EYESHOTS KICKOFF MEETING

7 – 8 March 2008

BOLOGNA

Partecipants:
Silvio Paolo Sabatini, Giorgio Cannata, Manuela Chessa, Andrea vanessa, Agostino Gibaldi **(UGE);** Fred Hamker, Markus Lappe, Mark Voss, Eckart Zimmermann, Katharina Havermann **(UMU);** Claudio Galletti, Patrizia Fattori, Nicoletta Marzocchi, Rossella Breveglieri, Michela Gamberoni **(UBO);** Angel del Pobil, Ester Martinez, Beata Joanna Grzyb, Eris Chinellato **(UJI);** Marc van Hulle, Nick Chumerin, Karl Pauwels **(KUL).**

Minutes compiled by Agostino Gibaldi

MINUTES RECORDING

FRIDAY 7th

09:30 Welcome and Introduction. Project overview – Silvio Sabatini

Introduction to perception/action and 3D awareness in the arm reaching space based even on motor information through:

- Embodiment of the head
- Fragmented vision
- Visuomotor description of space
-

Development of

- Anthropomorphic binocular head
- Space representation model
- Robot – man interaction model

Goals

- Active exploration
- Situation awareness of space
- Situation awareness of shared workspace

10:50 Models of Saccadic Motion – Giorgio Cannata

Development of the anthropomorphic binocular system in order to obtain a 3D exploration of space

Goals
- Tendon driven ocular system
- Saccades, smooth pursuit and vergence
- Integrated eye/arm system

Emulation of the physic structure
- Ocular movements
- Principles (Donder Law, Listing Law)
- Models

Geometry and Kinematic
- Rotational matrix R
- Quaternions
- Control signal and angular velocity
- Donder Law
- Listing Law and zero torsion
- Half angle rule

Eye anatomy
- Recti and obliqui muscles
- Soft pulleys
- Rotational bounds
- Model modified with three actuators
- Listing plane slant
- Coordinated ocular motion and Sherrington law

Q&A
Nick: What do you mean with orientation? *The gaze of a single eye*
Markus: The rotation vector V always lies on the listing plane? *Yes, it is a presumption*
Marc: Do the muscle / actuator in the model only contract? *No, they are coupled and when a tendon is pulled the coupled one is released*
Nick: In the eye we have tree degrees of freedom, why in the model only two? *While the head is fixed, I should have no torsion, so the third degree of freedom is not necessary; it become important if I include head movements*
Angel: Is there in the eye a bounded point? *No, and obliqui muscles help to keep it in his position*
Fred: Torsion of the eye in the model will improve only dynamic or even fixation?

11:30 Cortical Architectures for Dynamic Stereopsis – Silvio Sabatini

Computational theory of stereo vision and disparity detectors

Depth cues

- Personal space (retinal disparity ➔ depth perception)
- Action space
- Vista space

Plenoptic function:

- $P(x, y, z, V_x, V_y, V_z)$ Image system representation of a camera moving with pure translation (V_x, V_y, V_z)
- $P(x, y, z, V_x, V_y, V_z, V_A, V_E, V_T)$ Image system representation of a camera moving with translation and rotation (V_A, V_E, V_T)

1st Order Measures

- Position disparity

2nd Order Measures

- Motion disparity (motion in depth and interocular velocity difference)
- Orientation disparity

Visual signals change due to eye rotation and influence disparity detectors: visual feedback or preventive tuning to modify disparity detectors

Available information from disparity detectors

- Absolute disparity (area V1)
- Edges of disparity (area V2)
- Relative disparity (area V2)
- Orientation disparity (area V3)
- Higer order disparity (area V3)
- Disparity gradients (area V3)
- Slanted suface space orientation (area V4)
- Joint stereo motion processing (area V5)

Disparity detectors

- 6 types (TE, TI, TN, TF, NE, FA)
- Simple cells (pass band filters)
- Complex cells (energy model)

Multichannel architecture: spatial scales in order to avoid ambiguities and false matches and to obtain larger disparities

Ocular movement strategies: vergence loop

Horizontal disparity and vertical disparity

Goals

- Disparity coding
- Log polar mapping
- 2nd order disparities
- Achieve measures in the plenoptic space

Q&A

Fred: 2nd order measures derive from rotation or translation? *Only rotation.*
Fred: Are simple cells monocular? *No, they are binocular*
Giorgio: What is refinement of the image? *Instead of act a worping on the image, I change the vergence*
Fred: Can I calculate vergence values from the image? *It could be a way to get a better fusion*

Andrea Canessa – Eye Torsion

Helmoltz Coordinates (3 rotations)

Listing law and false torsion

Listing law extension: L2 (separation of left and right eye)

Horizontal and vertical disparity patterns

Epipolar lines and points

Open Issues

- Torsion to register the binocular images
- Mean disparity decrease
- Adaptation related to the scene

Q&A

Markus: Do eyes use Listing law or L2 law? *A little bit of both*
Fred: Do you consider eye movements or fixation? *I consider fixation, because during eye movements Listing planes tilt*
Fred: L2 law is useful only for stereopsis or even for the eye movements? *It is useful to achieve image fusion easier*

15:00 Active Stereopsis – Marc Van Hulle

Network paradigms for vergence: vergence model control implemented with disparity detectors and attention signal

3D description through stereopsis: exploration strategy through visual fragments insensitive to small errors of version and vergence

Vergence strategy

- One degree of freedom: vergence angle
- Maximize complex cells answer through vergence
- Known initial pattern

Reflex like vergence

- No attention signal
- Training phase
- Testing phase

Voluntary exploration

- Saliency maps

Nick Chumerin – 3D Virtual Reality Simulator

- Robotic head model
- Environment model
- Ray-tracing

Use of synthetic image database to obtain reliable horizontal disparity, vertical disparity and optic flow

Model with supervised gaze change and automatic vegence change

Q&A
Giorgio: How much is important the focus accommodation achiving vergence? *In the model there is no focus accommodation, but it can be used as an information source*

Karl Pauwels – Visual Fragments

Interactive depth perception

- Delineate the fragments
- Integrate disparity estimates over vergence/version movements

    - Accumulate data about disparity estimates
    - Inhibit cells corresponding to reliable estimates
    - New vergence/version
    - Eye coordinates ➔ Head coordinates
    - Gain modulation
- Robustify the coordinate transformation
- Coarse-to-fine disparities
    - Limits of phase-based disparity detection
    - Averaging across spatial scales does not perform well for low frequency
    - Coarse errors can't be corrected at finer scales
    - "Fine" structures separated from background by big disparities can't be detected

Real time with Ge Force 8800 GTX graphical device
Q&A
Silvio: Why do you inhibit reliable cells? *In order to move the vergence to a plane where I don't have reliable information*
Fred: Whit a single vergence is not possible to describe an object in the best way

16:30 Discussion and Planning

Silvio: The influence of eye movements on visual processing. The problem is to measure visual changes due to eye movements: goal of obtaining not a disparity map, but a vergence control mechanism.

Marc: The vergence is a way to improve the disparity map, but in a population network the zero disparity cell is not know, because of possible vergence errors

Karl: Is it possible to obtain a vergence value map for interesting points in the scene?

Marc: In a fragment small eye movements are possible

Claudio: Visual fragment as a volume

Silvio: Importance of the transformation from image plane to cortical plane (log polar) in order to mimic natural foveation strategies

Fred: If I know an object, I need only 100 ms to recognize it or to compare it with another object, otherwise I need to explore it

Giorgio: What is the role of drifts and microsaccades in biologic system

Claudio: If I use a retinocentric coordinate system or a headcentric one, starting and arriving point are the same, the difference is in the strategy of eye movements

<u>Assumption list for Interactive Stereopsis</u>

- Fixed gaze: one parameter for the vergence control
- Log polar mapping: binocular images with space variant filters
- Input for the eye movements: vergence angle and rotational velocity
- Input: binocular images convolved with space-variant filters
- Investigate the possibility of incorporate the Lenet Paradigm in the disparity detector architecture
- Perspective of incorporating attention signals
  - Saliency maps: dynamic evaluation of the scene
  - Disparity information to guide exploration
  - Cooperative vergence-version strategies

<u>Assumption list for eye movements strategies and functional implication for visual processing</u>

- Saccade (minimize loop time)
- No smooth pursuit
- Vergence
- No torsions *but* tilt of the Listing's plane
- Include binocular coordination

SATURDAY 8th

09:15 Attention and 3D Object Recognition – Fred Hamker

<u>Hierarchy of object recognition</u>: dimension and complexity of the receptive field increase along the cortical pathway

- Simple cells (S1)
- Complex cells (C1)
- Composite feature cells (S2)
- Complex composite cells (C2)
- View tuned cells (designed on a particular object)

In the model target representation is compared with stimulus analysis

IMAGE ➔ LOCAL DETECTOR ➔ POOLING ➔ RECOGNITION
↑ ↓
←←←←←

Recognition of paper clip: beyond and before a certain distance between two identical objects, the recognition results decrease

Saliency map model of attention

```
                                                         → ORIENTATION →
IMAGE → MULTI SCALE FEATURES EXTRACTION ┤→   INTENSITY   → ├→
                                                          →    COLOR     →

→ COMBINATION → SALIENCY MAP FOR ONE FEATURE
```

Once the first object is detected, it is inhibited and start the search for the second

This model is useful for

- Guide toward interesting points in the scene
- Segment the scene in different objects
- Recognizing simple features

Attention modulation

```
STIMULUS WITH  → SALIENCY → CONVOLUTION WITH      →   SELECTION OF
TWO OBJECTS                 KNOWN MODEL FUNCTION      ONLY ONE OBJECT
```

Tree phases model:

```
OBSERVATION → → → → → → → → → → → → RESPONSE
             └ ← INPUT GAIN CONTROL ← ┘
```

Receptive fields (V1)

- Can contract to detect the stimulus at best
- With the same task they can move toward the stimulus
- Hebbian learning and anti-hebbian decorrelation

Goals

- Include in the model disparity features information
- Train V1 model for disparity
- Expand to higher features
- Implement attention dynamics
- Learn when to look to a particular object through reward

Q&A

Patrizia: In the three phases model, if I want to achieve grasping, should I include the dorsal stream?

Claudio: The retinocentric system is not enough to reconstruct the fragments and to achieve grasping, I need headcentric system

10:30 Perception across Eye Movements – Markus Lappe

Motor description of space

- Saccade coordinate system

- Scene learned not through features, but through motor parameters to reach the points of the scene
- Model / Resource / Workspace for the saccades
- 2 saccades system improves speed and doesn't use memory
- Eye movements in memory retrieval

Predicting attention and hand movements from gaze

- Eye / hand coordination; gaze direction anticipate hand movements
- Gaze acting VS gaze observing
- Attention cued by gaze

Spatial stability: when moving the eyes, retinal position of the objects change, how do I construct a world map?

- Using eye position signal to compensate
- Using the image matching features

Saccadic mislocalization

- Shift
- Compression
- Adaptation-induced shift

Saccade adaptation: If, during a saccade prom P to T1, I move the target to T2, the tested subject doesn't realize it and move eyes to T1. Over many repetition of the experiment, the subject moves directly to T2, adapting the saccade

Q&A
Patrizia: Is it possible that the 2 saccade system goes through the dorsal stream and does't use memory?
Giorgio: What does it happen if, after the adaptation occur, I don't move the target to T2 during the saccade? *There will be a new adaptation*
Silvio: Is there a relation between adaptation an eccentricity? *The adaptation gain increases with the saccade amplitude (saccade adaptation field)*
Claudio: The adaptation gain act on the saccade, not on the space perception, hence it gives a different motor control

12:00 Joint Visuo-Motor Features in the Parietal Cortex – Patrizia Fattori

Experiment: monkey trained to fixate to trace the receptive fields

V6 neurons

- Retinal dependent: receptive filed fixed respect to retina position
- Gaze dependent: spike frequency is modulated by gaze
- Gaze independent (Real Position Neurons): encoding of visual space

V6 properties:

- Integration of gaze information in order to build from the gaze dependent neurons a map with spike frequency / position gaze
- The real position neuron receptive field can be constructed with five gaze dependent neurons
- Arm movement related neurons in term of prediction
- Arm movement related neurons in term of position information of the arm itself

Planned experiments:

- 3D eye position fields
- 3D mapping of the visual receptive field
- 3D gain fields
- Active exploration of space through arm

Rossella Breveglieri - Neuronal circuitry in visuomotor integration

Reconstruction of the paths with neuronal tracer injection

V1	→				→	LIP
V2	→	→	V6	→	→	V4
V3	→				→	V5 (MST)
V3a	→					

Q&A

Markus: Is there a magnification factor in area V6? *We don't know, we register from a too small number of cells*

Nick: Is there a continuum gaze modulation? *Yes, there is.*

Angel: Are the borders of V6 receptive field sharp? *Yes, they are, there is no Gaussian envelop*

Silvio: Could be interesting to check if in real position neurons there is an invariant part of the answer, in order to search a possible relation between retinic and headcentric coordinates

Robert: Is there a latency within gaze dependent and real position neurons? *There are no measures.*

Markus: Gain fields are use for many purposes, why aren't they even used to construct real position cell? *Maybe because of computational cost: gain fields are to be updated, while real position cells are fixed and as a matter of fact they are very fast*

14:30 Vision-Based Reaching in Robotics and Neuroscience – Angel del Pobil

Visual guided reaching: in opening a handle bar, I want to maintain fixed the relative position between arm and handle, even if I move the cameras or the handle position

Why neurobiology

- Materia tissues
- Morphology (body structure)
- Ethology (behaviour)
- From ethology to neurophysiology

Head control for visuomotor coordination

- Vestibulo Ocular Reflex
- Stabilize head in space
- Mapping

Goals

- Multisensory representation of 3D space
    - Binocular visual cues
    - Oculomotor signals
    - Arm position
- Regular updating of representation

Eris Chinellato – Dorsal and Ventral Streams

Vision based grasping

- Slant estimation
- Orientation estimation

DORSAL STERAM (viewing)	VENTRAL STREAM (object)
LOCAL OBJECT FEATURES ↓ AFFORDANCE (graspable parts of the object) ↓ ACTION	GLOBAL FEATURES ↓ OBJECT CLASSES ↓ OBJECT IDENTITY

Generating visuomotor descriptors for reachable objects

- Retinocentric representation
- Effect based representation
- Distance / vergence representation

15:40 Discussion and Planning

<u>Group1: WP1 WP2 WP4 (Giorgio, Marc, Silvio, Markus)</u>

- Coordination of binocular eye movements and implications for visual processing
- Achieve vergence: learning of disparity detectors
  - learning detectors
  - the control perspective (which signals must be read out?)

<u>Group2: WP1 WP2 WP3 (Marc, Fred, Silvio)</u>

- Learning of feature-disparity selectors
- Image sequence for training
- Target selection and saliency map

<u>Group3: WP3 WP4 WP5(Angel, Markus, Fred, Patrizia)</u>

- Reference frame transformation
  - experimental
  - modelling
- Relevance of object feature properties for the reaching task
- Role of the reliable stereo cues

<u>Group4: WP1 WP2 WP5(Patrizia, Marc, Silvio, Giorgio, Fred)</u>

- Real position cells
- Reference frame transformation
  - experimental
  - modelling

<u>Benchmarking</u>

- Benchmark with standard platform
- Epipolar transformation issues
- Voluntary exploration
- Define a particular task to be accomplished