# Inferring "Dark Matter" and "Dark Energy" from Videos


Dan Xie[1], Sinisa Todorovic[2] and Song-Chun Zhu[1]

[1] University of California, Los Angeles, [2] Oregon State University


## Problem

Localizing functional objects in surveillance videos

- ❑ Functional objects can satisfy human needs:
  - hunger: food truck,
  - thirst: vending machine,
  - rest: bench,
  - cleanliness: trash bin.
- ❑ Functional objects hard to detect = "Dark matter"
- ❑ "Dark matter" attracts people to satisfy the needs
- ❑ People have intents to approach "dark matter"

"Dark matter" is at the ends of people's trajectories

Challenges:

- Tracking people in surveillance videos is noisy.
- Not all end points of the trajectories observed.

## Approach

Assumptions:

- Scene layout consists of:
  - Dark-matter locations,
  - Walkable areas,
  - Non-walkable areas + obstacles = Constraint map.
- People:
  - Familiar with the scene layout,
  - Move only to one goal "dark matter" at a time,
  - Take the shortest path to the goal avoiding obstacles.

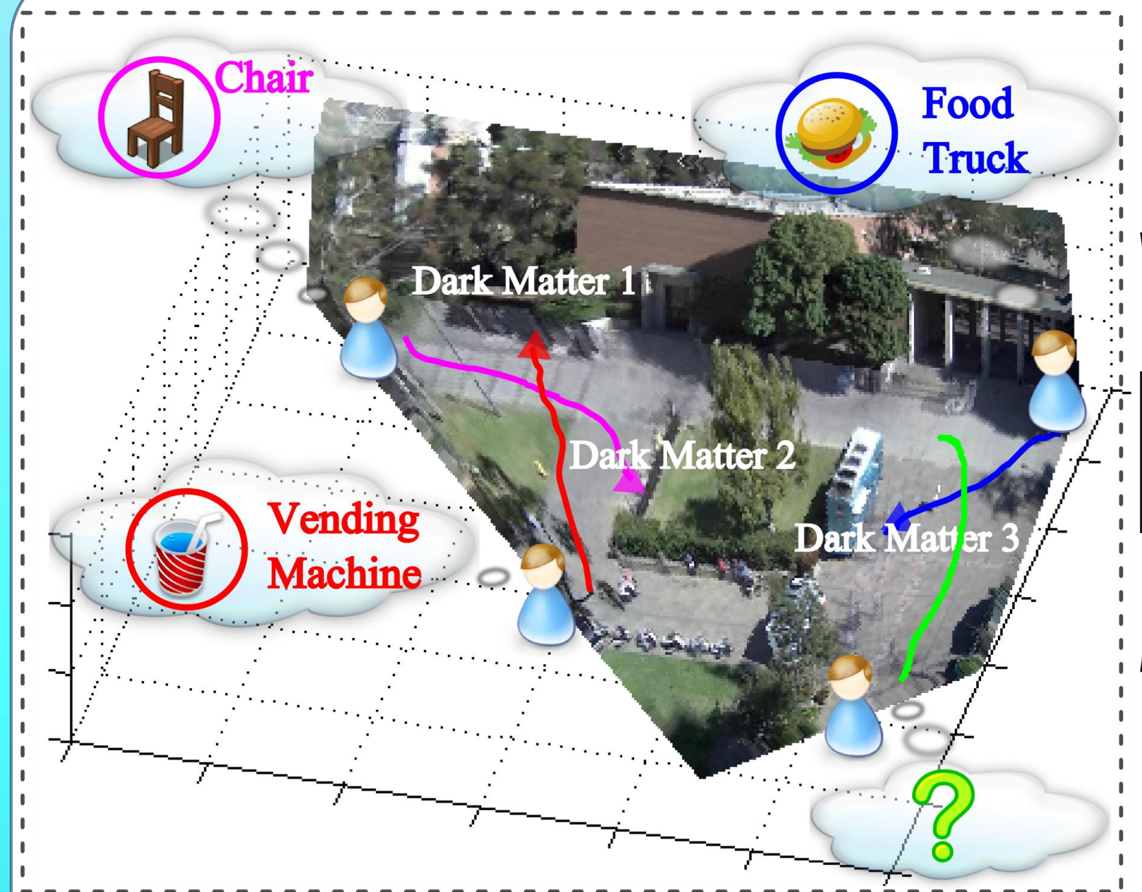
Allows a global estimation of the trajectories' end points

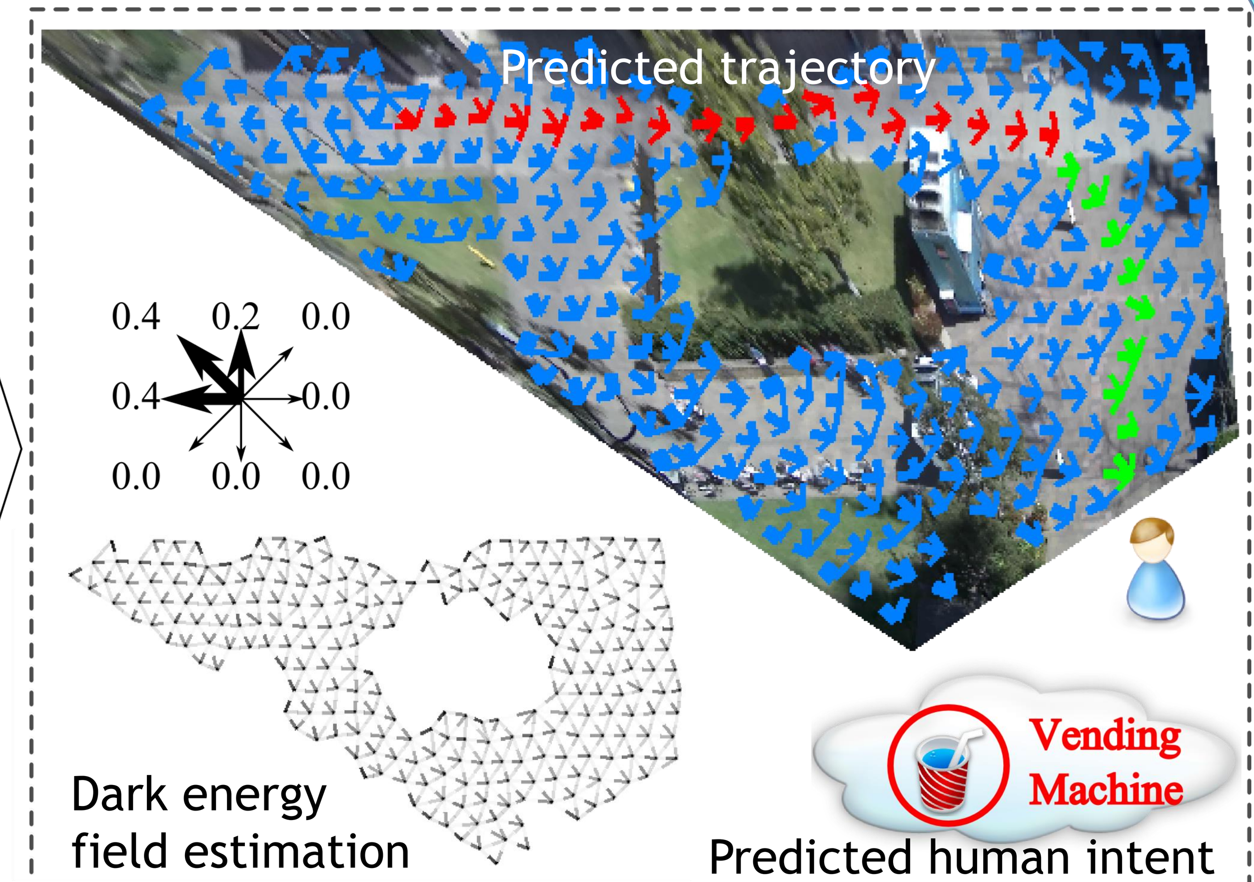
Given a video with partially observed trajectories of many people, use a Data-Driven MCMC to infer:

- ❑ Human mind = Intent to approach a particular "dark matter",
- ❑ Constraint map of the scene,
- ❑ "Dark energy" = Vector field that attracts/repels people
- ❑ End points of the trajectories = "Dark matter" locations.

## Contribution

Agent-based Lagrangian Mechanics cast within a Bayesian framework




## Modeling

Constraint Map: $P(C) \propto \exp[\underbrace{\beta \sum_{\mathbf{x}\in\Lambda, \mathbf{x}'\in\partial\mathbf{x}\cap\Lambda} c(\mathbf{x})c(\mathbf{x}')]}_{\text{Ising}}$

Human Goals: $P(R|S) = \underbrace{\prod_{j=1}^{N} \theta_j^{b_j}}_{\text{Multinomial}}$ $\quad b_j = \sum_{i=1}^{M} \mathbb{1}(r_{ij}=1)$, $\quad R = \{r_{ij}\}$

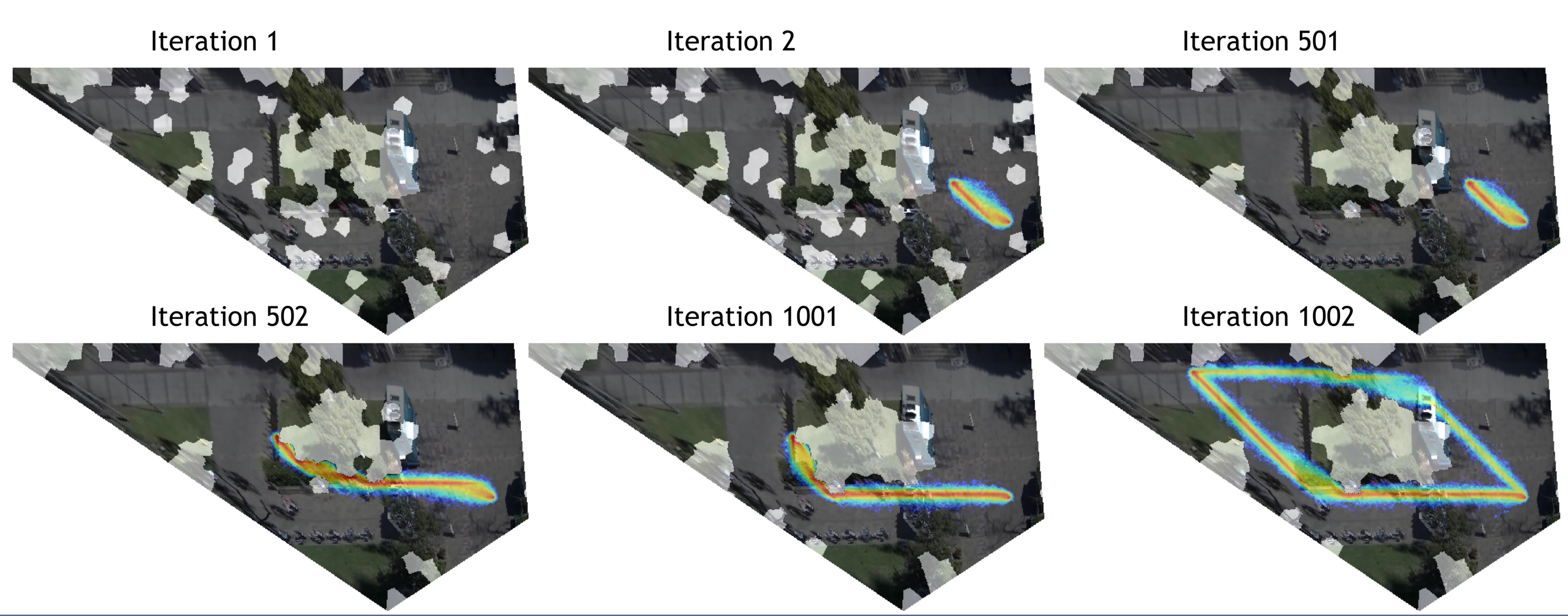
Dark Matter Locations: $P(S|C) \propto \underbrace{\frac{\eta^N}{N!} e^{-\eta}}_{\text{Poisson}} \underbrace{\prod_{j=1}^{N} \rho^{\frac{c(\boldsymbol{\mu}_j)+1}{2}} (1-\rho)^{\frac{1-c(\boldsymbol{\mu}_j)}{2}}}_{\text{Bernoulli}}$

Trajectory: $P(\Gamma_{ij}|C,S,r_{ij}{=}1) = P(\Gamma_{ij}|\vec{F}_{ij}(\mathbf{x})) \propto \exp[-\lambda \sum_{\mathbf{x}\in\Gamma_{ij}} ||\vec{F}_{ij}(\mathbf{x}(t))||\cdot||\vec{\Delta\mathbf{x}}(t)||]$

Posterior Distribution: $P(C,S,R,\Gamma|\underbrace{\Gamma^{(0)},I}_{\text{Partially observed trajectories, appearance features}}) \propto P(C,S,R) \prod_{i=1}^{M} \sum_{j=1}^{N} P(\Gamma_{ij}|C,S,r_{ij}{=}1)$

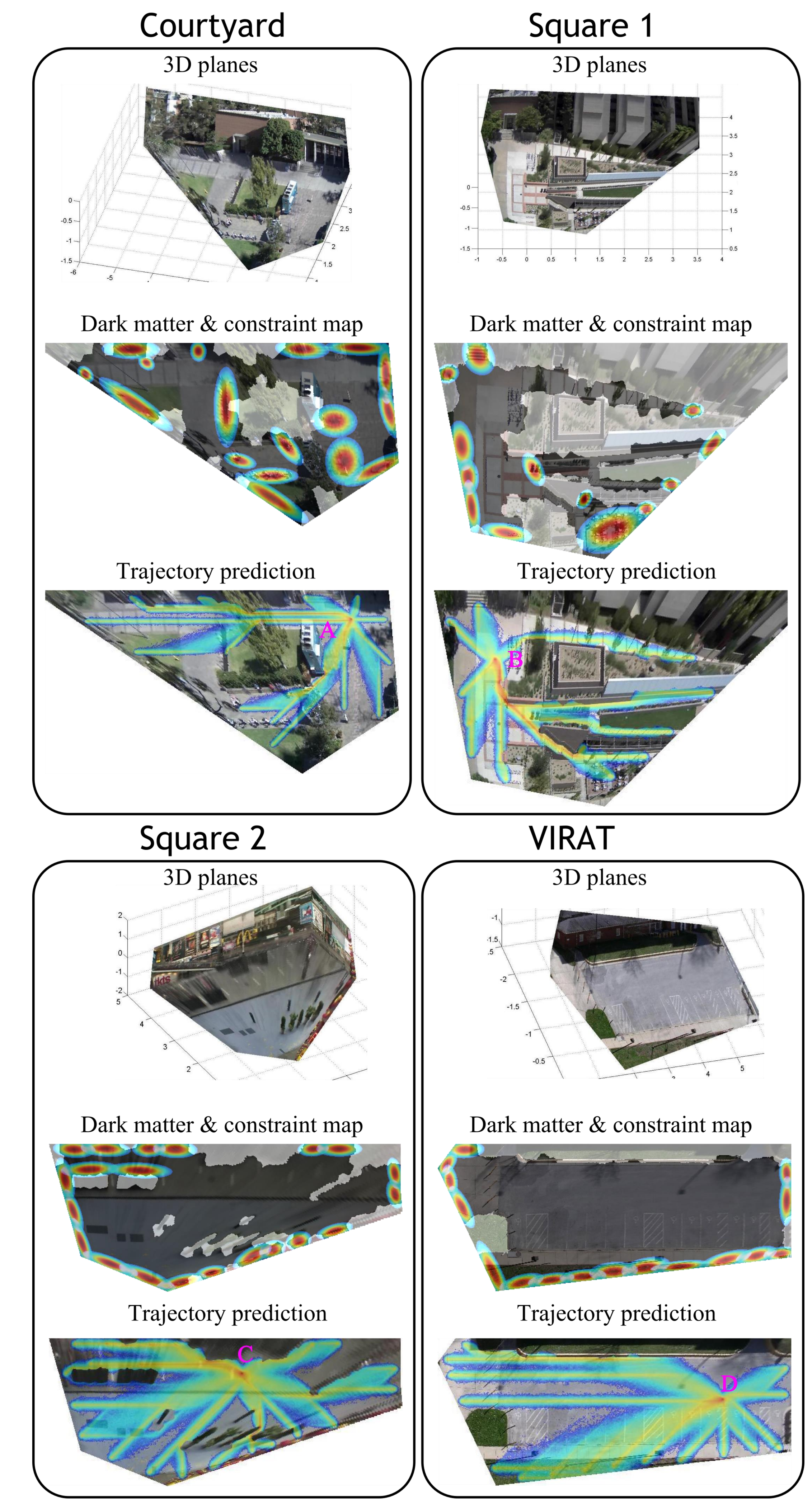
## MCMC Inference

MCMC samples of constraint map (gray), dark matter locations, human goals and trajectories (warm to cold colors).




## Experiments

Inference of 3D planes, dark matter locations, constraint map, and people's trajectories.




## Reference


[12] K. M. Kitani, B. D. Ziebart, J. A. Bagnell, and M. Hebert.Activity forecasting. In ECCV, 2012.

[23] Z. Tu and S.-C. Zhu. Image segmentation by data-driven Markov Chain Monte Carlo. TPAMI, 2002.

[26] Y. Zhao and S.-C. Zhu. Image parsing via stochastic scene grammar. In NIPS, 2011.


## Acknowledgement


DARPA MSEE FA 8650-11-1-7149 NSF IIS 1018751

ONR MURI N00014-10-1-0933 NSF IIS 1018490