# Combining Information Sources for Video Retrieval


## Lowlands Team

Thijs Westerveld

Tzvetanka I. Ianeva

Liudmila Boldareva

Arjen P. de Vries

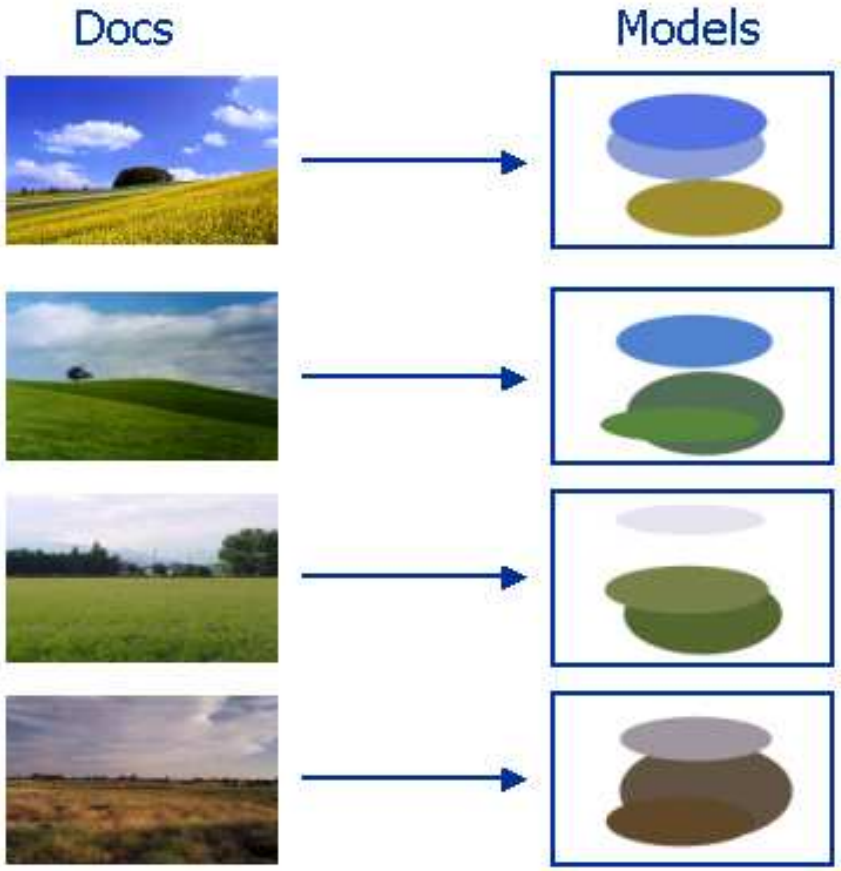
Djoerd Hiemstra


TRECVID workshop, 18 November 2003

# Introduction

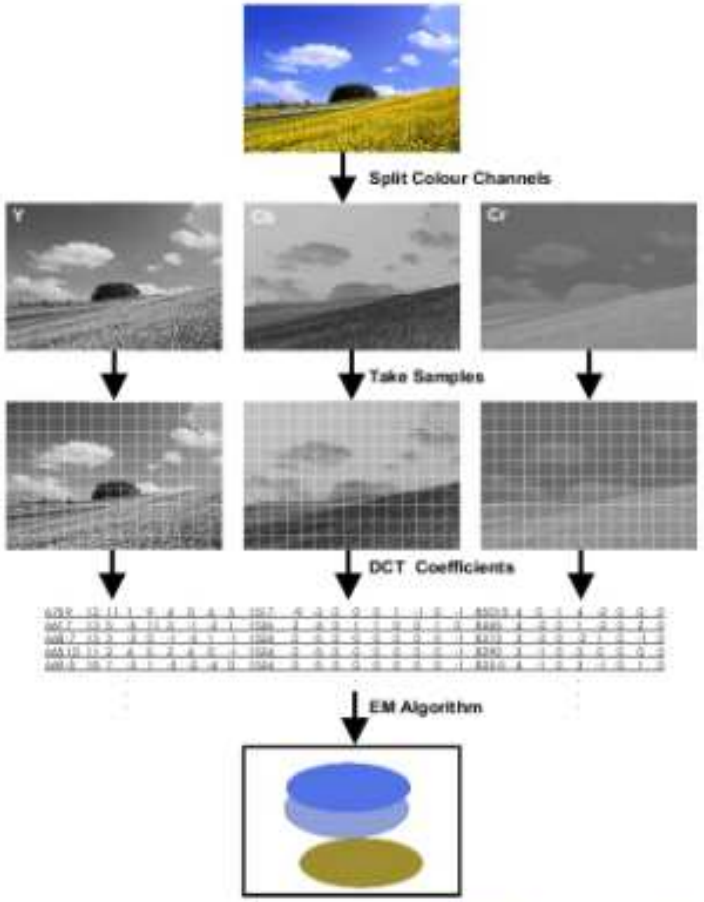
- Video Retrieval should take advantage of information from all available sources and modalities
  - …but so far ASR best for almost any query
- LL11@TRECVID2003: Combining information sources
  - Different models/modalities
  - Multiple example images
  - Model similarity and human-judged similarity

# Generative Probabilistic Model




- Indexing
  - Estimate Gaussian Mixture Models from images using EM
  - Based on feature vector with colour, texture and position information from pixel blocks
  - Fixed number of components

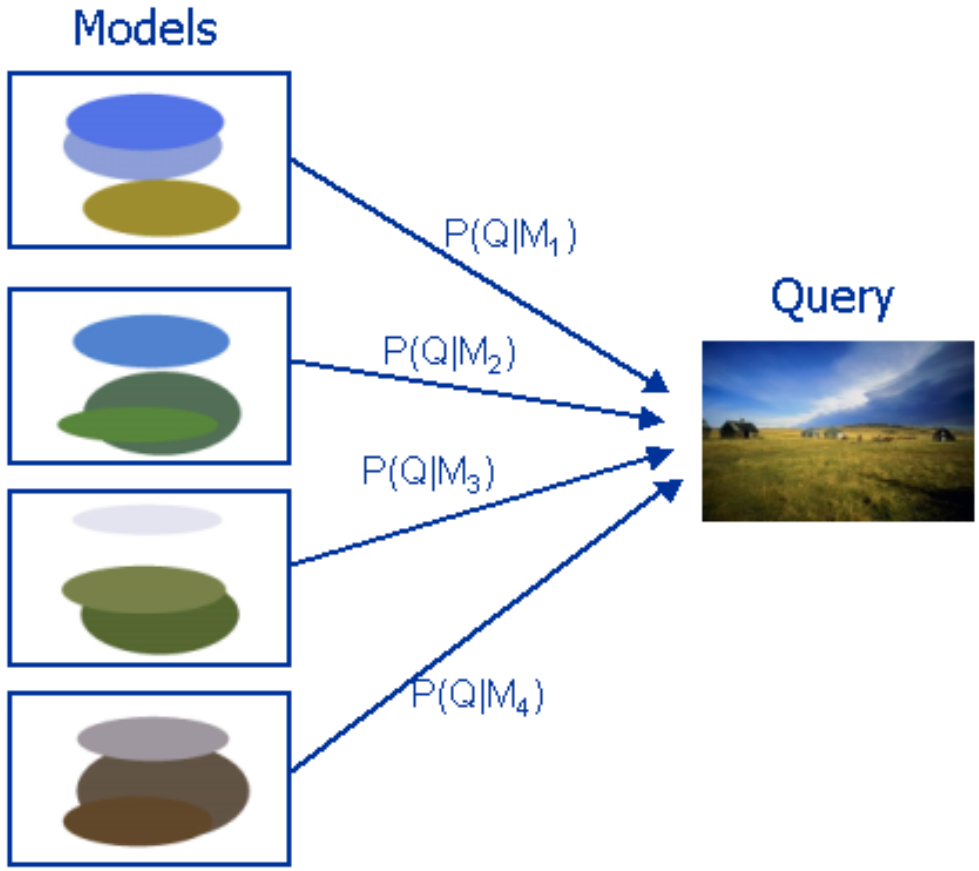
# Generative Probabilistic Model



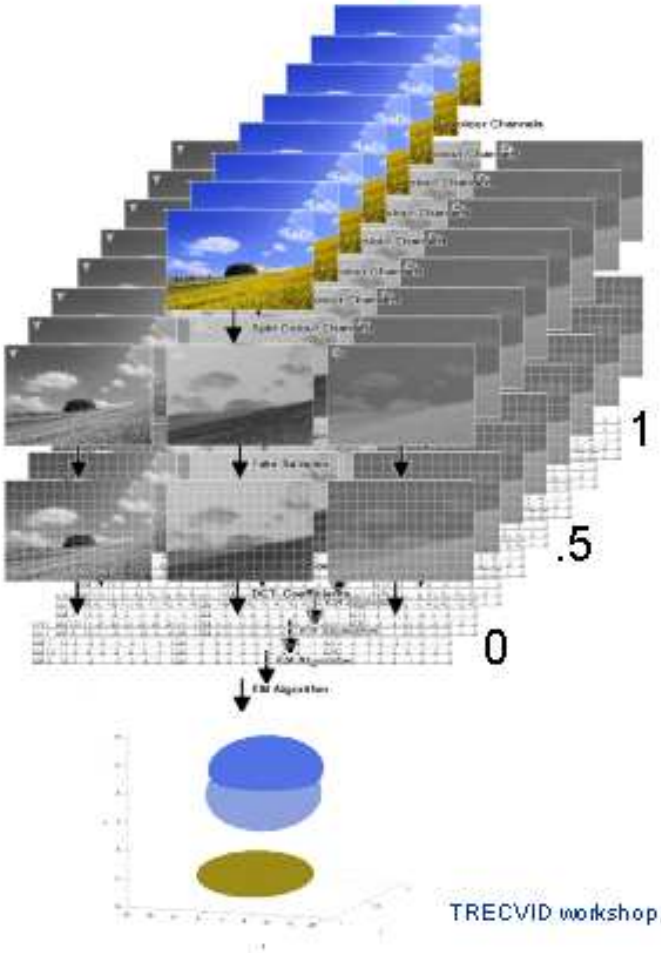

- Indexing
  - Estimate Gaussian Mixture Models from images using EM
  - Based on feature vector with colour, texture and position information from pixel blocks
  - Fixed number of components

# Generative Probabilistic Model

- Retrieval
  - –Calculate conditional probabilities of query samples given models in collection

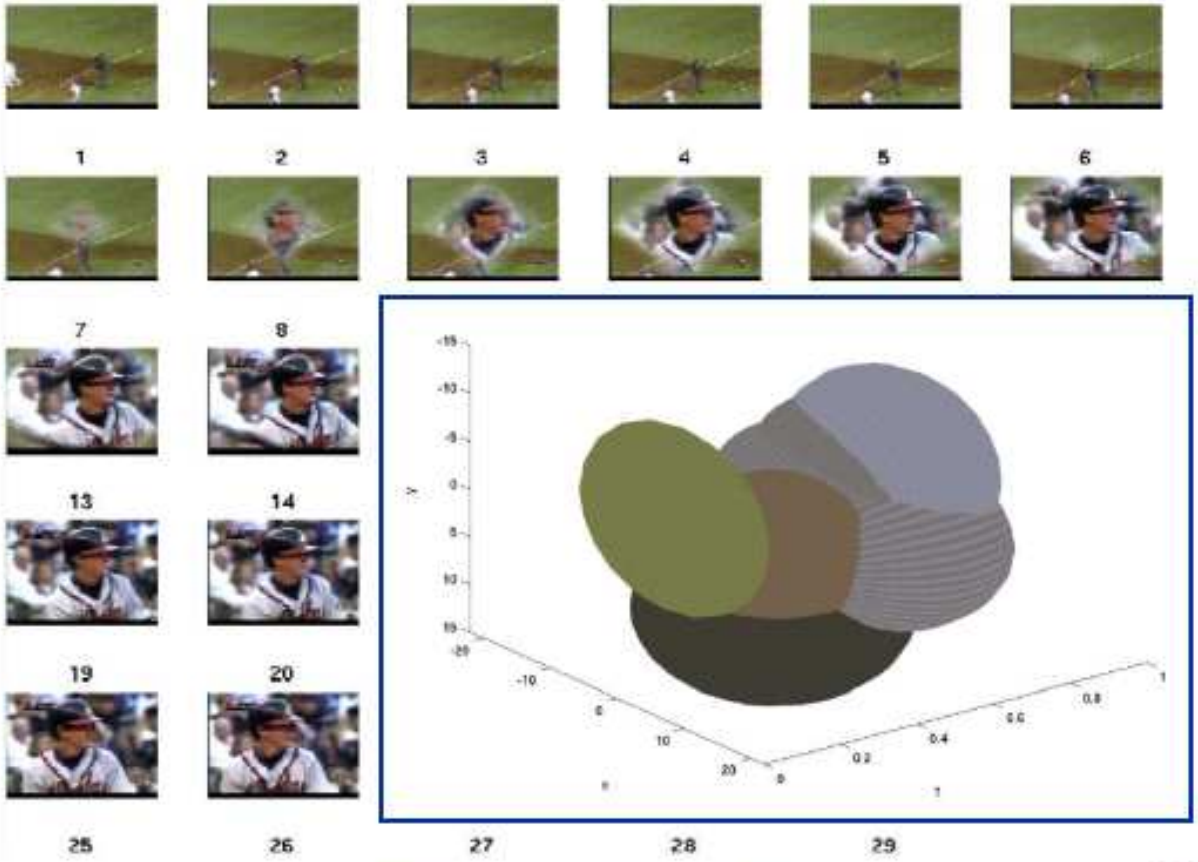


# Dynamic Model




- GMM from multiple frames around KF
  - Feature vectors for each frame
  - Add time info
  - EM
- Dynamic model capture spatio-temporal information

# Dynamic Model



# Experimental Set-up

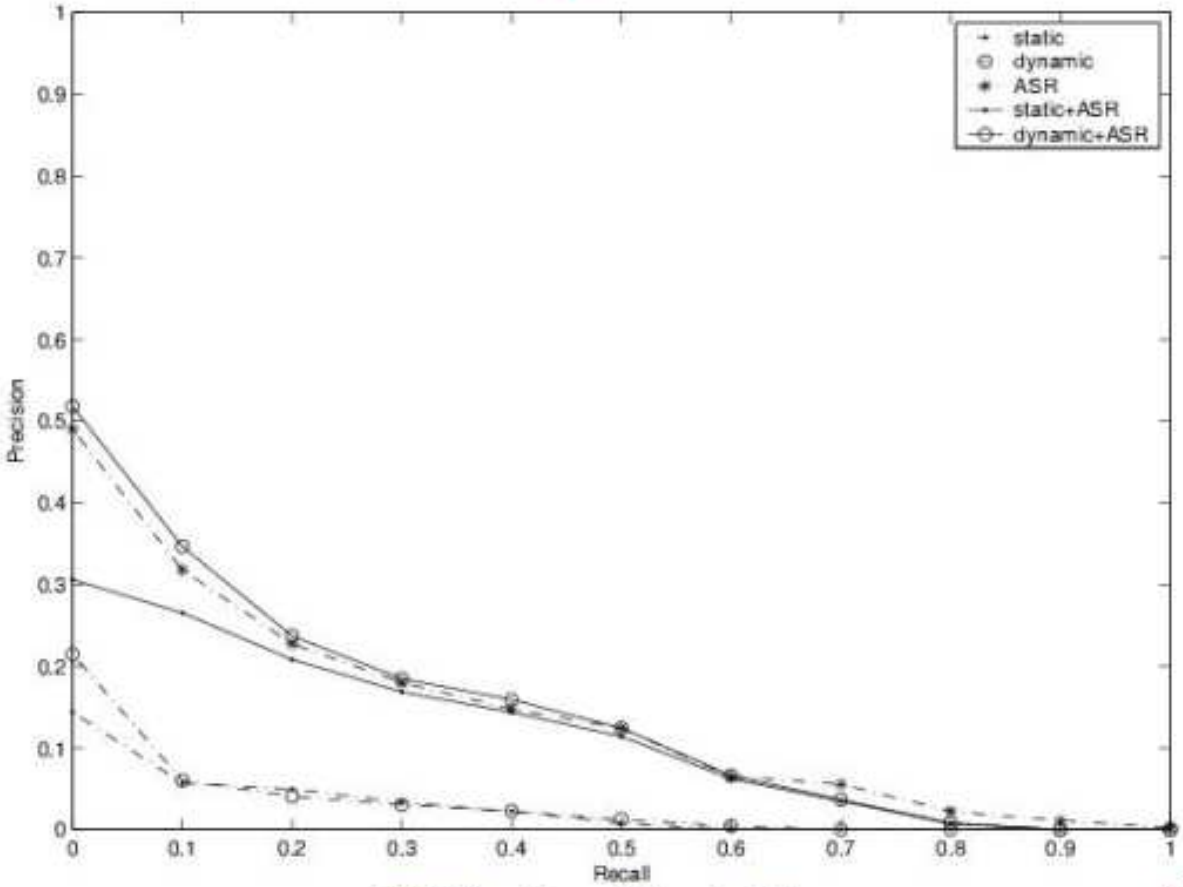
- Build models for each shot
  - Static, Dynamic, Language
- Build Queries from topics
  - Construct simple keyword text query
  - Select visual example
  - Rescale and compress example images to match video size and quality

# Combining Modalities

- Independence assumption textual/visual
  - $P(Q_t, Q_v|Shot) = P(Q_t|LM) * P(Q_v|GMM)$
- Strategy works well if both runs useful
  [CWI:TREC:2002]
- Dynamic run useful
- Static run not

| Run | MAP |
|---|---|
| ASR only | .130 |
| Static only | .022 |
| Static+ASR | .105 |
| Dynamic only | .022 |
| Dynamic+ASR | .132 |

# Combining Modalities



# Combining Modalities

- Textual

*"Dow Jones Industrial Average rise day points"*

+

- Visual

=

# Merging Run Results

- Combining (conflicting) examples difficult
  [CWI:TREC:2002]
- Single example → Miss relevant shots
- Round-Robin Merging
- MAP
  - All examples: .031
  - Combination .041
  - Single example: .022

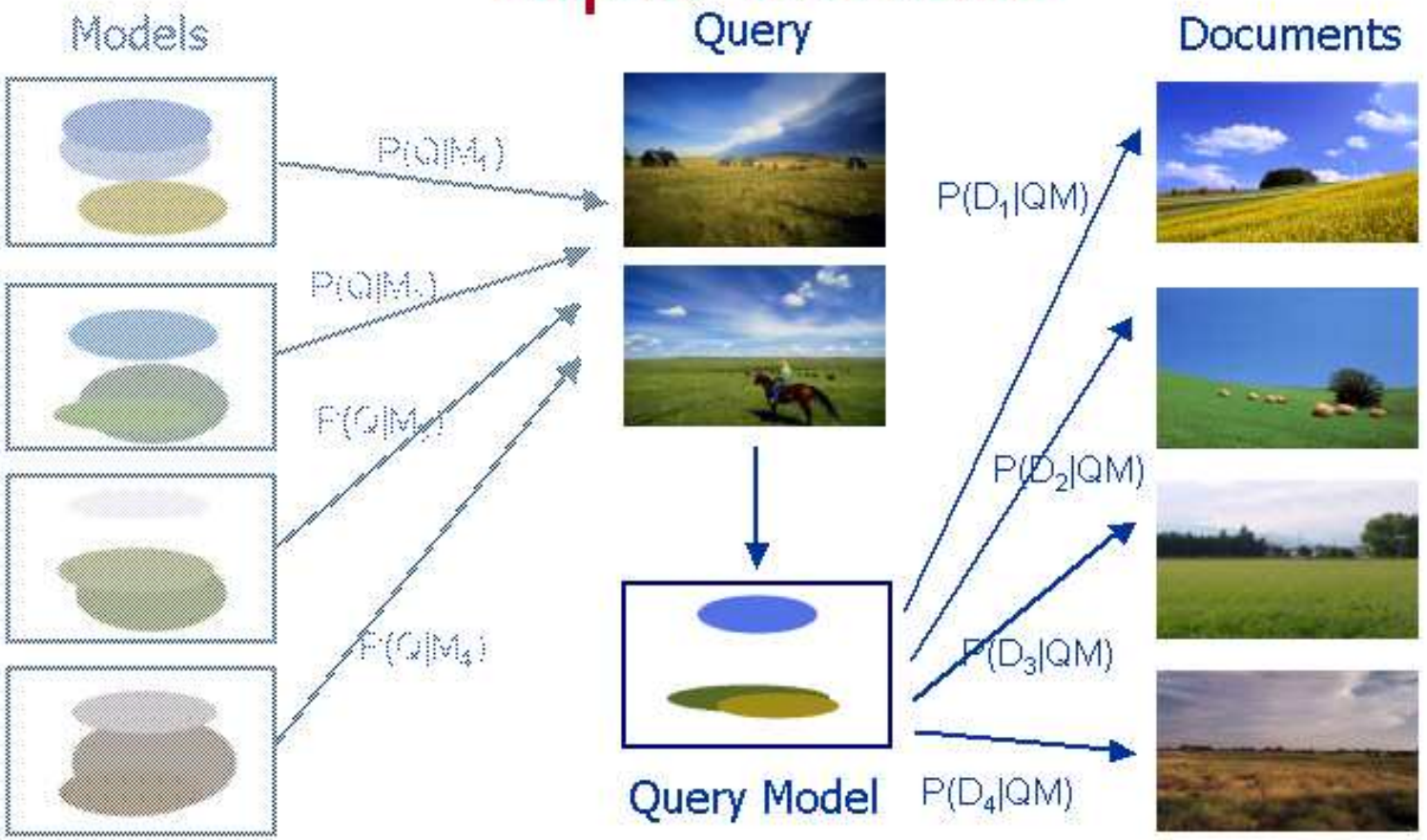
| | | Combined |
|---|---|---|
| 1 | 1 | 1 |
| 2 | 2 | 1 |
| 3 | 3 | 2 |
| 4 | 4 | ~~2~~ |
| 5 | 5 | 3 |
| 6 | 6 | 3 |
| 7 | 7 | ~~4~~ |
| 8 | 8 | 4 |
| 9 | 9 | . |
| 10 | 10 | . |

# Merging Run Results

Top 30 Results

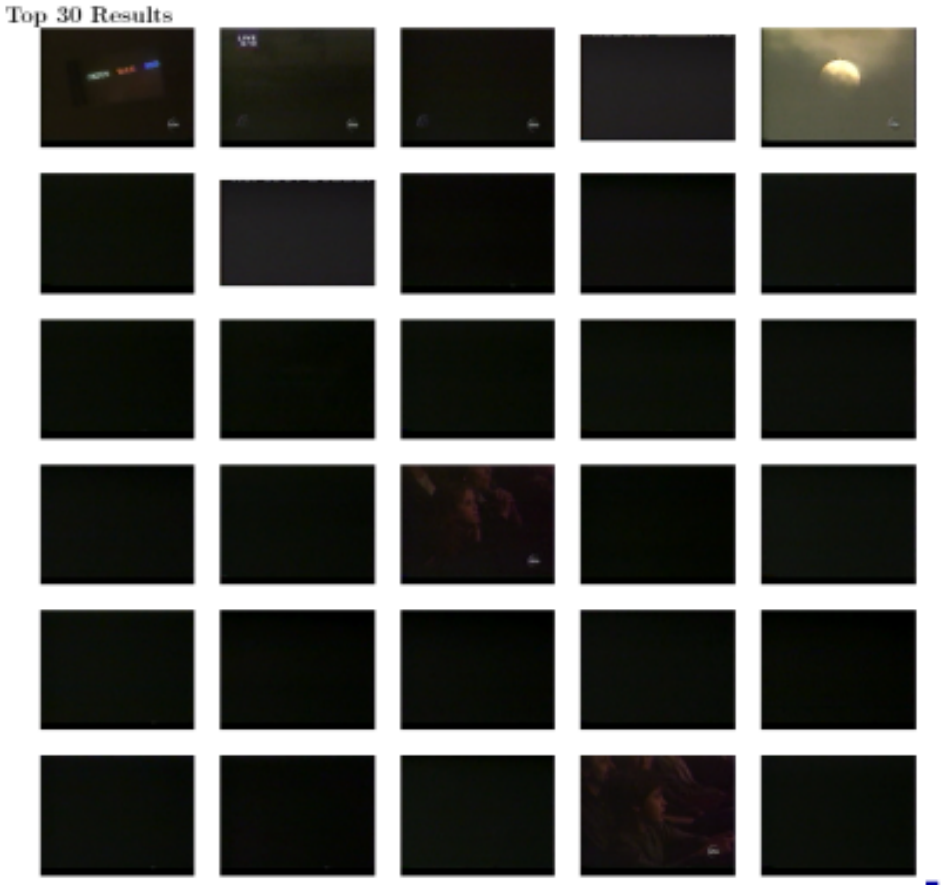
Top 30 Results

# Topic Models



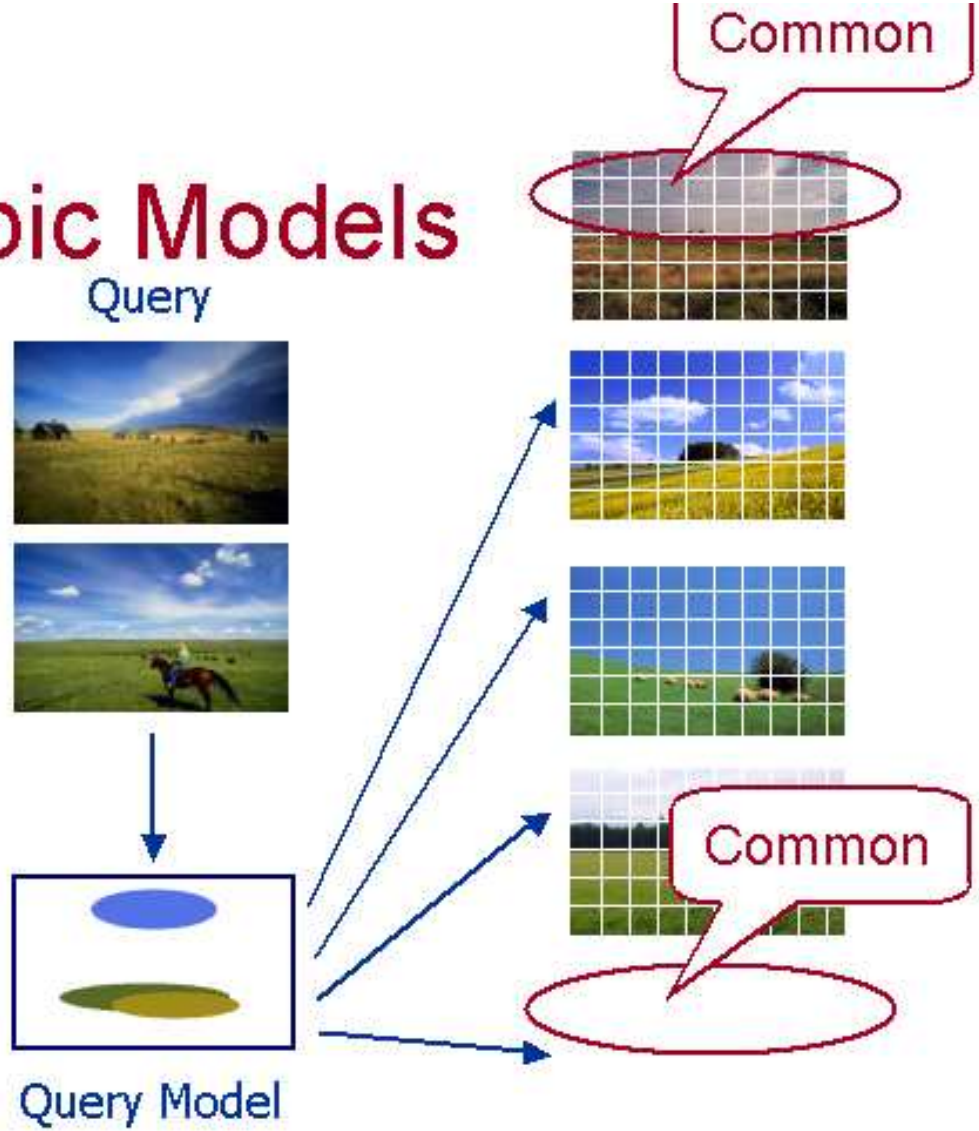
# Topic Models

- Disappointing results
- Problems with common samples

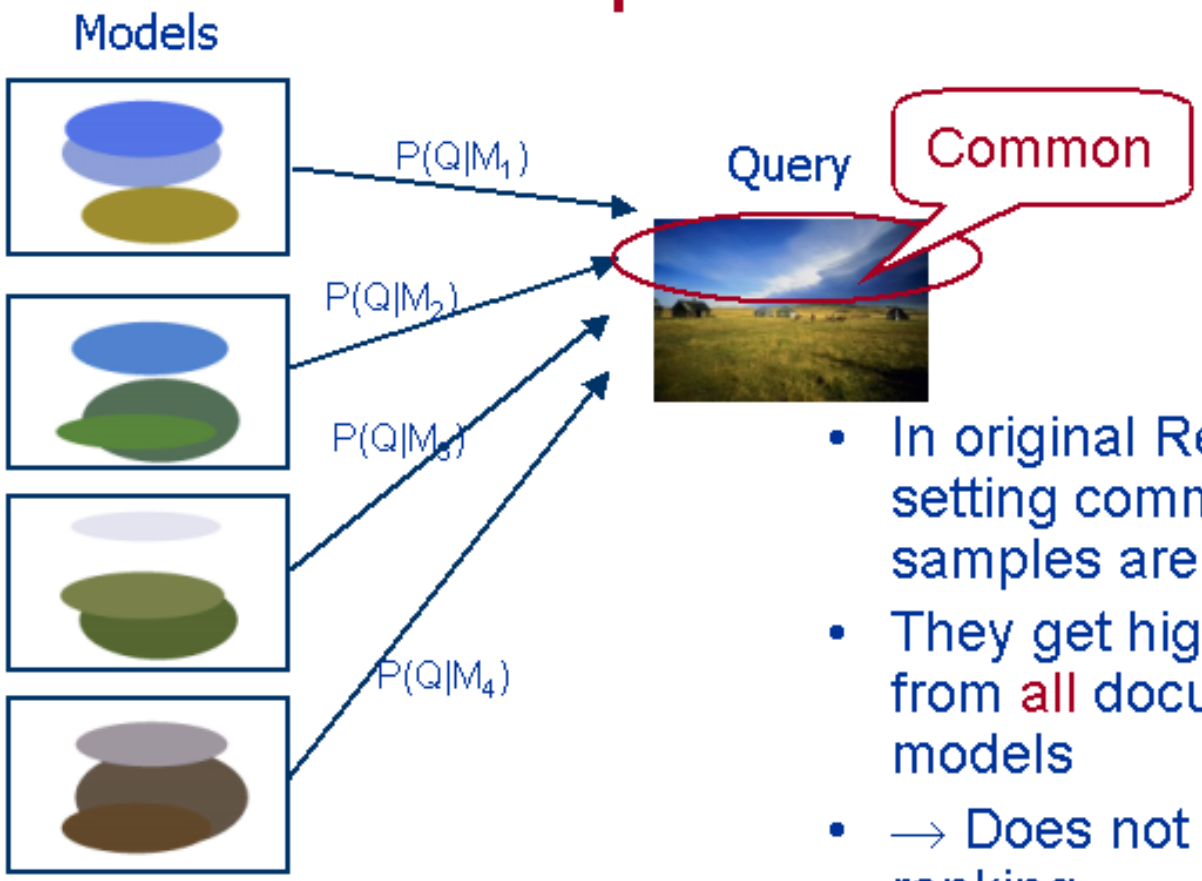


# Topic Models

- Disappointing results
- Problems with common samples



# Topic Models




- In original Retrieval setting common samples are no problem
- They get high scores from all document models
- $\longrightarrow$ Does not influence ranking