Reading the Foveal Window — architecture and tasks

Nov 6, 2026 · @Luiz Velho (with Claude Chat)

Goal

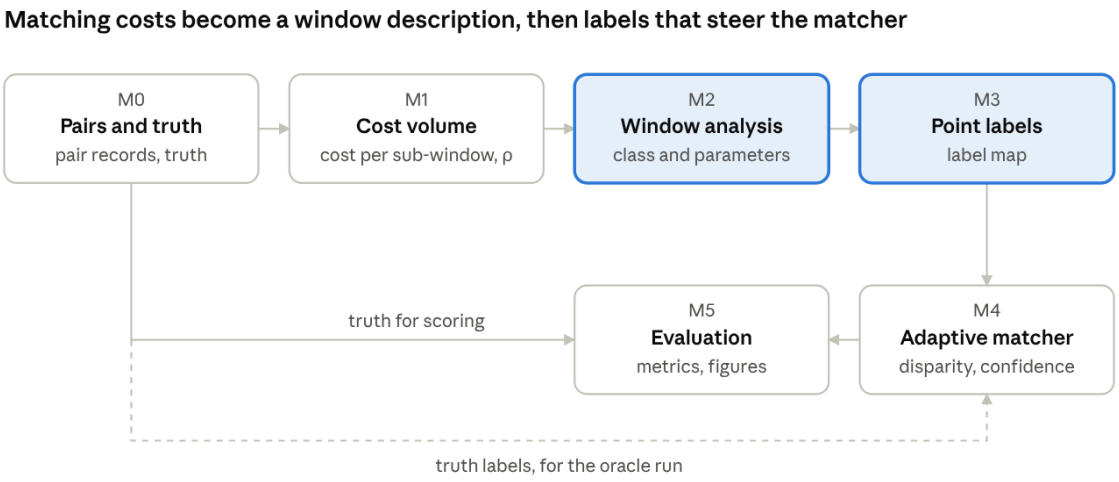
Build a module that reads the structure of the foveal window from the same matching costs the vergence search computes, and use what it reads to adapt the stereo matcher. The target is fewer errors at depth edges and half-occlusions, where the current block matcher has its floor (about 8% of the judged area).

The module describes the window at two levels. Globally, it says what the window contains: one frontal surface, one slanted surface, two layers with a boundary, repeated texture, low texture, or clutter. Locally, it labels each point, and the labels follow from the global structure: a point in the interior of a layer, in the band around a boundary, or in a half-occluded strip seen by one eye only.

How each part is computed is left to the team. This document fixes what the parts are, what they exchange, and what each must deliver.

Architecture

The pipeline runs on one foveal stereo pair at a time: costs, then window structure, then point labels, then a matcher that uses them. The two highlighted modules are the new part; the others exist in some form already.



architecture · 6 modules

The pair record also goes straight to M2 and M4. The dashed path is the oracle: truth labels stand in for M3's labels, to measure what perfect labels would buy.

The modules exchange six data objects, which the team should fix in week one so that work can proceed in parallel:

- **Pair record:** left and right foveal windows sampled in epipolar coordinates (θ, ϕ), with the fixation ω and its vergence.
- **Cost volume:** C_x(ρ) for each sub-window x over a band of inverse depth ρ.
- **Window description:** class, probabilities, and the parameters of that class.
- **Label map:** one label and probability per sample.
- **Matcher output:** disparity, confidence and occlusion mask.
- **Metrics table:** one row per look and label class.

What needs to be implemented

Six modules, each defined only by what goes in, what comes out, and what it must do. Methods, parameters and thresholds are the team's choice.

Module	Input	Output	Must
M0 · Pairs and truth	Blender renders of a scene (depth and position passes), or the saved fixation record	Pair records, plus truth label maps for each pair	Give every foveal sample its true layer, boundary band, and half-occluded strip with the eye that sees it
M1 · Cost volume	Pair record	Cost volume C_x(ρ): one cyclopean matching cost per sub-window x and inverse depth ρ	Use the symmetric (cyclopean) cost of the vergence search; cover a band of ρ around the fixation
M2 · Window analysis	Cost volume, pair record	Window description: class and its parameters (modes, slant, boundary line, depth jump, period)	Separate the six classes: frontal, slanted, two layers, repeated texture, low texture, clutter; give each a probability
M3 · Point labels	Window description	Label map: one label per sample, with a probability	Derive point labels from the window's structure (interior, boundary band, occluded strip and its eye, periodic, low texture)
M4 · Adaptive matcher	Pair record, optional label map	Disparity map, confidence, occlusion mask	Switch strategy by label; with no labels, reproduce today's matcher exactly
M5 · Evaluation	Matcher outputs, truth from M0	Metrics table and figures	Measure outliers at edges and occlusions, coverage and depth error; compare the matcher with and without labels

Once the window module works, a seventh step feeds the labels back into the active loop: boundaries and occluded strips go into the belief, and the gaze policy can move a fixation off an edge or onto the occluding surface.

Deliverables

The work is judged by three comparisons on the same set of pairs. The team sets its own success criteria, and writes them down before each run.

1. **Oracle.** M4 driven by truth labels from M0, against M4 with no labels. This shows how much perfect labels are worth. If they are worth little, stopping here is a valid result.
2. **Classifier.** Labels from M2 and M3, scored against truth from M0.
3. **Closed loop.** M4 driven by the estimated labels. The gap to the oracle shows what classification errors cost.

The team delivers the code for M0 to M5, the metrics table and figures for the three comparisons, and a short report of what worked and what did not.