EYESHOTS Meeting Castellon (UJI)

29. - 30. September 2008

Participants:

Silvio Paolo Sabatini, Giorgio Cannata, Manuela Chessa, Andrea Canessa, Agostino Gibaldi (UGE); Fred Hamker, Markus Lappe, Mark-André Voss, Katharina Havermann, Robert Volcic (UMU); Claudio Galletti, Patrizia Fattori, Nicoletta Marzocchi, Rossella Breveglieri, Michela Gamberoni (UBO); Angel del Pobil, Ester Martinez, Beata Joanna Grzyb, Eris Chinellato (UJI); Nick Chumerin, Karl Pauwels (KUL).

Minutes compiled by Mark-André Voss

14:40 Giorgio Cannata: Progress of work - WP 1 (UG-DIST)

Focus on task T1.1 - Binocular eye coordination

- study of geometric and kinematic effects of ocular postures and movements on vision and depth perception.

- eye rotations- control motion of the muscles of the eyes to go from one point in space to another, multiple solutions, additional constraints.

- constraints

- Local disparity - sensitivity measurement, is there a functional of disparity to torsion

- numerical analysis needed
- selection of variables and metrics
- visual based information

Silvio: important issue to couple the vergence and turning movements
dependent on size of saccade? Join two angles in a single value.

Giorgio: family of curves to be optimised.

Fred Hamker: time constant for beta faster than for vergence - any
Silvio: how can the trajectory result in a task like this. don't know if there are specific

Andrea: many studies are monocular.
Silvio: or old data. More efficient to use a combined gain.

15:15 Andrea Canessa: Influence of eye movements on disparity perception

- occular movements: Listings law - generalized Listing Law (L2)
  Torsion rotation

- disparity statistics

- Stereo-head comparison: pan tilt, ICub head - Listing, MacEye - Listing 2, modified MacEye
  A random dot world: random world - grid of fixation point

- SVD of disparity:

i=1:N
j=1:M
k=1:K

d_ij = disparity vector of the point q_ij
disparities are smoothed and resampled on retinal smaples s_k to obtain d_jk

create 2x2 matrix D_jk^s= d_jk(djk)^T and then averaged over fixation points to obtain D_k^s

Probability distribution:
  Pan tilt: traditional distribution, many horizontal, fewer vertical disparities
  "Iso-probability" contours
    Pan tilt: mean is not always in correspondence to the retinal center
    epipolar geometry and epipolar disparity - a link with the energy model:
      consider mean disparity pattern, we can see that it is not centered around the zero
    the optimal fixation surface

Results:
-conclusion: pan tilt constraints that may reduce disparity
- l2 has comporable behaviour but more complicated kinematics
- using pattern of non corresponding fixed points

- perspective statistics from natural environments
- try torsional movements may play role in perception of the surfaces of the fixated objects reducing the disparities

Silvio: common tilt constraint is desirable
Markus: what could the Torsion be useful for?

Camera no torsion - on the other hand the eye, probably the eye uses torsion to obtain something like pan tilt, but with lesser energy cost - also in pan tilt are components of torsion. use this kind of movement to use torsion but in a mechanically more efficient way.

15:42 Agostino Gibaldi: Disparity evoked vergence movements

- Vergence eye movements driven by a retinal disparity stimulus
    vergence models: image acquisition, controlled extraction, plant
    reliability of the disparity map: reliability depends on the range (only center accurate)
    vergence control directly by population of disparity detectors instead of the disparity map.
    Tuning curves with RDS
    Minimization of a target function
        sensitive to horizontal disparity but not to vertical disparity
    Biological plausibility: Takemura et al - 2001 weighted sum of the population response
        Semmlow el al 1986 - extraction of two different control signals
        Poggio et al. - 1988 - NE and FA detectors support coarse stereopsis while TN, T0 anf TF detectors support fine stereopsis
        dual mode controls
            near vergence
            very near vergence
        reliability of the control by LONG and SHORT

results:
    wide range of horizontal disparities corrected by vergence movements
    strong resistance to noise

Fred Hamker: why necessary to split into LONG and SHORT?
Karl: way to define a scale?
just one scale but a whole population of long and short disparity detectors.

vertical disparity effect
    - decrease the vergence response to the same horizontal disparity

LONG & SHORT effectiveness:
    LONG reliable for wide range of horizontal disparity and small range of vertical disparity
    SHORT opposite

Consideration Response separability

16:10 Nick: Network paradigm for intelligent vergence control

Problem describtion

Hypothesis: given stereo setup and its fixation point it is possible to estimate expected/prior disparity map patterns without decent information about scene structure.

Task: check hypothesis
- develop an approach to learn the function:
- fixation point -> disparity map

Input: head parameters

scene context

Output: d_x = d_x(x,y,X,Y,Z) and d_y = d_y(x,y,X,Y,Z)

Proposed approach
- setup realistic parameters, model them in 3D environment
- generate great number of 3D environments
- make "eyeshots"
- render disparity maps for each eyeshot

Head is fixed

baseline 70mm

focus length 6 mm

camera resolution 80x80

Workspace parameters
- width 1000mm
- height 1000mm
- depth 1000mm
- distance from head: 1500 mm
- fixation point

scene
- floor ceiling side walls back wall

For each FP we render:
- hor disp
- ver disp
- hits binary map (1 if disparity available at the pixel, 0 otherwise)

For each fixation point we accumulate:
- minHD, maxHD,minVD,maxVD
- hits
- sumHD, sumVD
- sumHD2,sumVD2
- histHD, histVD (20 bin)

Conclusion
- almost 5000 scenes have been processed
- obtained mean disparity maps are nonuniform
- disparity histograms are not flat

Future steps

check if results can improve disparity estimation
implement learning from example without the ground truth

16:30 Karl - Eyeshots WP 2 Task 2.2

rotation interpolation
eye orientations different from rectified position introduce (potentially) large orientation differences between left and right projections
in addition to this resulting from e.g. slanted surfaces
phase only robust to some extent

assuming orientation difference is known
orientation compensation
nearest neighbor selection of filter responses
linear interpolation of filter responses
compensate for sign-change of sine-filter

apply image rotation
compute phase difference at corresponding locations
spatial phase warping (bilinear interpolation)

figure: with higher rotation angle standard error gets higher for higher angles
with nearest neighbors it will control the error

gaze estimation:
compensate for small calibration errors
inaccurate motor feedback
estimate gaze on the basis of vector disparity deviations orthogonal to hypothesized epipolar lines

goals
improve disparity estimation
improve gaze estimate

Link to biological models is still there as we are not warping images
gaze estimation
compensate for calibration errors
e.g. assume rectified configuration

FOR all scales (coarse to fine)
UNTIL convergence
compensate for previous scale disparity vertical disparity, and orientation difference
update vector disparity
estimate 3D rotation (linear) from vector disparity component orthogonal
update geometry estimate

gpus
GPU vs CPU
optical flow very high speedup

16:50 Manuela Chessa, N. Chumerin: Learning mechanisms in the cortical architecture

Learning of disparity patterns due to the geometry of the scene and fixation point
Learning of tuning parameters (e.g. phase shifts) of the cortical-like architecture for disparity estimation

Active fixation in virtual and real environments (from range data)
3D range data
Matlab tool to obtain:
2D projections of the points into the image plane (stereoscopic image pairs)
Ground truth disparities (both horizontal and vertical)

BRID database, brwon
197 range images
operation range of the sensor 2-200 m
interior and exterior scenarios

Pros Cons
Pro: Matlab tool can be used with other available range databases
Con: occlusions occur by doing vergence movements.
many available databases have no color information
range data in peripersonal space are not available.
TODO: create our own database with a 3D laser scanner (distance between 30 cm and 2m)

C++ tool to obtain:
2D projection of virtual environments into the image plane
ground truth disparities (both horizontal and vertical)

parameters
3D position, textures, shape of the objects in the virtual environment
VRML models
illumination
interocular distance
3D fixation point

PRO no occlusion occurs
it is possible to have VRML and 3D models for a wide range of objects and scenes
CONs
scenes may be too simple (not photorealistic)
TODO
create more complex and textured scenes

to implement different eye movements (not only pan-tilt movements)

Fred Hamker: why not projection on a sphere?
Silvio: not meant to model the eye but a camera.

18:00 WP 4 - Sensorimotor integration

T4.1: Merging perception-related and action-related visual information
T4.2: Generating visuo-motor descriptors of reachable objects
T4.3: Constructing a global awareness of the peripersonal space

Dorsal stream (viewer centered) - ventral stream (object centered)

Interactions between the streams

Proposal of a multi-stage integration between the streams:

1. Ventral -> dorsal link supporting visual analysis through cognitive interpretations (e.g. use of perspective for parallel edge shapes)

2. Dorsal -> ventral link for simplifying and accelerating object identification through information on shape and affordances

The ventral stream influence on action execution is modulated by:
object familiarity and degree of confidence in the recognition process
availability and expected reliability

ANN simulation and experimental robot data

Probabilistic object classification
three possible classe (box, cylinder, sphere)
invariant to object pose and size
based on curvedness of visible object contour

classification at different angles
Markus Lappe: why not 50% probability between Cylinder and sphere?
because the data were not calibrated.

Surface and axis orientation selective neurons: SOS & AOS
CIP neurons in macaques are selective for orientation of visual features suitble for grasping, and for their proportions;

Humans fMRI studies indicate something similar

Activation modeled with 3 inhibition terms for symmetry, proportion and graspability:

R_SOS = 1 - ((a-b)/(a+b))^2 - 0.03 * (c)/(a+b) - 0.5 * (1)/(1+e^-0.04(c-H))

Three clusters, with clear distinction between elongated and flat shapes

Useful object characterization

The reaching paradigm Role of V6, V6A (vs V3A, CIP, AIP)

Analysis of single-cell data from UNIBO on reference frame transformation as performed by V6A trying to extract information on

- eye-arm coordination and reciprocal guidance
- modulation between retinal data, gaze direction and reaching

Model 2D data

Bodily self

the robot should have central representation of the position of body parts in space their configuration and hierarchical arrangement (Body SCHEMA)

robot should have a sense of embodiment

motivation to explore the environment

Development of the body knowledge for robots

piagets primary circular reactions

piagets secondary circular reactions ("making interesting sights last")

Active exploration of peripersonal space

Tension
Pleasure

agent maximizes its own pleasure

two main tensions:

- tension of playing
- tension of novelty-seeking

system switches between four different emotional states: flow, anxiety, relaxation, arousal

flow = repetition, arousal = looking for novell stimulus

video of grasping - object falling -

Task 4.1 basically done
Task 4.2, 4.3 working on

Important requirements

- suggestions feedback results from other groups and WP
- new stereo head (adapt model and code)
- more precise definition of experimental setups

18:30

Silvio: mention in the deliverable at 9th month: concepts are the same but from an experimental point of view it differs introduce switch off to the anxiety of the robot

18:35 Patrizia Fattori

Task 5.1 Role of visual and oculomotor cues in the perception of 3D space

Task 5.1 and 5.2 share initial milestones

electrophysiological recording on behaving monkeys
- binocular eye position recording
- 3D fixation

Eye recording system
- ISCAN ETL-200 Primate Eye tracker
  - video based
  - 120-240Hz temporal solution
  - high spatial solution (down to 0.06 degrees)
  - non invasive system

installation & testing has been done

integration with the experimental control and data acquisition

test the existance in the region we study of neural modulations related to the 3D position of the fixation point
verify which part of space is represented by parietal neurons

Silvio Q:Panel can be oriented in space?
A: No, it is fixed.

Experimental characterization of:
- 3D eye position fields of neurons of medial parieto-occipital cortex.
- 3D mapping of visual receptive fields
- 3D gain fields

Fred: how will the stimuli look like?
Patrizia: moving objects, static ones won't work.
Fred: What regions?
Patrizia: V6
Fred: what is the use of moving stimuli when we want to grasp something that is not moving?
Patrizia: its more for perception than vision.
Q: Are there connections to MT?
A: Many

Robotic arm Katana for moving objects in the visual space of the monkey.
System not huge and heavy (4.8 kg), enough precision

decision must be made for the purchase of a robotic arm - but at the moment not available

Task 5.1: role of visual and oculomotor cues in the perception of 3D space

V6a. Parieto-occipital sulcus: monkey grasping task.

discharges during the movement of the arm.

task in darkness - task in light
compare responses in the two cases

dark: proprioceptive guidance of arm movements: haptic, joint signals
light: visual and proprioceptive guidance of arm movements

Are reaching actions in different parts of the 3D space differently affect by the availability of visual information?

Silvio: have you tried to excite the cells in 3D with a light?
Patrizia: We planed to but we did not.
Silvio: Display with moving lights on a slanted screen as stimulus?
Patrizia: We can think of it.

Task 5.2: Link across fragments

active exploration of the peripersonal space through active ocular and arm movements
covert attention towards

experimental setup:

nine leds (three in three lines) at eyelevel at different deepths for the monkey to reach for
can be rotated

remark: build up something so that you know that the modulation is only caused by the vergence

silvio: can the monkey reach the leds when its slanted?
Patrizia: it should.

Markus: aren't there occlusions of the leds?
Patrizia: it should not matter, maybe if the hand of the monkey is still in a position where it occluds another led.

The change in direction should be excluded to be sure that the vergence modulates the neurons.

19:20 Markus Lappe

WP 5.3 Motor description of fragment location
WP 5.4 Predictive behaviour and cooperation in shared workspace

5.3
Saccade adaptation: during saccade the targed moves, after adaptation the eye movement goes directly to the moved target

saccadic adaptation induces a shift in perception

Current work: time course of shift during adaptation
re-evaluation of shift during fixation (ongoing)
reactive vs. scanning saccades (done)
model of adaptation and induced shift (starting)

Idea of WP is to look at connection of saccade targeting and visual space

what area in the brain tells us about positions in the space?

reactive saccades - elicited by flashed targets
scanning saccades - complex scene and you look at different objects in the scene

adaptation to reactive targets much higher than at scanning saccades

after a series of saccades in the scanning scenario only the last saccade gets adapted

saccades to reflexive targets get adapted in contrast to scanning targets

Q: is there a timing constraint?
A: No there is not.

Magnitude of adaptation and mislocalization reactive/scanning
adaption
mislocalization: flashed bars/stationary bars - adapted to reactive the mislocalization of flashed targets is great and for stationary bars is small.
mislocalization transfer: flashed bars/stationary bars

transfer between different saccade types is asymmetrical
transfer of reactive saccade adaptation to space perception is stimulus dependent

19:45 Katharina: A closer look into saccadic adaptation as sensorimotor process

Optimal controlled system with internal feedback

Basic differential equation
noise feedback

saccadic system: without external feedback

sensitivity against disturbances (fatigue effects, neural noise etc.)

control mechanism needed

model: inverse controller (IC) and forward model (FM)

IC performs motor process, realized via optimal control

FM of target and plant predict state and give feedback via Kalman filter

Inverse controler - generating saccades via optimal control

forward model - internal feedback via kalman filtering

predictor <--> corrector

aim of model: dissociation of sensory and motor components of saccadic adaption in FM model possible

Silvio: at the moment all experiments are done in frontoparallel plane. Does adaptation also work in 3D?

Markus: I think there was one experiment where they did this in 3D but it is very difficult to measure, because the indication of the perceived position is very difficult.

Patrizia: is it worth looking into velocity and if velocity is connected to adaptation.

Markus: From our data I think there is no connection.

19:55 Robert: Extraction of gaze direction

identify as soon as possible the end-location of the pointing movement

the observer is able to use the gaze direction of the other person to anticipate the pointing movement with a saccade.

predictable vs unpredictable movements

the gaze is directed to the end-location of the own and the other persons movements

no monitoring of the other person's gaze or hands

20:05 END