**Hysteresis in stereoscopic fusion: A phenomenon with a surprising amount of depth**

Majed Samad, Edoardo Balzani, Nathanael Larigaldie, Gabor Lengyel, Jean-Paul Noel

If each of our eyes sees a line and these are sufficiently close, we can fuse them into a 3D percept. How the brain solves the correspondence problem that results in a stereoscopically fused image has been the subject of intense investigation (Poggio & Poggio, 1984). Despite this research, the full dynamics of this process are still not fully understood. In a landmark study from the 1960s, it was observed that Panum's fusional area—the region of depth where fusion of stereo-presented images occurs—can be extended under some circumstances (Fender & Julesz, 1967). In particular, when the two images have been fused in the past, and if they are subsequently separated symmetrically, they are more likely to continue being fused. Across a series of experiments, Fender and Julesz carefully recorded the point at which the stereo images fused, measured as these stereoscopically presented lines moved laterally along the stereo displays either towards the temples or the nostrils. They found that the fusion point extends further when the lines started fused and is reduced if the lines started separated. They surmise this phenomenon indicates the existence of a cortical mechanism that allows stereoscopically fused images to remain fused based on the recent history of the stimulus.

Such a phenomenon suggests a cortical mechanism that uses statistical inference to allow fusion to be maintained even when lower-level mechanisms would not. This "cortical registration module" builds upon the work of the labeling and correspondence modules by using prior knowledge to help determining which parts of the images in the two eyes should be fused, in particular when retinal disparity cues cannot be used alone to solve this problem. We frame this as a causal inference problem. Causal inference is a principled statistical method for determining whether signals had a common cause in the environment or not. In this context, if the two retinal images start to move away from each other, they may represent a single line moving away from the screen in depth, or else represent the fact that they are two separate lines on the screen. As distance from the screen increases, the probability for common cause decreases.

We model this mechanism by making use of a switching Kalman filter (SKF), which is an extension of the traditional Kalman filter (Murphy, 1998). Specifically, the SKF considers a set of different linear models in parallel, infers the probability of being in any of them, and then takes a linear combination of them in accordance to that inference. In our implementation, we considered two causal structures for the world: **a)** stereoscopically presented lines are not fused and therefore move horizontally across the retinas, and **b)** that they are fused and therefore move in the depth dimension. The model maintains both representations in parallel, while constantly assigning a probability of being in one causal state or the other. Perceived depth and corresponding horizontal position of the lines on the screen are non-linearly related. This gives rise to different dynamics for movement in depth and the horizontal movement, which suggests that the switching Kalman filter would be well suited to model the situation.

First, we validated our implementation of the switching Kalman filter by using a damped-oscillator mechanical simulation and showed that it could reliably switch between damping and amplification regimes. We then simulated the experiment and represented the line's motion in the two reference frames (depth and horizontal) and used that as the input to the model. The results show that (1) the SKF model can track the changes in depth and the corresponding horizontal position of the lines with the appropriate causal structure model (**Fig. 2A, top and middle rows**), (2) the fusion point differs depending on the direction of motion of the line, and specifically on the history of the stimulus (**Fig. 2A, bottom row, also see the black vertical dotted lines in the top and middle rows**), and (3) the model can capture the empirical data from (Fender & Julesz, 1967) (**Fig. 2A, bottom row, dotted and solid lines**). Further, we tested the dependence of this behavior on a prior for continuity (i.e., the world is continuous and the latent causal structures do not change rapidly), which is implemented in the SKF as a transition probability for switches from one state to the other. We find that the larger this prior, the more robustly the model can capture the hysteresis phenomenon (**Fig. 2B**). We also tested the model's behavior under various levels of noise corrupting the stimulus inputs. We find that increasing the magnitude of the noise tends to decrease the hysteresis effect.

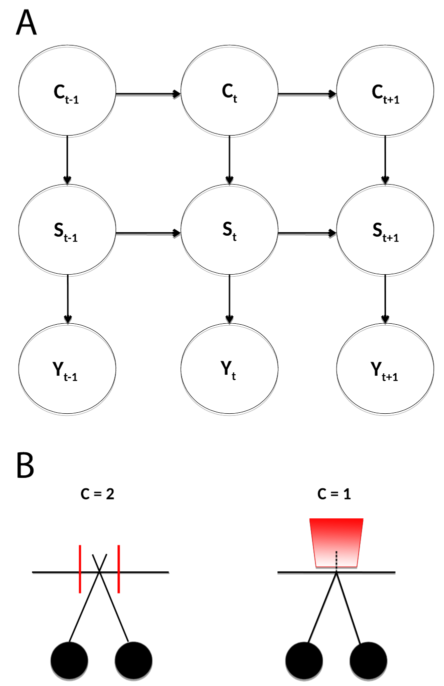
Our results suggest that while our brain represents multiple hidden states of the external world at the same time, vision is attuned to the statistics of the environment favoring continuity which enables us to maintain a coherent visual percept and avoid rapid switches between hidden states.

## References

Fender, D., & Julesz, B. (1967). Extension of panum's fusional area in binocularly stabilized vision. *Journal of the Optical Society of America*, *57*(6), 819–830.

Murphy, K. P. (1998). Switching kalman filters.

Poggio, G. F., & Poggio, T. (1984). The analysis of stereopsis. *Annual review of neuroscience*, *7*(1), 379–412.




C

1. Switch (C)

$$C \sim Bernoulli(p)$$

2. States (S)

Hidden cause 1 ($C = 1$)

$$\hat{S}_t^1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & \triangle t \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} disparity \\ depth \\ \triangle disparity \\ \triangle depth \end{bmatrix} + \mathcal{N}(0, Q_1)$$

Hidden cause 2 ($C = 2$)

$$\hat{S}_t^2 = \begin{bmatrix} 1 & 0 & \triangle t & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} * \begin{bmatrix} disparity \\ depth \\ \triangle disparity \\ \triangle depth \end{bmatrix} + \mathcal{N}(0, Q_2)$$

3. Observations (Y)

$$Y_t = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} * S_t^{1(2)} + \mathcal{N}(0, R)$$

Figure 1: **A.** The generative model of the switching Kalman filter (SKF). **B.** The two hidden causal structures in the SKF model. (C=1) The lines are fused moving in depth. (C=2) The horizontal position of the lines are changing. **C.** The linear dynamical system. *1.* The discrete switch variable $C_t$ which specifies which matrix to use at time $t$ in the transition model. *2.* The transition models for the two hidden causal structures: (1) Depth is changing and disparity is fixed (C=1), or (2) Disparity is changing and depth is fixed. *3.* Observation model. The two inputs of the model are changing in depth and in disparity.

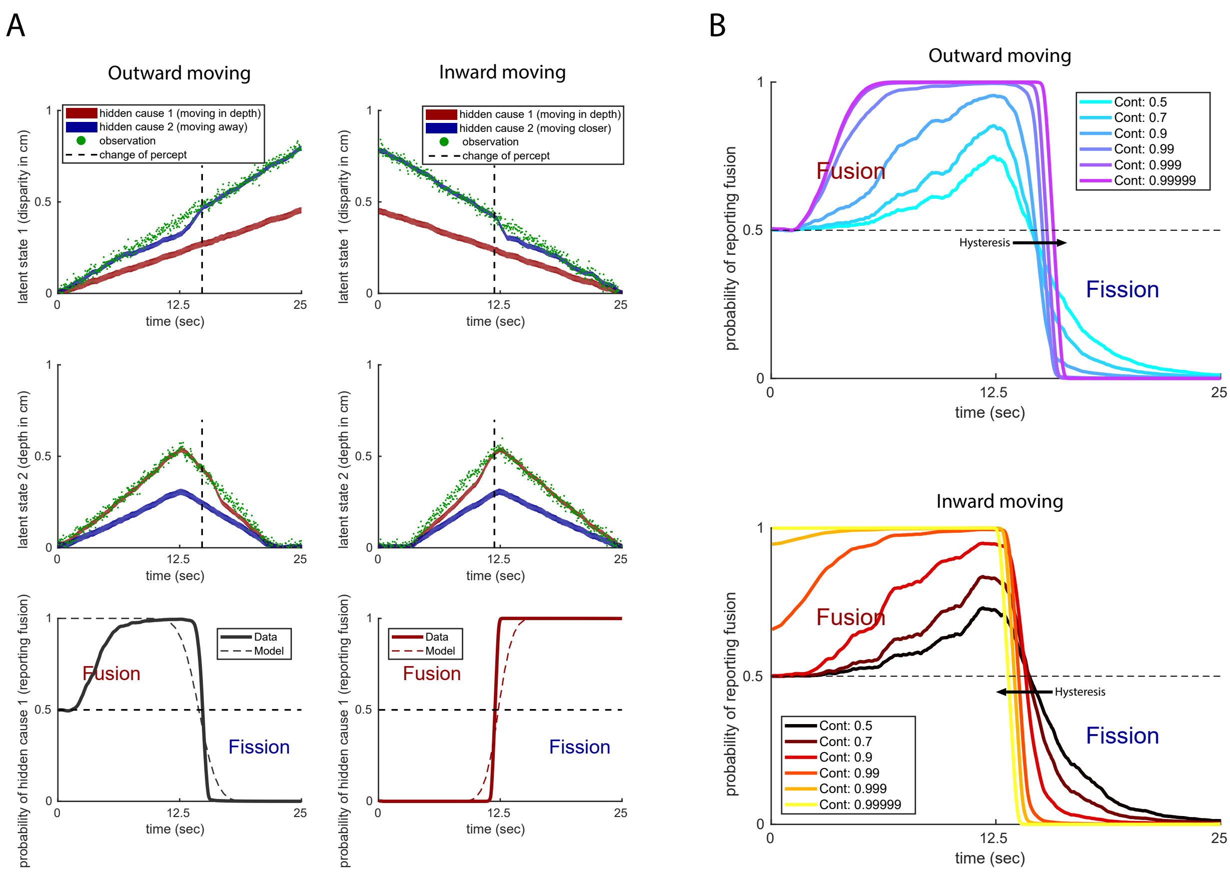



Figure 2: **A.** *Top:* Shows the latent state for disparity ($y$-axis) as a function of time ($x$-axis). *Middle:* Shows the latent state for depth ($y$-axis) as a function of time ($x$-axis). Green dots show the observations (the horizontal movement of the lines in the top row, and movement in depth in the middle row). Red shaded area represents the standard deviation of the model's predictions in case of the first hidden causal model (moving in depth). Blue shaded area represents the standard deviation of the model's predictions in case of the second hidden causal model (changing disparity). Black vertical dotted lines show the points at which the lines fuse or separate. *Bottom:* $Y$-axis shows the inferred probability of the hidden causal models ($P(C = 1)$) as a function of time. Dotted lines represent the empirical data from (Fender & Julesz, 1967) re-plotted. Solid lines show the model's prediction. Black horizontal dotted lines show the 0.5 probability that represents the change point between fusion and fission in perception. **B.** Shows that the hysteresis effect increases as the prior probability on continuity (i.e. staying in the previous hidden cause model) increases.