Autonomy, Learning, Information and Vision in the Atmosphere

Sai Ravela
Massachusetts Institute of Technology
People and Organizations

- Sai Ravela
  - Piyush Tagade, Hansjeorg Seybold, Clea Denamiel (Postdocs)
  - Randy Westlund (Engineer)
  - N. Roberts, E. Ismail, C. Wheatley, K. Wolfenberger, L. Wright, J. Castillo, K. Horne (Students)

- Han-Lim Choi (Cooperative Sensor Planning, Sampling)
  - Su-Jin Lee, Jung-Su Ha

- Kerry Emanuel (Clouds), Jon How (Methodology)

- Joaquin Salas (IPN, Logistics), Baruc Flores, Othon Gonzalez (Itzamna Aero)

- Earle Williams (Volcano), Dan Rubenstein (Animal Tracking), Dave Pieri (NASA Volcano monitoring) and Ru-Shan Gao (Atmospheric Sensing)

- CENAPRED (Mexico), Mpala Research Center (Kenya)
Environmental Systems Science

- In addition to ICCS

- DDDAS Workshop @ ACC 2015
  - Singla, Bernstein, Ravela and Darema
  - Controls in very large state-spaces
Localizing and Mapping Coherent Fluids

• Climate Application:
  • Shallow Cumuli, The big bad player in global warming
    • Model Parameterization, Verification, Database
    • Exactly why and how shallow clouds precipitate

• Natural Hazards:
  • Thermals, plumes, storms

• Localized features – Coherent Fluids
  • Organized, transient, tracer evidence.
  • Inference needs special attention!

• Simultaneously Localize Fluid and Map: DDDAS

Zhang and Klein, JAS 2013
Tracers and their Mutual Information

Just fly around? Eulerian and Lagrangian Sampling
Physics-based dimensionality and nonlinearity challenges
Fluid SLAM: Dynamic Data-Driven Approach

- Numerical Prediction
- Nowcast
- Plan
- Resample
- Data Assimilation

Graph:
- Forecast Lead Time (hours)
- Forecast Skill
- Numerical Models
- Extrapolation
- Blending
- Perfect
- Little

Diagram:
- DA
- LAM
- Target
- Map
- Nowcast
- Plan
- Track
- Remote
- Insitu

Ravela 2012, Ravela 2013
Fluid-SLAM: Elements of DDDAS Methodology

• Complex Physics ➔ Big Model: how to constrain?
  • A new technique for multi-objective non-Gaussian estimation in non-linear filtering (Ravela et al.)
    • Tractable Information theoretic inference is feasible, provides a uniform framework
    • Ensemble Learning provides a robust inference framework

• Big Models ➔ Reduced Models: what about model error?
  • Dynamically deformable models for linear reduced and nonlinear manifold techniques (Ravela et al.)
  • Biorthogonal field equations (Choi et al.)
  • Coherent random field expansions (Ravela et al.)
  • Ensemble Learning based nonlinear filtering for compensating for Model Error

• Big Data ➔ Learned Models: where is the fundamental physics?
  • A method to learn models from data using physically-based priors (Ravela et al.)
Fluid-SLAM: Elements of DDDAS Methodology

• Estimation ➔ Planning ➔ Control ➔ Sensing ➔ Prediction: where is the efficiency?
  • A tractable information theoretic framework is developed (Ravela et al.)
  • Game-theoretic cooperative sensor planning (Choi et al.)
  • Mutual information for windowed forecasting (Choi et al.)
  • Periodic trajectory for persistent informative sampling (Choi et al.)

• Coherent Geophysical Fluids: how to deal with nonlinearity and dimensionality?
  • Pattern perspective essential for coherent fluids
    • Feature and Deformation scale space.
    • Deformable States, Models and Manifolds
  • New methods for data assimilation, ensemble analysis, coherent random fields, principal appearance and geometry modes, deformable reduced models, reduced deformable models, coherent super-resolution, bias estimation (Ravela et al.)
Fluid-SLAM: Elements of Application and Experimental Program

- Application System: Fluid-SLAM for coherent geophysical fluids
  - Thermals: Atomic Atmospheric Structures
  - Plumes: Natural Hazards
  - Cumuli: Climate Processes

- Experimental Framework
  - Laboratory Studies
    - The first fully coupled differentially heated rotating annulus system (Ravela et al. 2008).
  - Portable, scalable autonomous field instruments
    - Hobby aircraft matched to problem scale
    - **Our approach is being copied quickly!**
      - The first demonstration of autonomous adaptive sampling for plumes (Ravela et al. 2013).
      - Rapid logistical advances for field experiments.
Some hypotheses can be validated in Laboratory

The first coupled lab-numerical observing system, enabled by cheap fiber optic rotary joint.

Narasimha et al., PNAS, 2011
Focus on Field Experiments

Rapid progress made by 2012 using hobby aircraft, first group to deliver hobby-based adaptive sampling.
Variety of Onboard payload modules, here a side-looking camera on final

- X8 has spiral mode issues, resolved with new design.
  - Itzamna: X8 → I9 & I9W
- Key issues
  - Multiple aircraft operator efficiency,
  - Multi aircraft coordination
  - Recovery and turn-around time

sUAS includes Glider trajectory planning

Continued...
Mexico Field Experiments: Quick Bootstrap

Current flight configuration: **Two remote, two in-situ.**

Visit us at AGU

Please visit [http://caos.mit.edu](http://caos.mit.edu) for archive.
The *X8-based autonomous plume hopper*, first introduced in 2012 for cooperative autonomous observation
Itzamna i9 based on MIT S9TWL modification in public domain used for experiments. Current prototyping includes the use of decelerons/duckerons instead of twin-tail configuration.
Plume Hopping: Front Localization BY MCMC

The first autonomous “MCMC-GP Plume Mapper” by hobby aircraft.

---

Ravela et al. 2012, Ravela et al., 2013
Periodic Trajectory for Persistent Informative Sensing

• Goal: Reduce uncertainty by persistent operation of sensors

\[
\begin{align*}
\min J_1 &= \lim_{T \to \infty} \frac{1}{T} \int_0^T U(s(t)) + u(t)^T R_c u(t) \, dt \\
\text{s.t. } \dot{x}(t) &= \begin{bmatrix} \dot{p}(t) \\ \dot{s}(t) \end{bmatrix} = \begin{bmatrix} u(t) \\ \text{vec}(\dot{S}(t)) \end{bmatrix}
\end{align*}
\]

- \( p(t) \): sensor dynamic state, \( s(t) \): Riccati state
- \( U(s(t)) \): uncertainty measure function

• Periodic Optimal Control Re-Formulation

\[
\min_{x(0), w, u(t)} \left. \right|_{x(0)=x(w)} J_2 = \frac{1}{w} \int_0^w U(s(t)) + u(t)^T R_c u(t) \, dt
\]

- Optimized \( x(0), w, u(t) \) \( \rightarrow \) periodic state/uncertainty dynamics

• Thm1: moving along POCP sensor trajectory \( \rightarrow \) convergence to POCP uncertainty trajectory

• Implication: may enable discussion on controllability of information state

Ha and Choi, 2014
Working Hypotheses: Topics of Research in 2014

• Products: Over 12 papers, please visit essg.mit.edu

• Portable, scalable autonomous field instruments
  • Hobby aircraft are matched to the scale of problems of interest

• Ensemble Learning Compensates for Model Error
  • Hold the rush to nonlinear filtering!

• Tractable Information theoretic inference is feasible
  • Uniform framework needed in the face of nonlinearity and dimensionality

• Pattern perspective essential for coherent fluids
  • Feature or Deformation scale space.
  • *Deformable States, Models and Manifolds (DyDESS 2014)*
Ensemble Learning in Data Assimilation

A new framework to incorporate ensemble-based inference with multiple priors
A new framework for multi-objective variational inference
Better treatment of uncertainty in proposal distributions
Better treatment of model error
Ensemble Learning: Multi-Objective Assimilation

• Motivation
  • Non-Gaussian estimation heightens sensitivity to model error.
    • EnKF quickly outperforms GMM filter in the presence of bias
    • Large-scale variance reduction cannot be ignored in the presence of model error (here, bias)
  • In general,
    • The prior distribution becomes imperfect
      • Thus, imperfect objective
    • Multiple objectives may be of interest
      • Multiple priors (proposals) will be needed

\[ p(x) \sim \sum_{m=1}^{M} \alpha_m N(x; \mu_m, P_m) \]
Ensemble Learning

- Proposition: True distribution unknown
  - Several proposals depending on the objective
- Test case: Large scale variance reduction and maximal posterior mode
- Approach: Boosted Mixture Ensemble Filter (B-MEnF)
  - Combine new ensemble mixture filter (MEnF) with EnKF
  - Using stacked and cascade generalization
- Mixture Ensemble Filter
  - A direct update in particle space without synthesis of moments
  - Direct fast smoother form
  - Efficient

\[
p(x) \sim \sum_{m=1}^{M} \alpha_m N(x; \mu_m, P_m)
\]

\[
x_e^a \equiv x_e + \sum_{m=1}^{M} \{K(P_m)(d_{em} - H \omega_{em} x_e)\}
\]

\[
A^a = A^f \sum_{m=1}^{M} \Xi_m \circ (W_m^N)^T = A^f \Xi
\]
Ensemble Machine: Boosted Mixture Ensemble Filter and Smoother

• Advantages
  • Non-Gaussian Inference
  • Direct State Adjustment
  • Negligible Resampling
  • Compact Ensemble Transform
  • Fast Smoother Form

• An inference counterpart to the “multi model ensemble”
Boosted Mixture Ensemble Filter and Smoother
Working Hypotheses: Topics of Research in 2014

• Products: Over 12 papers, please visit essg.mit.edu
• Portable, scalable autonomous field instruments
  • Hobby aircraft are matched to the scale of problems of interest
  • Our approach is being copied quickly!
• Ensemble Learning Compensates for Model Error
  • Hold the rush to nonlinear filtering!
• **Tractable Information theoretic inference is feasible**
  • Uniform framework needed in the face of nonlinearity and dimensionality
• Pattern perspective essential for coherent fluids
  • Feature or Deformation scale space.
  • *Deformable States, Models and Manifolds (DyDESS 2014)*
Tractable Information in Inference

A new framework tractable information theoretic variational inference
Better treatment of nonlinear filtering, among many others
Demonstration in data assimilation, cooperative sensor planning
Tractable Information Theoretic Approach

Mutual Information Filter (MuIF): Non-Gaussian prior and likelihood, Maximization of Kapoor’s quadratic mutual information

- Shannon Entropy $\mathcal{H}(X) = - \int f_X(x) \log(f_X(x)) dx$

- Mutual Information $\mathcal{I}(X; Y) = \mathcal{H}(X) - \mathcal{H}(X | Y)$

- Inference: Maximization of Mutual Information

- Many Applications
  - Sensor Planning, Data Assimilation, Control, Model Reduction, Adaptive Sampling, Causal Graphs

Tractable?
Tractable Approach: Quadratic Mutual Information

• We use Generalization of Shannon Entropy: Renyi Entropy
  • \( \alpha \)-Information Potential:
  • Others: Havrda-Chardat, Behara-Chawla, Sharma-Mittal
  • Kapur establishes equivalence

• Special Interest:

• Kapoor’s result:

• Mutual Information via Information potentials
  • Kernel density estimate
  • Reinterpretation using RKHS

• Maximization of MI: *Steepest Gradient Ascent*

\[
V_\alpha(X) = \int (f_X(x))^\alpha \, dx
\]
\[
R_\alpha(X) = \frac{1}{1 - \alpha} \log V_\alpha(X)
\]
\[
R_2(X) = - \log (V_2(x))
\]

\[
D_e(f_X, g_X) = \int (f_X(x) - g_X(x))^2 \, dx
\]

\[
\mathcal{J}(X; Y) = D_e(f_{XY}(x, y), f_X(x)f_Y(y))
\]
Application Results: Lorenz 95

• Lorenz 95 Model
\[ \frac{dx_i}{dt} = x_i \ 2x_i \ 1+x_i \ 1x_{i+1} \ x_i+u \]

Maximize conditional mutual information with parameterization:
\[ F(x, z) = x + K(z - Hx) \]

Tagade and Ravela, 2014
Information coupling makes decentralized sensor planning difficult
  – Sequential greedy often works well, but not really decentralized

Potential Game achieves cooperation by designing local utility appropriately
  – Plus, good recent progress on Nash-finding learning schemes available

Our suggestion: max own info gain conditioned on others decisions

⇒ PG with potential = global objective
  – Joint Strategy Fictitious Play used for learning

\[ U_i(Z_{s_i}, Z_{s_{-i}}) = \mathcal{I}(V; Z_{s_i} | Z_{s_{-i}}) \]
Mutual Information for Windowed Forecasting

- Smoother form works again:
  \[ \mathcal{I}(\mathcal{V}_{[T_i,T_f]}, \mathcal{Z}_{[0,\tau]}) = \mathcal{I}(X_\tau; \mathcal{Z}_{[0,\tau]}) - \mathcal{I}(X_\tau; \mathcal{Z}_{[0,\tau]} | \mathcal{V}_{[T_i,T_f]}) \]
  - Now looking at entropy of \( X_\tau \)
  - Conditioning on verification path is not trivial b/c (fictitious) **noise-free measurement** of \( V \) involved

- From [Bucy67; Mehra68; Fujita70]
  - Perfect measurement of process \( V \)
  - noise-free obs of \( V \) at initial time + running observation of \( dV \)
    \[ P_{X|V}(T_i) = [P^f(T_i)^{-1} + S^b(T_i)]^{-1} \]
  - Two-filter form of fixed interval smoothing readily applicable

Choi and Ha, (under review)
Working Hypotheses: Topics of Research in 2014

- **Products:** Over 12 papers, please visit essg.mit.edu
- **Portable, scalable autonomous field instruments**
  - Hobby aircraft are matched to the scale of problems of interest
  - Our approach is being copied quickly!
- **Ensemble Learning Compensates for Model Error**
  - Hold the rush to nonlinear filtering!
- **Tractable Information theoretic inference is feasible**
  - Uniform framework needed in the face of nonlinearity and dimensionality
- **Pattern perspective essential for coherent fluids**
  - Feature or Deformation scale space.
  - *Deformable States, Models and Manifolds (DyDESS 2014)*
Statistical Inference for Coherent Fluids: Feature-based

Useful in phenomenology, and simulation

Coherent structures provide a rich description of phenomenology, and simulation. They are sparse.

Location-Scale-Shape (Kinematic features) identified as extrema of invariant surfaces (not ad-hoc) are augmented with amplitude (dynamic features)

We call this a LASSO Perceptual Organization of coherent fluids

\[ u + \frac{1}{2} (u^2)_x + u_{xxx} + \gamma u = \eta f(x) \]
Feature-based Framework in Inference

- LASSO can be used to solve inference problems.
- Analysis Step: Extract LASSO features
- Inference Step: Low-dimensional non-Gaussian.
- Synthesis: Numerical fluid from sparse LASSO
- A perceptual analog of many analysis-synthesis schemes, e.g. wavelets
LASSO Open Loop

Tagade et al., 2014
Analysis, Inference and Synthesis: Closed Loop

Tagade et al., 2014
Statistical Inference for Coherent Fluids: Non-Parametric Approach

LASSO fails for ill-defined features, complex shapes, highly nonlinear deformations. Non-parametric method needed, we propose a Diffeomorphic Scale Space.
Supervised Inference: Data Assimilation by Field Alignment

\[ P(X, q \mid Y) \propto P(Y \mid X, q) P(X \mid q) P(q) \]

Note: Not just translation!
Stochastic optimization
Sparse Observations OK
Nonlinear deformations OK
Complex features OK

Dynamical Balance

Stochastic Optimization
Deformation Scale Space: SCA

• Homotopy on Spectral basis
  • A parsimony approach to deformation
  • Improvement over diffeomorphic approach
  • Uses Gabor filters in deformation space
    • Localized and therefore “portable”
Application to Nowcasting: A Reduced Deformable Model with Growth/Decay

VIL Statistics, 120 Minute Forecast, 2010-06-30

120 MINUTE FORECASTS
Application to Estimating Bias in Numerical Weather Prediction

A strong bias component exists in deformation space (Joint work with NOAA)

Forecast Error Decomposition (FED)
Application to Fluid Super-Resolution/Downscaling

A new approach to robust downscaling using non-parametric deformable coherent fluids

Use of Latent Semantic Hashes for retrieval of exemplars

Use of dynamically deformable fluid patches for downscaling

Combines learning/information retrieval with fluid dynamics. Extends example-based methods:

Ravela & Freeman, 2006
Non-parametric Deformation-Amplitude Downscaling for Coherent Fluids

- Completely Non-parametric, assumption on upscaling-downscaling functions
- Physics/Dynamics is implicit in association between “low-res” and “hi-res”
- Arbitrary information transfer
- Using Deformation-Amplitude Space allows great variability with small training set

Training Data: Low and High resolution Snapshots

Test Low Resolution Fluid Patch

Search DB: Using LSH

Retrieve similar Low-Res patches Associated Hi-Res patches

Dynamically Deform Ensemble of Low-Res Patches to maximize mutual information with Test patch

Apply deformation, growth/decay to retrieved hi-res patches to produce downscaled/super-resolved ensemble

T=24000
Methodological Evolution

- With or without correspondence
- Alternative measures (e.g. mutual information)
- Turbulent motion model
- Common basis with object recognition

- Simplest explanation
- Better Selectivity and Invariance tradeoff
Unsupervised Inference: Field Coalescence

• Field coalescence uses EM to solve for an “N-body” problem to discover the unknown mean

\[ P(\bar{X}, \{q\}|\{X^f\}) \]
Application to Nowcasting

• Blending from two sources by joint amplitude-position coalescence
Incoherent Reduction of Coherent Fluids

Standard statistics and expansions do not work for coherent fluids

The ordinary KL-expansion of coherent fields produces incoherent realizations

The basis becomes a “generic prior” and loses representation of coherence
Coherent Random Fields

Field Coalescence with KLT helps constitute Principal Appearance and Geometry modes and a Coherent Random Fields model.

The random realization using the new Coherent Random Field model, even just the leading mode, preserves coherence and the feature statistics (that were not used to construct the random field model).
POD (and other manifold techniques) fail terribly in the presence of model error. For example, change in boundary. These errors manifest in deformation and amplitude spaces.

- Traditional DA in amplitude space for reduced variable is insufficient, the basis is incorrect.

- Deformability is introduced as an adaptive mechanism for linear, nonlinear and manifold reduction when dealing with coherent fluids.
Dynamically Deformable Reduced Models

• Deformability as a model error compensating mechanism
• Non-parametric representation via auxiliary deformation vector field
Example

Make time predictions using PAG

Measurements are obtained

Dynamically Deform PAG basis

Normal DA and transform original PAG predictions
A Perceptual Approach to Coherent Fluids enables DDDAS

**Traditional Model**
- Physical Space
- Primitive Equations
- Lagrangian Features (e.g. LCS)
- Eulerian Features
- Polynomial Expansions
- POD
- Complex numerical model

**Our Model**
- Feature Space
- Apparent Dynamics
- Generalized Features
- Learned Expansions
- Principal Appearance and Geometry Modes (PAG)
- Simple Data-driven model
- Deformable Reduced Models (new)!
Summary

• Hobby aircraft deliver performance, we’ve applied and adapted designs for cooperative autonomous observation

• Three methodological advances in 2014
  • Multi-objective estimation is essential to deal with model error, ensemble learning framework is proposed
  • Tractable information theoretic approach demonstrated on data assimilation
  • A non-parametric pattern theoretic framework is introduced to deal with
    • Model error in reduced models of coherent structures
    • New random field model for coherent fluids
    • New appearance and geometry modes for coherent features
    • Alignment and Coalescence models for supervised and unsupervised inference.