

Stat232A: Statistical Modeling and Learning in Vision and Cognition

Winter, 2020, Stat 232A-CS266A, MW 5:00-6:15 pm, Boelter Hall 5273

[\[syllabus.pdf\]](#)

Course Description

This is the first of a series of graduate level courses (Stat232A, B and C) that introduce the principles, theories, and algorithms for modeling complex patterns in vision and cognition. Stat 232A is focused on images models in David Marr's paradigm. More specifically we study two classes of statistical models: 1) Descriptive models (Markov random fields, Gibbs distributions); and 2) Generative models (sparse coding, auto-encoding). We start with one layer models to study the principles and models to gain in-depth understanding, and then move to multi-layered structures (compatible with Deep Neural Networks) to scale up performances.

We will also study the interactions and integration of these models, and develop a general unified theory for pursuing statistical models over a series of probabilistic families. The course also teaches a common framework for conceptualizing stochastic patterns and for statistical knowledge representation. Although the lectures will mostly focus on visual patterns in images and videos, and the methodology should be generally applicable to a wide range of applications, such as biologic patterns, network traffic modeling, material science, artificial intelligence, cognitive modeling, and autonomous robots, etc.

Prerequisites : Basic statistics, linear algebra, programming skills for a project.

Textbook [\[draft book pdf\]](#)

- *Computer Vision: Statistical Models for Marr's Paradigm*, Draft book by Song-Chun Zhu and Yingnian Wu, 2020. We will print a hardcopy for you.

Instructor

- Prof. [Song-Chun Zhu](#), email: sczhu at stat.ucla.edu, 6-8693, BH 9404. **Office Hours: Monday 1:00-2:00pm.**
- Teaching Assistant, Erik Nijkamp. email: erik.nijkamp@gmail.com, BH 9406. **Office Hour: Wednesday 3-5:00pm.**

Grading Plan

Two homework assignments HW1 [10 %] -- HW2 [10%]	20%
Small projects and exercises <ul style="list-style-type: none"> • 1. Natural image statistics, scale invariance and image completion by PDE [10%] • 2. Sampling the Julesz texture ensemble [10%] • 3. Multi-grid sampling of DeepFRAME model [10%] • 4. Alternative back-propagation for hierarchical models [10%] 	40%
Final Exam: Monday, March 16, 11:30 AM - 2:30 PM	40%

List of Topics

Chapter 1 Introduction to Knowledge Representation, Modeling and Learning [\[Lecture 1.pdf\]](#)

1. Towards a unified representation of commonsense knowledge from pixels to minds
2. Modeling principles: Compositionality, reconfigurability, functionality and causality
3. Definition and Representation of concepts: Sets and statistical models
4. Examples and demos: Regimes of representations

Chapter 2 Empirical Observations: Image Space and Natural Image Statistics [\[Lecture 2.pdf\]](#)

1. Empirical observation I: filtered responses
2. Empirical observation II: scaling properties
3. Empirical observation III: patch frequency (structural and textural patches).

4. Empirical observation IV: information scaling and regimes of statistical models.

Chapter 3 Classical Markov and Gibbs Random Fields

1. Markov random field theory
2. Gibbs fields: Ising and Potts models
3. The equivalence of Gibbs and MRF (Hammersley-Clifford theorem)
4. Early Markov random field models for images
5. Maximum Entropy and Maximum Likelihood Estimation
6. Variations of likelihood: pseudo-, patch-, partial-likelihood
7. From Gibbs distributions to PDEs in image processing

Chapter 4 FRAME Model and Julesz Ensemble

1. FRAME model and texture modeling
2. Pythagorean theorem and information projection
3. Minimax entropy learning and feature pursuit
4. Julesz ensemble
5. Ensemble equivalence theorem
6. Ensembles in statistics mechanics
7. Other examples on general prior, shape, curves, Gestalt field etc

Chapter 5 DeepFRAME Model

1. Posing hierarchical Convolutional Neural Net as a unfolded Generalized Linear Model
2. Defining DeepFRAME model
3. Sampling and Learning the DeepFRAME model
4. Examples on synthesizing images, videos and 3D shapes.
5. Hopfield auto-encoder
6. Multigrid sampling
7. Adversarial interpretations
8. Landscape and short-run MCMC

Chapter 6 Classical Generative Models

1. Frame theory and wavelets
2. Design of frames: image pyramids
3. Over-complete basis and matching pursuit
4. Markov tree and stochastic context free grammar

Chapter 7 Sparse coding, Textons, and Active Basis Models

1. Learning sparse coding from natural images
2. Tangram models and hierarchical tiling
3. Textons and image dictionary
4. Sparse FRAME model
5. Active basis model
6. Examples

Chapter 8 Hierarchical Generative Models

1. Factor analysis and auto-encoder
2. Hierarchical Factor analysis with Convolutional Neural Nets
3. Learning by Alternating back-propagation
4. Distinguishing appearance and geometry
5. Short-run inference dynamics

Chapter 9 Information Scaling, Scale Space and Imperceptibility

1. Image scaling, perceptual entropy, and imperceptibility
2. A continuous entropy spectrum and transition of model regimes
3. Perceptual scale space
4. Perceptibility, metastability, and energy landscape

Chapter 10 Integrated Models: Descriptive + Generative

1. Primal sketch model as low-middle representations for generic images and video
2. 2.1D sketch or layered-representation
3. Mixed random fields
4. 2.5D sketch representation
5. Shape from stereo
6. Shape from shading

Chapter 11 A Tale of Three Families: Descriptive, Generative and Discriminative Models

1. Cooperative learning with descriptive and generative models
2. DDMCMC: Discriminative models driving generative models in inference
3. Divergence triangle: integrating variational and adversarial learning