Title: **Unifying binocular, spatial, and spatio-temporal frequency integration in models of MT neurons.**

Authors: Pamela M. Baker and Wyeth Bair, Dept. of Biological Structure, Univ. of Washington

Our goal is to build a model of MT neuronal responses that can be compared with single-unit data from nearly all currently existing MT stimulus protocols to unify existing findings across labs and offer predictions for future empirical studies. Many widely used models of MT are unable to do this because they omit essential features: specifically, many models (1) do not handle binocular stimuli; (2) do not explicitly incorporate spatial integration; or (3) do not sample broadly across the spectrum of spatial and temporal frequencies. To achieve a more unified model, we have combined critical features of existing models with key elements including an image-computable front end, diverse V1 channels for direction, SF, TF, ocular dominance, disparity and spatial location, and physiological mechanisms including normalization, motion opponency and surround suppression (Fig. 1). We enhanced our recent MT model framework, which accounted for responses to dichoptic plaids and 3D motion[1], to include spatial integration, and found that mechanisms important for explaining neuronal responses to dichoptic stimuli are also critical for explaining spatial integration in MT[2,3] (Fig. 2). We demonstrate how mechanisms present in early cortical levels (V1) are critical for selectivities typically thought to be computed only at deeper stages (MT). We have further extended our models to include multiple spatiotemporal frequency (STF) channels to model responses to Type II plaid stimuli and novel random-line plaids[4]. Our current model, which includes static form channels, now reproduces physiological data on Type II plaids[5,6]. These broad results demonstrate the utility of taking a unified modeling approach to understand functional circuitry in the visual system.

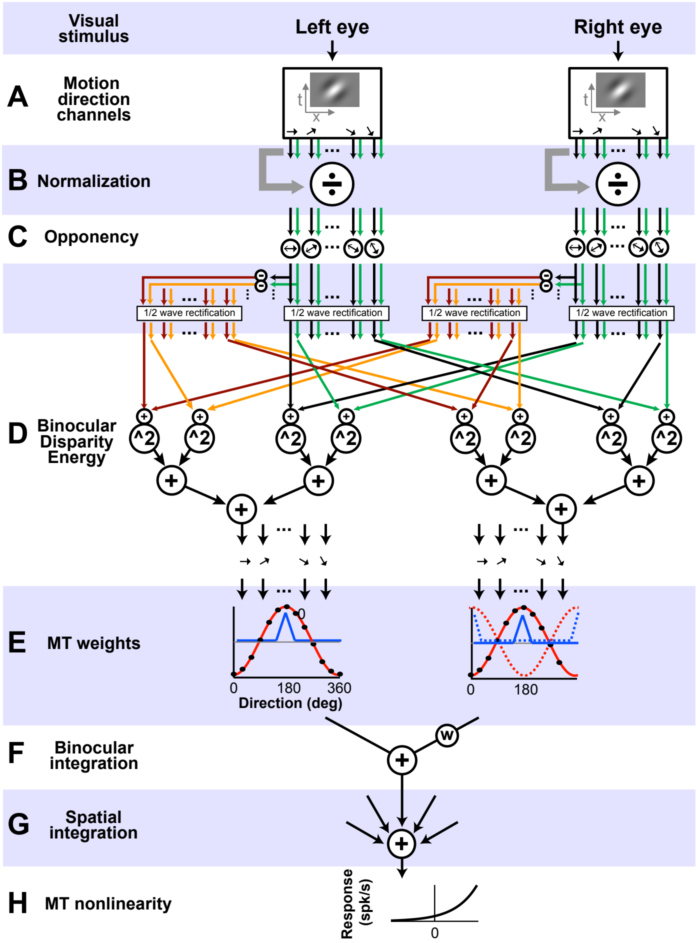

Figure 1. Our MT model begins with 3D (x,y,t) linear Gabor motion energy units at mutiple preferred ST and TFs. (A). It includes V1 normalization (B) and opponency (C), linear MT weights (E), spatial pooling (G) and a final nonlinearity (H). The MT weights can be set to create a “pattern” cell (red lines) or a “component” cell (blue lines). Binocular integration occurs in two stages: there is an (optional) V1 binocular integration stage that may include binocular disparity computation (1D) and binocular MT pooling (1F).

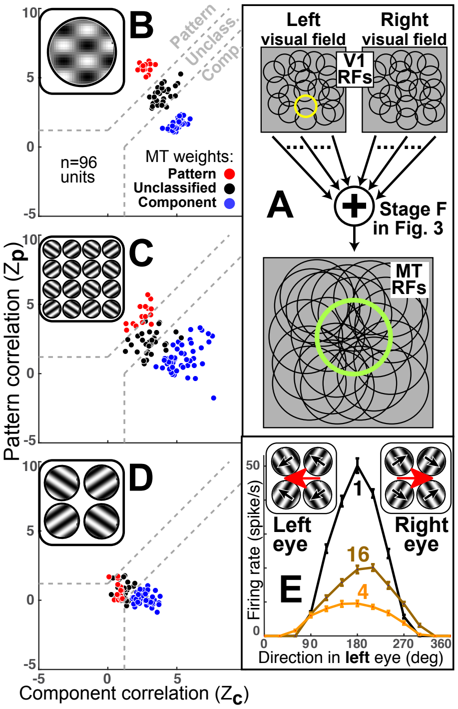

Figure 2. Probing spatial integration with pseudoplaid stimuli. **(A)** A large MT RF (green circle) is created by integrating V1 RFs at 20 locations in each eye (example marked in yellow). The MT population (large black circles) vary randomly in their choice of V1 inputs and their spatial location in the visual field. For simplicity, a cyclopean field is shown for MT, but MT units may have eye biases and different RF maps in the two eyes. **(B)** Each MT unit in our population model was tested with a plaid stimulus (icon) that filled its RF. For each unit (n=96) a point is plotted to show the pattern vs. component correlation. **(C,D)** The analysis in (B) was repeated with the stimulus grating components separated in space in a 4x4 and 2x2 grid, following work of Kumbhani et al. and Majaj et al. **(E)** We have used binocular variations of this stimulus with **1**, **4** and **16** patches to probe the spatial integration of 3D motion (icons show n=4). Note, stimulus spatial frequency not shown to scale.

References:

1. Baker and Bair (2016) J Neurosci 36:6563-82.
2. Majaj et al. (2007) J Neurosci 27:366-70.
3. Kumbhani et al. (2015) J Neurophys 113(7):1977-88.
4. Quaia et al. (2016) J Neurosci 36:3903-18.
5. Khawaja et al. (2013) J Neurophys 110:63-74.
6. Wallisch and Movshon (2011) SFN 577.16

Acknowledgments: We thank Christian Quaia and Bruce Cumming for useful scientific discussion and sharing unpublished results.

Funding: NSF CRCNS grant IIS-1309725