IJCNN2008 – WCCI2008
Hong Kong, June 1-6, 2008

Detection of Propagating Phase Gradients in EEG Signals using Model Field Theory of Non-Gaussian Mixtures

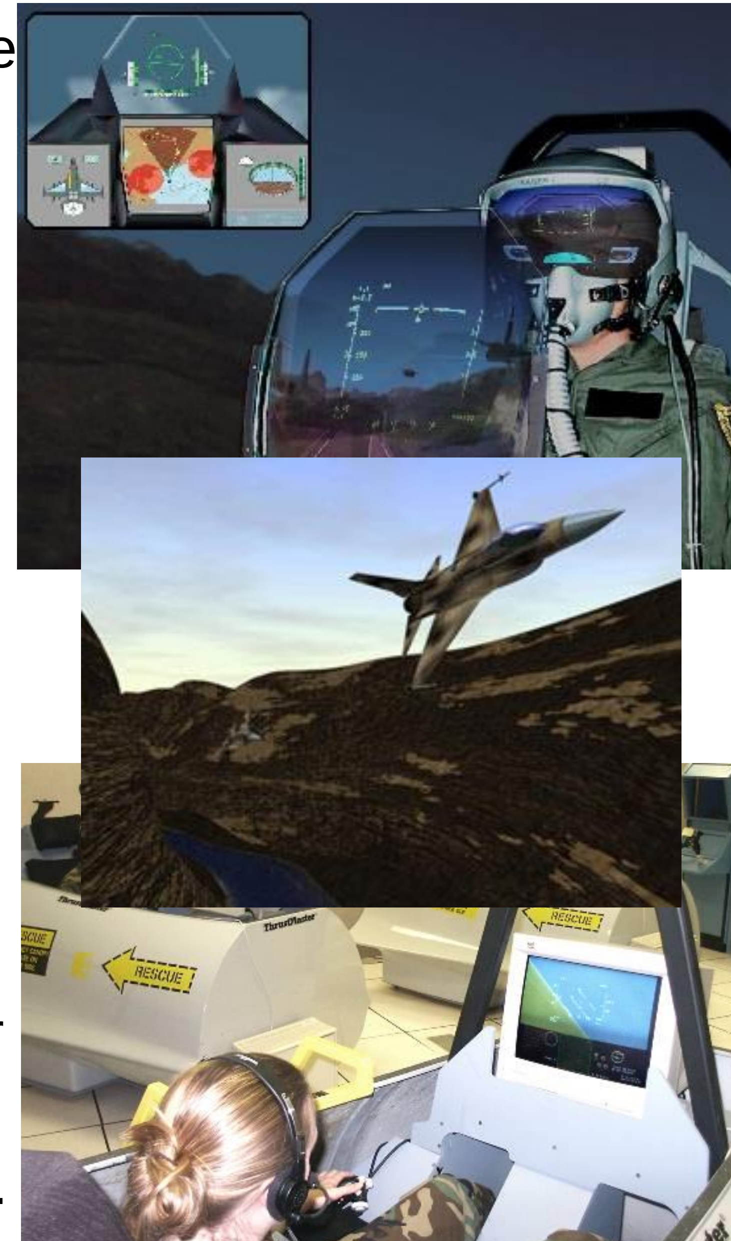
Robert Kozma(1,2), Leonid Perlovsky(1), Jaisantosh Ankishetty(2)

(1) Air Force Research Laboratory, HAFB, MA
(2) Computational Neurodynamics Lab, University of Memphis, TN, USA

AFRL Human Effectiveness

- Human effectiveness strategic directions are focused on
  - helping the information warrior think, decide, and act in new ways,
  - reducing decision times,
  - improving decision quality through improved human-system interfaces and processes,
  - protecting all airmen in all offensive and defensive environments.

- Some Relevant Tasks:
  - Enable warfighters to train as they fight by advancing education and training technologies and methods to provide required mission competencies for the expeditionary aerospace force.
  - Enable improved decision effectiveness for all warfighters by advancing cognitive modeling science, task critical information portrayal, and decision support technology.

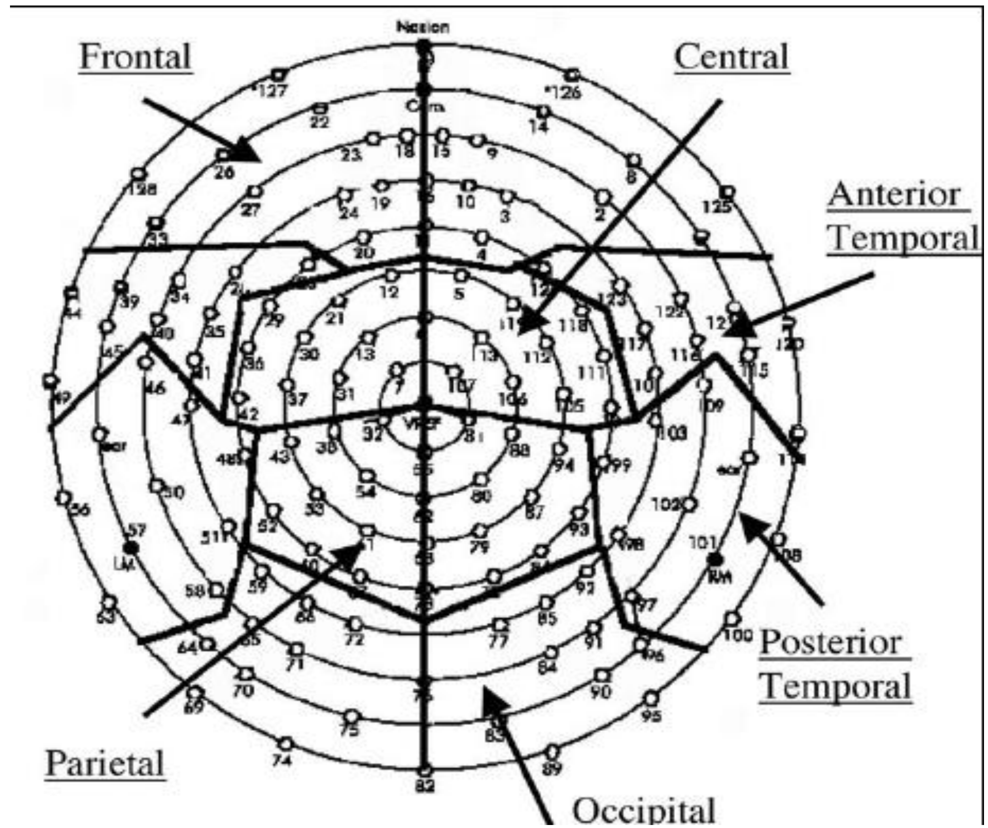


Cognitive Modeling & Experiments

- Understanding human cognitive processing
  - Perception-action cycle
  - Multi-sensory modalities
- Experimental Techniques
  - Task critical information displays
  - Decision support systems
- Human-computer & Brain-computer interfaces
  - Scalp EEG arrays
  - Challenges in interpreting and modeling multi-channel recordings
  - High noise, high clutter problems

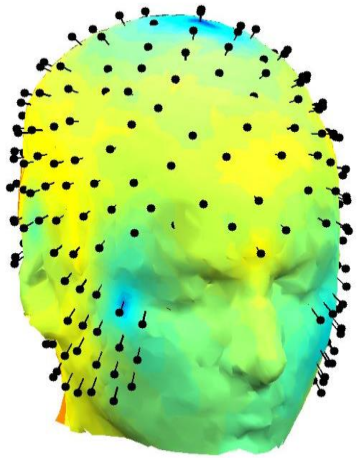
Noninvasive Scalp EEG

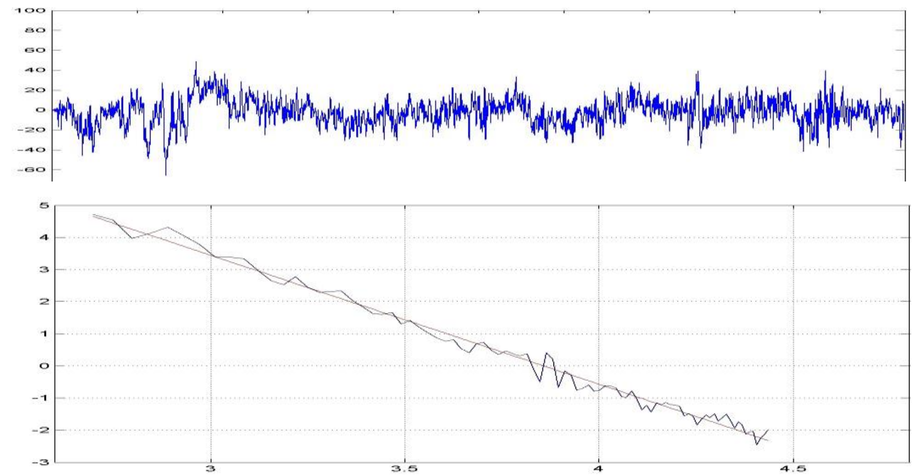
- **Scalp EEG**
  - Non-intrusive, but very difficult to look beyond the skull

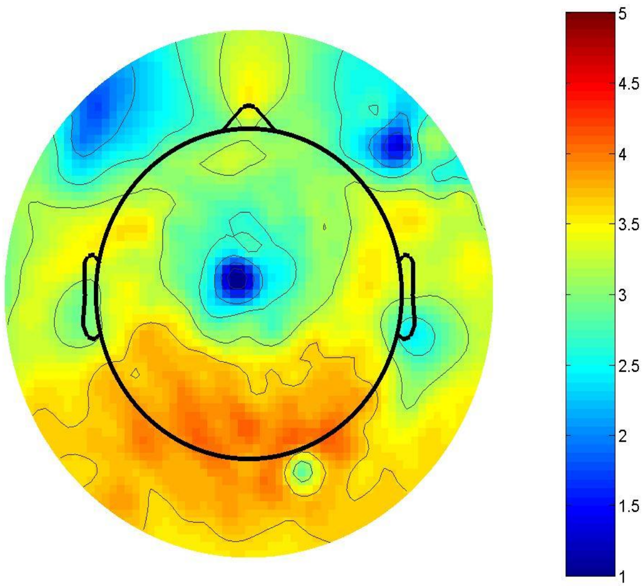


Electrode placement display featuring areas of the brain (EGI, OR)

Cognitive Features of EEG







- EEG data example – Estimating PSD slope
- Correlation with cognitive state
  - Cognitive Activity
  - Resting
  - Sleep
  - Medical conditions

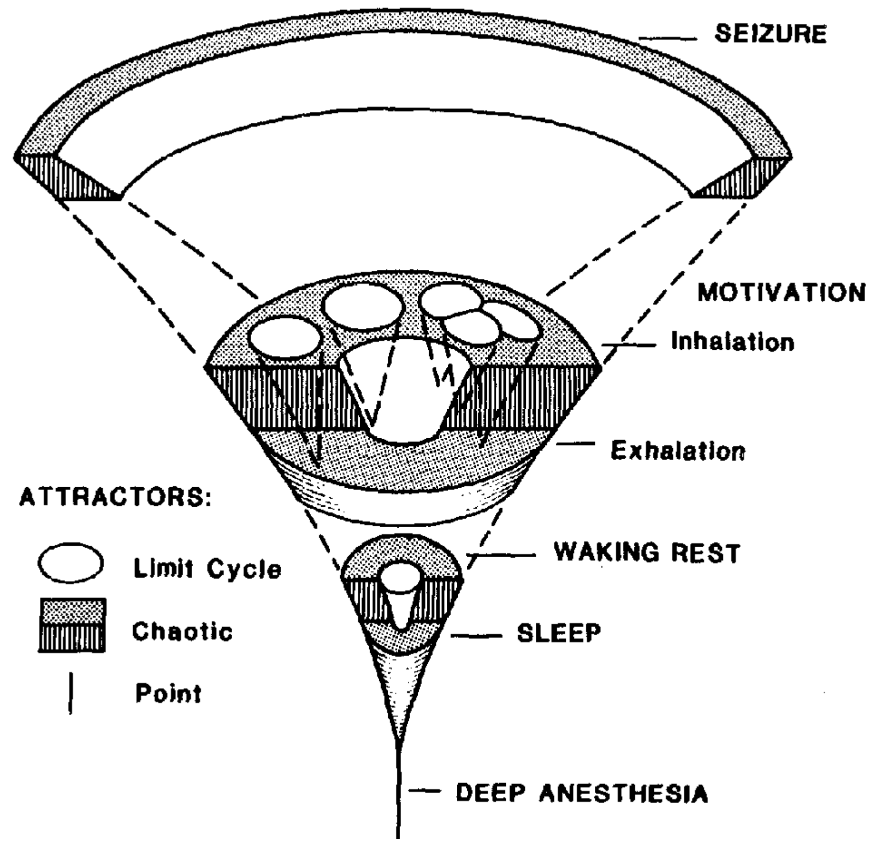
Dynamic Systems Phenomenology of Cortical Neurodynamics (Freeman)

The system maintains a state space dominated by a high-dimensional, flexible, evolving attractor landscape.

Input is by waves at the mesoscopic level from cortices that overlap but need not be synchronous.

Operation is by global phase transitions induced aperiodically by spatial integration of the wave packets. The transitions lead to hemisphere-wide spatial amplitude modulation patterns.

Output is by the spatio-temporal integration at subcortical targets simultaneously of the covariant fraction of the total variance of hemispheric neural activity.



Relevant Concepts in Physics and Chaos

- Scale free networks (Barabasi, Albert)
- Small worlds (Watts, Strogatz)
- Chaotic Intinerancy (Kaneko, Tsuda)
- Metastability (Kelso, Haken)
- Frustrated Chaos (Bersini)

Phase Cones in EEG

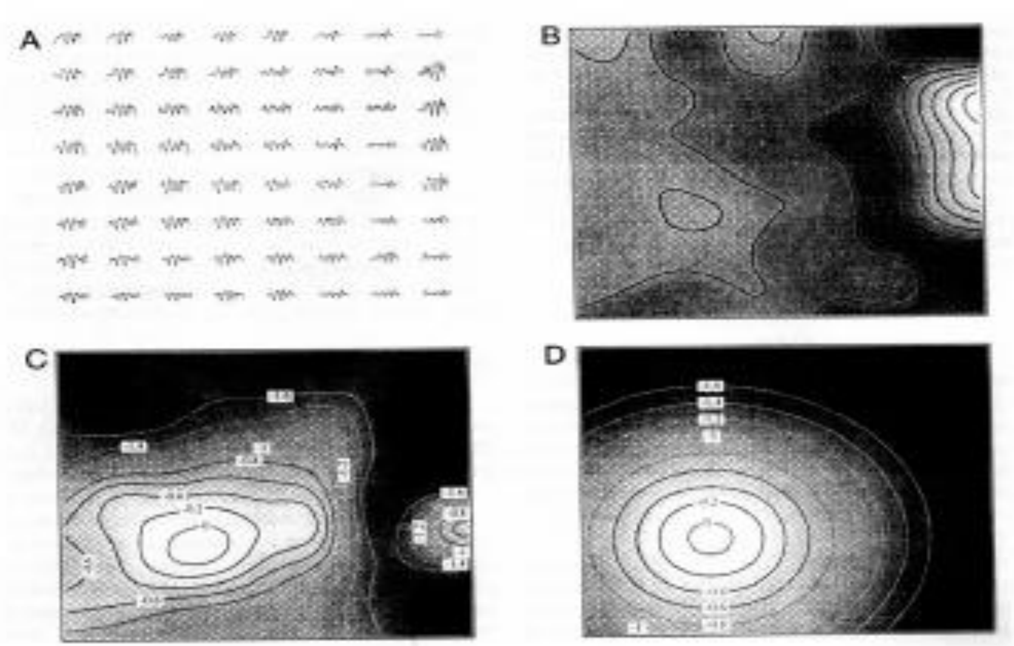


Fig. 2. Illustration of the onset of phase cones in the rabbit olfactory bulb, using an array of 8x8 intracranial electrodes. Schematic illustration of the observed amplitude modulation patterns are shown on panel A. Panels B, C, and D show the process of identification of the phase cones during the experiments (Barrie et al, 1996)

Amplitude modulation in rabbit EEG

Traditional approach

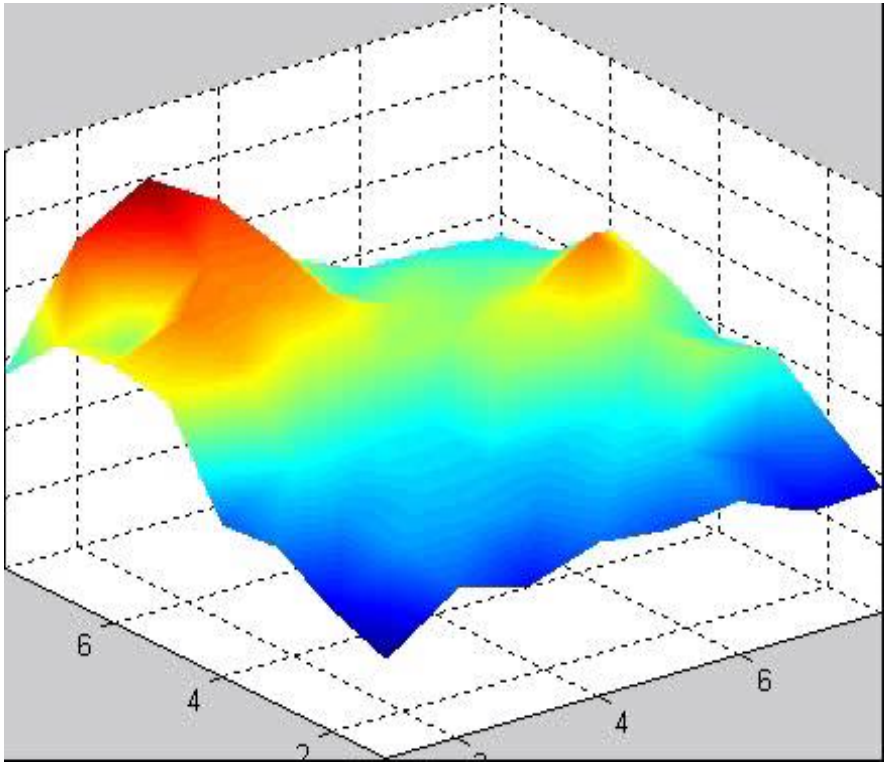
- Numeric fit of dominant phase cone (Freeman et al, 1996+)
- Major properties revealed but tedious

Dynamic Logic

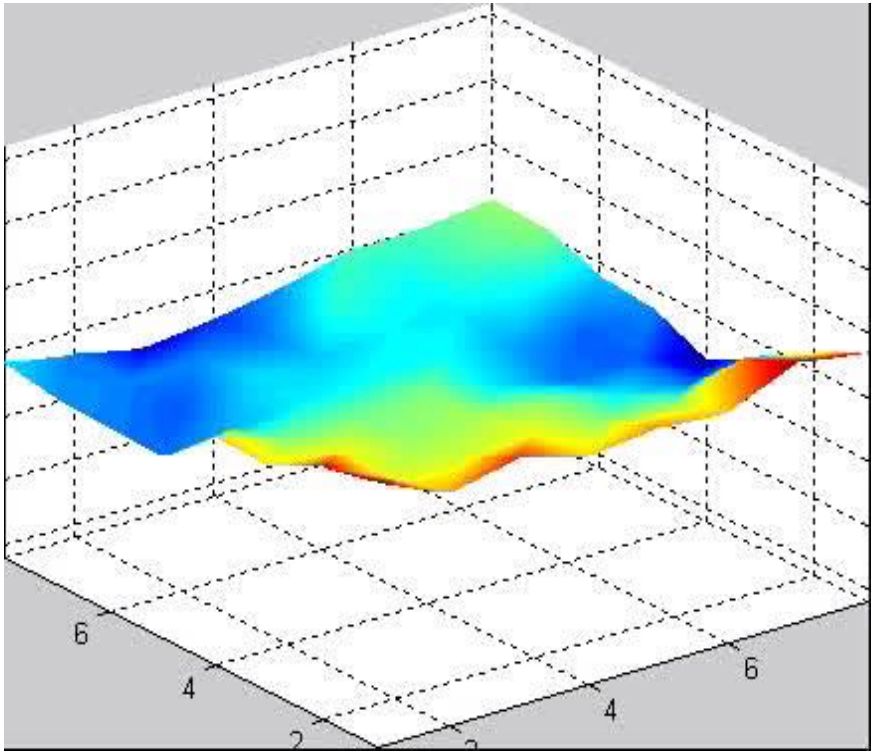
- A potential robust method of detecting multiple phase cones
- Present work

Rabbit Intracranial EEG Data

Visual EEG Analytic Amplitude

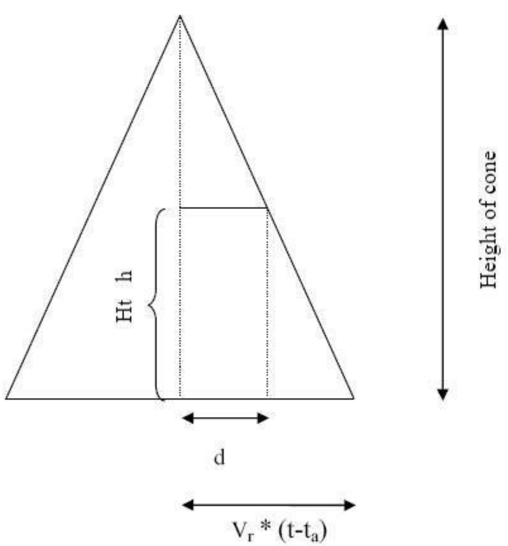


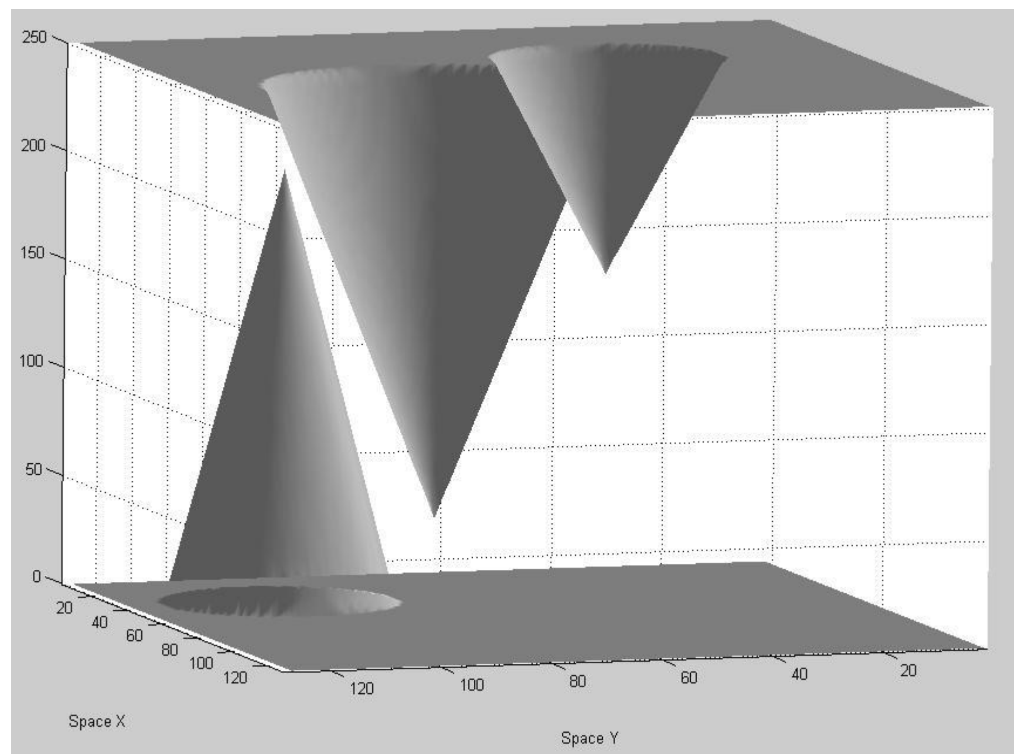
Visual EEG Analytic Phase



Model of EEG Phase Cones

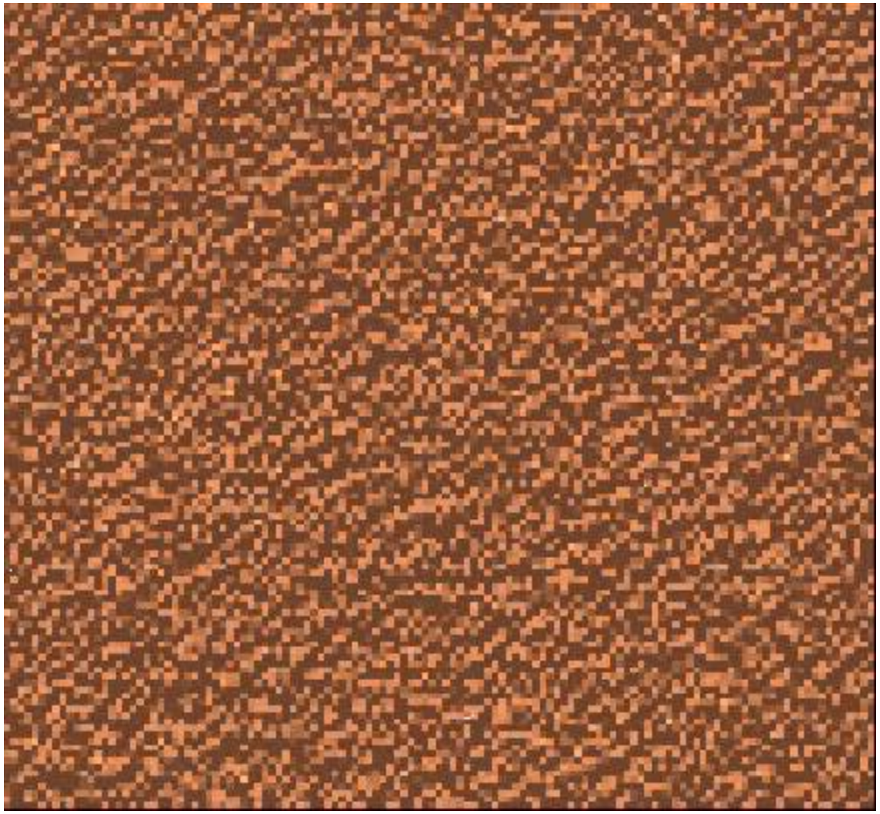
- **Multiple phase cones simulated**
  - Simultaneously coexisting/overlapping cones
  - Increasing/decreasing in time
  - Positive and negative slopes
  - Explosions and implosions
- **Previous work:** Gaussian mixture → **Now:** non-Gaussian



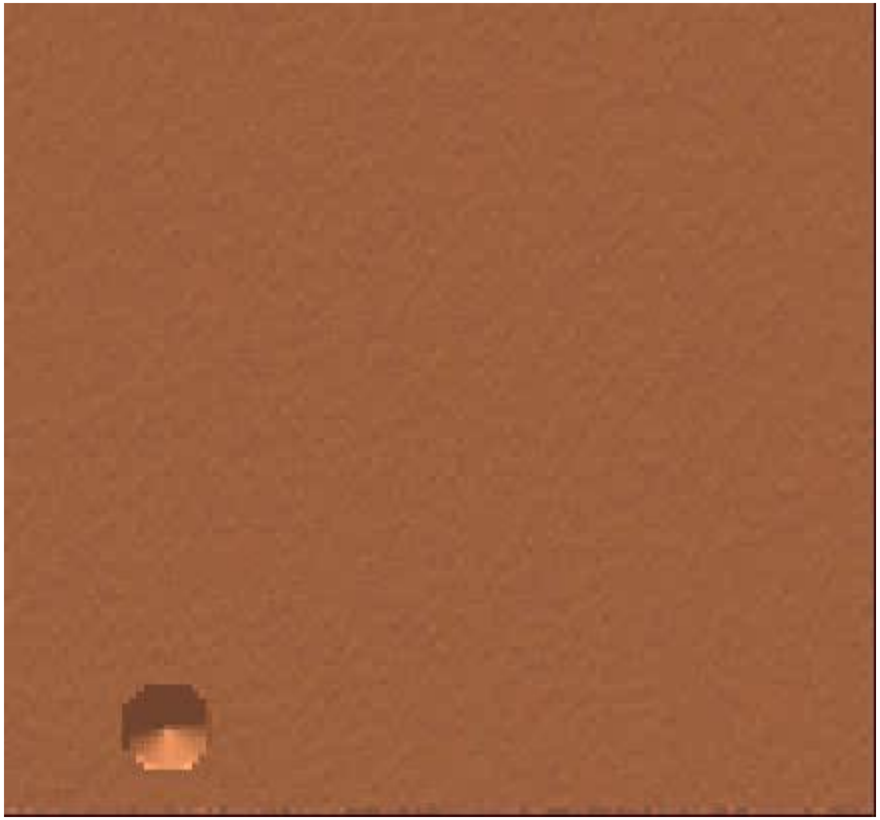


Simulated Phase Cone Data

Noisy Data Simulation



Cones without Noise



MODEL FIELD THEORY

- **Perlovsky in past 20+ years**
  - Orders of magnitude improvement in difficult pattern recognition problems with high level of noise and clutter
- **Signals and Concept-Models**
  - signals **X**(n), n = 1,…,N
  - parameters $\Theta=[\Theta_1, \Theta_2, ..., \Theta_s]$,
  - Consider k models, k =1 clutter (noise), the rest
  - $P(X|\ \Theta_k)$, k=2,…, K , components with their own parameters
- **Improve object-models while understanding signals**
  - associate samples *n* with models *k* and find parameters $\boldsymbol{\Theta}_k$
  - Model complexity increases with understanding data and avoids combinatorial explosion
- **Learning instinct**
  - Maximize similarity between signals and models
  - L: likelihood of similarity
  - LL: log-likelihood

Top-down and bottom-up dynamics in visual recognition

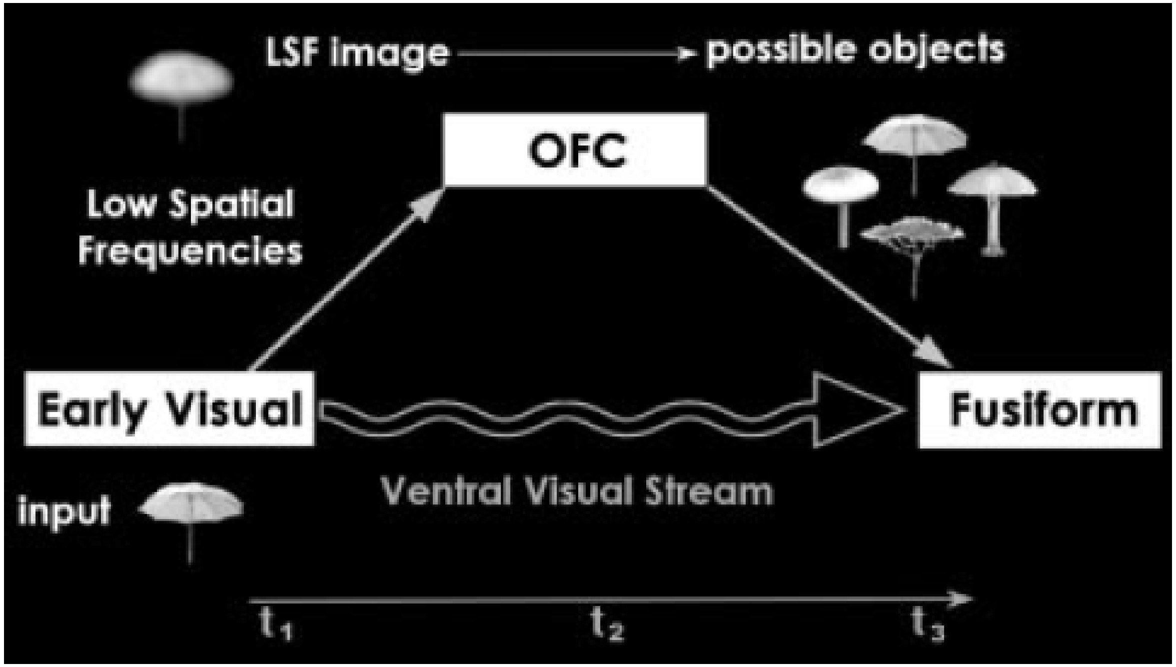


Fig. 1. An illustration of the proposed model. A LSF representation of the input image is projected rapidly, possibly via the dorsal magnocellular pathway, from early visual cortex to the OFC, in parallel to the systematic and relatively slower propagation of information along the ventral visual pathway. This coarse representation is sufficient for activating a minimal set of the most probable interpretations of the input, which are then integrated with the bottom-up stream of analysis to facilitate recognition.

Bar et al. PNAS, 2006.

MODEL FIELD THEORY

- Log-likelihood or mutual relevance:

$$LL_E = \sum_{n \in [N,T]} p_0(X_n) ln \sum_{k=1}^{K} r_k(t) p(X_n|\Theta_k).$$

  - $p(X_n|\Theta_k)$ – pdf of kth mixture component
  - $p_0(X_n)$ – distribution of data points
  - $r_k(t)$ – relative weight of component k
- Optimize parameters →maximize LL
- Normalization of total energy, introduce Lagrange multipliers and define association probabilities

$$\sum_{k=1}^{K} \sum_{t=1}^{T} r_k(t) = E \qquad P(k|n) = \frac{r_k(t)p(X_n|\Theta_k)}{\sum_{k=1}^{K} r_k(t)p(X_n|\Theta_k)}.$$

- Max LL

$$\frac{\partial F}{\partial \Theta_{k,i}} = -\sum_{T}\sum_{N} P(k|n)\{\frac{\partial ln(r_k(t))}{\partial \Theta_{k,i}} + \frac{\partial ln(p(X_n|\Theta_k))}{\partial \Theta_{k,i}}\} - \lambda \sum_{T} \frac{\partial r_k(t)}{\partial \Theta_{k,i}} = 0$$

Optimize MFT Parameters:
Component power is independent of parameters

The probability distribution function $p(X_n|\Theta_k)$ for component k is assumed of the following form:

$$p(X_n|\Theta_k) = \frac{1}{\sqrt{(2\pi)^3|C_k|}} e^{\{-\frac{1}{2}|X_n - M_n^k|^T C_k^{-1}|X_n - M_n^k|\}}. \qquad (21)$$

Here X_n denotes the experimental (measured) data points in space and time, and M_n^k is the model output for component k. C_k is the covariance matrix. Eq.(21) incorporates the assumption that the difference between data and model, i.e., the error of the model approximation is normally distributed. Based on the Gaussian approximation on the error, Eq. (20) can be simplified. It concerns the derivative of a quadratic form, which is evaluated using the corresponding matrix identity [4]. Substituting the derivative of the quadratic form, we obtain:

$$\sum_T \sum_N P(k|n) C_k^{-1}(X_n - M_n^k) \frac{\partial M_n^k}{\partial \Theta_{k,i}} = 0. \qquad (22)$$

Optimize MFT Parameters:
Component power DOES depend on parameters

- General form of LL equation

$$\sum_T \sum_N P(k|n)\{C_k^{-1}(X_n - M_n^k)\frac{\partial M_n^k}{\partial \Theta_{k,i}} + \frac{\partial ln(r_k(t))}{\partial \Theta_{k,i}}\} = -\lambda \sum_T \frac{\partial r_k(t)}{\partial \Theta_{k,i}}.$$

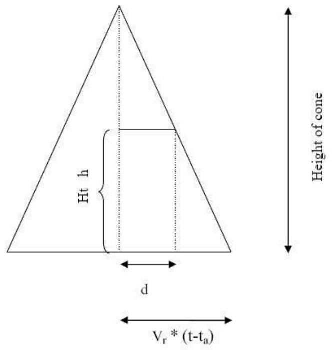
- Here Lagrange parameter is given by

$$\lambda = -\sum_{N,T} P(k|n) \frac{\frac{\partial ln(r_k(t))}{\partial \Theta_{k,i}} + \frac{\partial ln(p(X_n|\Theta_k))}{\partial \Theta_{k,i}}}{\sum_K \partial r_k(t)/\partial \Theta_{k,i}}.$$

Solution of MFT equation: Non-Gaussian Cones

- Model equations for components
  - Cone evolution equation (cylindrical)

$$M_x^k = v_b t - \frac{v_b}{v_r}\sqrt{(\xi - \xi_A)^2 + (\eta - \eta_A)^2}.$$



  - Update of apex spatial coordinate

$$\xi_A = \frac{< (X_n - M_n^k)\xi >_{k,c}}{< (X_n - M_n^k) >_{k,c}}. \qquad \eta_A = \frac{< (X_n - M_n^k)\eta >_{k,c}}{< (X_n - M_n^k) >_{k,c}}.$$

$$< * >_{k,c} = \sum_T \sum_N P(k|n) \frac{*}{\sqrt{(\xi - \xi_A)^2 + (\eta - \eta_A)^2}}.$$

  - Update of time initiation

$$< \frac{1}{t - t_A} >_k = \frac{v_B C_k^{-1}}{3} < (x_n - M_n^k) >_k$$
$$- \lambda D \tau_{max}^3.$$

MFT Algorithm

- **Start with a set of signals and unknown models**
  - Parameter values Θ_m
- **Improve parameter estimation**
  - Iteratively learn the parameters
  - Learn signal-contents of objects
- **Continue iterations**
  - Similarity of model and data increases on each iteration
  - Until convergence

Conclusions

- Initial results promising
  - Good estimation of parameters, modify energy function
- Future of applying MFT
  - Non-intrusive cognitive monitoring
  - Novel and until now unreachable & unseen details of cognitive processing
- MFT is cognitively motivated:
  - brains likely exhibit iterative learning a formulated in MFT
  - EEG analyzed by MFT algorithm can identify such behavior