Multi-sensory cue integration with reliability encoding, using Line Attractor Dynamics, searching for optimality

Mohsen Firouzi [1,2,3]; Cristian Axenie [1]; Jörg Conradt [1,2,3]

[1] Neuroscientific System Theory, Technische Universität München;
[2] Bernstein Center for Computational Neuroscience, Munich;
[3] Graduate School of Systemic Neurosciences, Ludwig-Maximilians- Universität München;

Introduction

A key requirement for any systems, including biological or man-made systems is their capability to estimate physical properties of the real world through partially reliable observations to interact properly with their environment. Apart from intrinsic variability of neural activity in the brain, accessible sensory cues are often uncertain and ambiguous. The human brain can combine these noisy and partially reliable pieces of information to optimally estimate the state of the world and gives us a coherent representation of the environment in order to efficiently handle cognitive tasks [1]. Psychophysical experiments show that human performs Bayesian model averaging over sensory observations based on relative variability of stimuli [1] [2]; but underlying neural mechanisms to encode and internalize the relative reliabilities of cues is still addressed as a central issue [2].

Neural activity modulation like Gain Field Modulation, is a well-known mechanism in brain to highlight information under specific internal or external constraints e.g. attention-based modulation of striate cortex by higher cortical feedbacks [3]. In this work in a sensory convergence problem and using an attractor network, we have investigated if modulating neural activity according to relative reliability of connected cues can bias the dynamics of the network in favor of more reliable cue, and consequently integrate cues in an optimal fashion. We have evaluated our methodology in a three-modal heading estimation experiment using an omni-directional mobile robot [4]. We have compared the outcome of the network with Maximum-Likelihood- Estimator (MLE) and it is shown that the network can realize a near optimal solution for reliability based multi-sensory cue integration.

Network Architecture

❖ The Attractor Network consists of three input populations each for single input cue, and three intermediate populations of multi-modal neurons (**Fig.1**).
❖ Input cues are encoded by the activity of input populations through overlapping Gaussian tuning curves.
❖ All Input neurons are reciprocally connected to intermediate neurons using von-Mises function (**Fig.1**).
❖ Modulating input populations strongly effect the dynamics of the network (**Fig.2**).

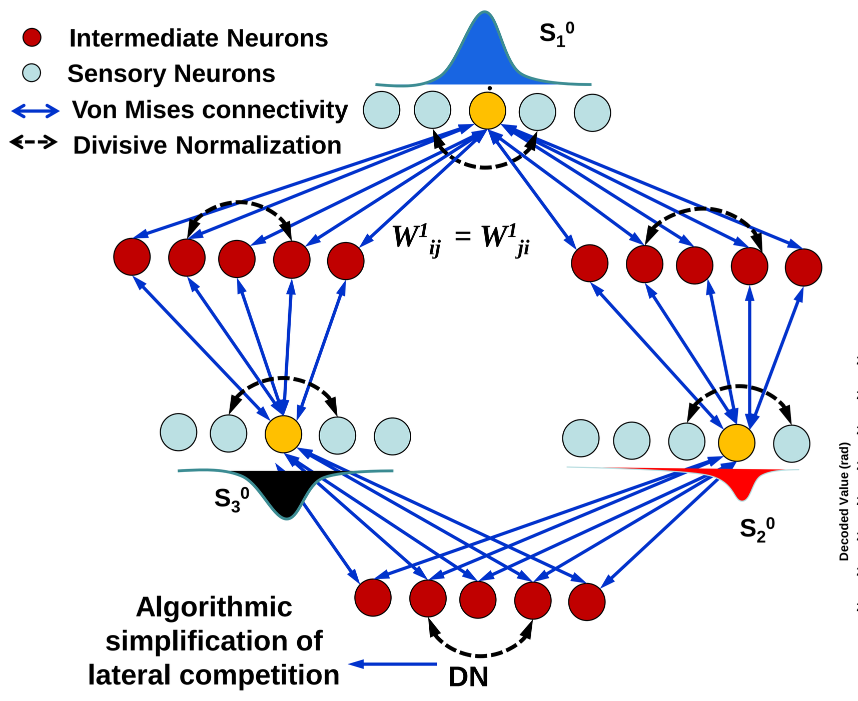

Fig.1 General architecture of Attractor Network in a sensory convergence scenario

$$W_{ij}^{k} = e^{\frac{\cos[2\pi/N](i-j)-1}{\sigma_1^2}}$$

von Mises Function
N : number of neurons
σ : width of projection

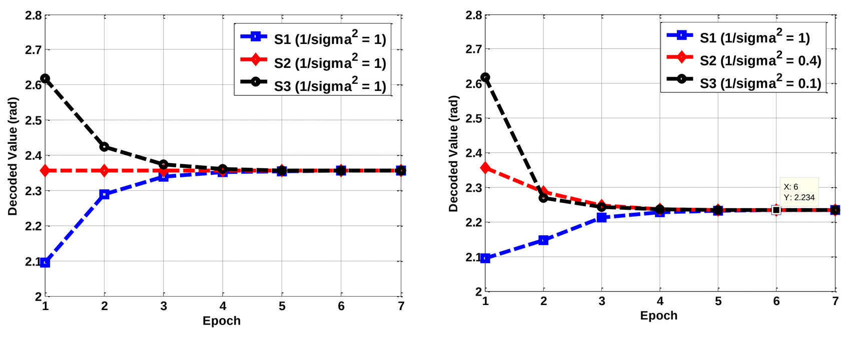

Fig.2 Dynamics of encoded cues, Left: with equal variability, Right: with different variability

Encoding and Network Dynamics

$$P(r_i \mid x) = \frac{[\Phi_i(\kappa,x)]^{r_i}}{(r_i)!} e^{-\Phi_i(\kappa,x)}$$ → *Poisson variability of i^{th} Neurons activity (r_i) in response to x stimulus*

$$\Phi_i(\kappa,x) = \kappa e^{-\frac{|x-x_i^c|}{2\sigma^2}} + v$$ → *i^{th} Neuron tuning curve function centered at x^i_c (Neuron Selectivity and mean value of Poisson variability)*

- ○ ***x:*** Input variable (stimulus) subjected to encoding by population vector *R*
- ○ $\boldsymbol{r_i}$**:** Spike number per sec (firing rate) for *i* th neuron.
- ○ ***v:*** Spontaneous activity rate.
- ○ ***k:*** Intensity of activity (stimulus strength)
- ○ $\boldsymbol{x^c_i}$**:** Preferred value for *i* th Neuron of Population Vector of *R*

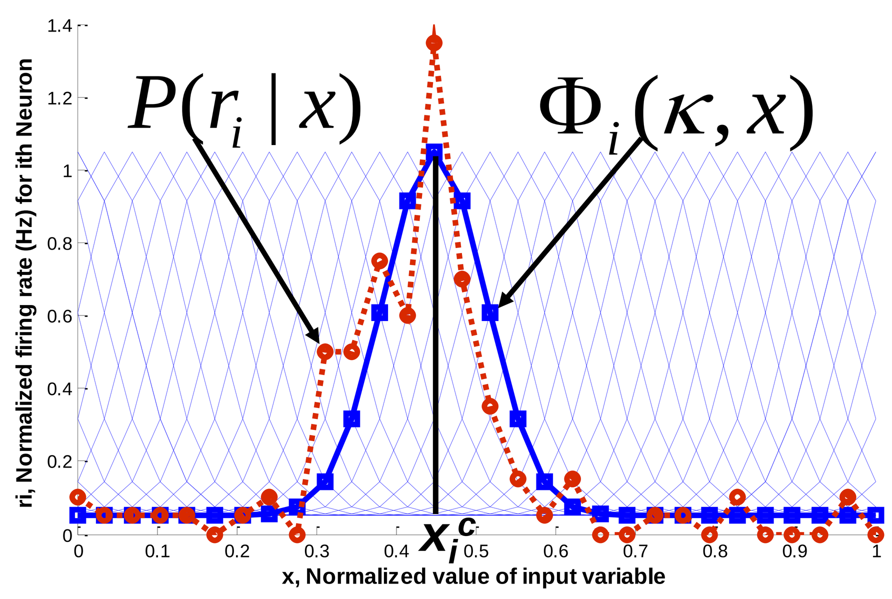

Fig.3 Red: The activity of i^{th} neuron in response to stimuli x, Blue: i^{th} neuron tuning curve centered at x_i^c

$$l_j(t) = \sum_k W_{kj}^1 r_k^1(t) + \sum_k W_{kj}^2 r_k^2(t) + \sum_k W_{kj}^3 r_k^3(t)$$

$$A_j(t+1) = \frac{(l_j)^\alpha}{\beta + \mu \sum_j (l_j)^\alpha}$$

$$r_i^{n\{=1,2,3\}}(t+1) = \frac{[\sum_j W_{ij} A_i(t+1)]^\alpha}{\beta + \mu \sum_k [\sum_j W_{kj} A_j(t+1)]^\alpha}$$

- ○ ***α:*** Divisive Normalization Power
- ○ ***β*:** Divisive Normalization Bias
- ○ $\mathbf{A_j}$**:** Activity of j^{th} neuron of intermediate layer
- ○ $\boldsymbol{W^n_{ij}}$**:** Synaptic weight between *i* th input neuron of *n* th input population and *j* th intermediate neuron

Multi-sensory Heading Estimation Experiment

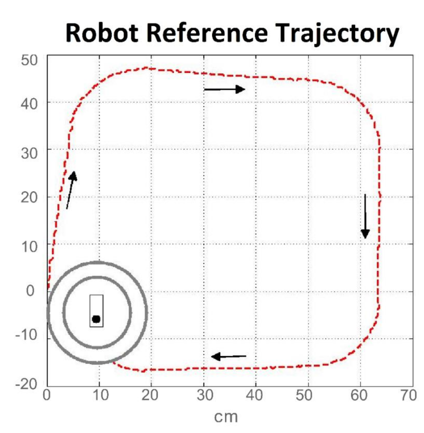

Fig.4 The reference trajectory of the omni-directional mobile Robot

Real time reliability map construction

$$\mu_k^{t+1} = \mu_k^t + \frac{S_k^{t+1}}{n} - \frac{S_k^{t-n}}{n}$$

$$z_{1,2,..,n}^{t+1} = S_{1,2,..,n}^{t+1} - \mu_{1,2,..,n}^{t+1}$$

$$\sigma^2(t+1) = \frac{n\sum_{k=1}^{n}(z_k^{t+1})^2 + (\sum_{k=1}^{n} z_k^{t+1})(\sum_{j=1}^{n} z_j^{t+1})}{n(n-1)}$$

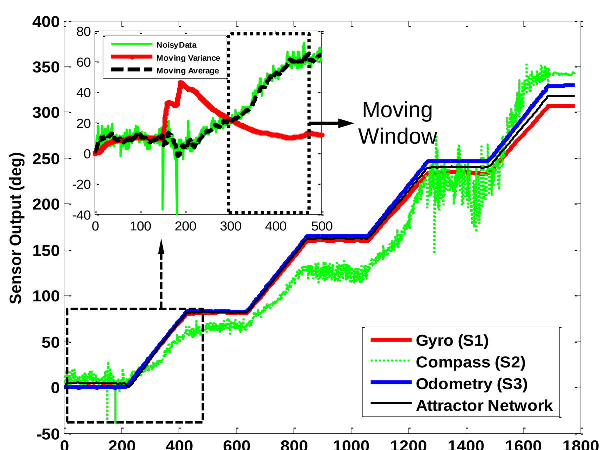

Fig.5 The observed Heading angel for each sensors (Gyroscope; Odometer; Compass) during experiment

$$S_k \sim N(\mu=\theta,\sigma_k); \quad MLE = \arg\{\max P(\theta \mid s_1,s_2,s_3)\} = \frac{\sum_i \frac{1}{\sigma_i^2} s_i}{\sum_i \frac{1}{\sigma_i^2}}$$

Normal Distribution

θ: Heading Angel,
$\mathbf{S_k}$ **: $\mathbf{k^{th}}$ sensory observation**

Assume external variability is governed by an Independent Gaussian Process

❖ MLE is chosen as ground truth (optimal estimator).
❖ Sensors are assumed to be governed by an Independent Gaussian process.
❖ Input populations are modulated (Initialized) according to momentary reliability map (**Fig.6-Bottom**).
❖ When network 's dynamics becomes stable, then relaxed activity is decoded as estimated value.
❖ The absolute error of MLE, Attractor Network and voting algorithm [5] are compared (**Fig.6-Top**).

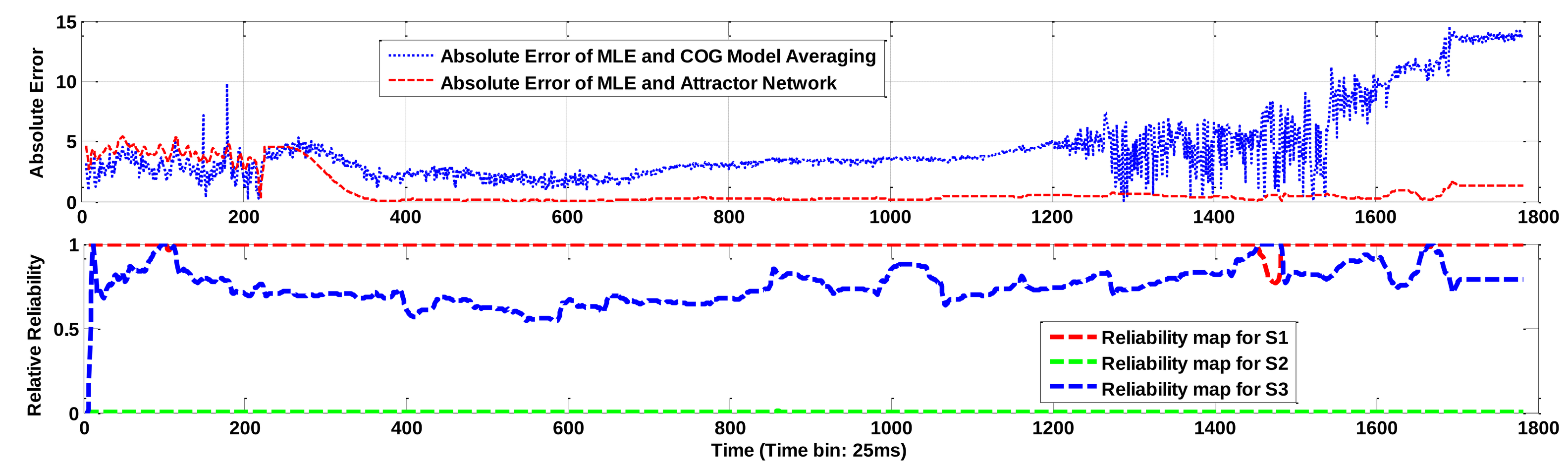

Fig.6 Top: Absolute error between MLE and modulated Attractor Network vs. COG vote-based integration [5]; Bottom: Constructed real-time relative reliability map

Conclusion and Remarks

❖ From Psychological experiments it is clear today that human observer performs a near-optimal Bayesian model averaging over partially reliable noisy sensory observations to achieve an estimation of the observed cues [2] .
❖ In this work we have investigated how possibly modulating neural activity in an Attractor Network can describe reliability-based weighted cuing. The Attractor Network consists of three input populations each encoding single cue and three intermediate populations conducting the Dynamics of the network (**Fig.1**).
❖ The outcome of proposed methodology is compared with MLE estimator and vote-based integration algorithm [5] and it is demonstrated that the network can integrate input cues in near optimal fashion if it is modulated according to relative reliability of connected cues (**Fig.6**).

Selected publications:

[1] M. O. Ernst; H. H. Bülthoff, “*Merging the senses into a robust percept*”, Trends on Cognitive Science, Vol.8, pp 162-168, April 2004.
[2] T. H. Weisswang; C. A. Rothkopf; T. Rodemann; J. Triesch, “*Bayesian Cue Integration as a Developmental Outcome of Reward Mediated Learning* ”, PLOS one, July 2011.
[3] J. W. Bisley, *“The neural basis of visual attention”,* The Journal of Physiology, 589, 49-57, January 1, 2011
[4] M. Firouzi; S. Glasauer; J. Conradt, “*Flexible Cue Integration by Line Attraction Dynamics and Divisive Normalization*”, 24th International Conf. of Artificial Neural Network, Hamburg, Germany, Sept 2014
[5] J. Triesch; C. Von der Malsburg, “*Democratic Integration: Self-Organized Integration of Adaptive Cues*”, Neural Computation, Vo.13 Issue.9, pp 2049 – 207, 2001

TECHNISCHE UNIVERSITÄT MÜNCHEN

INSTITUTE OF AUTOMATIC CONTROL ENGINEERING

Research Group: Neuroscientific System Theory (NST)