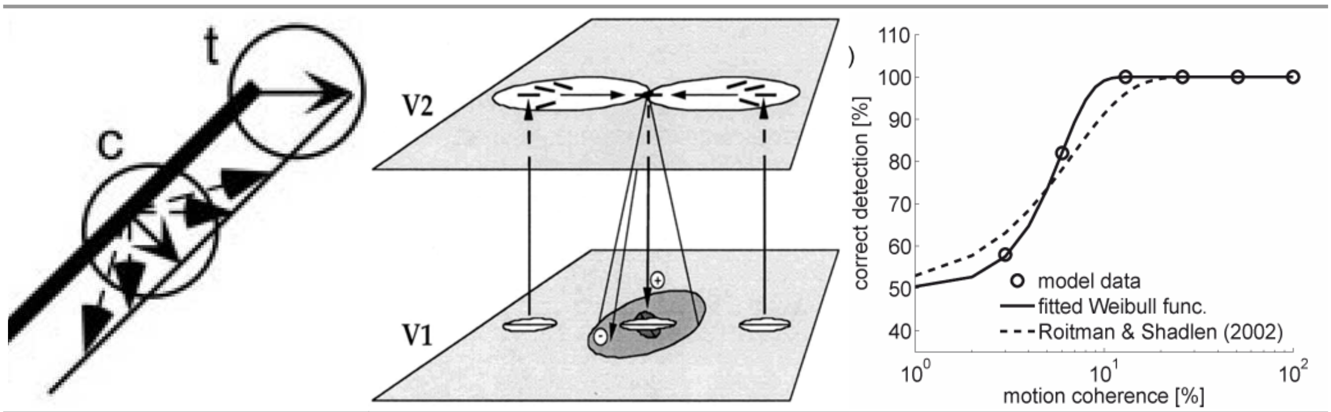


Neural Mechanisms of Form and Motion Detection and Integration - Biology meets Machine Vision

Heiko Neumann & Florian Raudies

Ulm University, Inst. Neural Information Processing
Boston University, Center for Computational Neuroscience & Neural Technology

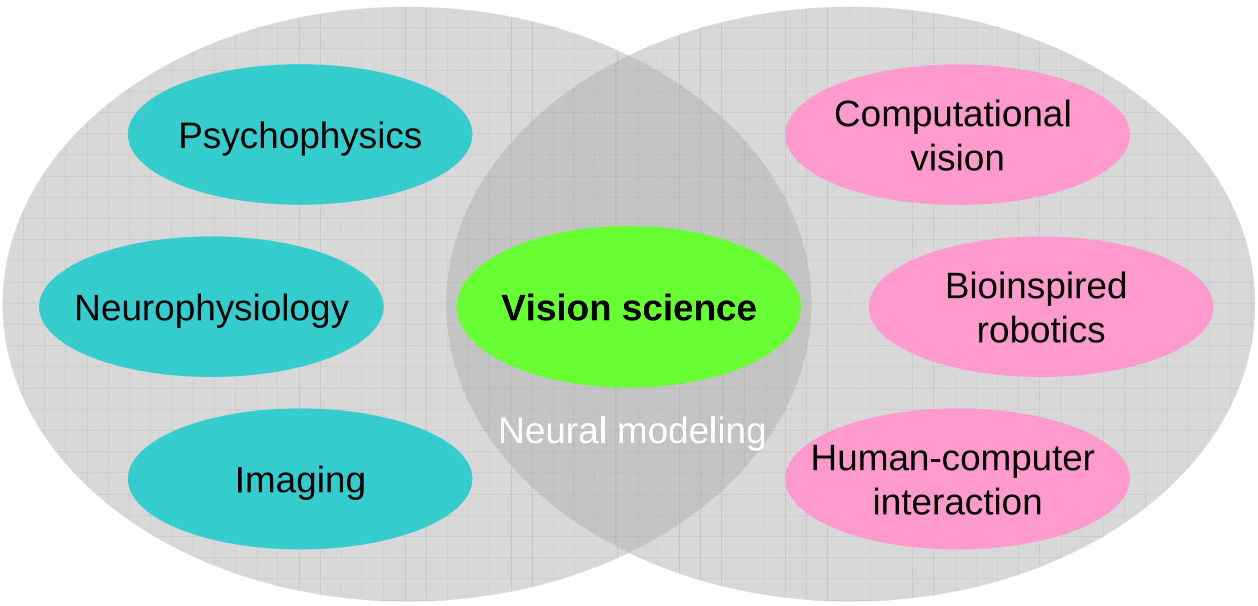
ECCV 2012 • Workshop on *Biological and Computer Vision Interfaces* •
Florence / Italy • Oct. 12, 2012

What we are doing …

Neural computation and the role in computer science

How does the brain control behavior?

How can technology emulate biological intelligence?



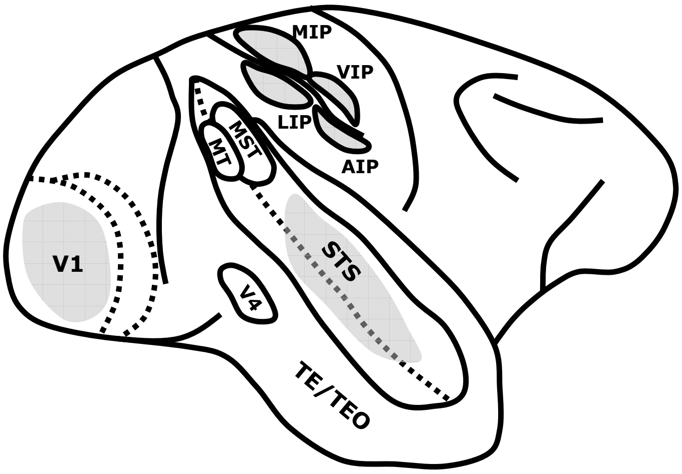
Neuroscience

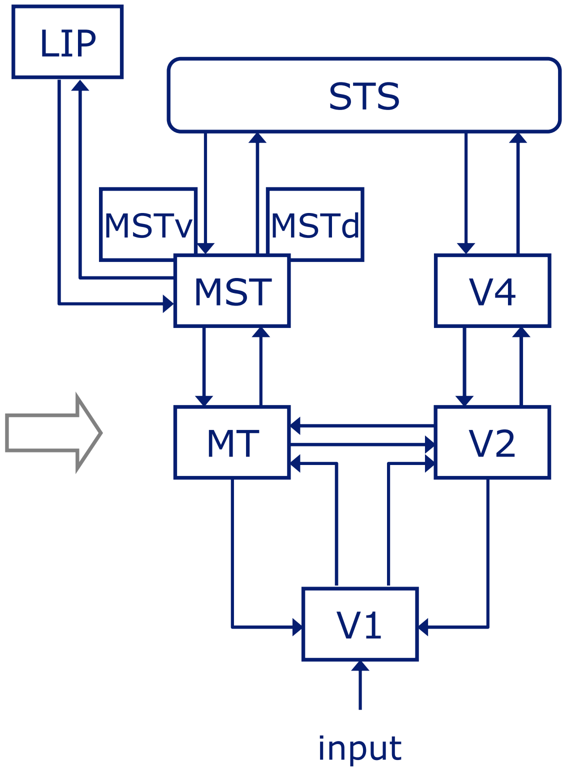
Computer Science & Engineering

Introduction and motivation

Brain processing is **interactive**

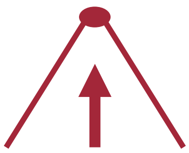
Visual system – structure and function





Different processing **principles** are identified in the brain

- Bottom-up (feed forward) processing

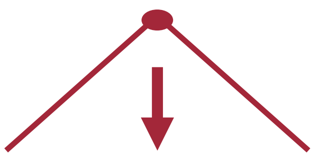


convergence / integration

- Lateral processing

integration / message passing

- Top-down (feedback) processing

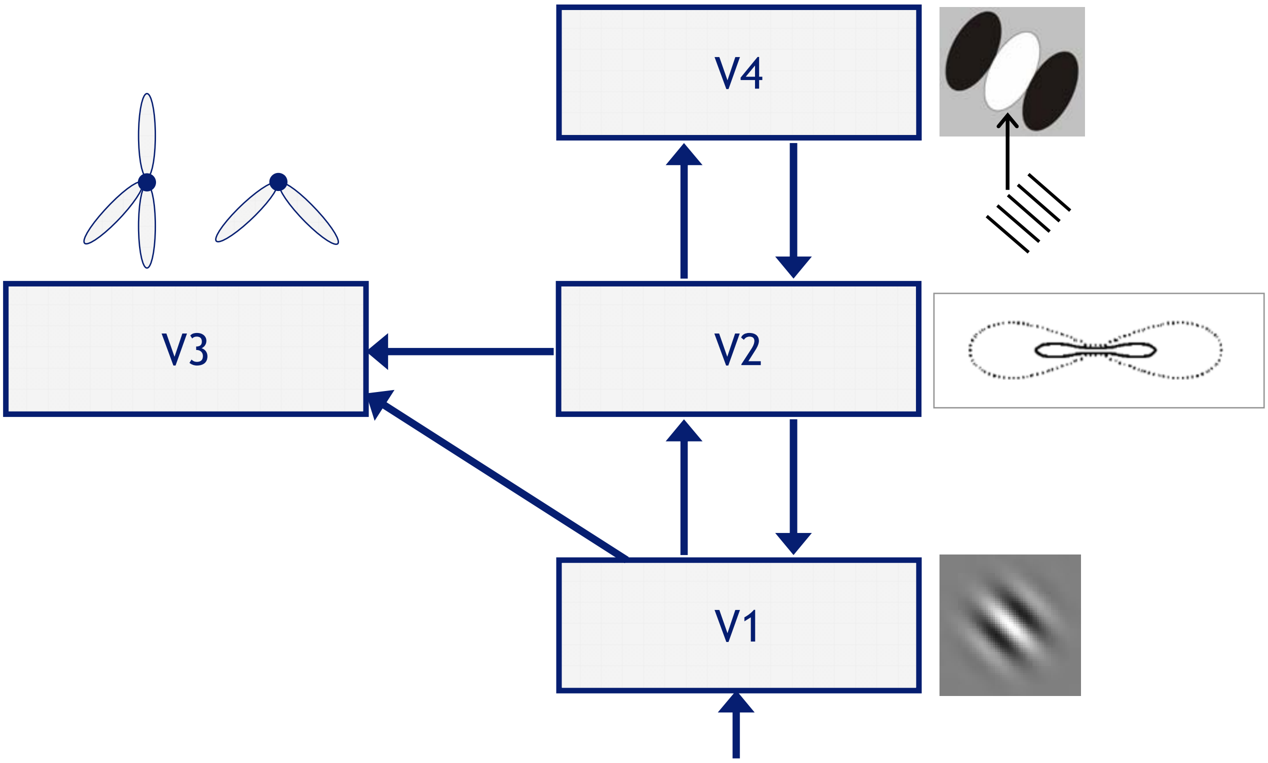


context / modulation / prediction

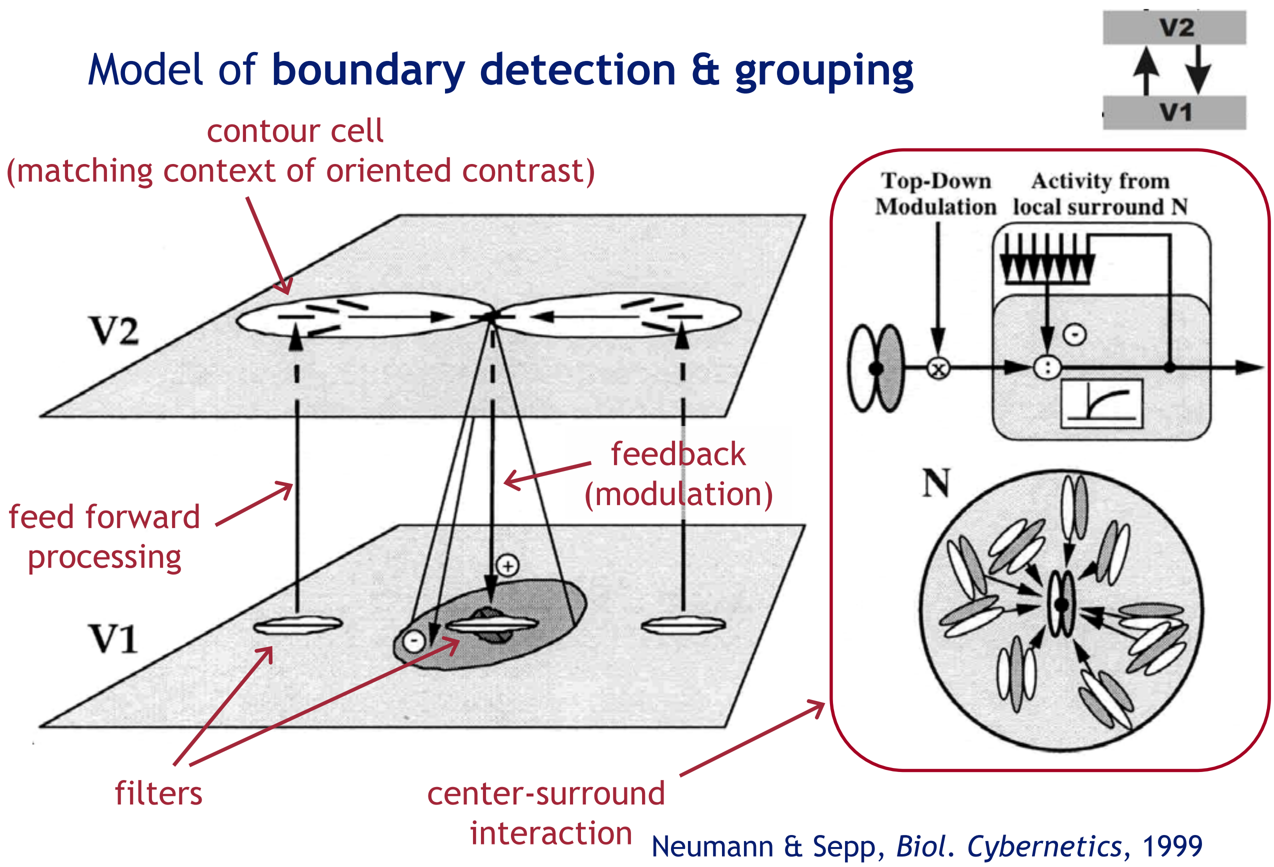
Yet little is really known what the role of feedback and the distributed computation is – top-down processes coordinate and bias local activity across lower-level regions based on global, contextual information

Form processing

Hierarchical form and shape boundary computation

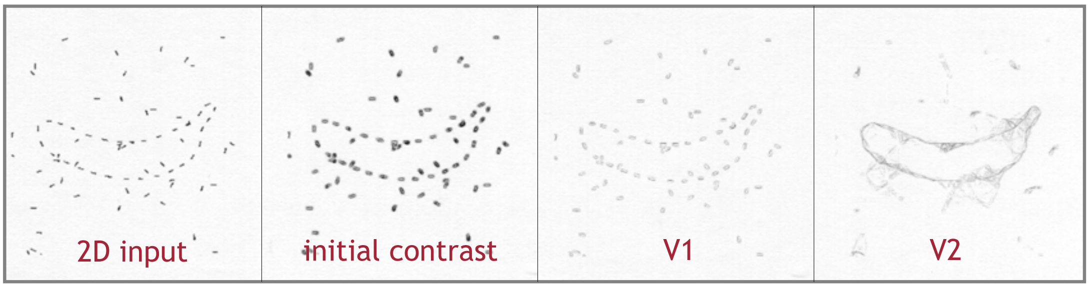


Model of **boundary detection & grouping**

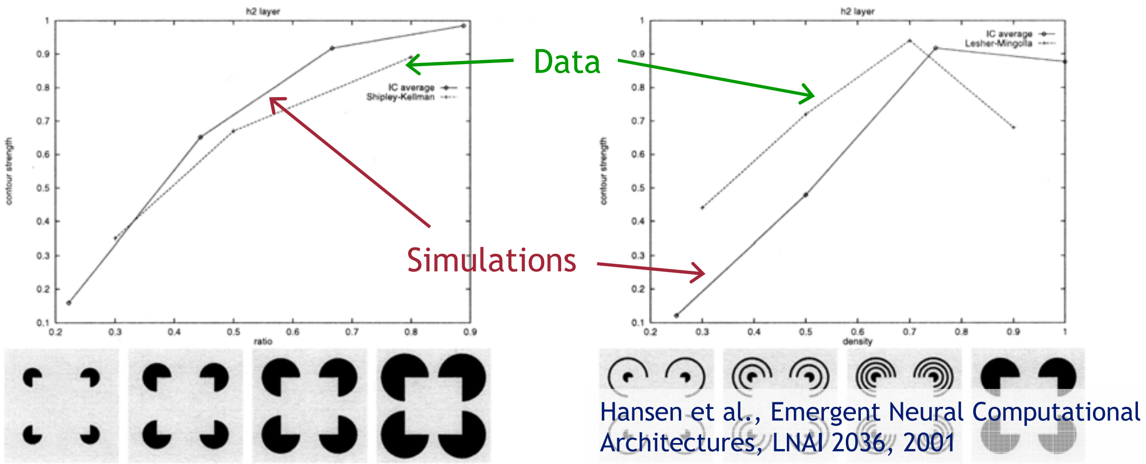


Neumann & Sepp, *Biol. Cybernetics*, 1999

Some computational results



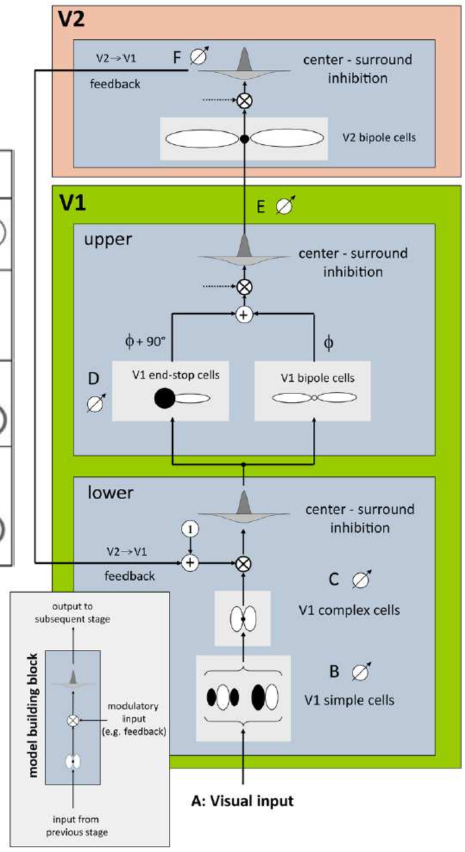
Neumann & Sepp, *Biol. Cybernetics*, 1999



Hansen et al., Emergent Neural Computational Architectures, LNAI 2036, 2001

Junctions can be read-out from distributed response maps in V1/V2

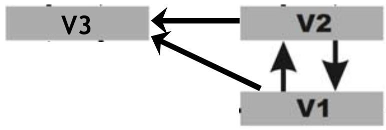
model area	cue type / structure / cell type	2D - corner	3D - corner	transparency	occlusion	contour	illusory contour
	structure	L	Y, Ψ	X	T		
V1	end-stop	2	3	·	1	·	·
V1	bipole	2	3	2	2	1	·
V2	bipole	·	·	2	1	1	1



Specific activity combinations

Visualization as likelihood map

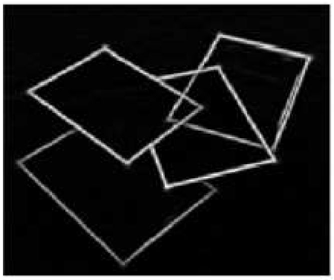
Weidenbacher & Neumann, *PLoS ONE*, 2009



Some computational results

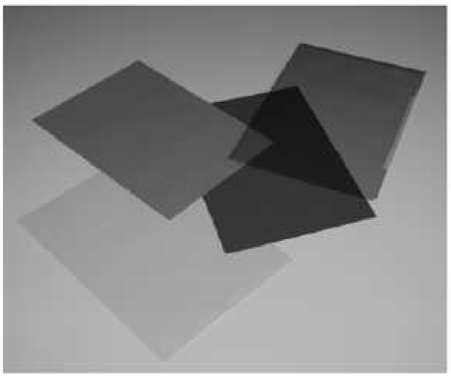
Green: L-junctions
Red: T-junctions
Blue: X-junctions

V1 bipole + V1 end-stop cells

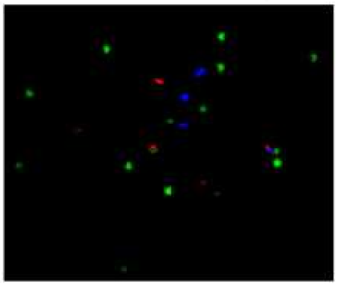


V2 bipole cells

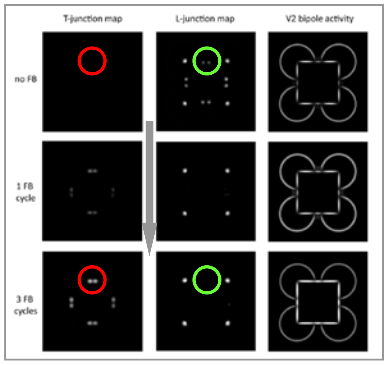
Stimulus



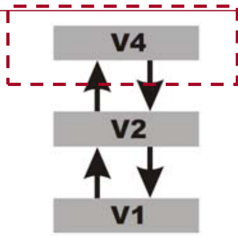
Likelihood feature map



Detected corners



Weidenbacher & Neumann, *PLoS ONE*, 2009

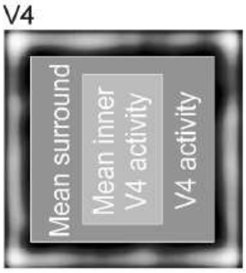


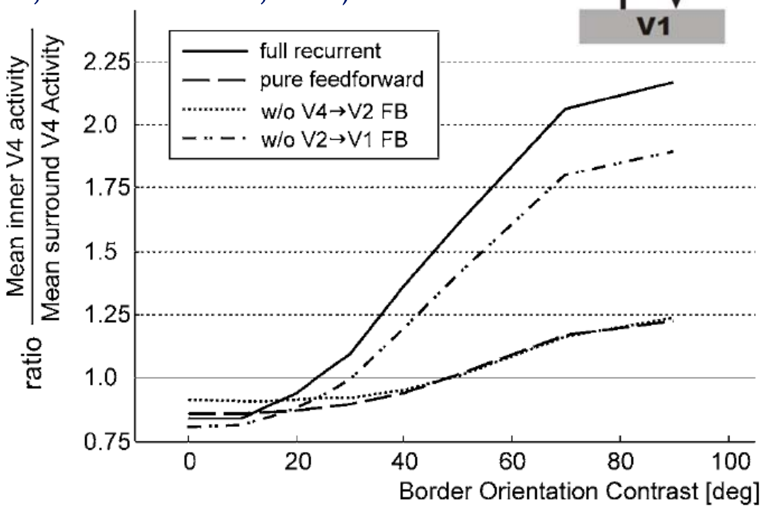
Feedback is used in **texture segregation**

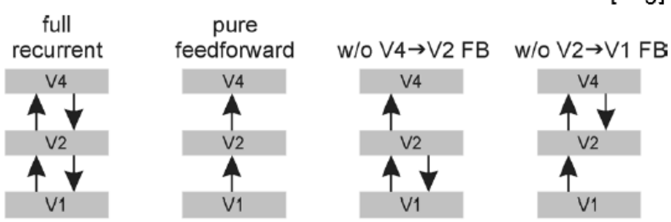
(compare stimuli H.C. Nothdurft, *Vision Research*, 1985)

Stimulus (example):
BN = 20°

V4







Thielscher & Neumann,
Neuroscience, 2003; *Spatial Vision*, 2005

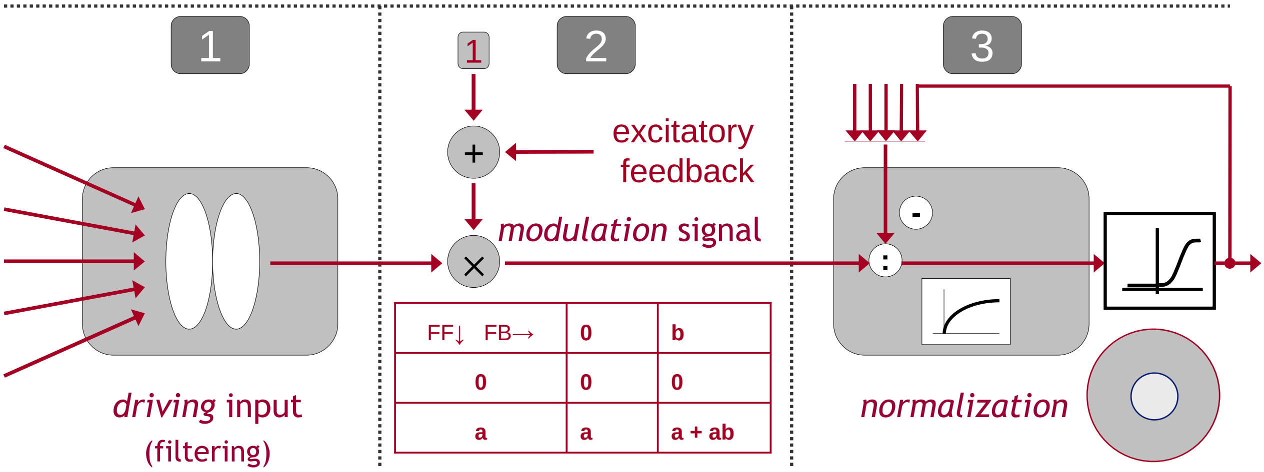
Generic neural model – columns and areas

Processing cascade: **Feedforward & feedback interaction**

FF = driver, FB = modulator
Experimental evidence (Hupé *et al.* 1998; Bullier 2001) and theory (Crick & Koch 1998)

FB is excitatory (in early visual cortical stages)

Withdrawal of FB … leads to less responsiveness to target object and higher response to background (similar to *biased competition* in attention – normalization model)



FF↓ FB→	0	b
0	0	0
a	a	a + ab

(**modulation**: linking inputs, Eckhorn et al., *NECO*, 1990)

(**normalization**: Carandini & Heeger, *Science*, 1994)

Gradual activation – membrane potential & firing rates

Response (non-) linearities (compare Carandini et al., *J. Neurosci.*, 1997)

- Driving feed-forward activation, filtering, and modulating feedback

center-surround filtering; lateral interaction; modulation (via feedback)

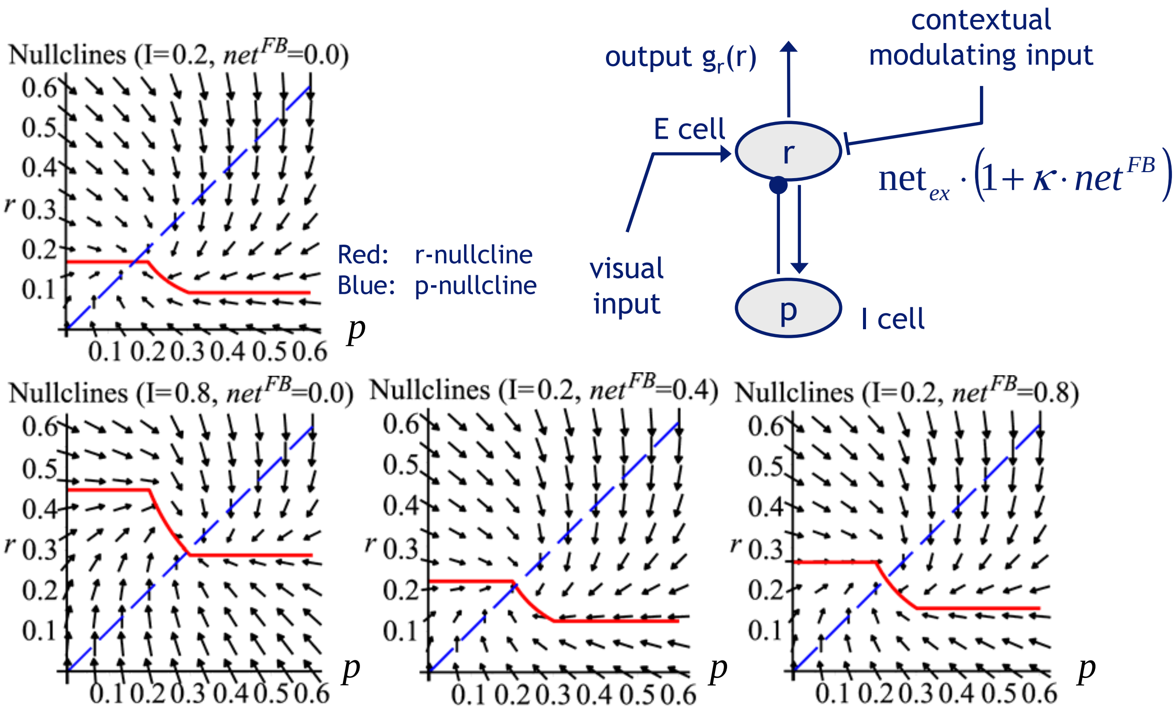
$$\tau \frac{dr_{\mathbf{v}}(\mathbf{x},t)}{dt} = -\alpha \cdot r_{\mathbf{v}}(\mathbf{x},t) + (\beta - r_{\mathbf{v}}(\mathbf{x},t)) \cdot \left[u_{\mathbf{v}}(\mathbf{x},t) + \gamma_{lat} \cdot \left\{ g_r(r_{\mathbf{v}}) \overset{\mathbf{x},\mathbf{v}}{*} \Lambda_\sigma^+ \right\}(\mathbf{x},t) \right] \cdot (1+\gamma \cdot \mathrm{net}^{FB}) - (\lambda + r_{\mathbf{v}}(\mathbf{x},t)) \cdot \left\{ u_{\mathbf{v}} \overset{\mathbf{x},\mathbf{v}}{*} \Lambda_\sigma^- \right\}(\mathbf{x},t) - \gamma_p \cdot r_{\mathbf{v}}(\mathbf{x},t) \cdot g_p(p_{\mathbf{v}}(\mathbf{x},t))$$

- Inhibitory pool activity (normalization)

$$\tau_p \frac{dp_{\mathbf{v}}(\mathbf{x},t)}{dt} = -\alpha_p \cdot p_{\mathbf{v}}(\mathbf{x},t) + \beta_p \cdot \left\{ g_r(r_{\mathbf{v}}) \overset{\mathbf{x},\mathbf{v}}{*} \Lambda_p^- \right\}(\mathbf{x},t)$$

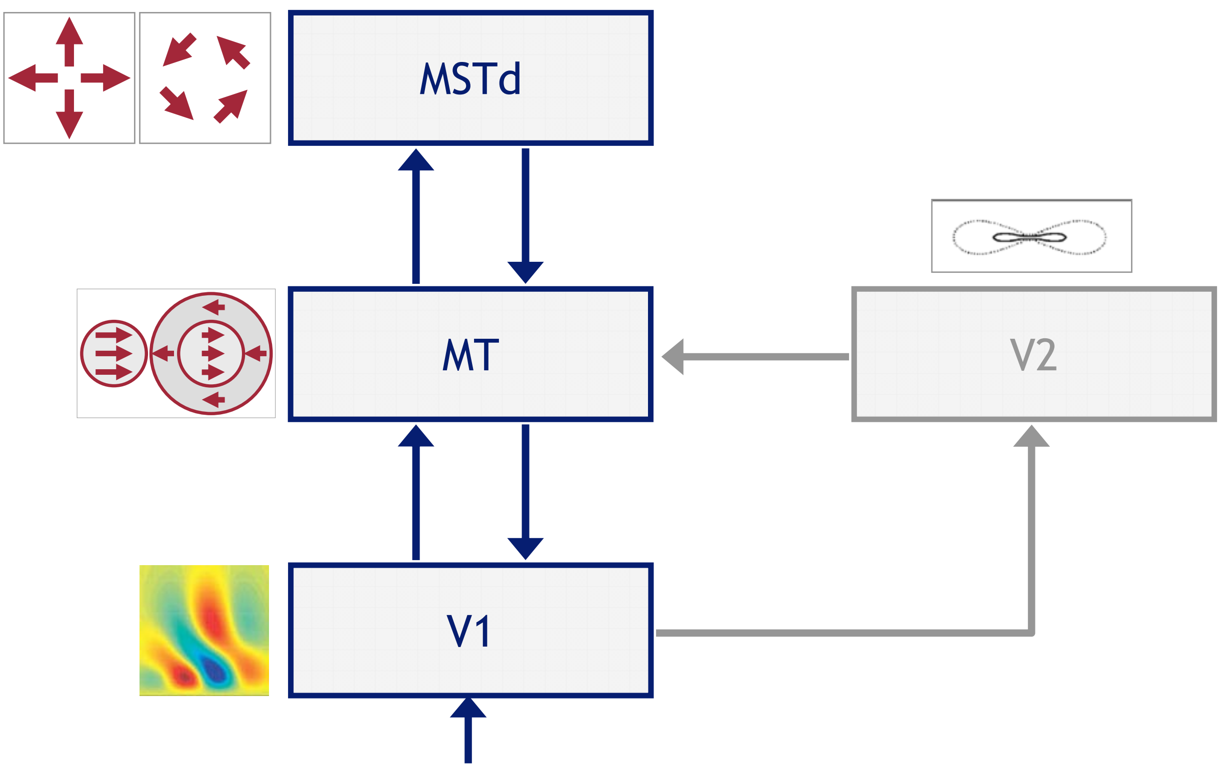
(Bouecke, Tlapale, Kornprobst & Neumann, *EURASIP JASP*, 2011)

Reduced columnar model – excit.-inhibit. (E-I) pairs for given feature
(compare L. Zhaoping, *Curr. Op. Neurobiol.*, 2011)



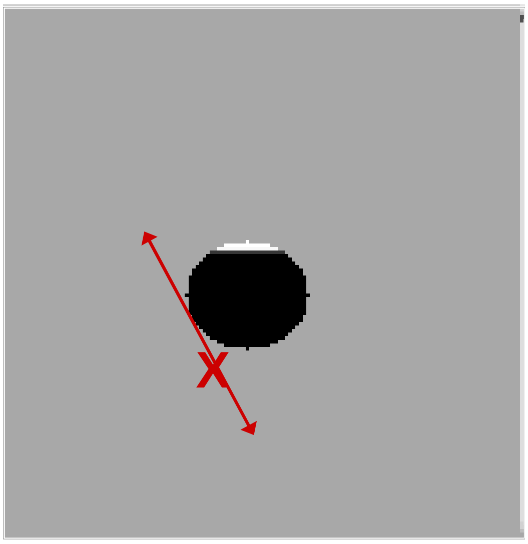
Motion processing

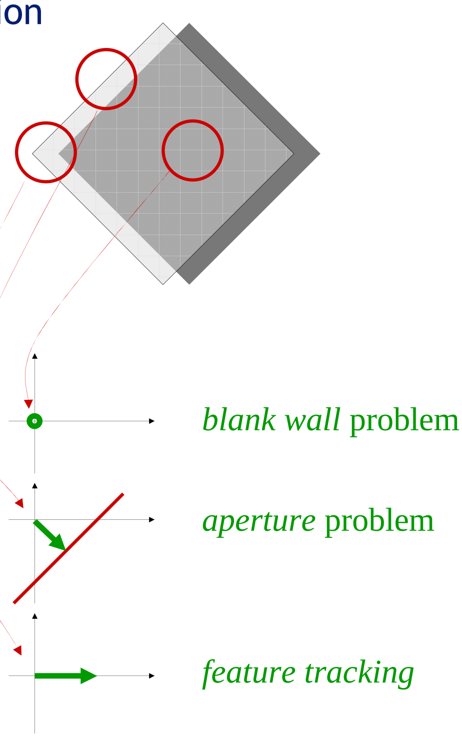
Hierarchical motion computation



Motion analysis – initial motion detection is ambiguous

The aperture problem of motion detection

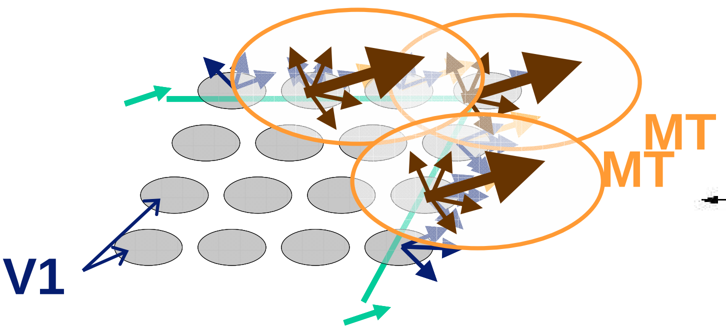


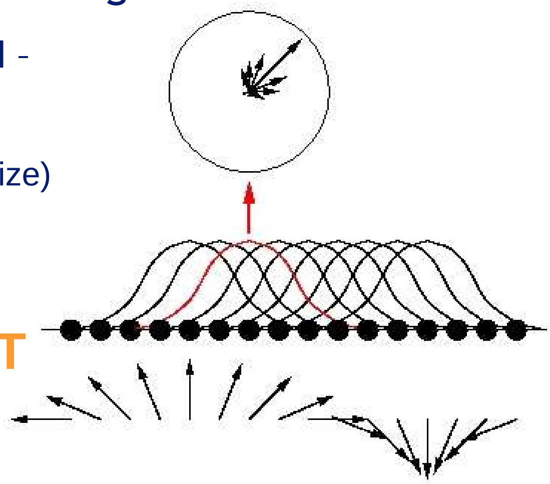


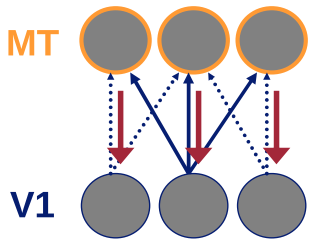
The brain needs to solve a binding problem

Modeling the integration of motion signals in area MT

- Local motion signals are integrated - summation of activities
- Large receptive fields (≈ 8 × V1 RF size)







- Recurrent V1 – MT interaction

  Disambiguation of visual motion in V1

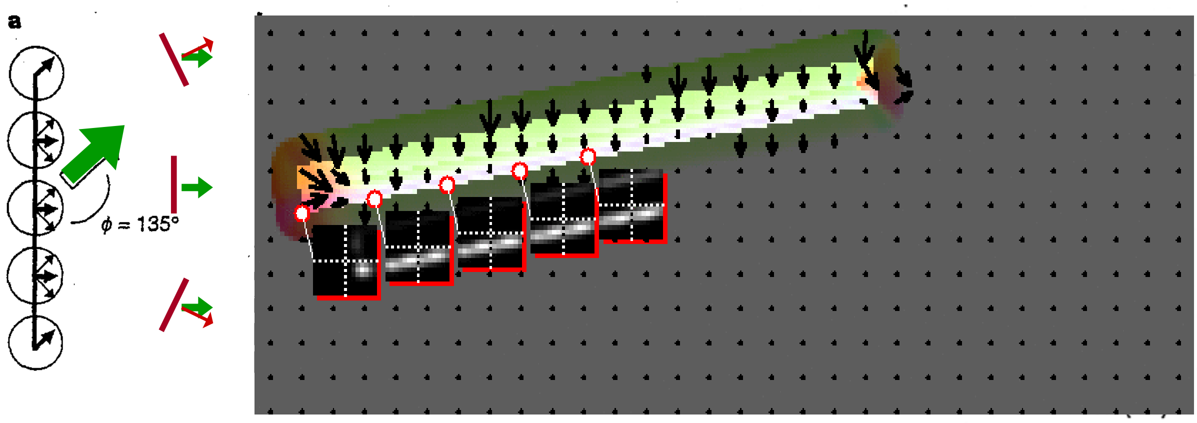
  Filling-in of disambiguated motion signals

Bayerl & Neumann, *NECO*, 2004

The brain solves the aperture problem dynamically

Temporal dynamics of area MT

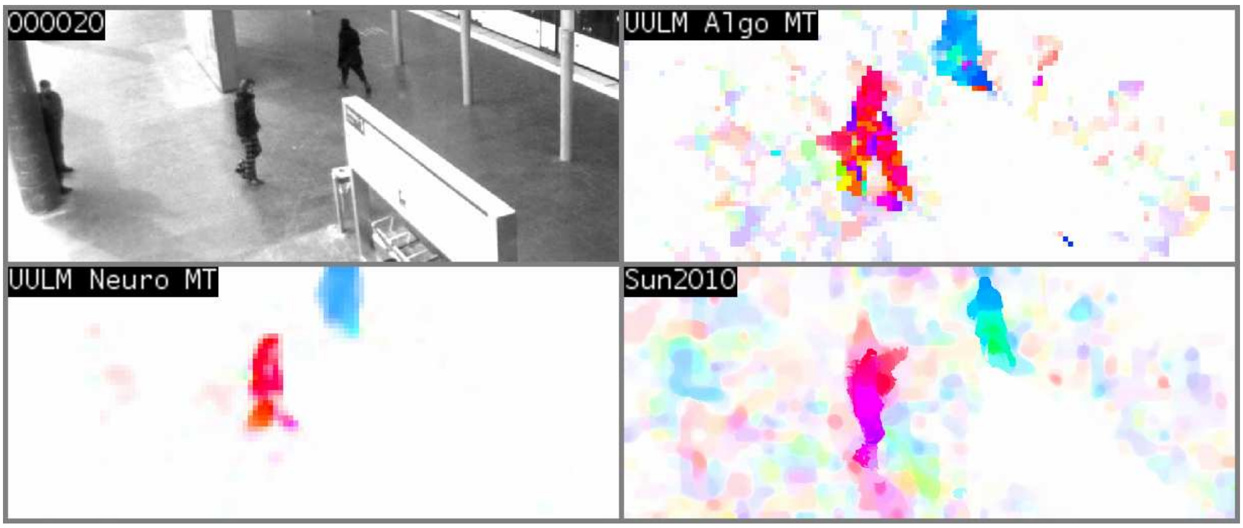
- After 60ms: MT cells respond to motion perpendicular to a contour (component response)
- After 150ms: MT cells indicate the actual stimulus direction (pattern response)



Bayerl & Neumann, *NECO*, 2004 Pack & Born, *Nature*, 2001

Neural models successfully process real-world sequences

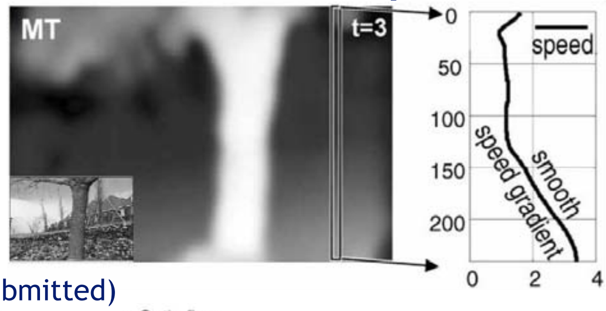
Action videos (EU SEARISE, joint INRIA/UUlm modeling)



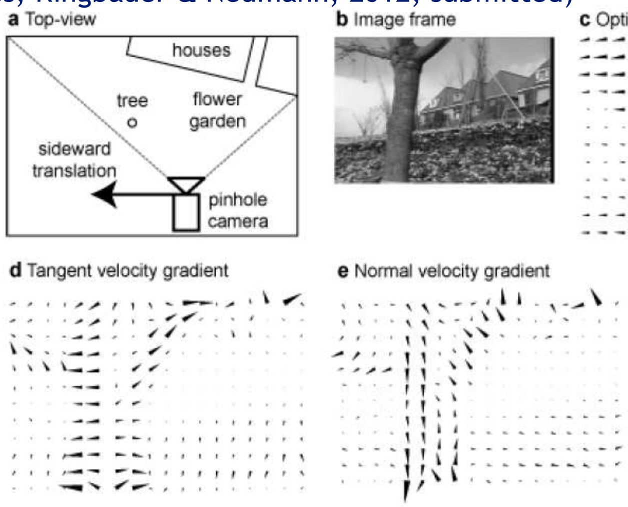
- short-range scenario – platform scene with high temporal resolution
- full neural model (UUlm/INRIA), motion algorithm (UUlm), Sun et al., CVPR'10

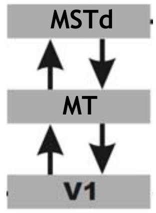
Motion gradients are represented in MT and beyond

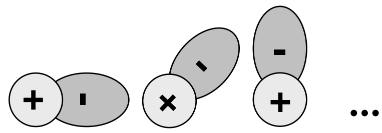
Example case *flower garden* seq.
– V1-MT motion integration
(Bayerl & Neumann, *NECO*, 2004)



Motion gradients – MSTd
(Raudies, Ringbauer & Neumann, 2012, submitted)







Occurrence of **motion (semi-) transparency**

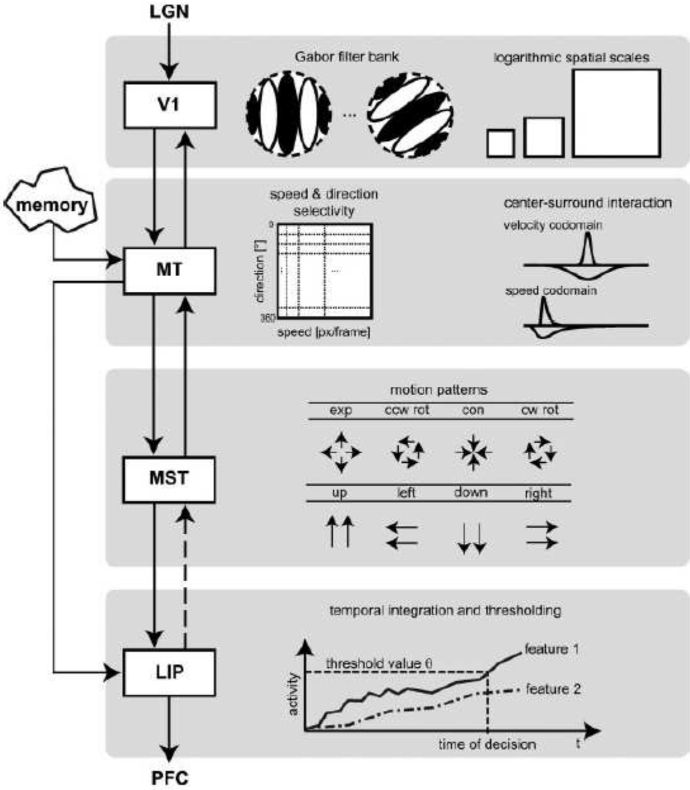
real motion transparency

semi-transparent motions

Interdigitating net motion signals appear to be integrated separately

Shibuya crosswalk, Tokyo
http://www.youtube.com/watch?v=4RYYHckgyUA

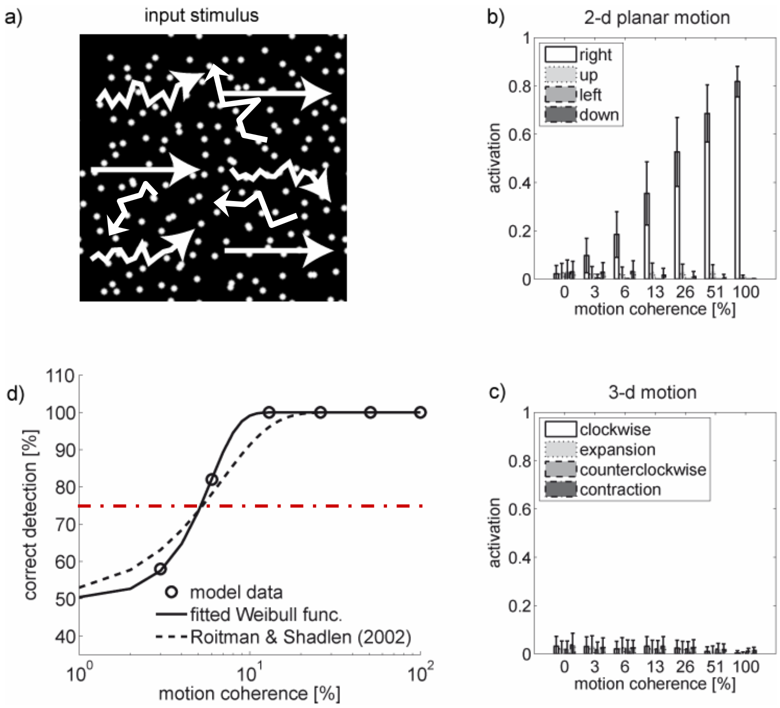
Motion representation in model cortical hierarchy



- **Necessary conditions** for perceiving multiple velocities at single locations: Define center-surround interaction in velocity space
- **Sufficient conditions**: Include global motion pattern responses

Raudies, Mingolla & Neumann, *NECO*, 2011

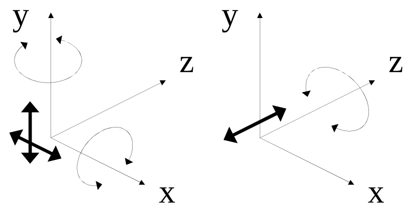
Decide about motion at different coherence levels



Rightward motion for n% of all dots (random selection in each frame), other dots appear at random positions

Motion pattern cell activities

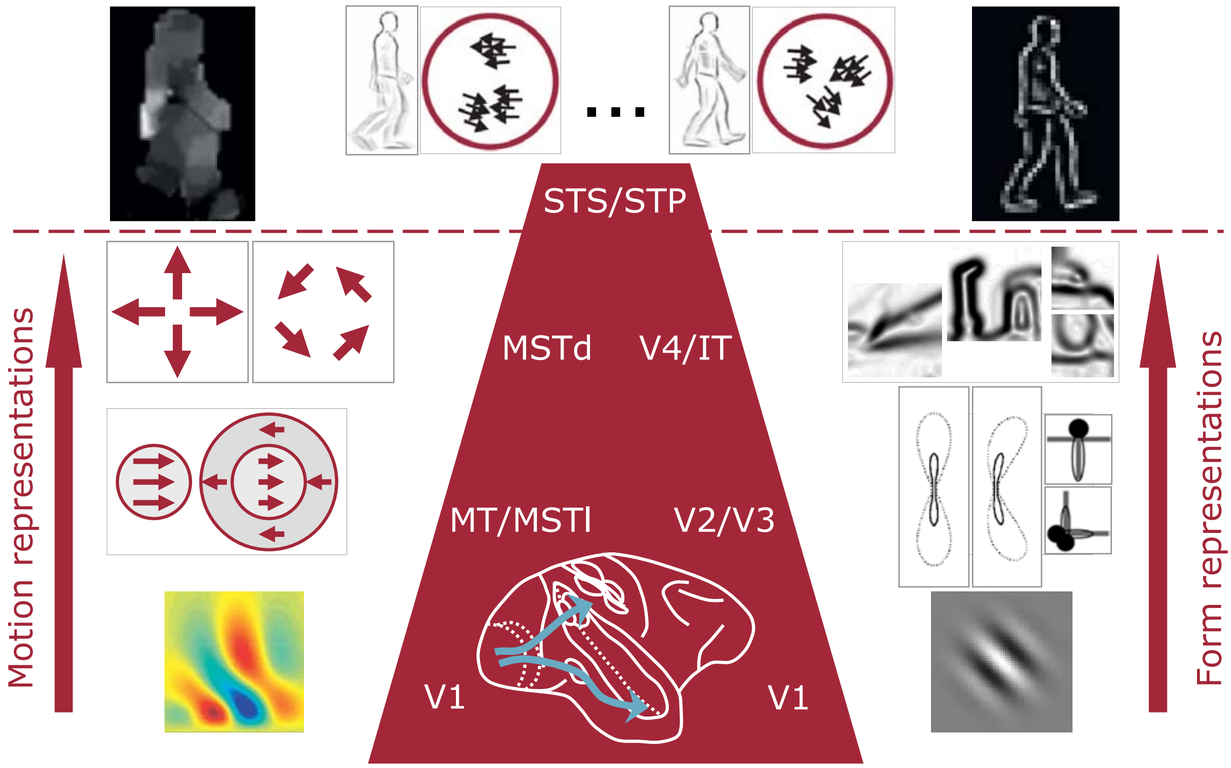
- 2D planar motion
- 3D pattern motion – EXP, CON,
- ROTcw, ROTccw



Result: Psychometric function fitted to decisions (temporal integration of signals from motion patterns & threshold function)

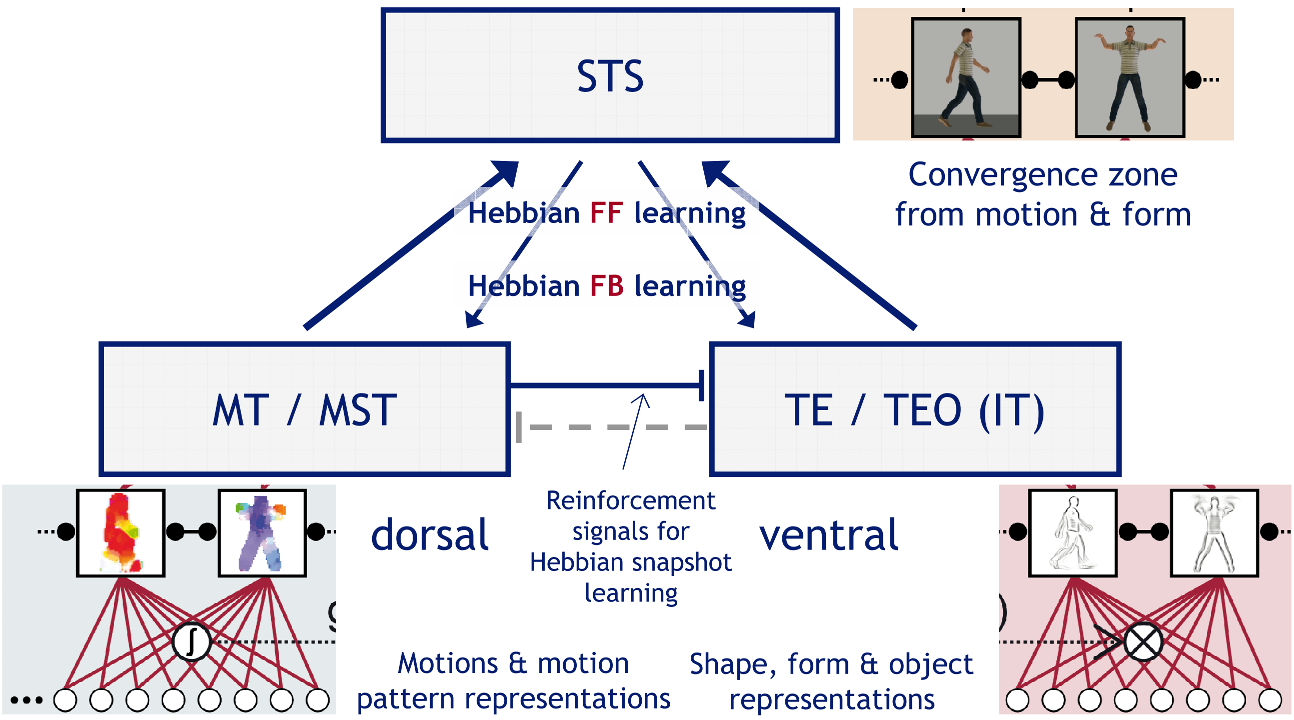
Roitman & Shadlen, *J Neuroscience*, 2002   Raudies, Mingolla & Neumann, *NECO*, 2011

Modeling **hierarchies and representations** in cortex



Model architecture for biological motion analysis

Biological motion is represented in neural hierarchy



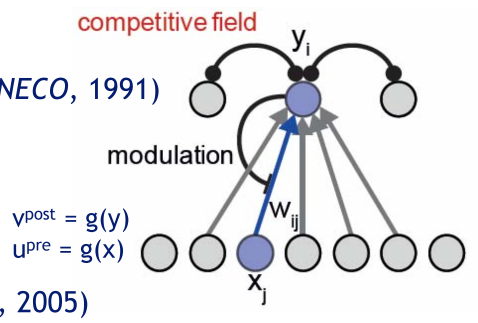
Hebbian learning of **motion and form prototypes**

Learning of prototype representations in form and motion pathway

$$\Delta w_{ji}^{FF,s} = \eta_s \cdot \bar{v}_i^{post}\left(u_j^{pre} - \bar{v}_i^{post} w_{ji}^{FF,s}\right)$$

with s ∈ {form, motion} and trace rule (Földiák, *NECO*, 1991)

$$\bar{v}_i^{t} = (1-\lambda)\cdot \bar{v}_i^{t-1} + \lambda \cdot v_i^{t} \quad (0<\lambda<1)$$



Learning of form prototypes is gated by reinforcement signal from motion energy (inspired by AGREL; Van Ooyen & Roelfsema, *NECO*, 2005)

$$\Delta w_{ji}^{FF,f} = \eta_f \cdot \boxed{g(m_c)} \cdot \bar{v}_i^{post}\left(u_j^{pre} - \bar{v}_i^{post} w_{ji}^{FF,f}\right),\quad m_c = \int_\Omega u_\phi(\mathbf{x}) \cdot \Lambda(\mathbf{x})\, d\phi\, d\mathbf{x}$$

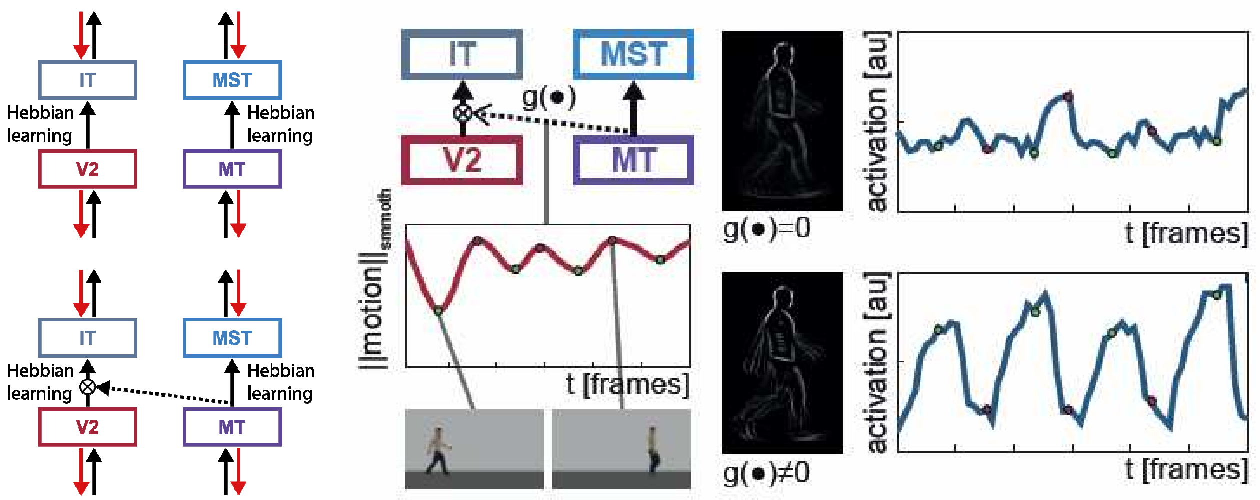
Hebbian learning of **sequence-selective prototypes**

- **Feedforward** connections are learned (instar)
  convergent connections IT → STS & MST → STS (Oja’s rule)
- **Feedback** connections are learned (outstar)
  divergent connections STS → IT + MST (Grossberg rule)

Form prototypes are **snapshots** of articulated poses

Hebbian learning (with trace) in **form** pathway

… and incl. **reinforcement** signal from **motion** pathway

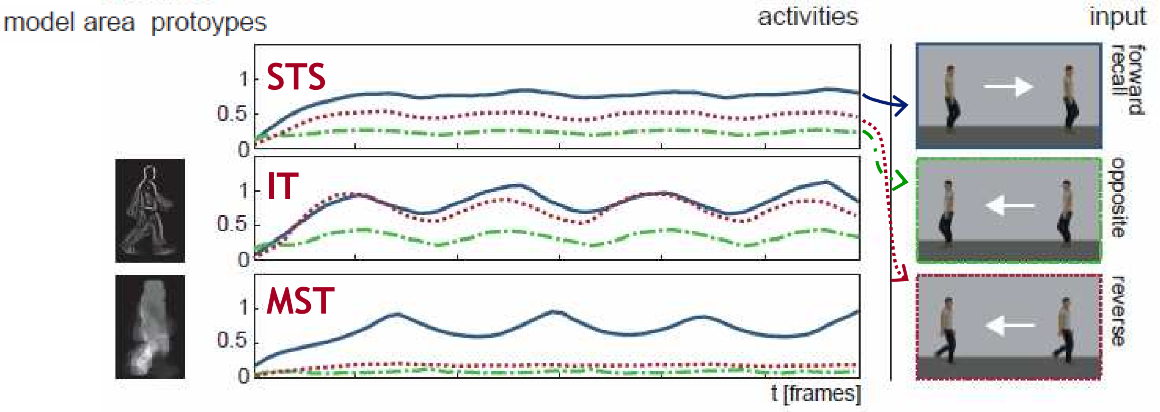


Main observations:

- Automatic selection of key poses (possible for static (ambivalent) poses)
- Reinforcement of learning inspired by AGREL

Sequence selectivity of STS neurons

(Perrett et al., *J. Exp. Biol.*, 1989; Oram & Perrett, *J. Neurophysiol.*, 1996)



- Probing sequence-selective representations in STS
  - Recall walking to the right (forward training sequence)
  - Walking to the left (opposite movement)
  - Walking backwards from right to left (reverse movement)
- STS neurons are driven by snapshots (form) & motion patterns

Layher & Neumann, *JoV* (abstracts), 2012; Layher et al., *ICANN'12*, LNCS 7552, 2012

Summary and conclusion

Form and motion processing – same generic principles

- Boundary grouping, corner/junction readout, texture boundary detection
- Motion integration, gradients, transparent motion segregation

3-stage cascade of columnar model architecture

- Filtering – linear/non-linear
- Modulating feedback
- Center-surround pool normalization

Biological inspiration for computational vision

- Building blocks for composition of system's components
- Enables context-information to bias early processing by feedback
- Unsupervised learning intermediate level representations, e.g. for biological motion analysis

Thanks to …

The visionaries @ UUlm & alumni

- Tobias Brosch
- Olman Gomez
- Luma Issa
- Georg Layher
- Stefan Ringbauer
- Stephan Tschechne
- Pierre Bayerl
- Cornelia Beck
- Jan D. Bouecke
- Silvia Corchs
- Florian Raudies (@BU)
- Axel Thielscher
- Ulrich Weidenbacher

Collaborators

- Yiannis Aloimonos & Cornelia Fermüller (U Maryland)
- Martin Giese & Dominik Endres (U Tübingen)
- Mark W. Greenlee (U Regensburg)
- Pierre Kornprobst (INRIA, Sophia-Antipolis)
- Ennio Mingolla (Northeastern Univ.)
- Pieter Roelfsema (Netherlands Institute of Neuroscience, NIN)

Sponsoring agencies

- EU, FP 6/7
- German Federal Ministry of Research and Technology (BMBF)
- German Research Foundation (DFG)

Thank you for your attention !

Website:

http://www.uni-ulm.de/in/neuroinformatik/mitarbeiter/h-neumann.html