#### Learning Better Object Models using Video Data Patrick Li, Inmar Givoni, Brendan Frey

#### Motivation

Training on a collection of static monocular images is unnatural.

Labelled Training Images are hard to get. And the lack of is becoming a problem.

There is a wealth of video data available.

# First Attempt: Learning Bags of Features Models for Image Classification

#### Goal:

Represent Objects as Bags of SIFT Features

Use unsupervised learning to learn models of objects

Use learned models for image classification

# Image Classification

#### **INPUT:**



**OUTPUT:** 



#### **TRAINING:**



"Boat"

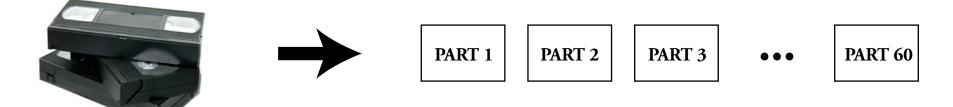
"Car"

"Sofa"

•••

## Overview of the Technique

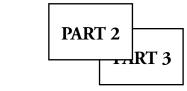
Unsupervised Training from Video



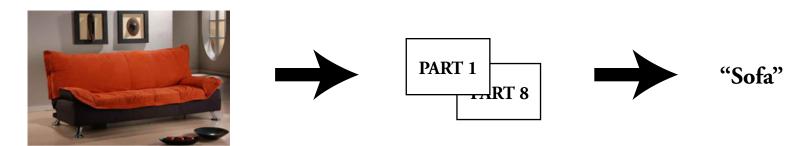
Supervised Training on Labelled Images



"Cow"

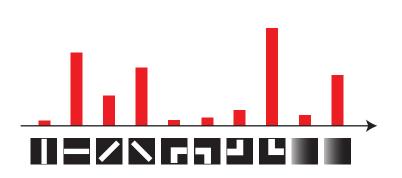


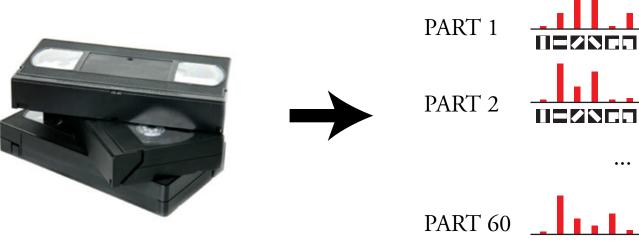
Testing



## Bags of Features Models

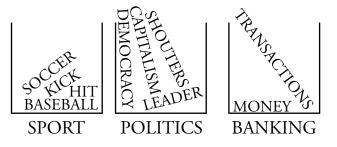








#### Latent Dirichlet Allocation for Topic Modelling



Convex Clustering with Exemplar-Based Models

Danial Lashkari Polina Golland

Computer Science and Artificial Intelligence Laboratory Massachusetts Institute of Technology Cambridge, MA 02139 {danial, polina}@csail.mit.edu

#### Abstract

Clustering is often formulated as the maximum likelihood estimation of a mixture model that ceptains the data. The bid appendix models used to solve the resulting optimization problem is inherently a gradient-descert method and is sensitive to initialization. The resulting solution is a local optimum in the neighborhood of the initial geness. This sensitivity to initialization presents a significant challenge in clustering inguing data sets into maxary clusters. In this spent, we present a dif-ferent approach to approximate mixture fitting for clustering. We introduce an exemption-based likelihood function that approximate the exact likelihood. This formulation leads to a convergence to the globally optimal models. The resulting clust-wich guaranteed more minimization problem and an efficient algorithm with guaranteed process prove the globally optimal models. The resulting clust-ecomplies that minimizes the vergat distance and the information-theoretic cost of mapping. We present experimently results likelihood to mixe model clustering. ring is often formulated as the maximum likelihood estimation of a mixture

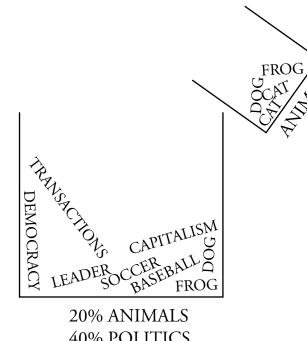
#### 1 Introduction

1 Introduction Clustering is one of the most basic problems of unsupervised learning with applications in a wide variety of fields. The input is either vectorial data, that is, vectors of data points in the feature space, are protouving data, the pairwise animitary or dissimity values between the data points. The determines the resulting clustering [1]. Intuitively, most methods seek, compare clusters of data points, namely, clusters with relatively small intra-cluster and high inter-cluster distances. Often approaches, such a Spectral Clustering [2], loak for clusters of more complex shapes bying on some low dimensional annialduk in the fratem space. These methods typically transform the data such set of the statement of the statement of the statement of the statement of the origin of the statement of the data statement of the statement is the statement of the statement of

on one remove one neero are encount compact-conservating teconiques such as e-means. The widely used 6th "emenses method is an instance of maximum likelihood fitting of a mixture model through the LM algorithm. Although this approach yields satisfactory results for problems with a small another of clastera at on a treatively task. Its use of a gradient descent algorithm for the scatch space grows, that is, the number of data points or clusters increases, it becomes harder the scatch space grows, that is, the number of data points or clusters increases, it becomes harder to large biological data sets such as gene-expression. Typically, one runs the algorithm many times with different random initializations and selects the best soution. More synderized initialization alloss of clustering for the different random initialization and selects the best soution. More synderizated initialization initialization and selects the best soution. More synderizated initialization alloss of the series of the seri methods have been proposed to improve the results but the challenge of finding go for EM algorithm remains [4].

We aim to circumvent the initialization procedure by designing a convex problem whose global optimum can be found with a simple algorithm. It has been shown that mixture modeling can

Single Document

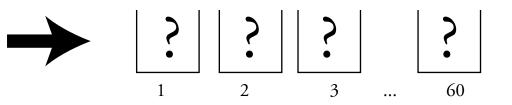


**40% POLITICS** 39% BANKING **1% SPORTS** 

#### Latent Dirichlet Allocation for Topic Modelling

|                                       | Convex Clustering with Exemplar-Based Models   |
|---------------------------------------|--|
|                                       | Danial Lashkari Pelina Golland<br>Computer Science and Artificial Intelligence Laboratory<br>Massachuseth Institute of Technology<br>Cambridge, MA 02159<br>{danial, polina}@cosail.mit.edu  |
| 1<br>C                                | Abstract<br>Clustering is often formulated as the maximum likelihood estimation of a mixture<br>model that explains the data. The EM algorithm widely used to solve the resulting<br>optimization, The resulting updaton is a board optimum in the method and<br>the initial games. This semilivity is initialization prevents a significant challenge<br>is the semilivity of the semilivity of the semilibrium of the semilibrium of<br>formst approach to approximate mixture litting for chardrage. We introduce an<br>exemptive based likelihood function that approximates the exact likelihood. This<br>formation leads to a search minimizing distance and the semilibrium of the set of<br>exemptive based is comparison to the set of<br>semigrates. The semicircle approximates that entering distance and the information-theoretic cost<br>of exemptive. Based as a probabilities temping of the data points to the set of<br>exemptive. The semicircle approximates is the artering distance and the information-theoretic cost<br>of exemptive. The semicircle approximates is a probabilities integrate of the data points to the set of<br>exemptive based is comparison with the conventional approach to mixture model<br>chartering. |
| · · · · · · · · · · · · · · · · · · · | 1 Introduction<br>Clustering is one of the most basic problems of unsupervised learning with applications in a wide<br>variety of fields. The input is either vectorial data, that is, vectors of data points in the feature<br>obscience of the clustering contact of the problem set of the set of the set of the<br>obscience of the clustering contact of the optimization algorithm employed to solve the problem<br>determines be resulting clustering [1]. Institutively, non-method sets data points, analy, clusters with relatively small intra-cluster and high intre-cluster datasets. Other<br>approaches, using a special clustering [2]. Institutively, non-method sets data points, and a special cluster gate large of data<br>points, namely, clusters with relatively small intra-cluster and high intre-cluster datasets. Other<br>approaches, using a special clustering [2]. Institutively, the state of the set on problem<br>also the manifold structures gate merged to point. The out cluster has also as a special cluster of<br>the state of the set of the state of the set on problem shares have a state and<br>be at the manifold structure compact-cluster data funding websites with a state of an a set on as.        |
|                                       | The widely used Soft Avagants method is an instance of maximum likelihood fining of a mininte<br>model through the IM algorithm. Albraugh this specific products further methods methods<br>with a small number of clusters and is relatively fast, in use of a gradient-descent algorithm for<br>minimization of a cost fraction with many local optimum makes it sensitive to initialization. As<br>the search space grows, that is, the number of data points or clusters increases, in becomes harder<br>to find a good initialization. This problem often annus in emerging applications of clustering for<br>large biological data sets such as gene-expression. Typically, one trans the algorithm many times<br>methods have been reproved to improve the results but the challenge off finding good initialization.   |

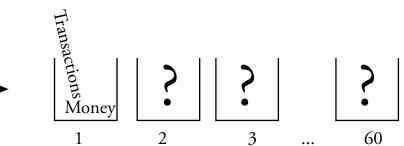
Corpus of Documents



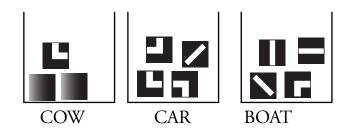
#### Latent Dirichlet Allocation for Topic Modelling

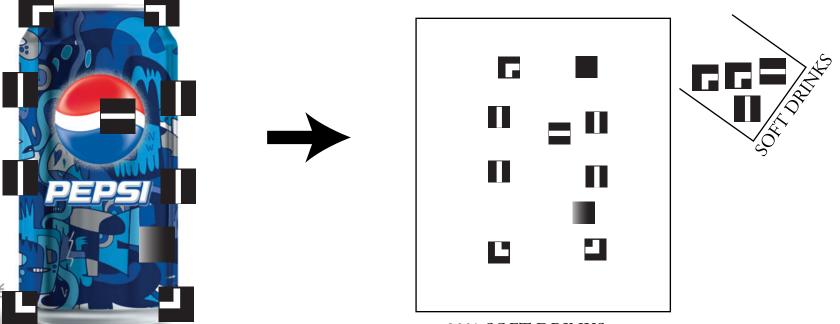


Corpus of Documents



# Latent Dirichlet Allocation for Object Modelling





90% SOFT DRINKS **10% CORPORATE LOGOS** 



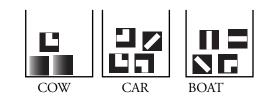
Single Image

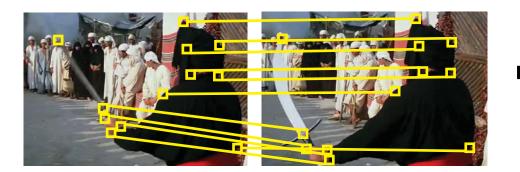
## Latent Dirichlet Allocation for Object Modelling



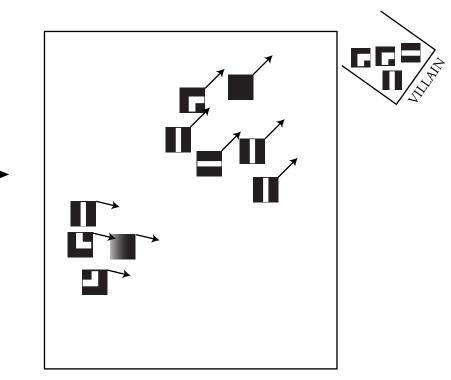
Image Collection

## Flow-LDA for Motion Modelling



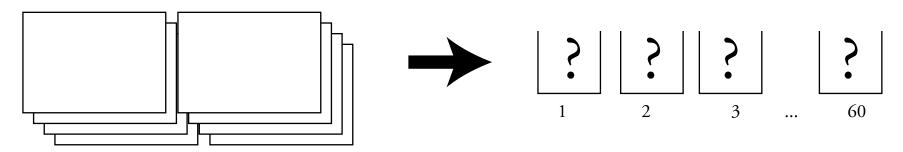


Pair of Consecutive Frame Pairs



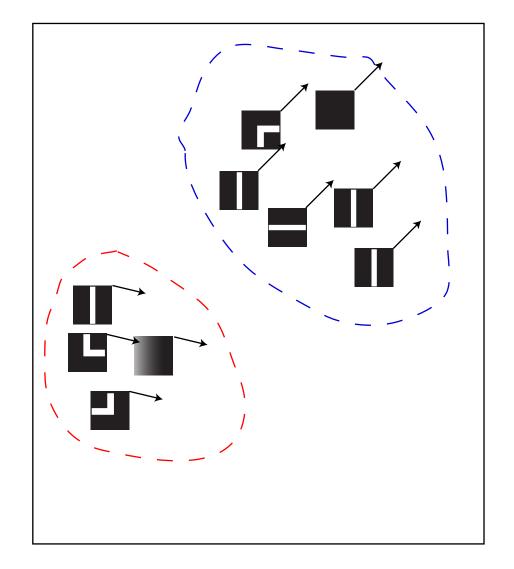
50% SWORD 50% VILLAIN

#### Flow-LDA for Motion Modelling

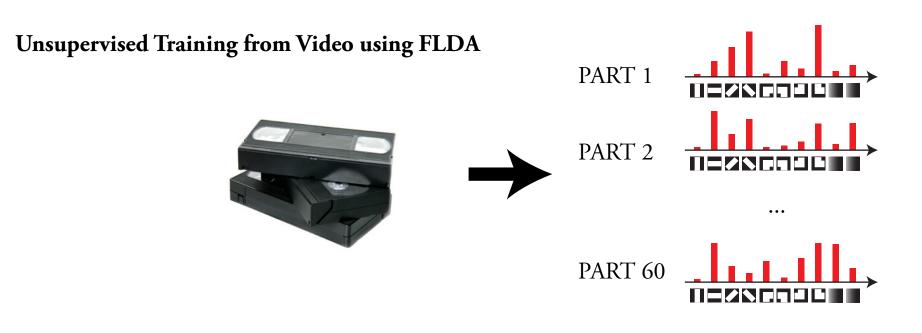


Frame Pair Collection

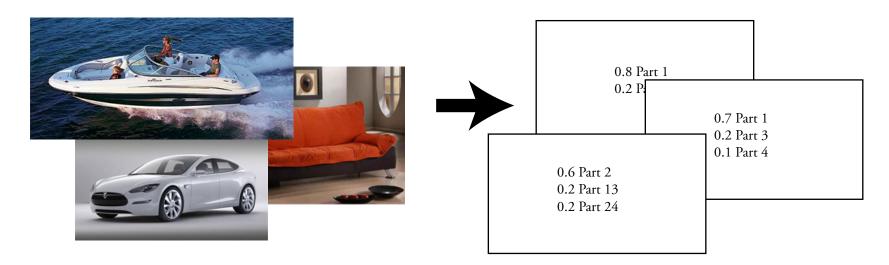
### Flow-LDA for Motion Modelling



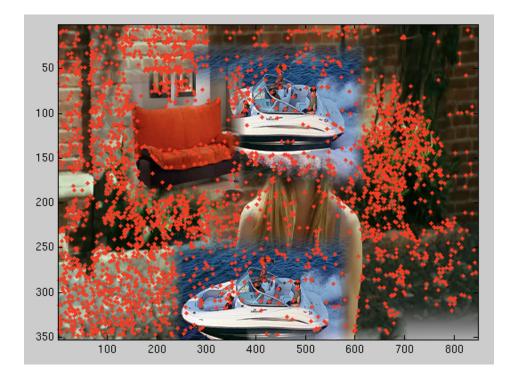
### Image Recognition



#### Training And Testing Images



#### Initial Results



Naive Guesser: 8.6% Error SVM trained on SIFT histograms directly: 8.6% Error SVM trained using LDA model (no motion): 5.6% Error SVM trained using FLDA model (motion): 3.7% Error

#### ... to continue

Experiment on Real Dataset

Go beyond Bags of Features models -Hierarchical Models -Account for Spatial Relations -Account for temporal relations between more than 2 frames

# Thank you!