



# Hamilton Institute

## Stochastic Spatio-Temporal Grammars for Images and Video

Jeffrey Mark Siskind

School of Electrical and Computer Engineering  
Purdue University

Friday, January 28<sup>th</sup>, 2005



### *Abstract*

Probabilistic Context-Free Grammars (PCFGs) induce distributions over strings. Strings can be viewed as observations that are maps from indices to terminals. The domains of such maps are totally ordered and the terminals are discrete. We extend PCFGs to induce densities over observations with unordered domains and continuous valued terminals.

We call our extension Spatial Random Tree Grammars (SRTGs). While SRTGs are context sensitive, the inside-outside algorithm can be extended to support exact likelihood calculation, MAP estimates, and ML estimation updates in polynomial time on SRTGs. We call this extension the center-surround algorithm. SRTGs extend mixture models by adding hierarchal structure that can vary across observations. The center-surround algorithm can recover the structure of observations, learn structure from observations, and classify observations based on their structure.

We have used SRTGs and the center-surround algorithm to process both static images and dynamic video. In static images, SRTGs have been trained to distinguish houses from cars. In dynamic video, SRTGs have been trained to distinguish events such as entering, exiting, picking up, putting down, sitting down, and standing up. We demonstrate how the structural priors provided by SRTGs support these tasks.

Joint work with Charles Bouman, Shawn Brownfield, Bingrui Foo, Mary Harper, Ilya Pollak, and James Sherman.

**Venue:** Seminar Room, Hamilton Institute, Rye Hall,  
NUI Maynooth

**Time:** 1.00 - 2.00pm (followed by tea/coffee)

Travel directions are available at [www.hamilton.ie](http://www.hamilton.ie)