. *Neural Comp.*, 16(10), 2004

### Motion-Induced Shift and Navigation in Virtual Reality

Bettina Friedrich,<sup>1</sup> Franck Caniard,<sup>2</sup> Astros Chatziastros,<sup>2</sup> Pascal Mamassian<sup>1</sup> and Ian M. Thornton<sup>2</sup>

<sup>1</sup>University of Glasgow, <sup>2</sup>Max Planck Institute for Biological Cybernetics, Tübingen

bettina@psy.gla.ac.uk

De Valois and De Valois [1] showed that moving Gabors (cosine gratings windowed by a stationary 2-dimensional Gaussian envelope) are locally misperceived in their direction of motion. In a pointing task, Yamagishi, Anderson and Ashida [2] reported even stronger visuo-motor localization error especially when participants had to make a speeded response. Here, we examined motion-induced bias in the context of an active navigation task, a situation in which perception and action are tightly coupled. Participants were presented with a birds-eye view of a vertically moving contour that simulated observer motion along a path. Observers centrally fixated while the path and a moving Gabor target were presented peripherally. The task was to follow the path with the moving Gabor, whose position (left/right) and direction (towards left/right) were varied in separate blocks. Gabor eccentricity was constant relative to fixation, with observers adjusting their simulated position with a joystick. Deviations from the path were analyzed as a function of Gabor direction. We found large and consistent misalignment in the direction of the moving Gabor, indicating that global position/motion judgments during action can be strongly affected by irrelevant local motion signals.

[1] De Valois, R.L. and De Valois, K.K. (1991): *Vis. Res.* 31(9), 1619–1626

[2] Yamagishi, N., Anderson, S.J. and Ashida, H. (2001): *Pro. R. Soc. Lond. B* 268, 973–977

We acknowledge the EC grant HPRN-CT-2002–00226

## Perception of Object-Movement in Goldfish (*Carassius Auratus*)

Martin Gehres and Christa Neumeyer

Universität Mainz

`martin.gehres@uni-mainz.de`

**Purpose:** Motion perception is a basic property of the visual system. Usually the optomotor response, a standard method for full field motion vision, is used to investigate the ability of motion detection. In goldfish motion vision measured with the optomotor response is colour-blind and dominated by the L-cone type [1]. Object motion as a movement of an object in front of a stationary background can also be perceived by animals. The two different kinds of motion seem to be processed by two distinct mechanisms [2][3][4]. To investigate the properties of objectmovement perception we trained goldfish to discriminate moving objects from non-moving objects. **Methods:** Four goldfish were kept separately in tanks (50 x 30 x 30 cm). Two quadratic openings (5 x 5 cm, 5cm apart) in a 'feeding plate' served as test fields, on which two scenes were shown using a Liquid Crystal Display. Food reward was given at the training stimulus. **First experiment:** Animals were trained on a moving black dot (diameter: 1 cm) on a white background vs. a uniform white background. Object movement vision was tested in transfer experiments: a moving dot was tested against a stationary dot, a moving black square against a stationary square, a stationary square against a stationary dot and a moving dot against a stationary dot with some unmoving dots in the background, in both cases. **Second experiment:** The fish were trained to detect a group of 20 moving dots (each dot 1 mm in diameter) within a background dot pattern vs. a stationary dot pattern. **Results:** All goldfish were able to discriminate the black moving dot from the white field. However, only one fish seems to have distinguished between a moving and stationary object independent of the object's shape in the transfer experiments. The others just showed a reaction to a black dot. They did not make a distinction between a black moving dot and a black stationary one. Also they did not react to a black square neither moving nor stationary. In the transfer test 'stationary square vs. stationary dot' they chose the stationary dot. In the second experiment all goldfish were able to distinguish between a moving and a stationary object. **Conclusion:** The random dot pattern is the appropriate method to test the perception of object-movement of goldfish. The fish were always able to detect the object motion, irrespective of one does not have to know whether the fish learned the shape of the moving object or the motion itself. Using this methods do not need to know which aspect of the training situation was associated with food reward. We are not able to say, whether the fish have a concept of 'moving vs. stationary' or not. This method is independent of such a concept, because the fish can see the object only in the moving scene. Now we can analyse further parameters of the perception of object-movement in goldfish such as contrast, speed and colour.

- [1] Schaerer, S. and C. Neumeyer (1996). "Motion detection in goldfish investigated with the optomotor response is "color blind"." *Vision Res* 36(24): 4025–34
- [2] Egelhaaf, M., K. Hausen, et al. (1988). "Visual course control in flies relies on neuronal computation of object and background motion." *Trends Neurosci* 11(8): 351–8
- [3] Frost, B. J., D. R. Wylie, et al. (1990). "The processing of object and self-motion in the tectofugal and accessory optic pathways of birds." *Vision Res* 30(11): 1677–88
- [4] Born, R. T. and R. B. Tootell (1992). "Segregation of global and local motion processing in primate middle temporal visual area." *Nature* 357(6378): 497–9

## Shifting the Onset of a Moving Sound Source: An Auditory Fröhlich Effect?

Stephan Getzmann

Ruhr-Universität Bochum

stephan.getzmann@rub.de

When observers are presented with a visual target in motion, they typically remember its onset position to be displaced in the direction of motion. This phenomenon, the Fröhlich effect, has been shown to depend on a number of variables, such as the target velocity, the predictability of the target onset position, and the psychophysical method used to determine the onset position: A strong Fröhlich effect usually occurred with relatively high target velocities, onset positions restricted to a small area of the visual display, or probe judgments [1,2,3]. The present study investigated the influence of motion on the perceived onset position in the auditory modality.

In a dark anechoic environment, an acoustic target (short noise pulses) appeared randomly at a peripheral or a central azimuthal position and moved from left to right or from right to left along the frontal horizontal plane (motion velocity 12 deg/s). Relative judgments were made to determine the onset position of motion: Listeners were asked to compare the onset position of the auditory target to a 2-s visual reference stimulus presented at the left or right of the auditory onset position. In a control condition, stationary target sounds were presented at the onset positions of the motion stimuli used in the main condition.

In general, compared with the positions of the stationary targets, the onset positions of the motion stimuli were localized as displaced in the direction of motion. The most prominent displacements ( $7.1^\circ$ , SE  $\pm 0.4^\circ$ ) occurred when the visual reference stimulus was presented after the auditory motion. With the reference stimulus presented before the auditory motion, the displacement was significantly reduced ( $1.9^\circ$ , SE  $\pm 0.5^\circ$ ). Moreover, displacements were stronger with peripheral ( $6.9^\circ$ , SE  $\pm 0.5^\circ$ ) than with central ( $2.1^\circ$ , SE  $\pm 0.6^\circ$ ) onset positions.

These findings suggest the existence of a potential analogue of the Fröhlich effect in the auditory modality. Although its origin is still unclear, errors in auditory short-term memory seem to contribute to the effect, which do not occur with judgments based on an immediate comparison of target onset position and reference stimulus. The increase of displacement with peripheral onset positions might be related to a general increase in localization error beyond the auditory median plane.

- [1] Kerzel, D. and Gegenfurtner, K. (2004): *Vis. Res.* 44, 577–590
- [2] Müsseler, J. and Kerzel, D. (2004): *Vis. Res.* 44, 2201–2206
- [3] Kerzel, D. (2002): *Exp. Brain Res.* 145, 340–350

## The Advantage of Left-Handers in Sports Due to a Perceptual Frequency Effect

Norbert Hagemann and Arndt Ziegler

Westfälische Wilhelms-Universität Münster

nhageman@uni-muenster.de

Left-handers occur at higher frequencies in many interactive sports, such as tennis, fencing or cricket [1]. The fact that left-handers do not occur more frequently in non-interactive sports contradicts the assumption of a general innate or neuropsychological superiority of left-handers in sports [2, 3]. However, left-handers rather seem to enjoy a negatively frequency-dependent strategic advantage compared to right-handers [4–6]. A sporting competitor playing against a left-hander is confronted with an unusual strategic play and angles of attack. We assume that the strategic or tactical advantage of left-handedness in interactive sports is partly due to a perceptual frequency effect of handedness. The perceptual frequency effect should result in better anticipatory skills of sporting competitors when facing right-handed opponents than left-handed opponents. Right-handed ( $n=54$ ) and left-handed tennis players ( $n=54$ ) (experts, intermediates and novices) participated in a tennis anticipatory test where they had to predict the directions of opponent's tennis strokes on a computer screen (point of impact in their own half). The 48 video clips of different strokes (serves, volleys and groundstrokes) of two right- and two left-handed male intermediate tennis players were temporally occluded at the racquet-ball contact. These video clips were also mirrored along the vertical axis, so that the same clip were visible as a stroke of a left-hander and a right-hander. The 2 x 3 (Handedness x Group) repeated measure ANOVA revealed a highly significant main effect for handedness of the player in the clip (independent of actual handedness),  $F(1, 102) = 83.74$ ,  $p < .01$ ,  $h^2 = .45$ , indicating that the direction of right-handed tennis strokes could be easier anticipated. This is evident in all three groups. The significant Handedness x Group interaction demonstrating that the experts have the most problems when facing left-handers,  $F(2, 102) = 3.77$ ,  $p < .05$ ,  $h^2 = .07$ . The results of this study suggest that the advantage of left-handers in interactive sports is at least partially accounted for by perceptual frequency effects.

- [1] Grouios, G. (2004): *Percept Mot Skills* 98, 53–66.
- [2] Raymond, M. et al. (1996): *Proc R Soc Lond B Biol Sci.* 263, 1627–1633.
- [3] Grouios, G. et al. (2000): *Percept Mot Skills* 90, 1273–1282.
- [4] Brooks, R. et al. (2004): *Proc R Soc Lond B Biol Sci* 271, 64–66.
- [5] Faurie, C. and Raymond, M. (in press): *Proc R Soc Lond B Biol Sci.*
- [6] Goldstein, S.R. and Young, C.A. (1996): *J Comp Psychol* 110:164–169.

**Spectral Sensitivity of Apparent Motion does not Follow  $V(\lambda)$  but  $[V(\lambda)]^2$  !**

Philipp Haselmayer, Christa Neumeyer and Christoph von Campenhausen

Johannes Gutenberg-Universität Mainz

neumeyer@uni-mainz.de

Purpose: Perception of movement is a basic property of visual systems. The phi-phenomenon or apparent motion is elicited by spatially separated light stimuli flashed with an appropriate time interval. This kind of stimulation should also be the adequate stimulus for the hypothetical Reichardt-Hassenstein motion detector. Yet there has been no examination of the spectral sensitivity in humans for the detection of apparent motion. To investigate the spectral sensitivity we used monochromatic stimuli of different wavelengths. The sensitivity function was compared with the human cone sensitivity functions and with  $V(\lambda)$ . In this way we could determine whether the phi-phenomenon is detected by the magnocellular or the parvocellular system. Methods: Three subjects were examined. The phi-phenomenon was investigated foveally by presenting two monochromatic lights (horizontal distance: 0.8 deg) of the same wavelength, sequentially flashing with a time delay of 15 to 60 msec. 15 wavelengths between 450 nm and 670 nm were tested each with 4–5 intensity steps 0.5 log units apart. For each monochromatic light stimulus the subjects were asked at which time difference (tested in steps of 2 msec) the flashes appeared synchronous or moving. In this way the time difference at which motion was just detected could be determined. Results: The time difference at which motion was perceived decreased with increasing intensity for all wavelengths. For a constant intensity (1011 and 1012 quanta/cm<sup>2</sup> sec) smallest time delays of about 20 msec were found in the wavelength range between 550 and 580 nm. The spectral sensitivity function, i.e. the amount of quanta/cm<sup>2</sup> sec necessary to obtain perceived motion for time differences of 23–27 msec, is maximal between 550 nm and 580 nm and corresponds in this respect to  $V(\lambda)$ . The shape of the function, however, is much more narrow. The data are fitted accurately by  $[V(\lambda)]^2$ . Conclusion: We assume that the surprising result that the spectral sensitivity of the phi-phenomenon does not follow  $V(\lambda)$  but its square is due to the multiplicative processing of the correlation as proposed in the Reichardt-Hassenstein motion detector.

**Global Motion Thresholds in Albinotic and Pigmented Ferrets (*Mustela Putorius Furo*)**

Daniela Hupfeld and Klaus-Peter Hoffmann

Ruhr-Universität Bochum

daniela.hupfeld@rub.de

The visual system of albino mammals shows various deficits of retinal differentiation and visual system development. One of the effects is usually a deficit in the optokinetic nystagmus (OKN). This reflex is elicited by full field visual movements and stabilises the image on the retina. Albino ferrets don't show an OKN at all, they are "optokinetic blind". One of the possible causes for this total lack of visual evoked OKN may be a lack of perception of the moving stimulus.

Are albino ferrets motion blind?

A visual discrimination task was performed with 17 ferrets: 4 wildtype pigmented (WT) and 6 albino (ALB) males, 3 WT and 4 ALB females. Using a 2AFC paradigm they were tested for global motion perception with moving random dot patterns (RDP). A coherent movement had to be distinguished from a dynamic noise, the choice of the coherent RDP was rewarded by food. Compared to males female ferrets show a significantly lower discrimination performance. Coherence detection was equally well possible for albino and wildtype ferrets of both sexes, thus albino ferrets are not motion blind.

Based on this result the signal strength (% of coherence) was reduced and individual thresholds were measured using the constant stimuli method. Female ferrets have significantly higher thresholds than males, but for both sexes a phenotype effect was found. Albino ferrets need about twice the signal strength necessary for WT ferrets.

Albino ferrets are not motion blind, but higher coherence thresholds reveal an impairment in global motion detection.

A cortical area with a high percentage of directional selective cells could be identified in the ferret by an electrophysiological study. This area, located at the posterior bank of the suprasylvian sulcus, is referred to as the posterior suprasylvian area (PSS). To evaluate the role of area PSS in global motion perception lesion experiments were done in one WT male, one WT and one ALB female. Postlesion performance showed a severe impairment of coherence detection ability. Control lesions (extrastriate areas) in three ferrets of corresponding sex and phenotype prove the specificity of the motion detection deficit to PSS lesions.

This work was supported by the DFG, SFB 509.

## Mechanisms of Spatio-Temporal Interpolation

Farid I. Kandil and Markus Lappe

University of Münster

kandil@uni-muenster.de

When a moving object is observed through a picket fence, in every instance of time only a small part of the figure can be seen [eg 1, 2, 3]. However, the brain seems to provide mechanisms ready to re-assemble these slit-form views into the view of a whole object again. More precisely, subjects perceive the object as if all parts of the object were presented simultaneously (temporal interpolation) and as if the parts were not presented successively at the same spots, namely within the narrow slits, but rather next to one another, that is also in-between the slits (spatial interpolation). This interpolation process is thought to be accomplished iteratively. First, the motion direction of the object has to be detected to decide in which direction to 'place' successive views next to one another. Using dichoptic as well as binocular masking displays, here we show that this object motion is identified by binocular rather than monocular motion detectors. Second, there has been a long-standing discussion on whether this motion information can directly feed the pattern perceiving mechanism or whether it rather has to trigger eye movements in the preferred direction. (Aside: When the eyes move with the object, successive views are not placed any more on the same spot of the retina but rather on neighbouring ones. This simple 'hard-ware' solution of the interpolation problem is called 'Retinal Painting Theory'). Using eye trackers, we found that people do perform smooth-pursuit eye movements in all trials for which they can report the object correctly, even if they report to have kept their eyes still. In a different experiment, smooth-pursuit eye-movements were triggered by external stimuli. When stimuli for which the object motion was otherwise ambivalent were viewed under these conditions, people were able to perceive the interpolated object and to respond accordingly. Third, much like in the motion task, we started investigating whether the form processing (assembling) mechanism is binocular or monocular by using dichoptic displays and others with binocular masking as well as again others with stereo-information and obtained preliminary results pointing rather to a binocular mechanism, hence partly contradicting earlier reports [2]. We also discuss our findings in the light of a recent report [4] and point out that the main difference between the displays does account for the diverging results. Moreover we discuss the implications our findings may have on modern models of motion detectors.

- [1] Burr, D.C. (1979) *Vis.Res.* 19, 835–837
- [2] Fahle, M. & Poggio, T. (1981) *Proc.R.Soc.B*, 213, 451–477
- [3] Morgan, M.J. & Watt, J (1983) *Vis.Res.* 23, 997–1004
- [4] Nishida, S. (2004) *Curr.Biol.*, 14, 830–836

Supported by EU Ecovision Grant

## Recognition of Biological Motion from Blurred Natural Scenes

Simone Kuhlmann and Markus Lappe

Westfälische Wilhelms Universität, Münster

nimonimo@gmx.de

Biological motion perception is often considered a template matching process. We are interested in the visual cues in this template. Studies of biological motion perception usually use Johansson's point-light displays. These stimuli are very abstract. To use more natural stimuli we recorded movies of different actions in natural scenes. By blurring the scenes we modified the visual cues, particularly the local form and motion information. Observers were asked to identify the action portrayed.

In a first experiment, we presented movies of a real walking person at various blur levels to a group of naive subjects. Even in highly blurred stimuli recognition rate was over 90%. Upside-down presentation reduced recognition rates massively similar to the classical Johansson display.

To measure the extend of static shape information in the template we compared recognition from complete action movies (100 frames, 25 frames per sec) to recognition from single still frames, both in eight different blur levels. For unblurred stimuli, action recognition was near perfect in both conditions. With increasing blur level, the decrease in recognition from still images was steeper than for complete action movies.

To investigate the contribution of global motion versus successive static shape information we compared stimuli consisting of three still frames showed separately (global shape change, no motion) to short movies consisting of three images in immediate succession (single shape and global motion, little shape change). For both conditions, recognition improved in comparison with the single still frame condition. In the more difficult blur levels, the short movies lead to better recognition than the separate still frames.

Finally, we constructed flicker movies to probe the role of local motion for shape segmentation. The flicker movies contained only the segmentation information of the motion of the short-movies, i.e., no information about direction and speed. We found an enhancement of the recognition rate in comparison with the single-frames, but recognition was below that of the other conditions.

Our results demonstrate that templates for biological motion recognition combine global form and motion cues. Reductions of local form and motion information by blurring can be compensated by global shape change and motion. Local motion information is also used for segmentation.

## Perception of Motion Induction for Naturalistic Images in the Human Visual Cortex

Scott McDonald and Zoe Kourtzi

MPI for Biological Cybernetics, Tübingen

`scott.mcdonald@tuebingen.mpg.de`

Our visual perception often differs from the physical reality of the natural world. We exploited this discrepancy, with a perceptual illusion known as motion induction, to examine motion processing of naturalistic stimuli in the human visual cortex. When an incoherently moving stimulus is placed in a coherently moving surround, observers perceive the incoherent pattern moving in the opposite direction to the surround. We used psychophysics and fMRI to investigate the neural basis of motion induction for naturalistic images. Specifically, we used natural texture stimuli with  $1/f$  amplitude spectra that consisted of a central region in a moving surround. The surround moved up or down with 100% coherent motion. We manipulated the coherence of the motion of the center: the center moved (up or down) at different coherence levels between random and 100%. When tested psychophysically, subjects misreported the direction of motion of the center. When the center had 0% coherence, observers perceived it moving in the opposite direction to the surround. The observers' point of subjective equality (PSE), i.e. when observers reported the center moving up and down an equal number of times, occurred when it had 30–50% coherent motion in the same direction as the surround. Based on the physical properties of the stimulus, we predicted that fMRI responses would be lower at 0% coherence than at the PSE where the motion coherence is higher. Alternatively the perceptual results predicted that the difference between these two conditions would be absent, or even reversed. That is; stronger fMRI responses would be observed at 0% coherence where there is more perceived motion than at the PSE. Our results in hMT+/V5 showed that fMRI responses correlate with the perceptual rather than the physical coherence in the stimulus. Our findings suggest that the motion induction effect can be mediated by a combination of motion coherence and motion contrast mechanisms in hMT+/V5.

## **A Motion Sensitive Area in Ferret Extrastriate Visual Cortex: An Analysis in Pigmented and Albino Animals**

Roland Philipp, Claudia Distler and Klaus-Peter Hoffmann

Ruhr-Universität Bochum, Germany

philipp@neurobiologie.rub.de

In an attempt to localize the neuronal substrate for motion analysis in the ferret (*Mustela putorius furo*), we extracellularly recorded from extrastriate visual cortex in 5 pigmented and 2 albino ferrets, each of them in 3–5 sessions. The animals were initially anaesthetized with 20mg/kg keta-mine and 2mg/kg thiazinhydrochloride, intubated through the mouth and placed into a stereotaxic frame. The craniotomy was performed under additional local anaesthesia with bupivacainhydro-chloride. During electrical recordings they were anesthetized with halothan (0.2%-0.4% in air as needed) and positioned stereotaxically by means of an implanted head post into the experimental setup. The visual stimuli composed of a random dot-pattern could be moved in a circular pathway and as expanding or contracting radial motion simulating self-motion (optic flow). Peri stimulus time histograms (PSTH) were generated online and stored in a PC for offline analysis.

Altogether 235 neurons were recorded in pigmented, 132 neurons in albino ferrets. In 92% of the neurons in pigmented and in 72% of the neurons in albino ferrets stimulation in the preferred direction elicited at least twice as many spikes than stimulation in the nonpreferred direction. Therefore, we tentatively call this motion sensitive area the posterior suprasylvian area (PSS).

Additionally, 192 neurons in pigmented and 116 neurons in albino ferrets were tested with the optic flow stimuli. 35% of the neurons were classified as expansion cells and 13% as contraction cells in pigmented ferrets. In albino ferrets only 20% of the neurons were classified as expansion and 12% as contraction cells. Thus expansion cells were significantly more frequent than contraction cells only in pigmented ferrets.

Electrical stimulation experiments in the ipsiversive nucleus of the optic tract and dorsal terminal nucleus (NOT-DTN) of the pigmented ferrets have shown that PSS could be antidromically activated and thus exhibits projections to the NOT-DTN.

All these data suggest that PSS may play an important role in motion and direction processing in ferret visual cortex and, by its direct projection to the NOT-DTN, in the optokinetic reaction (OKR). Furthermore our data indicate a reduced direction selectivity in albino ferret PSS.

Supported by the DFG (SFB 509-A11).

## Apparent Speed in the Footstep Illusion is not Totally Contrast Dependent

Paolo Pretto<sup>1</sup> and Astros Chatziastros<sup>2</sup>

<sup>1</sup>University of Padua, Italy; MPI for Biological Cybernetics, Tübingen,

<sup>2</sup>MPI for Biological Cybernetics, Tübingen

paolo.pretto@tuebingen.mpg.de

Recently, the role of luminance contrast on the perception of velocity has been emphasized by the so called “footstep illusion” [1]: when a grey bar drifts steadily across a background of stationary black and white stripes, it appears to stop and start as its contrast to the background changes. When two bars of unequal grey values drift in parallel, an impression of alternating footsteps of a moving person can be elicited. This change in apparent speed has been attributed to the different instantaneous contrasts of the grey bar against the white and black areas of the background (Contrast Ratio explanation, [1]). We report a series of experiments with further variants of the footstep illusion to test the validity of the Contrast Ratio (CR) explanation. In an adjustment task, we presented a standard footstep illusion together with second comparison stimulus. A total of five participants were required to adjust the contrast of the background of the comparison stimulus in order to match the strength of the standard illusion. The moving bars were presented in six luminance levels in the range between 6 and 103 cd/m<sup>2</sup>. For each luminance level we computed the CR by the Weber fraction, following the original study. We compared the CR values of the standard and comparison stimulus and expected to obtain comparable CRs within matched illusion pairs. In the first experiment we found a discrepancy between predicted and obtained CR values. The strength of the illusion was matched at an increased CR, i.e. participants set the background stripes to higher contrast in order to obtain the same illusion. In a second experiment the moving bars were rendered transparent, with a transparency of 50%, and we obtained basically the same results as in the first experiment. In the last experiment we tested whether the footstep illusion can be elicited with isoluminant red/green patterns. Individual isoluminance levels were identified with flickering stimuli. Subjects rated an apparent velocity change of the isoluminant moving bars again by adjusting the contrast of the standard luminance version. The results show clearly that the footstep illusion occurs also with isoluminant patterns. Furthermore, we found that the adjusted CR was always near the maximal possible value, suggesting that participants matched the isoluminant illusion to the strongest available luminance illusion. The present data suggest that foreground and background stimuli contribute differently to the footstep illusion, a finding which is not predicted by the current Contrast Ratio account. Most notably, the existence of the illusion with isoluminant stimuli shows the clearest limitation of the Contrast Ratio explanation. That the strength of the isoluminant footstep illusion was comparable to the luminance version, suggests a crucial role of chromatic input in an adequate explanation of this phenomenon.

[1] Anstis, S. (2001): Perception 30(7), 795–794.

The authors wish to thank Prof. Dr. Heinrich H. Bülthoff and Prof. Sergio Roncato.

**Recognizing Emotions Expressed in Human Walking: Integration of Synergies**

Claire L. Roether and Martin A. Giese

ARL, Dept. of Cognitive Neurology, University Clinic Tübingen

`claire-louise.roether@student.uni-tuebingen.de`

One of the major hypotheses in the motor-control domain is that for the production of complex body movements not all degrees of freedom (i.e., joints) are controlled separately. Instead, subsets of degrees of freedom have been suggested to be controlled together, in synergies. Such synergies may have an influence on the visual perception of body movements. We studied the way in which information carried in the two synergies of arm and leg motion is integrated for recognizing different emotions expressed in human walking. We specifically tested whether information is combined in a statistically optimal fashion. To this end we recorded neutral and emotional (angry, happy, sad and fearful) human gaits using a 3-D motion-capture system. Modifying a technique of Giese and Poggio [1] we created stimuli by motion morphing between neutral walking and the four different emotional gaits. Stimuli could belong to three different categories: (1) full body emotional: all degrees of freedom morphed simultaneously; (2) emotional upper body: morphing arm movements only, always combined with the leg movement of the neutral walk; (3) emotional lower body: morphing leg movements only, combined with the arm movement of the neutral walk. Stimuli were presented as point-light walkers. It was found that the morphed stimuli look highly natural, even when emotional arm motion is combined with neutral movement of the lower body, or vice versa. The emotions differ in the amount of information carried by upper-body and lower-body synergies. Consistently with previous work, emotions become more easily recognizable when movements are caricatured in space-time. Exaggeration also increases the recognizability of emotions when it is applied to individual synergies only. Presently we focus on investigating the integration of information from the two synergies. We test whether the combination of information is accomplished in a statistically optimal way by fitting the data with Bayesian ideal-observer models.

[1] Giese, M.A. and Poggio, T. (2000): *International Journal of Computer Vision* 38(1), 59–73.

Supported by the Human Frontier Science Program, the Deutsche Volkswagenstiftung, and the Hermann and Lilly Schilling Foundation

## Categorization of Complex Dynamic Patterns in the Human Brain

Pegah Sarkheil,<sup>1</sup> Jan Jastorff,<sup>2</sup> Martin A. Giese<sup>2</sup> and Zoe Kourtzi<sup>1</sup>

<sup>1</sup>MPI, Tübingen, <sup>2</sup>University Clinics Tübingen

pegah.sarkheil@tuebingen.mpg.de

The ability to categorize actions is critical for interacting in complex environments. Previous studies have examined the neural correlates of categorization using static stimuli. The goal of our study was to investigate the neural substrates that mediate learning of complex movement categories in the human brain. We used novel dynamic patterns that were generated by animation of an artificial skeleton model and presented as point-light displays. We created prototypical stimuli that differed in the spatial arrangement of their segments and their kinematics. Intermediate stimuli between the prototypes were generated by a weighted linear combination of the prototypical trajectories in space-time. We compared fMRI activations when the observers performed a categorization vs. a spatial discrimination task on the same stimuli. In the categorization task, the observers discriminated whether each stimulus belonged to one of four prototypical classes. In the spatial discrimination task, the observers judged whether each stimulus was rotated (or translated) leftwards vs. rightwards. These tasks were matched for difficulty based on the observers' performance during a practice session. We observed significantly stronger fMRI activations for the categorization than the spatial discrimination tasks in the dorsal, inferior parietal and the medial, inferior frontal cortex, consistent with previous findings on the categorization of static stimuli. Interestingly, we also observed activations in visual motion areas (V3a, hMT+/V5), higher-order motion areas in the intraparietal sulcus (VOIPS, POIPS, DIPSM, DIPSA) and parieto-frontal areas (supramarginal gyrus, postcentral gyrus, ventral and dorsal premotor cortex) thought to be involved in action observation and imitation. These findings suggest that categorization of complex dynamic patterns may modulate processing in areas implicated in the analysis of visual motion and actions.

## Contrast-Sensitivity of Goldfish, Measured with the Optomotor Response, is Independent from Glycine- and NMDA-Receptor Mediated Actions

Ruth-Barbara Schmidt-Hoffmann and Carlos Mora-Ferrer

J.Gutenberg Universität Mainz

rutschmidt@gmx.de

**Purpose:** Different retinal transmitters are involved in retinal motion coding. Acetylcholine, GABA, and glutamate have been shown to be essential for the coding of full-field motion in the retina. A function of glycine for retinal motion coding in fish has not been described until now. The contribution of glycine and glutamate, acting through glycine receptors, and of glutamate, acting through NMDA-receptors (NMDA-R), can be evaluated using receptor-specific antagonists. Therefore, contrast-sensitivity of the optomotor response of goldfish was measured and the effect of the glycine-antagonist strychnine and the NMDA-antagonist AP7 was investigated. **Methods:** The photopic optomotor response to two stimulus velocities (3 and 10 rounds per minute [RPM]) was elicited utilizing paper cylinders (24 cm diameter) with sine-wave gratings (4 cm period) of differing contrasts (80, 60, 40, 20 and 10%). Contrasts were varied around the medium remission value of the paper pattern cylinders. Strychnine, a non-specific glycine antagonist, was injected into the vitreous of both eyes and used at intra-vitreous concentrations between 50 nM-5µM. AP7, a NMDA-R specific antagonist was used at intra-vitreous concentrations of 30 and 60 µM. **Results:** Results were similar for all Strychnine concentrations tested. For both pattern velocities the optomotor response decreased with contrast. The half-maximal response was elicited at about 25% contrast, irrespective of pattern velocity. Normalization of pre- and post-injection data for 3 and 10 RPM pattern velocity revealed nearly identical values for the optomotor response as a function of contrast. There is no significant difference between pre- and post-injection data. A slight difference could be observed between the contrast sensitivity functions at contrast 60% and 40% achieved with 3 and 10 RPM pattern velocity. The NMDA-R antagonist AP7 also had no effect on the contrast-dependent optomotor response of goldfish except for the highest contrast used. However, there was no difference between the post-injection data for the contrasts 40, 60, and 80%. This indicates, that the difference between pre- and post-injection data is too small to for a substantial impairment of the contrast-dependent motion perception. **Conclusion:** The contrast-sensitivity of the optomotor response is independent of pattern velocity. The half-maximal response is elicited by a contrast value of about 25%. This is similar to values obtained for guppy fish [1]. Neither glycine nor the NMDA-R contribute to the neuronal processing underlying contrast-coding nor to the directional-coding process.

- [1] Anstis, S., Hutahajan, P. and Cavanagh, P., Optomotor test for wavelength sensitivity in guppyfish (*Poecilia reticulata*), *Vision Res*, 38 (1998) 45–53.

## Does the Paradoxical Perception of Motion in Depth Reflect an “Optimal” Solution?

Andrew E. Welchman

MPI for Biological Cybernetics, Tübingen

`andrew.welchman@tuebingen.mpg.de`

How do we decide whether an object approaching us will hit us? The optic array provides information sufficient for us to determine the approaching trajectory of a projectile. However, when using binocular information, two previous studies [1,2] have reported that observers overestimate the angular approach of a presented motion-in-depth trajectory. Specifically, observers report that an object on a collision course with their head will miss it. From an evolutionary perspective this appears to make little sense. However, here we consider whether this perceptual phenomenon might actually reflect “optimal” behaviour by the visual system. Our experiment consisted of two stages. First, we recorded thresholds for detecting an increment in a distance moved in depth ( $\theta_d$ ) or a distance moved laterally ( $\theta_l$ ). Second, we recorded observers’ reports of the perceived trajectory of a point moving towards them in depth. Trajectories were chosen to span a range of angles with respect to the mid-sagittal plane and all trajectories consisted of lateral- and depth- components of motion. We then constructed a Maximum Likelihood model (cf. [3]) for weighting the lateral- and depth- components of the presented motion trajectory. Specifically, we broke down the presented trajectories into a component in depth and a lateral motion component. Then, using the  $\theta_d$  and  $\theta_l$  thresholds we weighted the two motion components in accordance with the observers’ sensitivity to each. Finally, using these weighted motion components we calculated a trajectory angle based on a simple trigonometric ratio. This simple model captured characteristic aspects of the observers’ behaviour, and in some cases a very good fit to the data. We suggest that at least part of the explanation for the apparent paradox in perceiving motion towards the head results from the discrepancy between sensitivity to lateral motion and motion-in-depth.

[1] Harris, JM & Dean, P (2003) *J. Exp. Psych*, 29, 869

[2] Welchman, AE, Tuck, V & Harris, JM (2004) *Vis. Res*, 44, 2027

[3] Clark, JJ & Yuille, AL (1990) *Data Fusion for Sensory Information Processing Systems*, Kluwer, Boston

Max-Planck Society

**Audio-Visual Perception of Self-Induced Apparent Motion**

Mikhail Zvyagintsev,<sup>1</sup> Hans Menning,<sup>1</sup> Krystyna Swirszcz,<sup>1</sup> Argiro Vatakis,<sup>2</sup> Zoe Kourtzi<sup>3</sup>  
and Klaus Mathiak<sup>4</sup>

<sup>1</sup>Univ. Tübingen, <sup>2</sup>Univ. Oxford, <sup>3</sup>MPI Biological Cybernetics, Tübingen, <sup>4</sup>Univ. Aachen

mikhail.zvyagintsev@med.uni-tuebingen.de

Apparent motion is the perception of the realistic smooth motion of an object which flashes or sounds first at one place and then at another. In a whole-head magnetoencephalography study, we assessed neural correlates of multisensory perception of apparent motion in 12 healthy volunteers. Two successive disks (1° diameter, 100 cm distance,  $\pm 6^\circ$  eccentricity, 67 ms duration, 67 ms ISI) were displayed simultaneously with auditory white noise signals (simulated by means of auditory virtual reality at the same locations). Conditions with self-induced and random direction were compared. During the first condition, the direction of apparent motion stimuli was determined by the button press of the subject and during the second, the stimulus direction was selected randomly. The time of stimulus onset was self-induced in both conditions. Subjects were instructed to determine direction of motion after every sweep. We recorded 4 sessions with 260 sweeps in each subject. Similar evoked response fields were observed for self-induced and randomized sequences up to 140 ms after stimulus onset. The pattern accorded to distributed neuromagnetic activity. Peaking at about 160 ms, the difference field between the predictable and un-predictable condition exhibited a bilateral dipolar field structure with a higher negativity for the unpredictable stimulus directions. This pattern accords to the N1 component of auditory evoked fields. Globally, apparent motion seems to engage areas related to auditory, visual, and motor processing and posterior parietal regions. As the main contrast of interest of our study, predictable versus un-predictable, i.e., self-induced vs. random, directions affected in first place auditory areas. In a similar vein, stimulus adaptation in the auditory domain has been observed during the perception of one's own speech. This specific mechanism seems to be of high neurobiological importance as impairments of these—mainly auditory—anticipatory projections have been suggested to be influential for hallucinations in schizophrenia.

DaimlerChrysler AG with DARPA AugCog, SFB 550 B1



## **Visual Perception: Spatial Vision**

## **Saccadic Compression with near Threshold Stimuli in Light and Dark Adapted Conditions**

Katharina Georg, Fred H. Hamker and Markus Lappe

Allgemeine Psychologie, Psychologisches Institut II, Westf. Wilhelms-Universität Münster

kathageorg@hotmail.com

A bar briefly flashed around saccade onset is perceptually mislocalized with a pattern resembling a compression of distance towards the saccade target. The strength of saccadic compression depends on the contrast of the perisaccadic stimulus. Since dark adaptation reduces luminance threshold we reasoned that the adaptation state may influence the strength of perisaccadic compression observed in dark experimental environments. We investigated the impact of stimulus luminance on the degree of mislocalization in light and dark adapted conditions. Five subjects conducted visually guided rightward horizontal saccades of 20° amplitude in absolute darkness. A light bar was flashed for 8ms at a random time between -150 to 150 ms relative to the saccade onset. The bar was flashed at one of four possible positions at -15°, -7.5°, 7.5° and 15° relative to the saccade target. Subjects had to indicate the perceived position of the flashed bar. The study was split into two experiments: In the first experiment subjects were dark adapt for about 25 minutes. Dark adaptation was maintained throughout the experiment. In the second experiment subjects stayed light adapted. They performed five blocks of 10 saccades in complete darkness, but each block was followed by a 90-second refreshment of the light adaptation. In both experiments two luminance levels of the bar were compared, one near to, the other clearly above detection threshold. When stimulus luminance was near threshold, bars were perceived closer together. This compression was strongest at the beginning of the saccade. In contrast, bars well above detection threshold were perceived almost perfectly at their true positions. These findings expand the results of previous studies which had shown that compression was stronger with low contrast stimuli [1]. Since absolute luminance thresholds were lower in dark adapted state, compression for stimuli with same luminance was weaker after dark adaptation. We suggest that the strength of perisaccadic compression is inversely related to stimulus visibility.

[1] Michels, L. and Lappe, M. (2004): *Vis. Res.* 44(20), 2327–2336

**Rapid Natural Scene Discrimination: Orientation and Context Effects.**

Nick Koechy

Universität Magdeburg

nkoechy@web.de

**Purpose:** Objects in natural scenes are recognized very rapidly and seemingly without effort. Here we probed the limitations of rapid scene processing and the role of context.

**Methods:** We investigated rotation- and context effects in rapid natural scene processing using a 2AFC-paradigm. In all experiments two photographs of natural scenes were presented masked to the left and right of the fixation. All scenes contained one clearly identifiable object embedded in a background. The target scene contained an animal and the subject's task was to indicate its location by a button press. Discrimination performance was measured in each condition as a function of scene-mask SOA. Psychometric functions were fitted to the proportion of correct responses to determine threshold-SOAs (75% correct). A) The first experiment investigated the effect of orientation on scene discrimination. Both photographs were presented at one out of three orientations (0, 90, 180deg), each at multiple SOAs. B) The second experiment aimed effects of orientation mismatch between object and background. Objects (Ob) and backgrounds (Bg) were rotated independently (Ob0deg/Bg0deg (all upright), Ob0deg/Bg90deg, Ob90deg/Bg0deg, Ob90deg/Bg90deg). SOA was the independent variable. C) The third experiment focused on semantic object-background interactions. Cut out objects were presented in semantically congruent, incongruent, and neutral phase randomised backgrounds and SOA was varied. The objects were embedded into scene backgrounds of appropriate scale and at physically plausible positions.

**Results:** A) Rotation of the full scene revealed that discrimination threshold SOAs for upright and for inverted scenes did not differ (32ms 0deg, 35ms 180deg). Rotations by 90deg significantly extend the threshold SOA (55ms 90deg). B) Rotation by 90deg compared to upright presentation has a significant detrimental effect on discrimination no matter if object, background, or both are rotated (36ms Ob0deg/Bg0deg; 57ms Ob90/Bg0deg; 61ms Ob0/Bg90deg; 64ms Ob90/Bg90deg). The thresholds of the three rotation conditions did not differ. C) Objects in semantically neutral and congruent backgrounds were discriminated equally fast (37ms congruent; 40ms neutral), while threshold was nearly doubled when they were semantically incongruent (77ms incongruent).

**Discussion:** Our data indicate that rotation affects the rapid processing of natural scenes, although inversions might be processed very efficiently. Object and background appear to be processed together even at very short processing intervals. The mismatch of perceptual and semantic features has a detrimental effect on rapid object processing. Our results further suggest that congruent backgrounds do not facilitate the processing of objects in the scenes, but incongruent backgrounds impair object processing.

**Time Will Tell: The Temporal Order of Events Determines Feature Integration**

Frank Scharnowski and Michael Herzog

Brain Mind Institute Lausanne

frank.scharnowski@epfl.ch

How features of an object are bound into a unique percept is one of the puzzling problems in the cognitive and neurosciences. In order to investigate the temporal dynamics of feature binding, we used a feature fusion paradigm: a vernier (V) was immediately followed by a vernier with opposite offset direction (AntiV). Because of the very short presentation times of V and AntiV, feature fusion occurs, i.e. only one vernier is perceived. We presented various sequences of Vs and AntiVs while keeping their total physical energy (duration x luminance) constant. Surprisingly, the contribution of each vernier to the fused percept depends not only on its energy but also on the temporal order of the elements. If, for example, a V was followed by an AntiV, the AntiV dominated the perceived offset (condition V—AntiV). This changed when the V was subdivided into two equal parts, of which one was presented before and the other after the AntiV (condition  $\frac{1}{2}$ V—AntiV —  $\frac{1}{2}$ V): none of the verniers dominated. In general, our results show that any level of performance can be achieved by arranging sequences of Vs and AntiVs appropriately—even though the total physical energy of V and AntiV is identical. We conclude that for a given physical energy of V and AntiV the temporal order of presentation determines the integration of features. Different positions within the temporal sequence of events are of different importance. It seems that later elements of the sequence influence the perceived offset more than earlier ones. Our findings provide evidence that the temporal order of elements is crucial for the integration of features. These results pose serious challenges for most models of feature processing, since they are mainly energy-based while ignoring temporal aspects.

## Character Recognition and Ricco's Law

Hans Strasburger

Univ. of Göttingen and Univ. of München

strasburger@med.uni-goettingen.de

The contrast threshold for the detection of patches of light depends upon stimulus size as described by Ricco's classical law of areal summation; the critical diameter within which Ricco's law holds increases with retinal eccentricity. Here we present an analogon of Ricco's law for the recognition of characters at low contrast, and describe its variation with retinal eccentricity. Methods: Michelson contrast thresholds for the recognition of singly presented digits were determined in a 10-afc, maximum-likelihood adaptive procedure (ML-Pest /  $R_{\text{contrast}}$  [1,2]), as a function of character size ( $0.2^\circ$  -  $5^\circ$ ), at 13 retinal eccentricities on the horizontal meridian up to  $36^\circ$ . Thresholds were converted to Weber contrast  $dL/L$  to be comparable to the classical formulations of Ricco's law. Log-log contrast-size functions were analysed with respect to maximum slope and slope of -2. Results: Stimulus size has a more pronounced effect on character recognition than it has on stimulus detection, such that the maximum slope of the (log-log) areal-summation function is much steeper than Ricco's (-2) slope. It ranges from -3 in the fovea to -7.5 at  $30^\circ$  eccentricity. At larger stimulus sizes there is a range at which Weber contrast threshold  $CW$  is proportional to stimulus area  $S$  (i.e. slope is -2); we denote this as the Ricco size range. The latter increases with retinal eccentricity at the same rate as receptive field size. Furthermore, the effect size  $CW \times S$  is a constant multiple of Spillmann's perceptive field size. The law will be formally related to that of Fischer & May for the cat [3]. Conclusion: Areal summation at the ganglion cell level does not predict areal dependency for character recognition. However, the dependency of the area-dependency function on retinal eccentricity is closely related to receptive and perceptive field size. It is well described by a compact set of equations that will be presented.

[1] Harvey, L.O., Jr. (1986): *Behav. Res. Meth. Instr. Comp.* 18, 623–632

[2] Strasburger, H. (1997): *Spatial Vision* 10, 495–498

[3] Fischer, B. and May, H. U. (1970): *Exp. Brain Res.* 11, 448–464

Study supported by the Deutsche Forschungsgemeinschaft, STR 354/3–1.

## A Neurocomputational Model of Perisaccadic Compression and Receptive Field Dynamics Based on Oculomotor Feedback

Marc Zirnsak, Fred H. Hamker, Dirk Calow and Markus Lappe

Allgemeine Psychologie, Psychologisches Institut II, Westf. Wilhelms-Universität Münster

zirnsak@psy.uni-muenster.de

Studies investigating perception around the time of saccades under illuminated conditions reported a mislocalization of flashed stimuli towards the saccade target, which is strongest around saccade onset [1, 2]. Recent findings show that ‘compression’ of visual space also occurs orthogonal to saccade direction [3]. Moreover, the pattern of mislocalization is not symmetric in space and depends on the eccentricity of the flash. So far there exists no explanation for both, the temporal and spatial pattern of compression (see [4] and [5] for partial accounts). Here, we provide a computational model for compression based on neurophysiological and neuroanatomical aspects of the visual system. Consistent with findings of Fischer & Boch [6] and Moore & Armstrong [7] we assume that oculomotor-related activity enters the visual pathways and modulates the stimulus driven activity [8]. This gain control shifts the population activity and causes flashes to be mislocalized towards the saccade target. In contrast to Ross et al. [9], who suggested that predictive remapping of receptive fields in the direction of the saccade, for instance in cortical area LIP, leads to compression, our model predicts a perisaccadic shrinkage and shift of receptive fields towards the saccade target as reported in V4 [10]. By incorporating cortical magnification and receptive field size into the model, we are able to quantitatively reproduce the overall spatial pattern of compression. Variations of magnification show that an anisotropic magnification is more consistent with the psychophysical data. Note, that we only assume local anisotropy, which is consistent with the findings of Adams & Horton [11] for V1. Furthermore, the model is able to at least qualitatively account for the time course of the mislocalization assuming that the saccade related activity builds up before the actual eye movement and reaches its maximum around saccade onset like movement cells of the FEF and SC do. To conclude, this is the first integrative account for explaining perisaccadic compression by a spatially selective gain control, which is able to replicate both the spatial and temporal pattern of the mislocalization. Moreover, we argue that phenomena of spatial attention, perisaccadic receptive field dynamics and compression of visual space can be unified by a spatial reentry theory.

- [1] Ross J. et al. (1997): *Nature* 386, 598–601
- [2] Michels L. and Lappe M. (2004): *Vis. Res.* 44, 2327–2336
- [3] Kaiser M. and Lappe M. (2001): 41, 293–300
- [4] Morrone M.C. et al. (1997): *J. Neurosci.* 17, 7947–7953
- [5] VanRullen R. (2004): *Biol. Cyb.* 91, 131–137
- [6] Fischer B. and Boch R. (1981): *Exp. Brain Res.* 44, 129–137
- [7] Moore T. and Armstrong K. M. (2003): *Nature* 421, 370–373
- [8] Hamker F. H. (2003): *J. Vis.* 3, 808–816
- [9] Ross J. et al. (2001): *Trends Neurosci.* 24, 113–121
- [10] Tolias A. S. et al. (2001): *Neuron* 29, 757–767
- [11] Adams D. L. and Horton J. C. (2003): *J. Neurosci.* 23, 3771–3789

# Index of Authors

- Ackermann, H., 60, 108  
Alter, K., 60  
Altmann, A., 122
- Bach, M., 110, 127, 133  
Backer, G., 66  
Barteit, D., 174  
Barth, E., 162  
Bausenhardt, K., 48, 55  
Bayerl, P., 62, 139, 174, 175  
Beecham, R., 58  
Behrend, K., 100  
Berger, D., 78  
Beykirch, K., 160, 167  
Biber, P., 65  
Birk, A., 64  
Bochev, V., 114  
Bock, S. W., 123  
Böhme, M., 162  
Braun, C., 101  
Braun, D., 79  
Bremmer, F., 83, 164–166  
Brouwer, A.-M., 161  
Bruhn, H., 61  
Bülthoff, H. H., 42, 74, 76, 78, 122, 125,  
135, 136, 157  
Bülthoff, I., 124  
Busch, A., 119
- Calmanti, C., 96  
Calow, D., 198  
von Campenhausen, C., 180  
Caniard, F., 74, 176  
Carbon, C.-C., 40, 41, 138  
Casana-Perez, S., 51  
Castrilln-Santana, M., 68  
Chatziastros, A., 76, 176, 186  
Clifford, C. W. G., 49  
Cooke, T., 125  
Cootes, T., 24  
Curio, C., 25
- Curio, G., 138
- Dahl, C., 115  
Deco, G., 119  
Dicke, P. W., 123  
Diehl, V., 107  
Dietrich, S., 60  
Dinkelbach, A., 50  
Dinse, H. R., 102, 112  
Distler, C., 185  
Dorr, M., 162  
Drewes, J., 88, 162  
Drewing, K., 80  
Duchaine, B., 45
- Eckhorn, R., 103  
Eggert, T., 51, 156  
Elliott, E. A., 94  
Ernst, M. O., 72, 73
- Fahle, M., 107  
Fallgatter, A., 129  
Fischer, T., 61  
Fleck, S., 65  
Fleming, R., 89, 135, 139  
Franke, U., 28  
Franz, G., 145  
Franz, V., 161, 163  
Franzkowiak, S., 112  
Freund, S., 84  
Friedel, E., 99  
Friedrich, B., 176  
Frintrop, S., 66
- Gail, A., 103  
Garsoffky, B., 116, 143  
Gavrik, V. V., 52  
Gegenfurtner, K. R., 79, 88, 148, 149, 161–  
163, 169, 171  
Gehres, M., 177  
Georg, K., 194

- Gerstmayr, L., 90  
 Getzmann, S., 178  
 Gharabaghi, A., 98  
 Giese, M. A., 32, 37, 131, 187, 188  
 Gillner, S., 144  
 Glaus, S. D., 142  
 Godde, 99  
 Goertz, R., 120  
 Golovina, E., 92  
 Graf, M., 126  
 de Grave, D., 163  
 Grill-Spector, K., 128  
 Grodd, W., 122  
 Grosjean, M., 35  
 Groß, C., 46  
 Grueter, M., 138  
 Grueter, T., 138  
 Gulan, S., 81  
  
 Hagemann, N., 179  
 Hamker, F. H., 194, 198  
 Hansen, T., 148, 149  
 Harris, J. A., 49  
 Haselmayer, P., 180  
 Hecht, H., 54  
 Hegner, 99  
 Heier, K., 150  
 Hein, C. M., 127  
 Hein, E., 53, 55  
 Heinisch, C., 112  
 Heinrich, S. P., 128  
 Helbig, H., 72, 126  
 Herrmann, M., 129  
 Hertel, R., 152  
 Herzmann, G., 44  
 Herzog, M., 196  
 Hill, N. J., 109  
 Hofbauer, M., 156  
 Hofer, F., 130  
 Hoffmann, K.-P., 82, 85, 181, 185  
 Horstmann, A., 82  
 Huff, M., 116, 143  
 Hupfeld, D., 181  
  
 Ilg, U. J., 81  
 Ischebeck, A., 60  
  
 Jäkel, F., 106  
 Jakob, P., 117  
  
 Jarchow, T., 142  
 Jastorff, J., 131, 188  
  
 Kahlesz, F., 26  
 Kalisch, T., 112  
 Kaminiarz, A., 164  
 Kandil, F. I., 182  
 Karim, A. A., 99  
 Karnath, H.-O., 122  
 Kerzel, D., 161, 169, 171, 172  
 Kiefer, M., 114, 118, 126  
 Kiper, D. C., 115  
 Kirchner, H., 132  
 Klaes, C., 102  
 Kleiser, R., 165, 166  
 Klingenhoefer, S., 83  
 Koch, R., 29  
 Koechy, N., 195  
 Konen, C. S., 165, 166  
 Kornmeier, J., 127, 133  
 Kourtzi, Z., 122, 131, 184, 188, 191  
 Krägeloh-Mann, I., 101  
 Kragic, D., 27  
 Kreiner, W. A., 134  
 Krieschel, S., 54  
 Krug, J., 160, 167  
 Krummenacher, J., 119  
 Kuhlmann, S., 183  
 Kuss, M., 106  
  
 Lange, C., 73  
 Lanwer, S., 65  
 Lappe, M., 182, 183, 194, 198  
 von Laßberg, C., 160, 167  
 Liebich, S., 118  
 Liepelt, I., 84  
 Lueschow, A., 138  
  
 Mahler, T., 62  
 Maier, S. J., 157  
 Mallot, H. A., 90, 144  
 von der Malsburg, C., 67  
 Mamassian, P., 176  
 Marconato, F., 101  
 Martinez-Trujillo, J., 34  
 Masseck, O., 85  
 Mathiak, K., 108, 191  
 Maurer, J. P., 110  
 McDonald, S., 184

- Menning, H., 191  
Michler, F., 103  
Miltner, W. H. R., 54  
Mora-Ferrer, C., 100, 189  
Muehlberger, A., 129
- Nagel, A., 86  
Neumann, H., 32, 33, 62, 139, 174, 175  
Neumeyer, C., 177, 180  
Niehaus, S., 86  
Nitschke, J., 93  
Nusseck, M., 61, 135
- Pavlova, M., 98, 101, 137  
Philipp, R., 185  
Pilz, K., 136  
Platz, T., 84  
Poggel, D., 96  
Pracejus, L., 79  
Pretto, P., 186
- Ragert, P., 102, 112  
Rasmussen, C. E., 109  
Reichelt, A., 145  
Riecke, B. E., 74  
Röll, B., 85  
Roether, C. L., 187  
Rohe, M., 164  
Rolke, B., 48, 50, 53, 55  
Rome, E., 66  
Rossmanith, N., 145
- Saenger, J., 56  
Sander, T. H., 138  
Sarkheil, P., 188  
Schäfer, B. M., 168  
Schankin, A., 57  
Scharnowski, F., 196  
Schilling, A., 22  
Schmidt, T., 86  
Schmidt-Hoffmann, R.-B., 189  
Schneider, W. X., 119  
Schöner, G., 112  
Schueler, A., 99  
Schulte-Pelkum, J., 74  
Schultz, S., 150, 151  
Schwan, S., 116, 143  
Schwaninger, A., 42, 130  
Schwarzer, G., 46
- Schweinberger, S., 43  
Seitz, R. J., 165, 166  
Seymour, K., 81  
Shenvi, G., 94  
Sim, E.-J., 114, 118  
Sireteanu, R., 120  
Sokolov, A., 101, 137  
Sokolov, A. A., 98, 137  
Spang, K., 107  
Spehar, B., 58  
Spering, M., 169  
Steckner, C., 75  
Stemme, A., 119  
Strasburger, H., 96, 197  
Straßer, W., 65  
Straube, A., 51  
Stritzke, M., 170  
Swirszcz, K., 108, 191
- Tanaka, J., 118  
Tanner, T. G., 109  
Tatagiba, M., 98  
Tegenthoff, M., 102  
Teichert, T., 103  
Terzibas, C., 78  
Tewes, A., 67  
Thiel, A., 120  
Thier, P., 123  
Thornton, I. M., 36, 132, 136, 176  
Thorpe, S. J., 132  
Trahms, L., 138  
Trenner, M., 84  
Treue, S., 34  
Tretwein, B., 96  
Trippe, R., 54  
Trommershäuser, J., 80, 170
- Ulrich, R., 48, 50, 53, 55, 93
- Vatakis, A., 191  
Vorberg, D., 20  
Vuong, Q. C., 68, 132
- Wachtler, T., 103, 152  
Wacker, A., 76  
Wallraven, C., 42, 125  
Walter, S., 148, 153  
Wascher, E., 56, 57  
Weber, J. E., 138

Weber, M., 62  
Weidenbacher, U., 139  
Weiss, A., 144  
Weiss, T., 54  
Welchman, A. E., 157, 190  
Westphal, G., 69  
White, B., 171  
Wichmann, F. A., 88, 106, 109  
Wiener, J. M., 145  
Wilimzig, C., 112  
Wittenberg, M., 103  
Wolf, M., 110  
Womelsdorf, T., 34  
Würtz, R. P., 67, 69

Ziegler, A., 179  
Ziegler, N. E., 172  
Zihl, J., 119  
Zirnsak, M., 198  
Zvyagintsev, M., 108, 191