Biologically Inspired Active 3D Vision — Course Syllabus

Oct 26, 2026 · @Luiz Velho

A PhD course on one thesis: vision is budgeted, active sampling. It joins two topics — biologically inspired computer vision, and active foveated reconstruction and understanding of 3D scenes — treating biology as a source of constraints and hypotheses, not a template to copy.

Course at a glance

Item	Detail
Level	PhD (open to advanced MSc students)
Format	12 weeks; two 2-hour sessions per week — a lecture and a lab
Institution	IMPA, Visgraf laboratory
Instructor	Luiz Velho
Prerequisites	Linear algebra, probability and estimation, an introductory course in computer vision or graphics, working Python
Tools	Python, Blender (Cycles), OpenCV, a learned stereo matcher (IGEV-Stereo or FoundationStereo), the `visgraf/fov-3d-vision` engine, an AI coding agent
Work	Weekly lab notebook, one paper presentation, a group research project

The course is organized in three parts: **I. The eye is not a camera** (weeks 1–4), **II. Two eyes** (weeks 5–8), **III. Looking to know** (weeks 9–12).

Learning goals

By the end of the course, students should be able to:

1. **Conceptual:** frame vision as sampling and control under a budget, and tell a moving camera from active perception.
2. **Technical:** work with verged binocular geometry (horopter, Listing and Donders laws, vertical disparity), non-uniform sampling (cortical magnification, log-polar and spherical lattices), classical and learned stereo, gaze policies, and fusion over repeated looks.
3. **Methodological:** compare methods at equal budget (rays, energy, latency), write decision rules before running experiments, and read psychophysics for its computational content.
4. **Practical:** build and run a closed-loop experiment in simulation and report it in paper form, whether the result is positive or negative.

A standing question runs through every session: *under which constraint is this biological feature optimal, and does that constraint hold for our system?*

Weekly schedule

Each week pairs a lecture with a lab on the `fov-3d-vision` engine; readings are listed by author and year (full references below).

Week	Lecture	Lab	Readings	Project milestone
Part I	**The eye is not a camera**			
1	Recorder vs sampler: why CV adopted the static pinhole; Marr's program; the 1987–1995 active vision wave and why it faded	Install the engine; render a scene conventionally and as a 360° panorama	Marr 1982 (ch. 1); Aloimonos et al. 1988; Bajcsy 1988; Ballard 1991	Groups formed
2	The retina and foveation: photoreceptor topography, cortical magnification, log-polar and space-variant sampling	Foveated rendering: the warp and its parameters (E2, e_max); render budget vs resolution	Curcio et al. 1990; Schwartz 1977; Traver & Bernardino 2010	
3	Eye movements: saccades, fixational movements, vergence; Listing and Donders laws; the mathematics of eye rotations	Gaze kinematics: fixation sequences on the head sphere	Rucci & Poletti 2015; Haslwanter 1995; Land & Tatler 2009	
4	Sampling under a budget: why foveation only pays when the budget binds; data processing inequality; "same posterior for less computation"	Equal-budget comparison: uniform vs foveated sampling of one scene	Bolduc & Levine 1998; Geisler & Perry 1998	Project proposals due
Part II	**Two eyes**			
5	Binocular geometry: verged eyes, Vieth–Müller circle, horopter, epipolar geometry on the sphere, vertical disparity	Verged fixation pairs; ground-truth correspondence and triangulation	Hartley & Zisserman 2004 (ch. 9–12); Mayhew & Longuet-Higgins 1982; Gårding et al. 1995	
6	Human stereopsis: stereoacuity, Panum's area, random-dot stereograms, da Vinci stereopsis, the disparity energy model	Half-occlusions and depth edges in rendered pairs: where truth has no match	Julesz 1971; Nakayama & Shimojo 1990; Ohzawa et al. 1990; Cumming & DeAngelis 2001	

Week	Lecture	Lab	Readings	Project milestone
7	Computational stereo: rectification, block matching, semi-global matching, the evaluation taxonomy; what rectification throws away	SGBM on foveal tangent-plane projections; error maps by kind (occluded, window, search)	Marr & Poggio 1976; Scharstein & Szeliski 2002; Hirschmüller 2008	
8	Learned stereo: cost volumes, iterative refinement, foundation models; failure at edges and occlusions	Learned matcher vs SGBM on the same foveal windows	Lipson et al. 2021; Xu et al. 2023; Wen et al. 2025	Midterm checkpoint: baseline results
Part III	**Looking to know**			
9	Active perception: gaze as a POMDP; information gain, coverage, inhibition of return; optimal search in humans	The active loop: run and compare policies at equal rays	Najemnik & Geisler 2005; Itti & Koch 2001; Bajcsy et al. 2018	
10	Belief and fusion: representations on the sphere, variance models, repeated looks, consensus vs averaging	Replay a saved long run; re-fuse with mean vs consensus	Velho TR-09-2026; Mishra et al. 2009	Pre-registered decision rules due
11	Applications: foveated rendering and gaze-contingent displays, the vergence–accommodation conflict, robot heads, event cameras	Project work	Guenter et al. 2012; Patney et al. 2016; Hoffman et al. 2008; Gallego et al. 2022	
12	Synthesis: what biology buys, and when	Project presentations	—	Final reports due

Project track

Students work in groups of 3–4 on one research question built on the `fov-3d-vision` engine and its saved records. A negative result reported well earns full credit.

Flagship project: classifying the foveal window. Reuse the nominal-vergence search to classify the fixated window. The aim is to separate local properties of the fixated point (near a half-occlusion) from global properties of the window (contains a depth edge). The matcher can then switch to an adaptive strategy where a single block matcher fails, at the roughly 8% of the area at depth edges and half-occlusions.

Alternative projects

- **Learned matcher on foveal windows:** IGEV-Stereo or FoundationStereo on rectified tangent-plane projections, judged against SGBM at equal rays.
- **The undecided map as a gaze policy:** fixate where consensus among looks is still undecided, compared with coverage-first.
- **Vertical disparity as a cue:** recover viewing distance and fixation error from unrectified verged pairs.
- **The warp as an action:** vary foveation parameters per fixation to break the fine-coverage ceiling.
- **A matcher on the native lattice:** stereo matching directly on non-uniform samples, without reprojection.

Working method

Each group runs its project with an AI coding agent, following a lightweight version of the method in (Velho TR 2026):

- a short `CLAUDE.md` with the project's contract and conventions;
- a decision log, one entry per decision;
- decision rules written before each experiment runs;
- commands classed by cost (interactive under 10 s, batch under 5 min, overnight);
- roles rotated within the group across implementation, review and reporting.

Milestones: proposal (week 4) → baseline results (week 8) → pre-registered decision rules (week 10) → presentation and final report (week 12).

Assessment

Component	Weight (%)	What is judged
Lab notebook	20	Weekly labs run and interpreted; one paragraph per lab on what the result means
Paper presentation	15	One reading from the list, presented with a critique: which constraint makes the idea work, and does it hold for engineered systems
Project proposal and checkpoint	15	A clear question, a baseline, and an equal-budget comparison defined in advance
Final report	35	Paper-style report of 8–10 pages; claims traceable to code and runs; decision rules stated before results
Final presentation	15	20-minute group talk, with each member presenting a part

Readings

Compiled without a lookup; check volume and page details before publishing.

Course notes (Visgraf technical reports)

- Velho, L. *Physiology of Foveated Human Vision: Retina, Fovea, and Parafovea from Biology to Computation.*
- Velho, L. *Eye Movements and Stereoscopic Depth Perception.*
- Velho, L. (2026). Experiments on Active 3D Vision. TR-09-2026.
- Velho, L. (2026). Personal Notes on a Methodology for R&D in Applied Computational Mathematics. TR-00-2026.

Part I — The eye is not a camera

- Marr, D. (1982). *Vision.* Freeman.
- Aloimonos, J., Weiss, I., Bandyopadhyay, A. (1988). Active vision. *IJCV* 1(4).
- Bajcsy, R. (1988). Active perception. *Proc. IEEE* 76(8).
- Ballard, D. (1991). Animate vision. *Artificial Intelligence* 48.
- Curcio, C. et al. (1990). Human photoreceptor topography. *J. Comparative Neurology* 292.
- Schwartz, E. (1977). Spatial mapping in the primate sensory projection. *Biological Cybernetics* 25.
- Bolduc, M., Levine, M. (1998). A review of biologically motivated space-variant data reduction models for robotic vision. *CVIU* 69.
- Traver, V. J., Bernardino, A. (2010). A review of log-polar imaging for visual perception in robotics. *Robotics and Autonomous Systems* 58.
- Haslwanter, T. (1995). Mathematics of three-dimensional eye rotations. *Vision Research* 35.
- Rucci, M., Poletti, M. (2015). Control and functions of fixational eye movements. *Annual Review of Vision Science* 1.
- Land, M., Tatler, B. (2009). *Looking and Acting.* Oxford University Press.
- Geisler, W., Perry, J. (1998). A real-time foveated multiresolution system for low-bandwidth video communication. *Proc. SPIE* 3299.

Part II — Two eyes

- Hartley, R., Zisserman, A. (2004). *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge University Press.
- Julesz, B. (1971). *Foundations of Cyclopean Perception.* University of Chicago Press.
- Marr, D., Poggio, T. (1976). Cooperative computation of stereo disparity. *Science* 194.

- Mayhew, J., Longuet-Higgins, H. C. (1982). A computational model of binocular depth perception. *Nature* 297.
- Gårding, J., Porrill, J., Mayhew, J., Frisby, J. (1995). Stereopsis, vertical disparity and relief transformations. *Vision Research* 35.
- Nakayama, K., Shimojo, S. (1990). Da Vinci stereopsis: depth and subjective occluding contours from unpaired image points. *Vision Research* 30.
- Ohzawa, I., DeAngelis, G., Freeman, R. (1990). Stereoscopic depth discrimination in the visual cortex. *Science* 249.
- Cumming, B., DeAngelis, G. (2001). The physiology of stereopsis. *Annual Review of Neuroscience* 24.
- Howard, I., Rogers, B. (2012). *Perceiving in Depth.* Oxford University Press (reference).
- Scharstein, D., Szeliski, R. (2002). A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *IJCV* 47.
- Hirschmüller, H. (2008). Stereo processing by semiglobal matching and mutual information. *IEEE TPAMI* 30.
- Lipson, L., Teed, Z., Deng, J. (2021). RAFT-Stereo. *3DV*.
- Xu, G. et al. (2023). Iterative geometry encoding volume for stereo matching (IGEV-Stereo). *CVPR*.
- Wen, B. et al. (2025). FoundationStereo. *CVPR*.

Part III — Looking to know

- Bajcsy, R., Aloimonos, Y., Tsotsos, J. (2018). Revisiting active perception. *Autonomous Robots* 42.
- Najemnik, J., Geisler, W. (2005). Optimal eye movement strategies in visual search. *Nature* 434.
- Itti, L., Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience* 2.
- Mishra, A., Aloimonos, Y., Cheong, L.-F. (2009). Active segmentation with fixation. *ICCV*.
- Guenter, B. et al. (2012). Foveated 3D graphics. *ACM TOG* 31(6).
- Patney, A. et al. (2016). Towards foveated rendering for gaze-tracked virtual reality. *ACM TOG* 35(6).
- Hoffman, D., Girshick, A., Akeley, K., Banks, M. (2008). Vergence–accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of Vision* 8(3).
- Gallego, G. et al. (2022). Event-based vision: a survey. *IEEE TPAMI* 44(1).