Cognitive Dynamic Systems

Simon Haykin
Cognitive Systems Laboratory
McMaster, University
Hamilton, Ontario, Canada

email: haykin@mcmaster.ca
Web site: http://soma.mcmaster.ca

Outline of the Lecture

1. Background Behind the Emergence of Cognitive Dynamic Systems (CDS)
2. Cognition
3. Highlights of Research into CDS in my Laboratory
4. Bayesian Filtering for State Estimation of the Environment
5. Cubature Kalman Filters
6. Feedback Information
7. Dynamic Programming and Optimal Control of the Environment
8. Summarizing Remarks on the Cognitive Process
9. Experiment on Cognitive Tracking Radar for Demonstrating the Power of Cognition
10. Final Remarks

    Last Note

1. Background Behind the Emergence of Cognitive Dynamic Systems (CDS)

Broad Array of Subjects that have prepared me for my current research passion: CDS

Signal Processing;

Control Theory;

Adaptive Systems;

Communications;

Radar; and

Neural Information Processing

Background (continued)

Two Seminal Journal Papers

(1) Simon Haykin, "Cognitive Radio: Brain-empowered Wireless Communications", IEEE Journal on Selected Areas in Communications, Special Issue on Cognitive Networks, pp. 201-220, February, 2005.

(2) Simon Haykin, "Cognitive Radar: A Way of the Future", IEEE Signal Processing Magazine, pp. 30-41, January 2006.

Background (continued)

Predictive Article[3]

"I see the emergence of a new discipline, called Cognitive Dynamic Systems, which builds on ideas in statistical signal processing, stochastic control, and information theory, and weaves those well-developed ideas into new ones drawn from neuroscience, statistical learning theory, and game theory. The discipline will provide principled tools for the design and development of a new generation of wireless dynamic systems exemplified by cognitive radio and cognitive radar with efficiency, effectiveness, and robustness as the hallmarks of performance".

---

3. Simon Haykin, "Cognitive Dynamic Systems", Proc. IEEE, Point of View article, November 2006.

2. Cognition

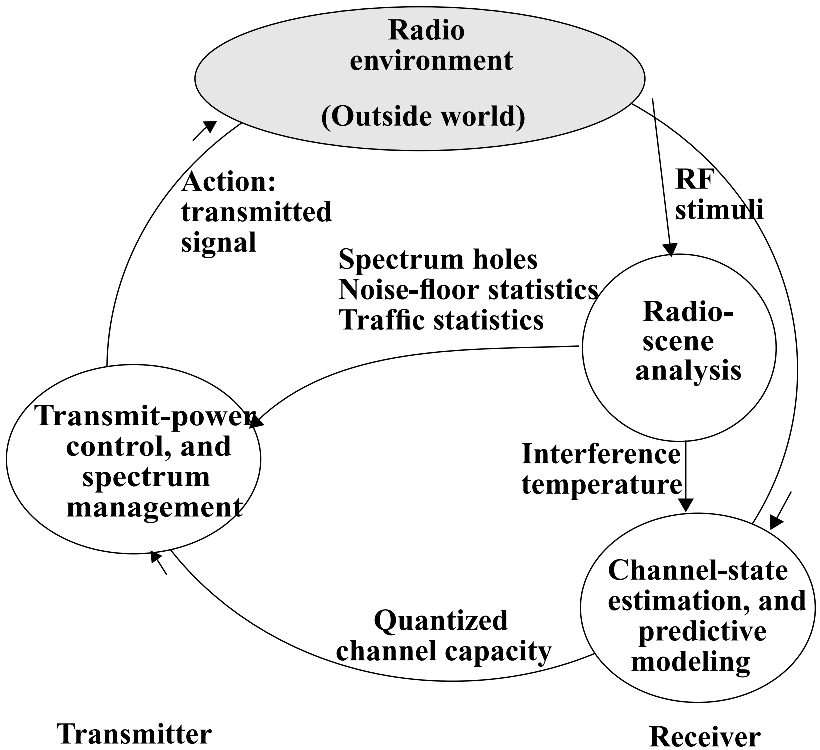

Figure 1: Information-processing Cycle in Cognitive Radio[1]

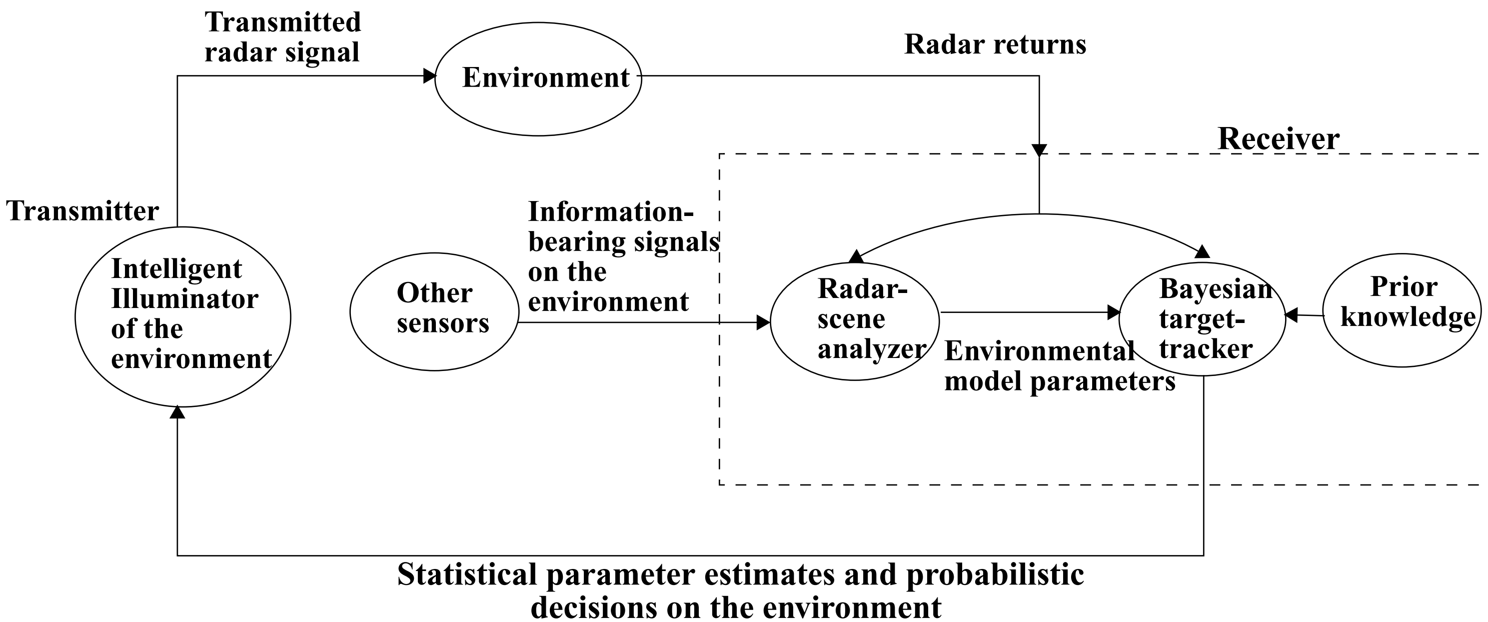

Figure 2: Block diagram of cognitive radar viewed as a dynamic closed-loop feedback system[2]

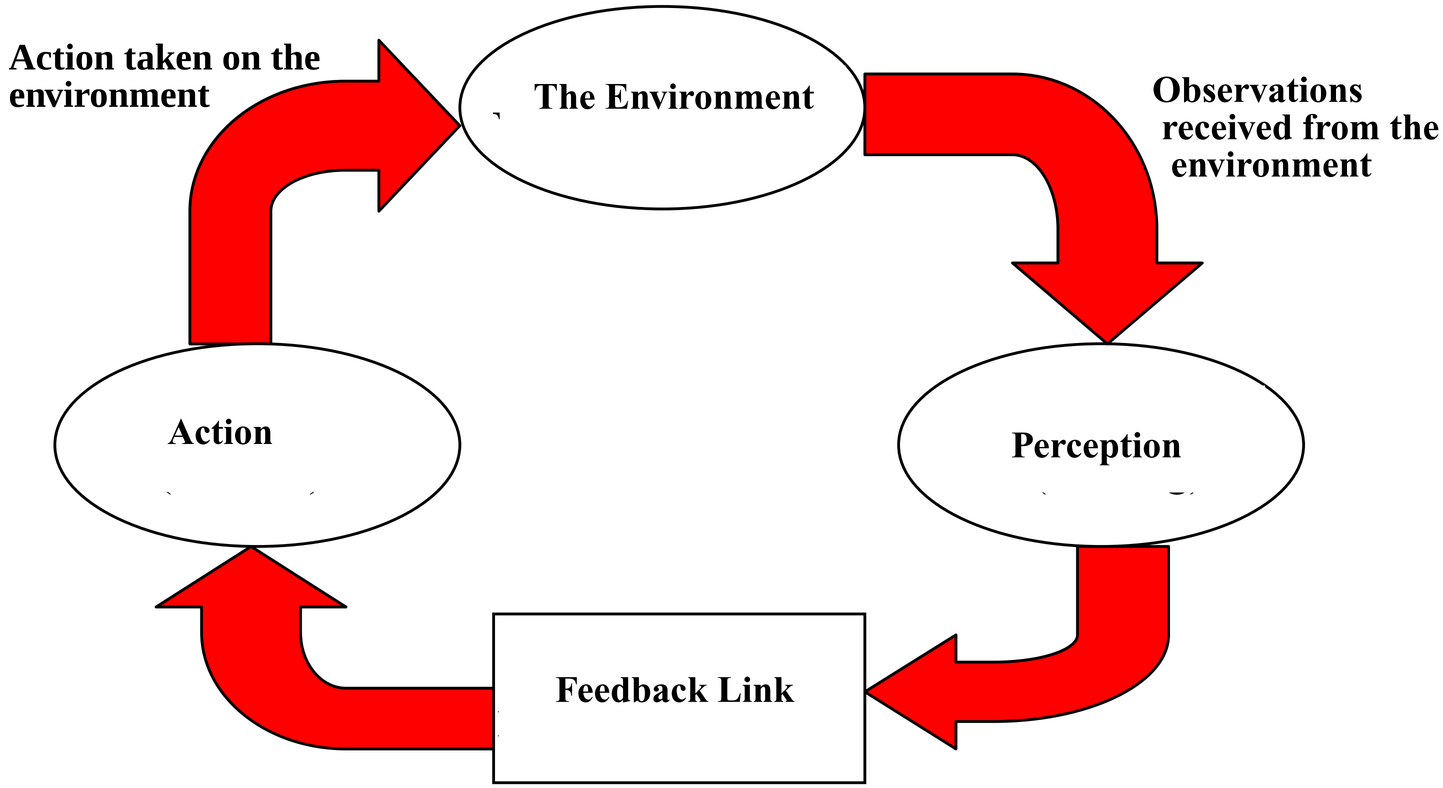

Figure 3: Graphical representation of perception-action cycle in the Visual Brain, (D.A. Milner and M.A. Goodale, 2006; J.M. Fuster, 2005)

3. Highlights of Research into CDS in my Laboratory

(i) Cognitive Radio

Self-organizing dynamic spectrum management for cognitive radio networks

The design of a software testbed for demonstrating this novel DSM strategy (using 5,000 lines of codes) has been completed, ready for experimentation; the strategy is motivated by Hebbian learning.

(ii) Cognitive Mobile Assistants

New generation of hand-held biomedical wireless devices used as aids for memory-impaired patients, and other related applications.

(iii) Cognitive Tracking Radar

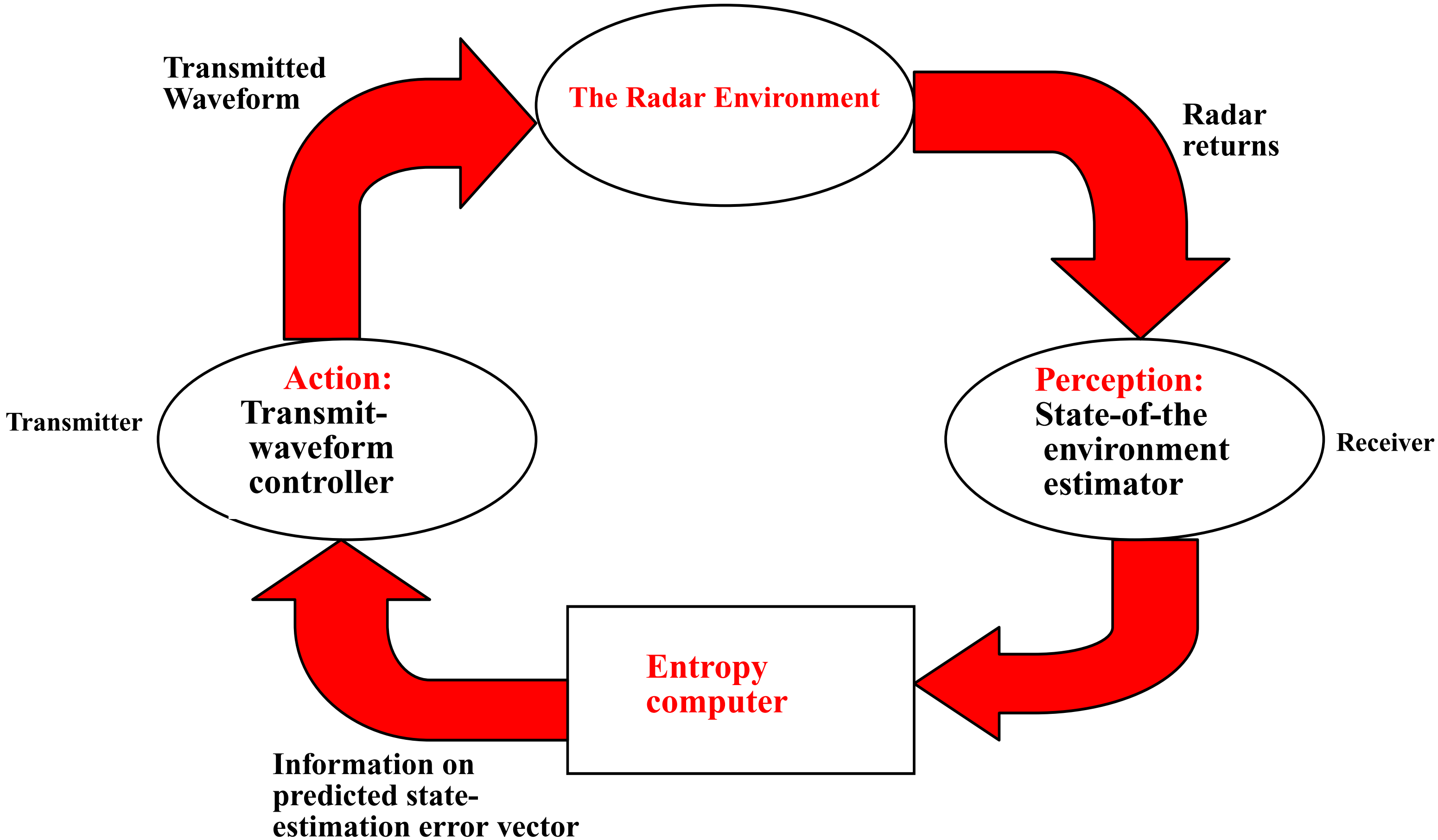

Figure 4: Cognitive information-processing cycle of tracking radar, revisited in light of the perception-action cycle in the visual brain

4. Bayesian Filtering for State Estimation of the Environment (Ho and Lee, 1964)

State-space Model

1. **System (state) sub-model**

$$\mathbf{x}_{k+1} = \mathbf{a}(\mathbf{x}_k) + \omega_k$$

2. **Measurement sub-model**

$$\mathbf{y}_k = \mathbf{b}(\mathbf{x}_k) + \nu_k$$

where k **= discrete time**

$\mathbf{x}_k$ **= state at time** k

$\mathbf{y}_k$ **= observable at time** k

ω_k **= process noise**

ν_k **= measurement noise**

Prior Assumptions:

- **Nonlinear functions $\mathbf{a}(\cdot)$ and $\mathbf{b}(\cdot)$ are known, with $\mathbf{a}(\cdot)$ being derived from underlying physics of the dynamic system under study and $\mathbf{b}(\cdot)$ derived from the digital instrumentation used to obtain measurements.**

- **Process noise ω_k and measurement noise ν_k are statistically independent Gaussian processes of zero mean and known covariance matrices.**

- **Sequence of observations**

$$\mathbf{Y}_k = \{\mathbf{y}_i\}_{i=1}^{k}$$

Up-date Equations:

1. Time-update:

$$\underbrace{p(\mathbf{x}_k|\mathbf{Y}_{k-1})}_{\substack{\text{Predictive}\\ \text{distribution}}} = \int_{R^n} \underbrace{p(\mathbf{x}_k|\mathbf{x}_{k-1})}_{\substack{\text{Prior}\\ \text{distribution}}}\underbrace{p(\mathbf{x}_{k-1}|\mathbf{Y}_{k-1})}_{\substack{\text{Old}\\ \text{posterior}\\ \text{distribution}}}d\mathbf{x}_{k-1}$$

where R^n denotes the *n*-dimensional state space.

2. Measurement-update:

$$\underbrace{p(\mathbf{x}_k|\mathbf{Y}_k)}_{\substack{\text{Updated}\\ \text{posterior}\\ \text{distribution}}} = \frac{1}{C_k} \cdot \underbrace{p(\mathbf{x}_k|\mathbf{Y}_{k-1})}_{\substack{\text{Predictive}\\ \text{distribution}}}\underbrace{l(\mathbf{y}_k|\mathbf{x}_k)}_{\substack{\text{Likelihood}\\ \text{function}}}$$

where C_k is the normalizing constant defined by the integral

$$C_k = \int_{R^n} p(\mathbf{x}_k|\mathbf{Y}_{k-1})l(\mathbf{y}_k|\mathbf{x}_k)d\mathbf{x}_k$$

Notes on the Bayesian Filter

(i) The posterior fully defines the available information about the state of the environment, given the sequence of observations.

(ii) The Bayesian filter propagates the posteriori (embodying time and measurement updates for each iteration) across the state-space model:

The Bayesian filter is therefore the maximum a posteriori (MAP) estimator of the state

(iii) The celebrated Kalman filter (Kalman, 1960), applicable to a linear dynamic system in a Gaussian environment, is a special case of the Bayesian filter.

(iv) If the dynamic system is nonlinear and/or the environment is non-Gaussian, then it is no longer feasible to obtain closed-form solutions for the integrals in the time- and measurement-updates, in which case:

We have to be content with approximate forms of the Bayesian filter

5. Cubature Kalman Filters[4]

Objective

Using numerically rigorous mathematics in nonlinear estimation theory, approximate the Bayesian filter so as to completely preserve second-order information about the state $\mathbf{x}_k$ that is contained in the sequence of observations $\mathbf{Y}_k$

In cubature Kalman filters, this approximation is made directly and in a local manner.

---

4. **I. Arasaratnam and S. Haykin, "Cubature Kalman filters", IEEE Trans. Automatic Control, pp. 1254-1269,** June 2009.

Steps involved in deriving the CKF:

In a Gaussian environment, approximating the Bayesian filter involves computing moment integrals of the form:

$$h(\mathbf{f}) = \int \underbrace{\mathbf{f}(\mathbf{x})}_{\text{Arbitrary nonlinear function}} \; \underbrace{\exp(-\mathbf{x}^T\mathbf{x})}_{\text{Gaussian function}} d\mathbf{x}$$

where x is the state.

CKF Steps (continued)

(i) Cubature rule, which is constructed by forcing the cubature points to obey symmetry:

Let $\mathbf{x} = r\mathbf{z}$ with $\mathbf{z}^T\mathbf{z} = 1$ for $0 < r < \infty$

The Cartesian coordinate system is thus transformed into a spherical-radial coordinate system, yielding

$$h = \int_0^\infty S(r) r^{n-1} \exp(-r^2) dr$$

where

$$S(r) = \int_{R^n} \mathbf{f}(r\mathbf{z}) d\sigma(\mathbf{z}),$$

$d\sigma(\mathbf{z})$ is an elemental measure of the spherical surface

and n is the dimension of vector x (state).

CKF Steps (continued)

(ii) Spherical rule of third-degree:

$$\int_{R^n} \mathbf{f}(r\mathbf{z})\, d\sigma(\mathbf{z}) \approx w \underbrace{\sum_{i=1}^{2n} \mathbf{f}[u]_i}_{\substack{\textbf{2}n\textbf{ cubature points resulting} \\ \textbf{from the generator }[u]}}$$

where w is a scaling factor.

CKF Steps (continued)

(iii) Radial rule, using Gaussian quadrature known to be efficient for computing integrals in a single dimension, which yields

$$\int f(x)\underbrace{w(x)}_{\text{Weighting function}}dx \approx \sum_{i=1}^{n} w_i f(x_i)$$

where

$$w(x) = x^{n-1}\exp(-x^2), \qquad 0 < x < \infty \text{ and } w_i = w(x_i)$$

and the integral is in the form of the well-known generalized Gauss-Laguerre formula.

Properties of the Cubature Kalman Filter

- ***Property* 1:** **The cubature Kalman filter (CKF) is a *derivative-free on-line sequential-state estimator*, relying on integration from one iteration to the next for its operation; hence, the CKF has a built-in *smoothing* capability.**

- ***Property* 2:** **Approximations of the moment integrals in the Bayesian filter are all *linear* in the number of function evaluations.**

- ***Property* 3:** **Computational complexity of the cubature Kalman filter grows as n^3, where n is the dimensionality of the state.**

- ***Property* 4:** **The cubature Kalman filter *completely preserves second-order information* about the state that is contained in the observations; in this sense, it is the best known information-theoretic approximation to the Bayesian filter.**

Properties of the Cubature Kalman Filter (continued)

- ***Property* 5: Regularization is naturally built into the cubature Kalman filter by virtue of the fact that the *prior* in Bayesian filtering is known to play a role equivalent to regularization.**

- ***Property* 6: The cubature Kalman filter *inherits well-known properties of the classical Kalman filter*, including square-root filtering for improved accuracy and reliability.**

- ***Property* 7: The CKF eases the curse-of-dimensionality problem, depending on how nonlinear the filter is:**

The less nonlinear the filter is, the higher is the feasible state-space dimensionality of the filter.

- ***Property* 8: The equally weighted cubature points provide a representation of the estimator's statistics (i.e., mean and covariance); computational cost of the CKF may therefore be reduced by modifying the time-update to propagate the cubature points.**

Hybrid CKF: Application to Tracking Coordinated Turns[5]

For a (nearly) coordinated turn in three-dimensional space subject to fairly small noise modeled by independent Brownian motions, we write the state equation

$$d\mathbf{x}(t) = f(\mathbf{x}(t))dt + \sqrt{\mathbf{Q}}d\beta(t)$$

where, in an air-traffic-control environment, the seven-dimensional state of the aircraft $\mathbf{x} = [\varepsilon, \dot{\varepsilon}, \eta, \dot{\eta}, \zeta, \dot{\zeta}, \omega]^T$ **with** ε, η **and** ζ **denoting positions and** $\dot{\varepsilon}, \dot{\eta}$ **and** $\dot{\zeta}$ **denoting velocities in the** *x, y* **and** *z* **Cartesian coordinates, respectively;** ω **denotes the turn rate; the drift function** $f(\mathbf{x}) = [\dot{\varepsilon}(-\omega\dot{\eta}), \dot{\eta}, \omega, \dot{\varepsilon}, \dot{\zeta}, 0, 0]^T$**; the noise term** $\beta(t) = [\beta_1(t), \beta_2(t), \ldots, \beta_7(t)]^T$**, involving seven mutually independent standard Brownian motions, accounts for unpredictable modeling errors due to turbulence, wind force, etc.**

---

5. I. Arasaratnam, s. Haykin, and T. Hurd, Cubature Filtering for Continuous-Discrete Nonlinear Systems: Theory with an Application to Tracking, submitted to IEEE Trans. Signal Processing.

The gain matrix $\mathbf{Q} = \mathrm{diag}([0, \sigma_1^2, 0, \sigma_1^2, 0, \sigma_1^2, 0, \sigma_2^2])$. **For the experiment at hand, a radar was located at the origin and digitally equipped to measure the range, *r*, and azimuth angle,** θ**, at a measurement sampling interval of *T*. Hence, we write the measurement equation:**

$$\begin{pmatrix} r_k \\ \theta_k \end{pmatrix} = \begin{pmatrix} \sqrt{\varepsilon_k^2 + \eta_k^2 + \zeta_k^2} \\ \tan^{-1}\left(\dfrac{\eta_k}{\varepsilon_k}\right) \end{pmatrix} + \mathbf{w}_k$$

where the measurement noise $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0},\mathbf{R})$ **with** $\mathbf{R} = \mathrm{diag}([\sigma_r^2, \sigma_\theta^2])$.

Data. $\sigma_1 = \sqrt{0.2}; \; \sigma_2 = 7 \times 10^{-3}; \; \sigma_r = 50\mathrm{m}; \; \sigma_\theta = 0.1\,\mathrm{deg};$

and the true initial state $\mathbf{x}_0 = [1000\mathrm{m}, 0\mathrm{ms}^{-1}, 2650\mathrm{m}, 150\mathrm{ms}^{-1}, 200\mathrm{m}, 0\mathrm{ms}^{-1}, \omega\ \mathrm{deg/s}]^T$.

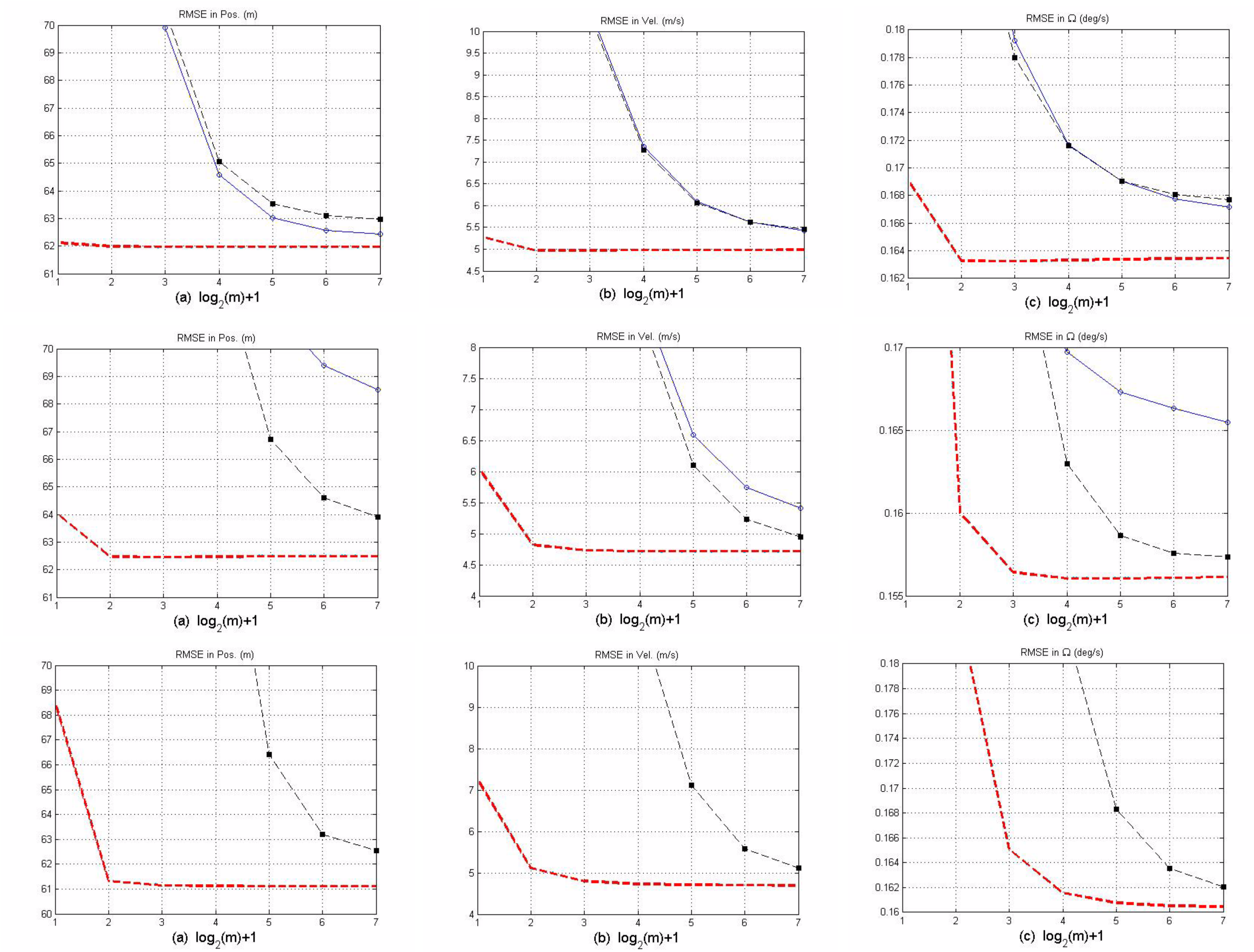

Figure 5: Accumulative RMSE Plots for a fixed sampling interval *T* = 6s and varying turn rates: first row, ω = 3 deg/s; second row, ω = 4.5 deg/s; third row, ω = 6 deg/s (Solid thin with empty circles-EKF, dashed thin with filled squares-UKF, dashed thick-hybrid CKF)

Matlab Codes

The Matlab codes for the discrete-time version of the CKF are available on the

Website: http://soma.mcmaster.ca

The Visual Cortex Revisited:

- **In Rao and Ballard (1997), the extended Kalman filter (EKF) was used to demonstrate that Kalman-like filtering (i.e., predictive coding) is performed in the visual cortex.**
- **The EKF relies on differentiation for its computation, whereas the CKF relies on integration.**
- **Since integration is commonly encountered in neural computations, it would be revealing to revisit the Rao-Ballard model using the CKF in place of the EKF.**

6. Feedback Information

(i) Principle of Information Preservation

In designing an information-processing system, regardless of its kind, we should strive to preserve the information content of observables about the state of the environment as far as computationally feasible, and exploit the available information in the most cost-effective manner.

(ii) Computation

At the receiver, the CKF computes the predicted state-estimation error vector.

With information preservation as the goal of cognitive processing, entropy of this error vector is the natural measure of feedback information delivered to the transmitter by the receiver.

For Gaussian error vectors, the entropy is equal to one-half the logarithm of determinent of the error covariance matrix, except for a constant term.

7. Dynamic Programming and Optimal Control of the Environment[6]

Design of the transmitter builds on two basic ideas:

(i) Bellman's dynamic programming and its approximation.
(ii) Library of linear frequency-modulated (LFM) waveforms with varying slopes, both positive and negative.

Given the feedback information about the state of the environment that is delivered by the receiver, an approximate dynamic programming algorithm (e.g., *Q*-learning, least squares policy iteration) in the transmitter updates selection of the current LFM waveform so as to reduce the entropy of the predicted state-error vector.

---

6. Dimitri Bertsekas, "Dynamic Programming and Optimal Control", Vol. 1 (2005); and vol. 2(2007), Athena Scientific.

8. Summarizing Remarks on the Cognitive Process

Each cycle of the cognitive process in radar consists of two updates:

(i) Transmitted waveform-update.
By delivering feedback information about state pf the the radar environment, the receiver reinforces the action of the transmitter by adapting it to update selection of the transmitted LFM waveform.

(ii) Feedback information-update.
The transmitter, in turn, reinforces the action of the receiver so as to update the entropy of the feedback information, viewed as the cost-to-go function of the dynamic programming algorithm.

This cycle of joint-reinforcement continues, back and forth, until the radar achieves its ultimate target objective.

9. Experiment on Cognitive Tracking Radar for Demonstrating the Power of Cognitive Process

Object Falling in Space, where the dynamics change as the object reenters the atmosphere

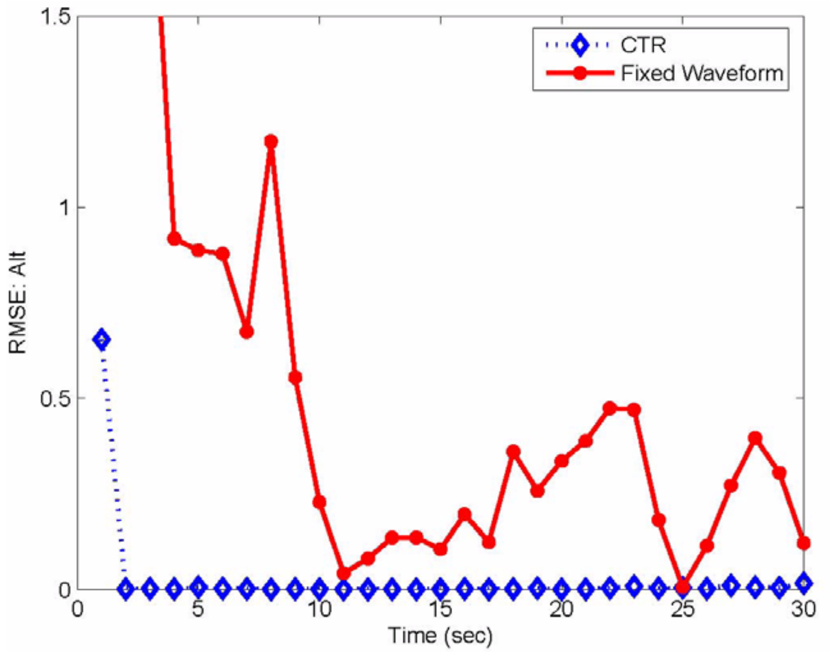

Figure 6: RMSE of altitude

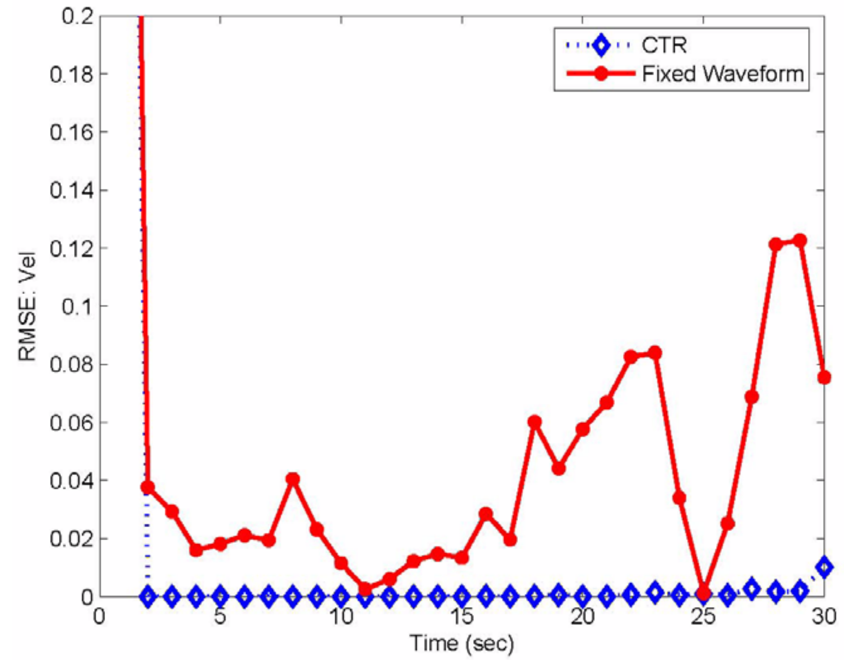

Figure 5: RMSE of velocity

10. Final Remarks

Emboldened by my extensive work done on Cognitive Radio for the past four years and the exciting experimental results presented in this lecture on Cognitive Radar, I see breakthroughs in designing new generation of engineering systems that exploit cognition exemplified by

- **Cognitive radio networks for improved utilization of the electromagnetic spectrum**
- **Cognitive mobile assistants for a multitude of biomedical and social-networking applications**
- **Cognitive radar systems with significantly improved accuracy, resolution, reliability, and fast response**
- **Cognitive energy systems for improved utilization and integration of different sources of energy**

Final Remarks (continued)

Simply stated:

Cognition is a transformative software technology, which is applicable to a multitude of engineering systems, old and new.

Cognitive Dynamic Systems and Neuroscience

The study of cognitive dynamic systems is motivated by ideas drawn from cognitive neuroscience, particularly, the visual brain.

Just as we learn from ideas basic to the human brain, it is my belief that the study of cognitive dynamic systems (particularly, cognitive radar) from an engineering perspective may well help us understand some aspects of the brain.

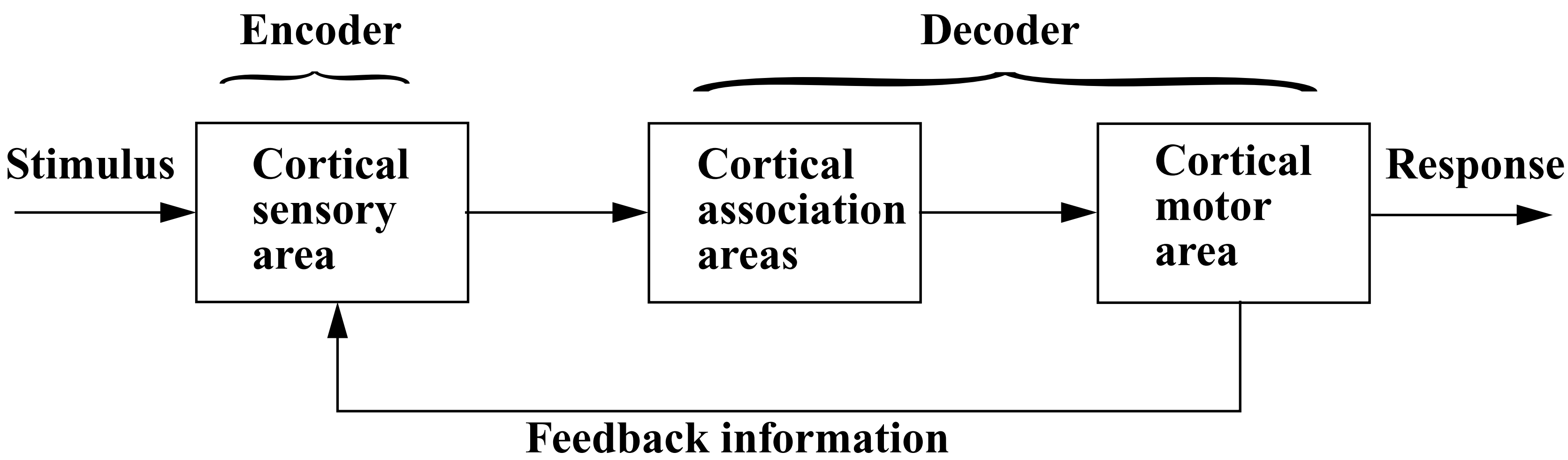

Figure 7: Block-diagram representation of processing stages in perceptual tasks (Adapted from C. Speidemann, Y. Chen, and W.S. Geisler, chapter 29 in M.S. Gassaniga, editor-in-chief, The Cognitive Neurosciences, 4th Edition, 2009).

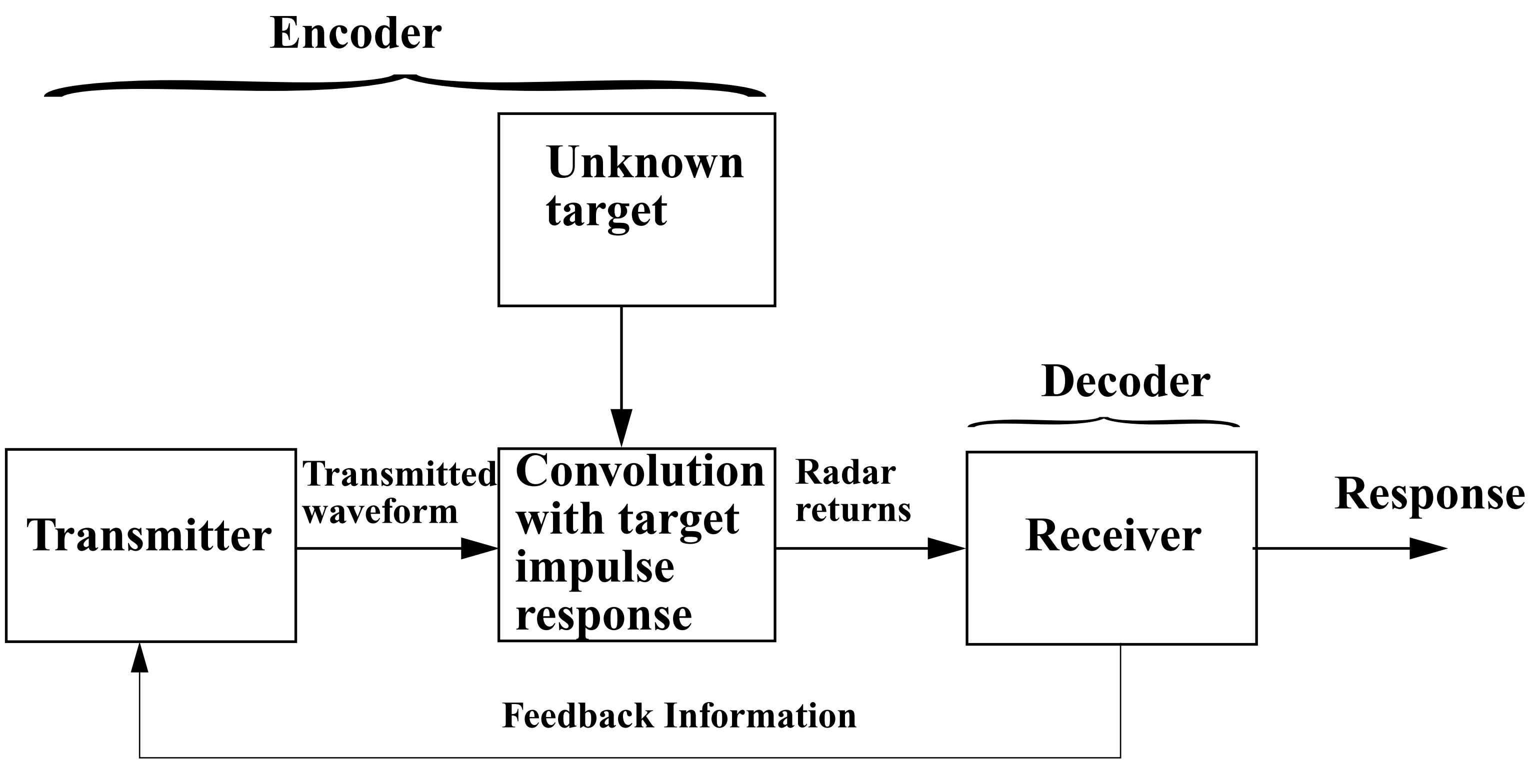

Figure 8: Block-diagram representation of processing stages in radar system

Last Note

The complete set of slides for this lecture is downloadable from our website:

http://soma.mcmaster.ca