HAVSS Summer School

Topic Models and Temporal Activity Mining

Rémi Emonet – 2012-10-05

Content of the Lecture

- Introduction to topic models
  - basics
  - examples
  - extensions
- Example with audio data
- Temporal topic modeling

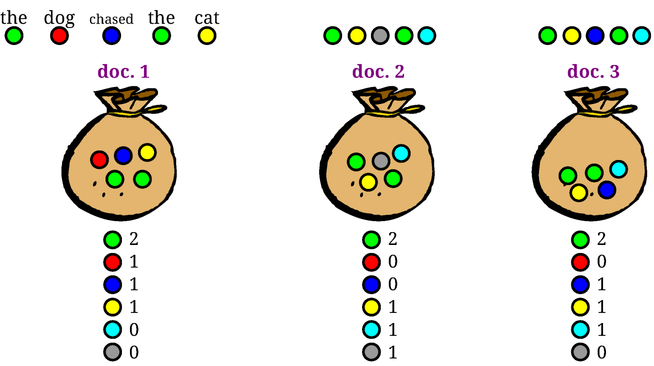
Origins of Topic Models

- Corpus of text documents
  - set of documents
  - documents made of words
- Goal
  - understand what documents are about
  - find “topics” shared by documents
  - do soft clustering of documents
  - unsupervised co-occurrence finding

Digression on Collections

- Sequence
  - [the, dog, chased, the, cat]
  - ordered, possible duplicates
- Set
  - {cat, chased, dog, the}
  - unordered, uniqueness
- Bag
  - (cat, chase, dog, the, the)
  - unordered, possible duplicates
  - {(cat,1), (chase,1), (dog,1), (the,2)}

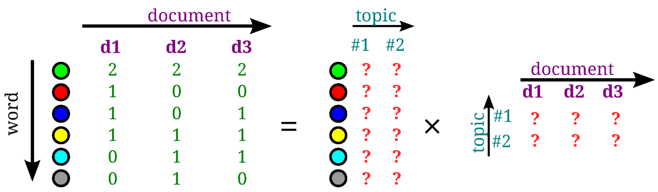
Topic Models: Bag of Words



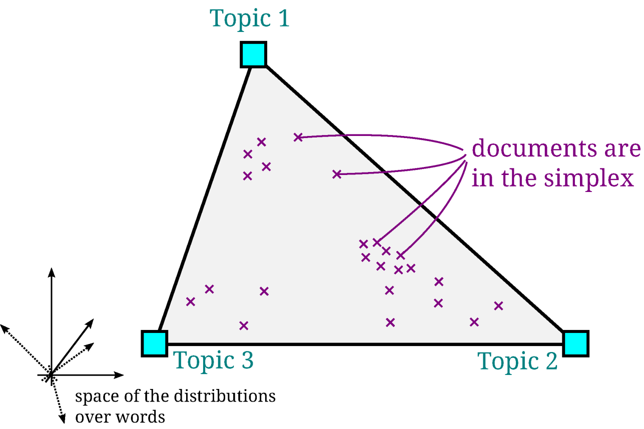
- Input = bags of words
- “Bag of Words” representation
  - document = bag

Topic Models: Matrix View

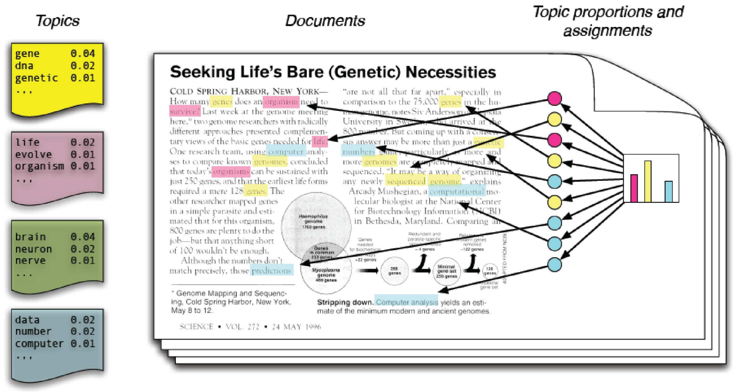
- Probabilistic Latent Semantic Analysis (PLSA)
  - matrix decomposition
  - non-negative
  - probabilistic interpretation: $p(w|d) = \sum_z p(w|z)p(z|d)$



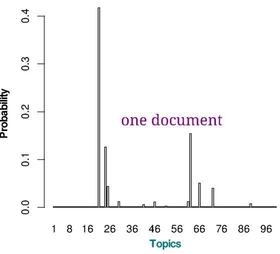
Topic Models: sub-simplex view



Another View



Example on Text Documents



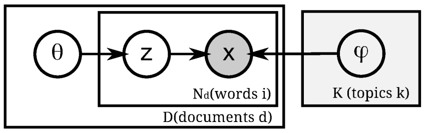
Four most probable topics in this document

"Genetics"	"Evolution"	"Disease"	"Computers"
human	evolution	disease	computer
genome	evolutionary	host	models
dna	species	bacteria	information
genetic	organisms	diseases	data
genes	life	resistance	computers
sequence	origin	bacterial	system
gene	biology	new	network
molecular	groups	strains	systems
sequencing	phylogenetic	control	model
map	living	infectious	parallel
information	diversity	malaria	methods
genetics	group	parasite	networks
mapping	new	parasites	software
project	two	united	new
sequences	common	tuberculosis	simulations

most probable words in each topic

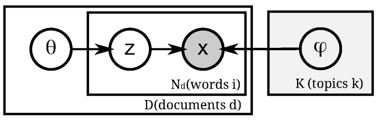
- PLSA
  - captures co-occurrence
  - handles polysemy
  - handles synonymy
- "Names" of topics are artificial

PLSA: Graphical Model



- **P**robabilistic **L**atent **S**emantic **A**nalysis
  - observations: x_{di}, a given word in a document
  - latent variables: z_{di}, the topic index of each observation
  - parameters: $\varphi_k = p(w|z=k)$ and $\theta_d = p(z|doc=d)$
- Generative process, $\forall d, i$:
  - draw z_{di} from $\text{Categorical}(\theta_d)$
  - draw x_{di} from $\text{Categorical}\left(\varphi_{z_{di}}\right)$
- Likelihood: $\prod_d \prod_i p(w = x_{di}|z = z_{di}) p(z = z_{di}|d)$

PLSA: Inference

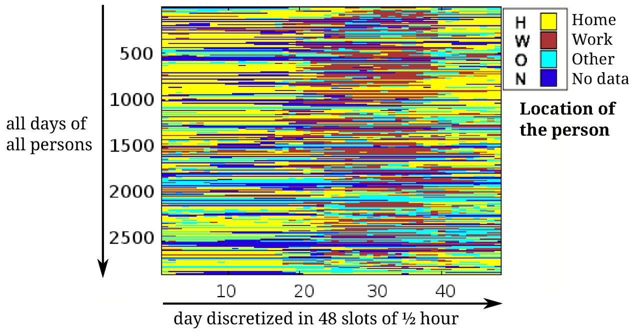


- Reminder
  - observations: x_{di}
  - latent variables: z_{di}
  - parameters: φ_k and θ_d
- An EM Algorithm can be derived
  - E: $\forall d, i$, compute the distribution (table): $p\left(z_{di}|x, \varphi^{t-1}, \theta^{t-1}\right)$
    $\forall d, i, \quad p(z_{di}|\ldots) \propto \theta_d^{t-1}(z_{di})\, \varphi_{z_{di}}^{t-1}(x_{di})$
  - M: find the new best parameters:
    $\left(\theta^t, \varphi^t\right) = \text{argmax}_{\varphi,\theta}\left(q\left(\theta, \varphi \middle| \theta^{t-1}, \varphi^{t-1}, z\right)\right)$

Application to Various Datatypes

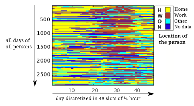
- PLSA on text
  - documents = bags of words
  - output: topic = co-occurring words
  - output: per doc. topic distribution
- On other datasets
  - need to define a vocabulary
  - need to define the documents

Ex. 1: Human Routines from Cell Phone Data (Farrahi, ISWC2008)



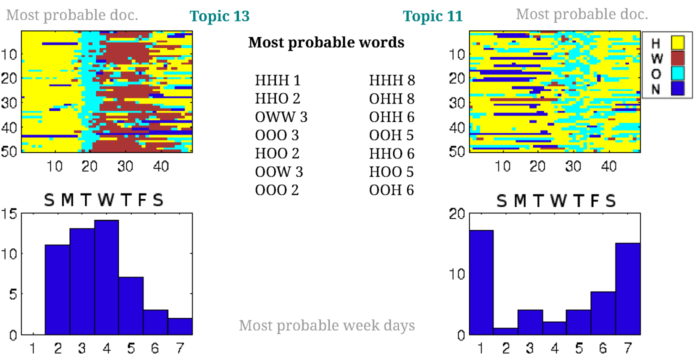
- Raw input: location of people during the experiment
- Goal: find daily routines

Vocabulary: Routines from Cell Phone Data



- Document: one day of one person
  - routines across persons
  - identity is lost
- Vocabulary
  - 8 timeslots: 0-7, 7-9, 9-11, 11-14, 14-17, 17-19, 19-21, 21-24
  - trigram of locations
  - word = trigram "+" one of the eight timeslots
  - e.g., HHH1, being at home for 1.5 hour before 7AM (slot 1)
  - vocabulary size: $4^3 \times 8$

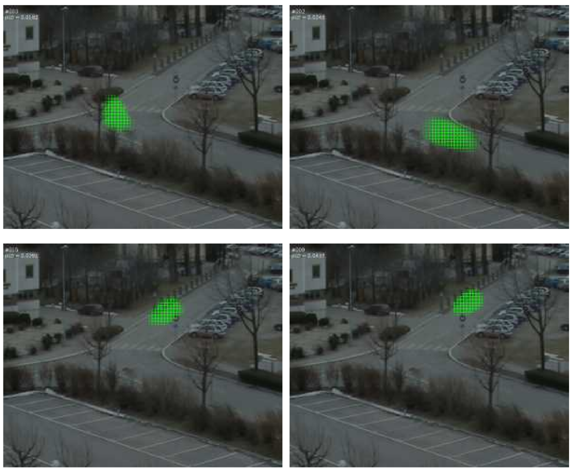
Results: Routines from Cell Phone Data



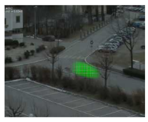
Ex. 2: Mining Visual Activities in Traffic Scenes

Mining Visual Activities in Traffic Scenes

- Vocabulary: localized motion
  - define region in the image (e.g. 75 of them)
  - word = presence of a motion pixel in a region



Mining Visual Activities in Traffic Scenes



- Vocabulary: localized motion
- Document: temporal window
  - accumulate motion over a temporal window (e.g., 5 seconds)
  - ignore temporal ordering within the window (bag of words)
  - 20 second long windows ⇛ what does a document contain?

Mining Visual Activities in Traffic Scenes

- Results?
  - 20 second long windows
  - what 2 topics would you expect?
- Viewing results
  - what happens with win. of 20 second, 2 topics?
  - what happens with win. of 10 second, 2 topics?
  - what happens with win. of 10 second, 3 topics?
  - what happens with win. of 5 second, 3 topics?
  - what happens with win. of 5 second, 5 topics?

PLSA: Summary

- Probabilistic Latent Semantic Analysis
  - inputs a set of documents, each being a bag of words
  - does co-occurrence analysis
  - finds topics defined as distribution over words

PLSA: Summary

- Probabilistic Latent Semantic Analysis
  - inputs a set of documents, each being a bag of words
  - does co-occurrence analysis
  - finds topics defined as distribution over words
- Comments
  - can be solved with EM (with pros and cons)
  - need to fix K, the number of topics
  - vocabulary definition? Documents definition?
  - bag
  - easy multi-modality

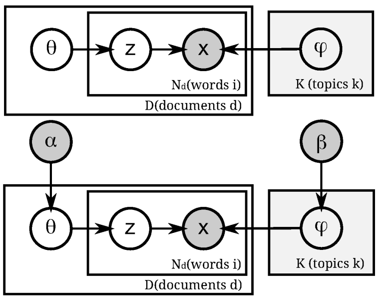
HAVSS Summer School

Beyond PLSA on Audio Data

by Bertrand Ravera – 2012-10-05

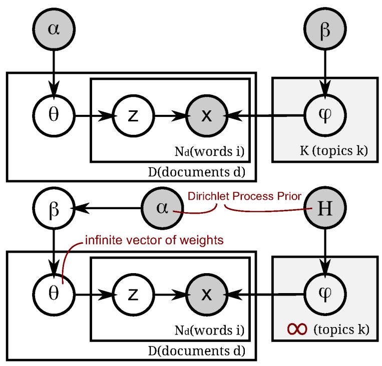
Extensions: LDA

- Motivation
  - PLSA is not fully generative
  - PLSA has no prior on θ and φ
- Latent Dirichlet Allocation
  - adds prior
  - fully generative
  - inference scheme
    - e.g., variational inference, Gibbs sampling (MCMC)
    - use of conjugate prior, Dirichlet/Categorical



Extensions: HDP-LDA

- Motivation
  - LDA needs K, the number of topics
  - need to remove “stop words“ (appearing too often)
- Hierarchical Dirichlet Process
  - “non-parametric” method
  - cleanly models a $K = \infty$
  - finds the “best” K
  - better handles high-frequency words



SUMMARY

- Topic models
  - unsupervised mining of “themes”
  - document = bag of words
- Evolutions
  - non-textual documents
  - mixed feature types
  - various models
- vocabulary definition? Documents definition?

HAVSS SUMMER SCHOOL

Temporal Activity Mining

Rémi Emonet – 2012-10-05

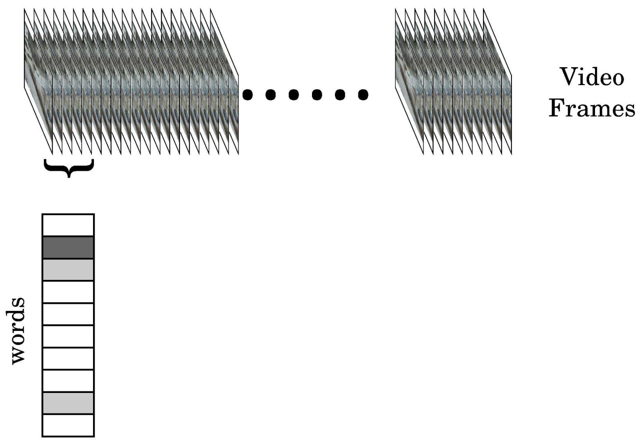
(RERE)SUMMARY

- Topic models
  - unsupervised mining of “themes”
  - document = bag of words
- Evolutions
  - non-textual documents
  - mixed feature types
  - various models
- vocabulary definition? Documents definition?

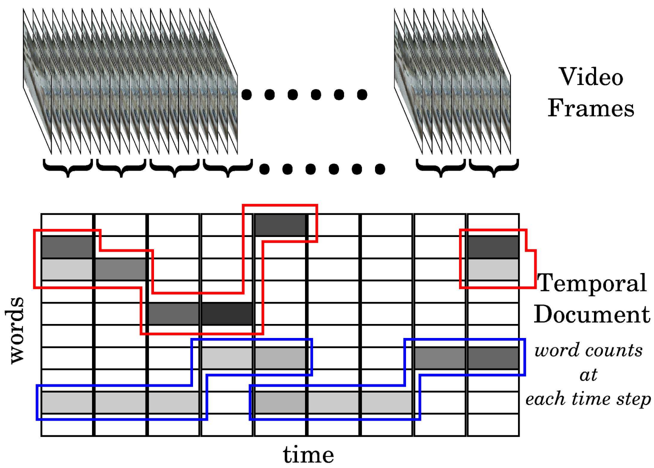
EXTENSIONS FOR TEMPORAL MODELING

- At vocabulary level
  - similar raw observations at different time ⇒ different word
- On top of the topic model
  - HMM over topic distributions (Hospedales, ICCV2009)
  - drifting topics
- Within the model
  - topic ⇒ motifs: “PLSM”
    (Varadarajan, BMVC2010, Emonet, CVPR2011)
  - with HMM and local rules: “MERM”
    (Varadarajan, CVPR2012)

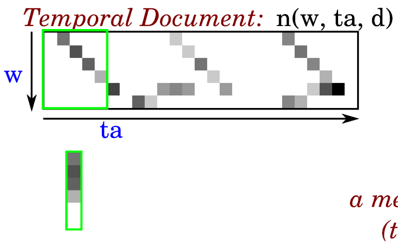
TEMPORAL DOCUMENTS



TEMPORAL DOCUMENTS

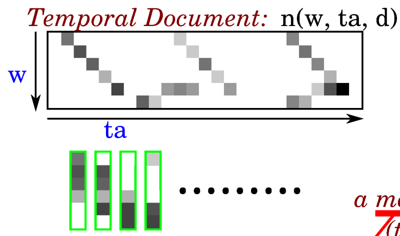


PLSA for Temporal Data



Documents from
a merging sliding window
(then use PLSA, etc.)

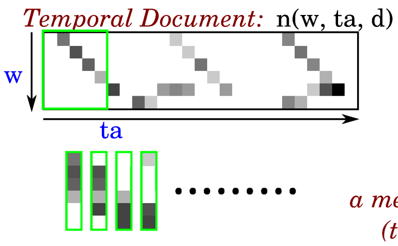
PLSA for Temporal Data



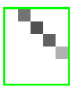
Documents from
a merging sliding window
(then use PLSA, etc.)

2 problems
- documents are not aligned
- sequence information is lost

PLSA for Temporal Data



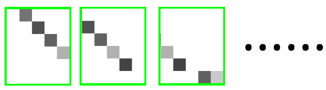
Documents from
a merging sliding window
(then use PLSA, etc.)



Documents from
a sliding window,
words include time
(then use PLSA, etc.)

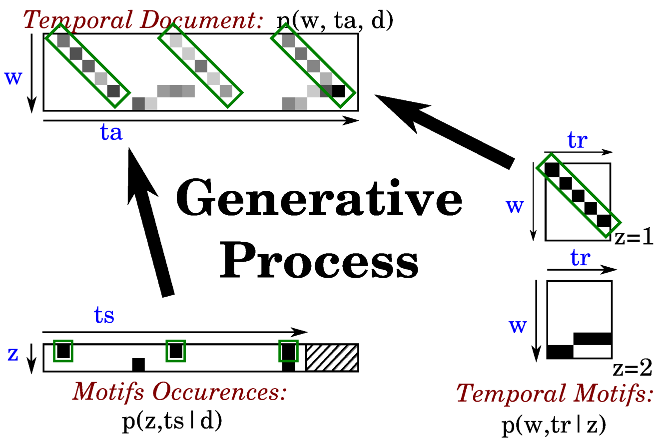
PLSA for Temporal Data

- sequence is preserved
- ok if document can be synchronized (e.g. traffic lights)

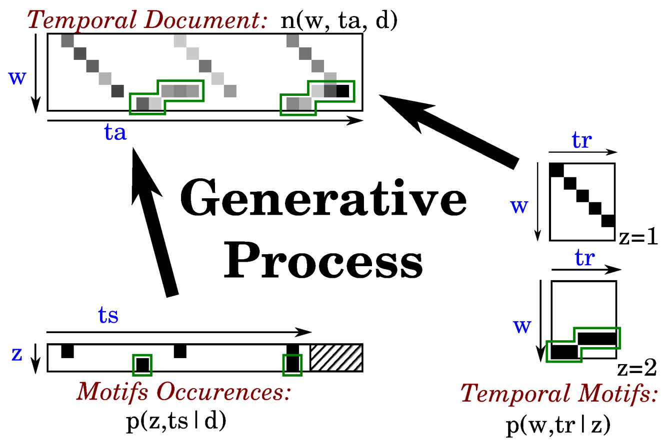


Documents from
a sliding window,
words include time
(then use PLSA, etc.)

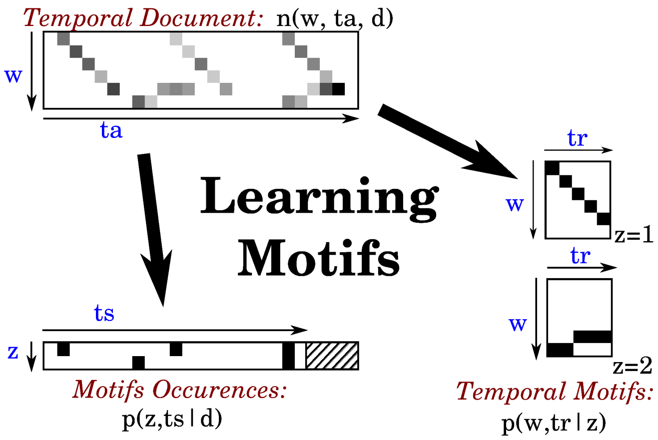
Probabilistic Latent Sequential Motifs



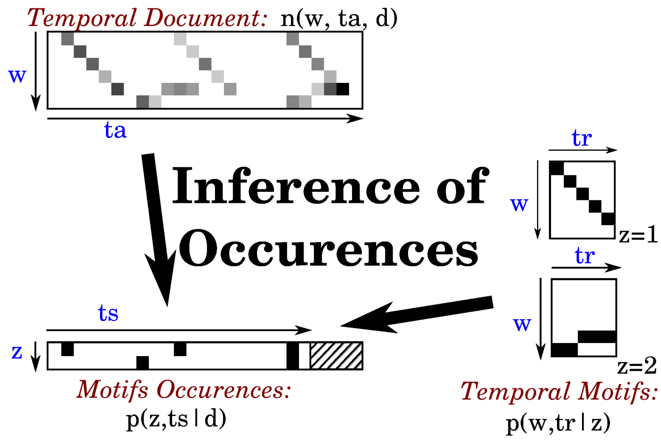
Probabilistic Latent Sequential Motifs



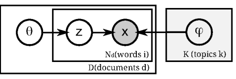
Probabilistic Latent Sequential Motifs

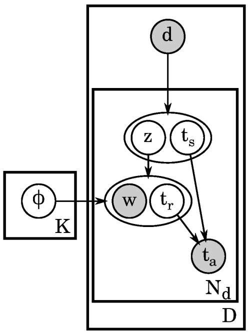


Probabilistic Latent Sequential Motifs



PLSM Graphical Model

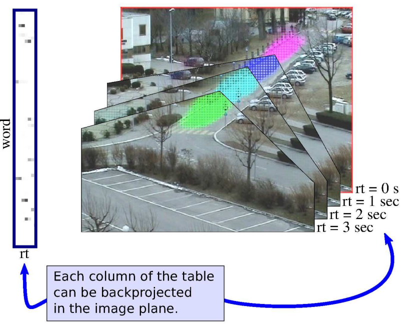




PLSM Results: Representation

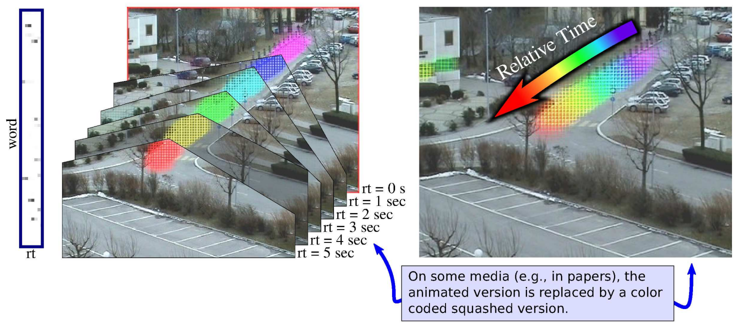
A motif recovered by the model is a plain probability table.
The vertical axis is the vocabulary.
The horizontal one is the time (relative to the motif beginning).

PLSM Results: Representation



Each column of the table can be backprojected in the image plane.

PLSM Results: Representation

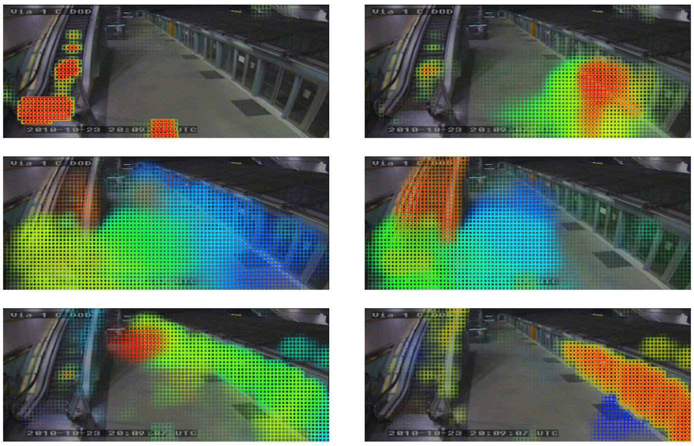


On some media (e.g., in papers), the animated version is replaced by a color coded squashed version.

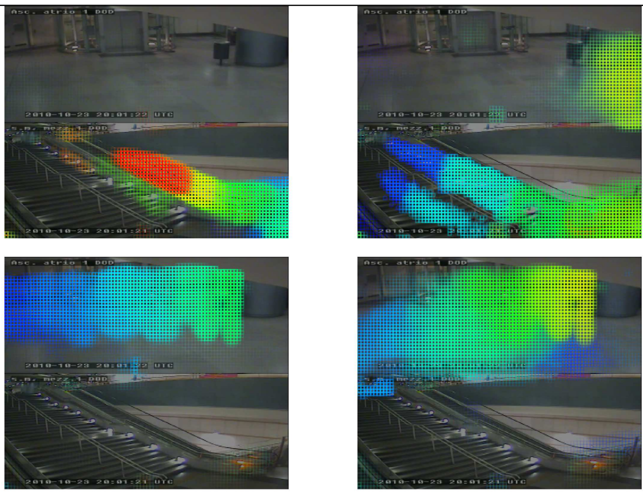
PLSM Results

- Traffic scenes
  - "rue"
  - "Kuettel"
  - "MIT"
- Metro station
  - Single camera
  - Multiple cameras
  - More cameras

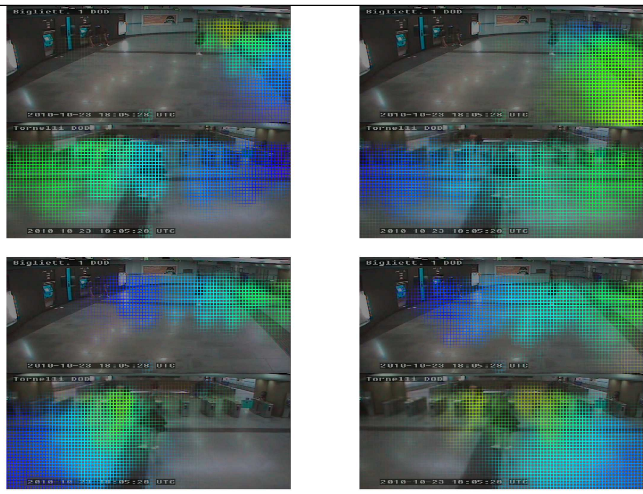
Example Motifs: platform



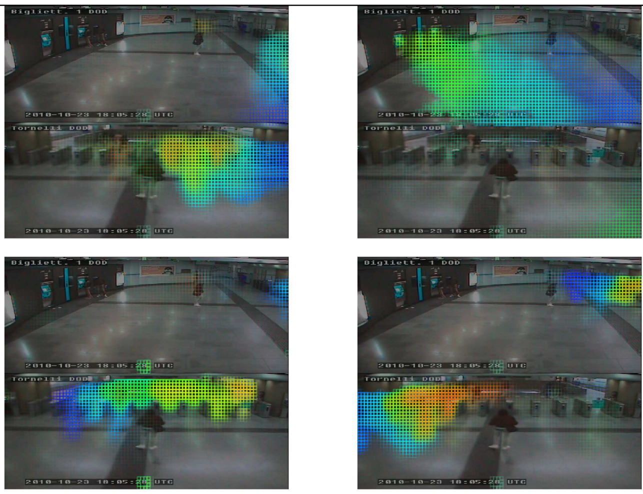
Example Motifs: mezzanine



Example Motifs: ticket hall



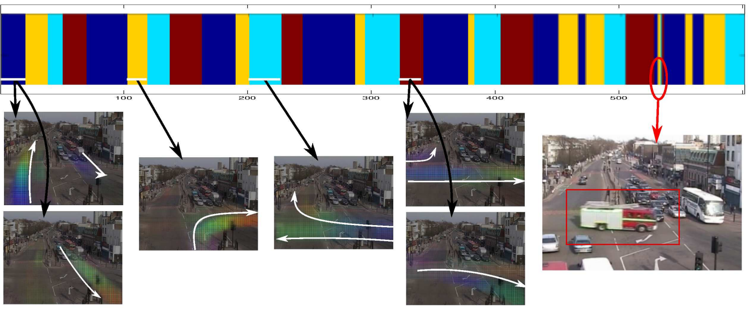
Example Motifs: ticket hall



PLSM Extensions / Limitations

- Absence of sparsity
- Fixed number of motifs
- Fixed motif duration
- No scene level cycle modeling

HMM Based Temporal Modeling



More

- About EM for PLSA?
- About Dirichlet Process?
- About Gibbs Sampling?
- About HMM/HSMM (semi-markov)?

Questions?