# Interaction with an autonomous agent

Jim Little
Laboratory for Computational Intelligence
Computer Science
University of British Columbia
Vancouver BC Canada

# Surveillance and recognition

Jim Little

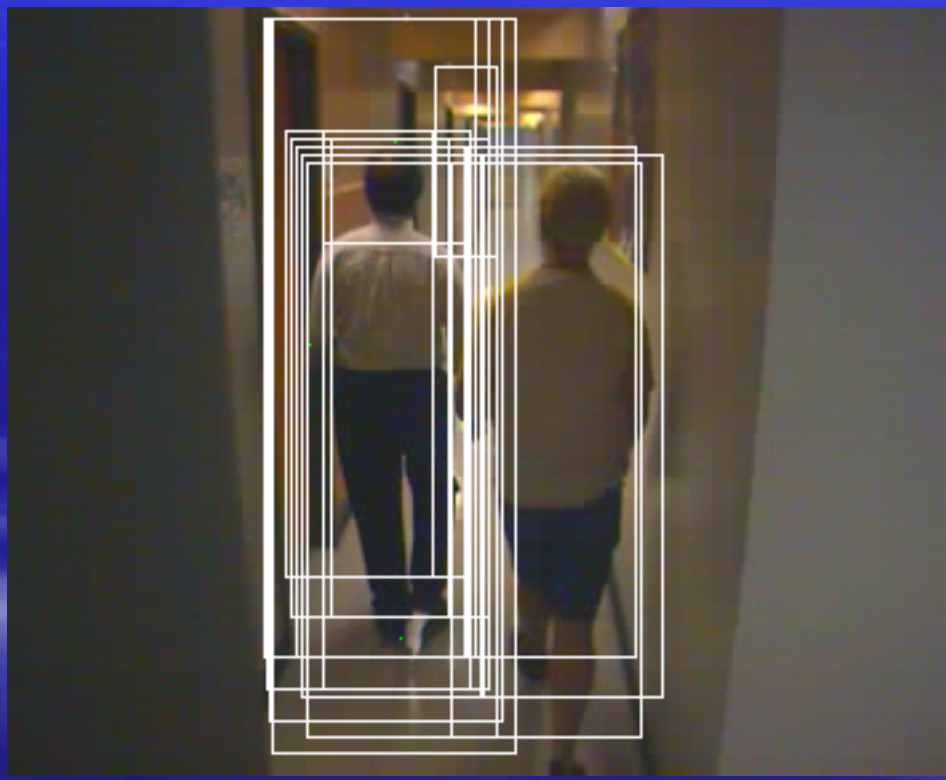
Work by Jim Clark and his students (McGill University)

# Multiple Camera Area Surveillance Techniques

- To distinguish "normal" people, objects, and activities from anomalous ones, and alert a security agent in the case of anomalous conditions.
- Our current projects are:
  - ***Fast people detection*** in corridor images, to be used as input for activity recognition.
  - ***Similarity filter*** development, for knowing when a given scene has been seen before.
  - ***Context-based*** object and scene recognition algorithms.

# People Detection Algorithm

- uses JPEG encoded images from a network camera. A *large* set of image features is computed, *without* the need for complete JPEG decompression.
- A support-vector machine (SVM) is used to classify image regions into people or non-people regions.
- A related project is developing an FPGA hardware implementation.



# Similarity Filters

- Design temporal filters which signal when a current input vector (e.g. image features) matches those present during one or more periods of time in the past.
- The similarity filter can be used in surveillance to detect anomalous scenes (e.g. a new person).
- Problem: avoiding the need to store all previous inputs. Our approach gets around this by determining a set of subspaces which represent prototypical inputs. These are subject to compression, leading to further reduction in storage needs.
- Remaining problems include: knowing when to create new prototypes, how to do the compression (signature analysis) and what image features are best.

# Context-based Object and Scene Recognition

- use low-level, low-resolution global image data ("*gist*") to provide statistical models for the identity of an object and the class of scene that it inhabits.
- we use Strongly-Coupled data fusion techniques, wherein Bayesian inference modules interact via modulation of prior statistical models and likelihoods.
- We are studying the dynamics of the interaction, which could possibly result in "rivalry" or multi-stable perceptions.
- We are working on a low-level version of this approach, wherein the perception of low-level features such as intensity and color are driven by statistical models of the surround

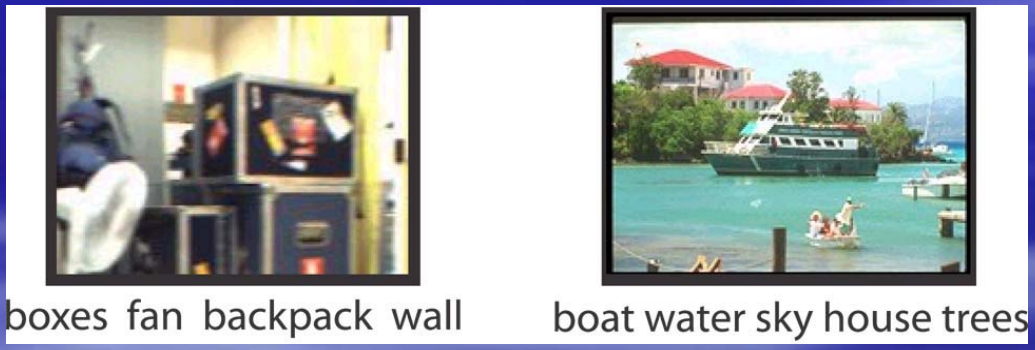
# Associating words with objects in images

Jim Little

Work by Nando de Freitas and
Peter Carbonetto

# Statistical Translation for Object Recognition

- statistical model for learning the probability that a word is associated with an object in a scene.
- learn these relationships *without* access to the correct associations between objects and words.




- a Bayesian scheme for automatic weighting of features (e.g., colour, texture, position) improves accuracy by preventing overfitting on irrelevant features.

# Contextual Translation

- Poor assumption: all the objects are independent in a scene.
- Our more expressive model takes context into account.
- Use loopy belief propagation to learn the model parameters. –
- On our Corel data set (www.cs.ubc.ca/~pcarbo), we achieve almost 50% precision.

# Object Recognition for Robots

- Our scheme is not real-time because of the expensive segmentation step.
- BUT: our contextual translation model + a fast, crude segmentation results in equal or better precision! Moreover, object recognition is more precise because the segments tend to be smaller.

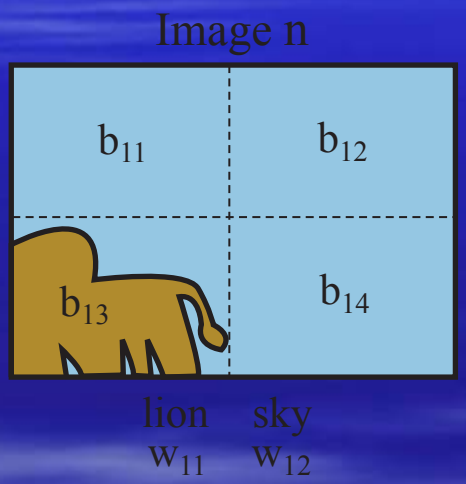
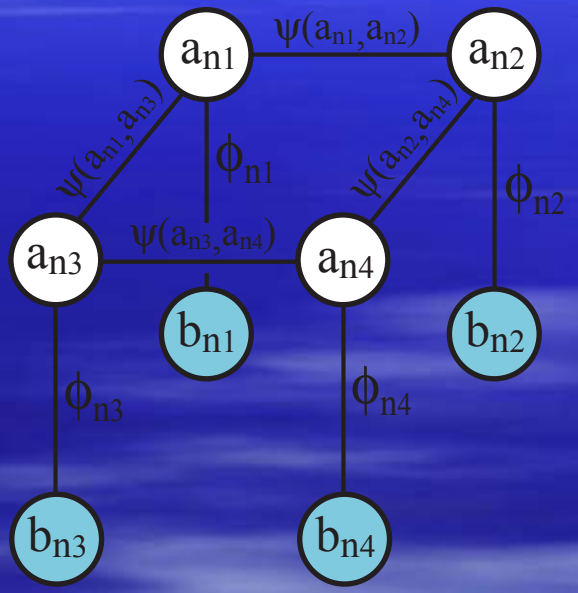


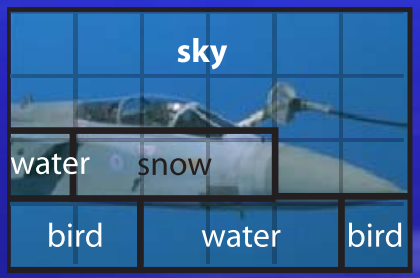
Image n
b11
b12
b13
b14
lion
sky
w11
w12


an1
an2
an3
an4
ψ(an1,an2)
ψ(an1,an3)
ψ(an2,an4)
ψ(an3,an4)
ϕn1
ϕn2
ϕn3
ϕn4
bn1
bn2
bn3
bn4

ORIGINAL

TRUE RESULT


tIND ANNOTATION




tMRF ANNOTATION


ORIGINAL

TRUE RESULT


tIND ANNOTATION


tMRF ANNOTATION

ORIGINAL

TRUE RESULT


tIND ANNOTATION

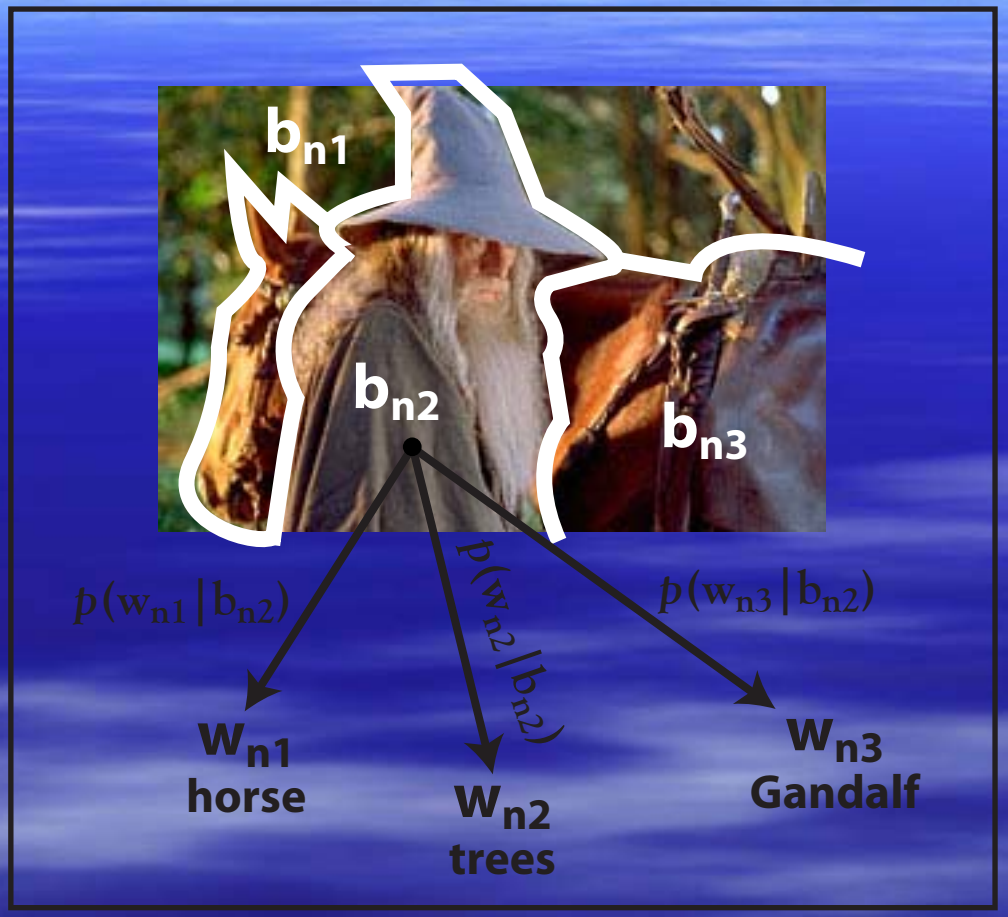

tMRF ANNOTATION

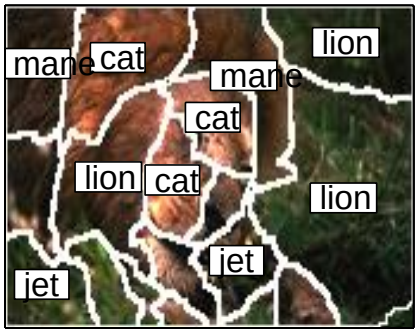
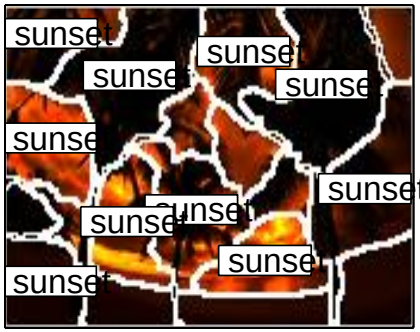
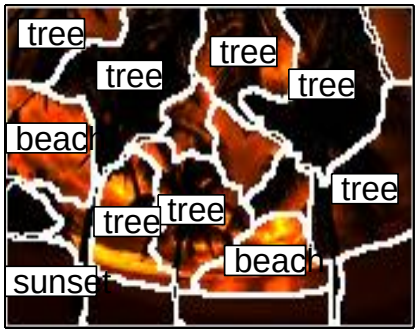
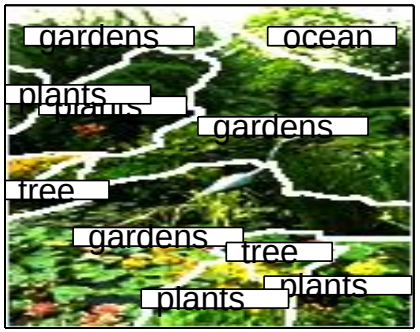

ORIGINAL

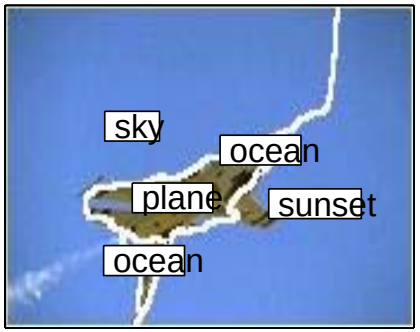
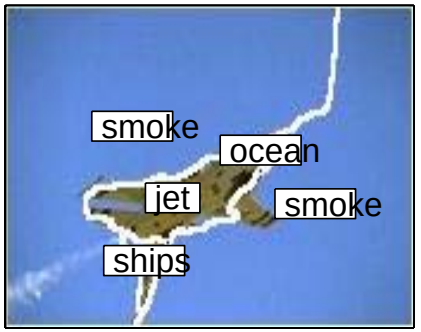
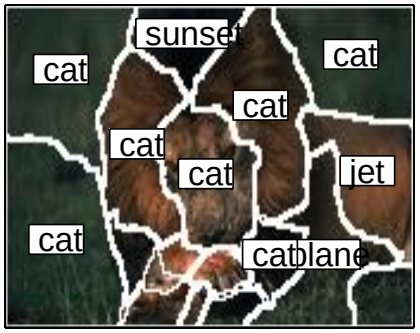
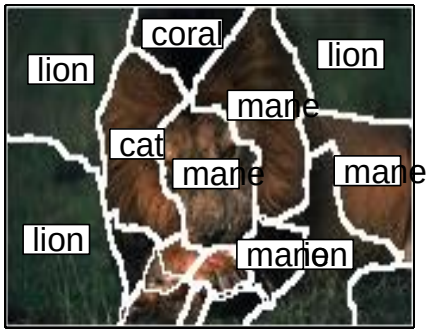
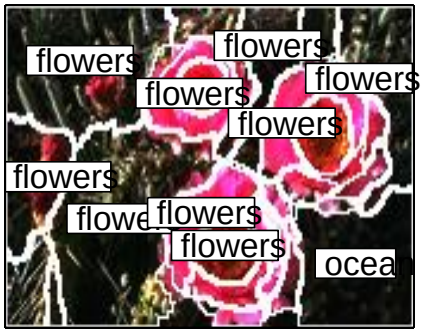
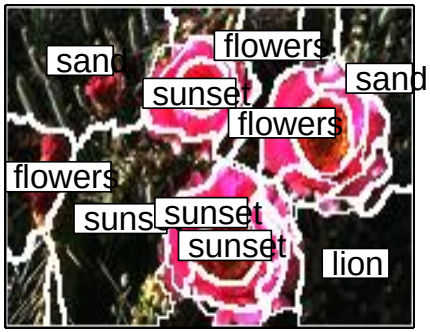
TRUE RESULT

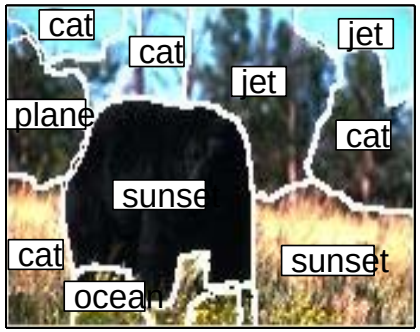
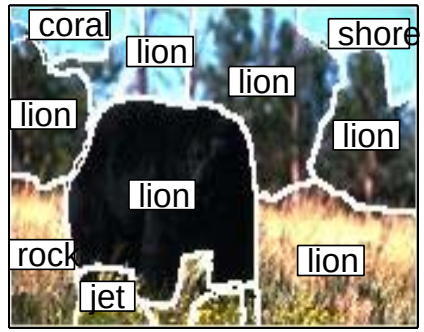
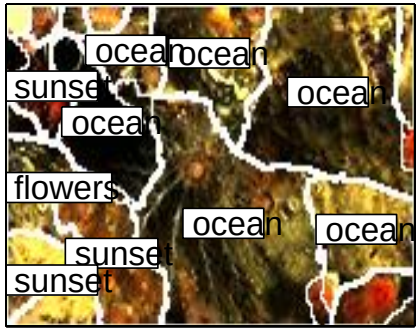
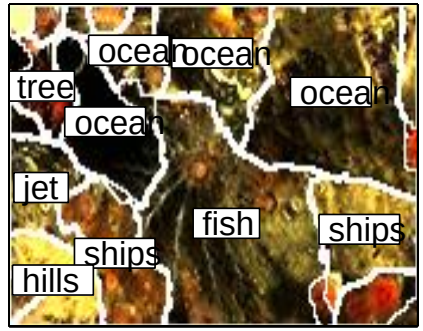

tIND ANNOTATION

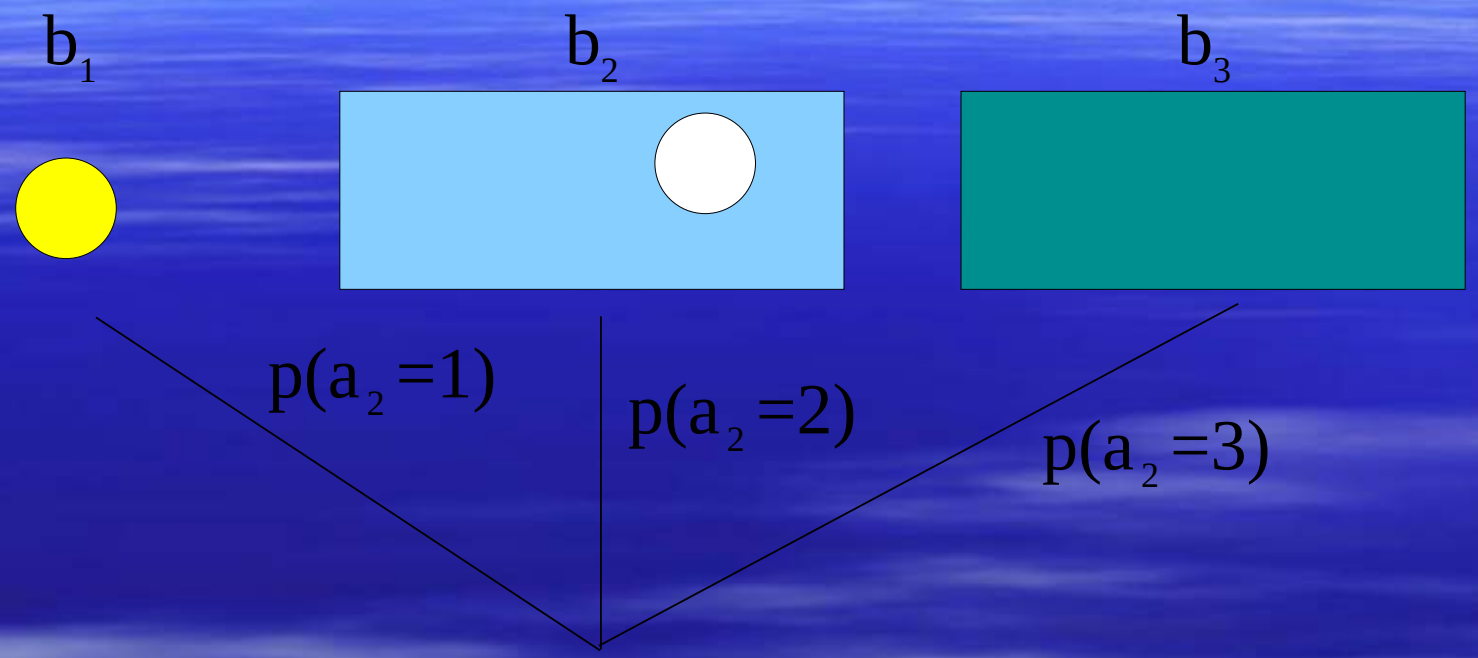

tMRF ANNOTATION

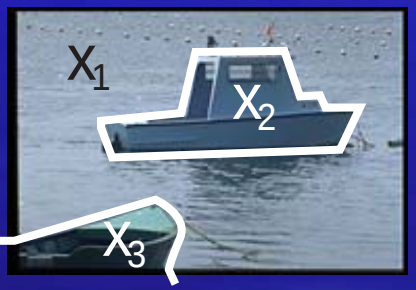

$b_{n1}$
$b_{n2}$
$b_{n3}$
$p(w_{n1}|b_{n2})$
$p(w_{n2}|b_{n2})$
$p(w_{n3}|b_{n2})$
$w_{n1}$ horse
$w_{n2}$ trees
$w_{n3}$ Gandalf

| Original | ML | MAP |
| --- | --- | --- |
|  |   |   |
|  |   |   |
|  |   |   |

| Original | ML | MAP |
| --- | --- | --- |
|  |  |  |
|  |  |  |
|  |  |  |

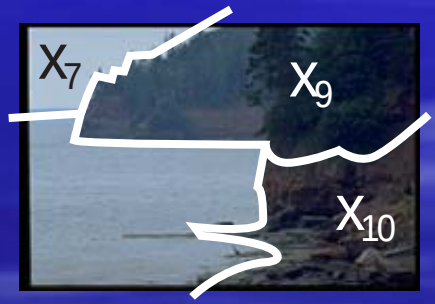
| Original | ML | MAP |
| --- | --- | --- |
|  |  |  |
|  |  |  |
|  |  |  |


b1
b2
b3
p(a2 =1)
p(a2 =2)
p(a2 =3)
sky
w1
sun
w2
sea
w3




water boat


sky water rock




water rock sky trees


trees boat ($X_1$