

The background of the slide is a faded image of the Stanford University Main Quad. It features the central building with its large stained-glass window and the surrounding colonnades. The title text is overlaid on the central part of the image.

Stanford I2V: A News Video Dataset for Query-by-Image Experiments

André Araujo, J. Chaves, D. Chen, R. Angst, B. Girod

Stanford University

Motivation

Example: Brand Monitoring



Logo or product

Retrieval



***NBC, 11/18/2014,
7:35:33 PM***

Motivation

Example: Content Linking

Armstrong's Critics Largely Unmoved by Interview

By IAN AUSTEN
Published: January 18, 2013

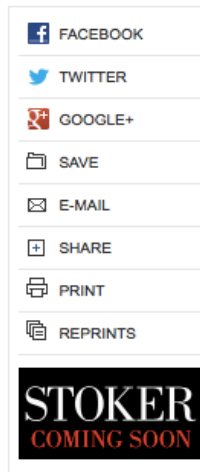
Whatever its impact on the broader public, the first of [Lance Armstrong](#)'s two nights of televised confession appeared to have little positive effect on the cycling and antidoping communities.

 [Enlarge This Image](#)



George Burns/Harpo Studios, via Getty Images
The first of Lance Armstrong's two nights of televised confession appeared to have little positive effect on the cycling and antidoping communities.

Several members of both groups faulted Armstrong for the vagueness of his confession, particularly around sensitive matters, and its lack of apology, particularly toward people he had attacked for telling the truth in the past. Many characterized Armstrong's interview with [Oprah Winfrey](#) as being more self-serving than revelatory.



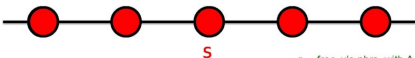
***KDTV, 01/18/2013,
6:41:45PM***

Motivation

Example: Lecture search

Infinite Graphs

- Consider **infinite graph G**
 - (but each node has finite number of neighbors!)
- We say that a finite set S causes a cascade in G with **threshold q** if, when S adopts A , eventually **every node in G adopts A**
- Example: **Path**
If $q < 1/2$ then cascade occurs



v chooses A if $p > q$
 $q = \frac{b}{a+b}$


$p \dots$ frac. v 's nbrs. with A
 $q \dots$ payoff threshold

13/12/2014 stanford, Stanford I2V: Social and Information Network Analysis, http://i2v.stanford.edu/ 43

Retrieval

Infinite Graphs

- Consider **infinite graph G**
 - (but each node has finite number of neighbors!)
- We say that a finite set S causes a cascade in G with threshold q if, when S adopts A , eventually **every node in G adopts A**
- Example: **Path**
If $q < 1/2$ then cascade occurs



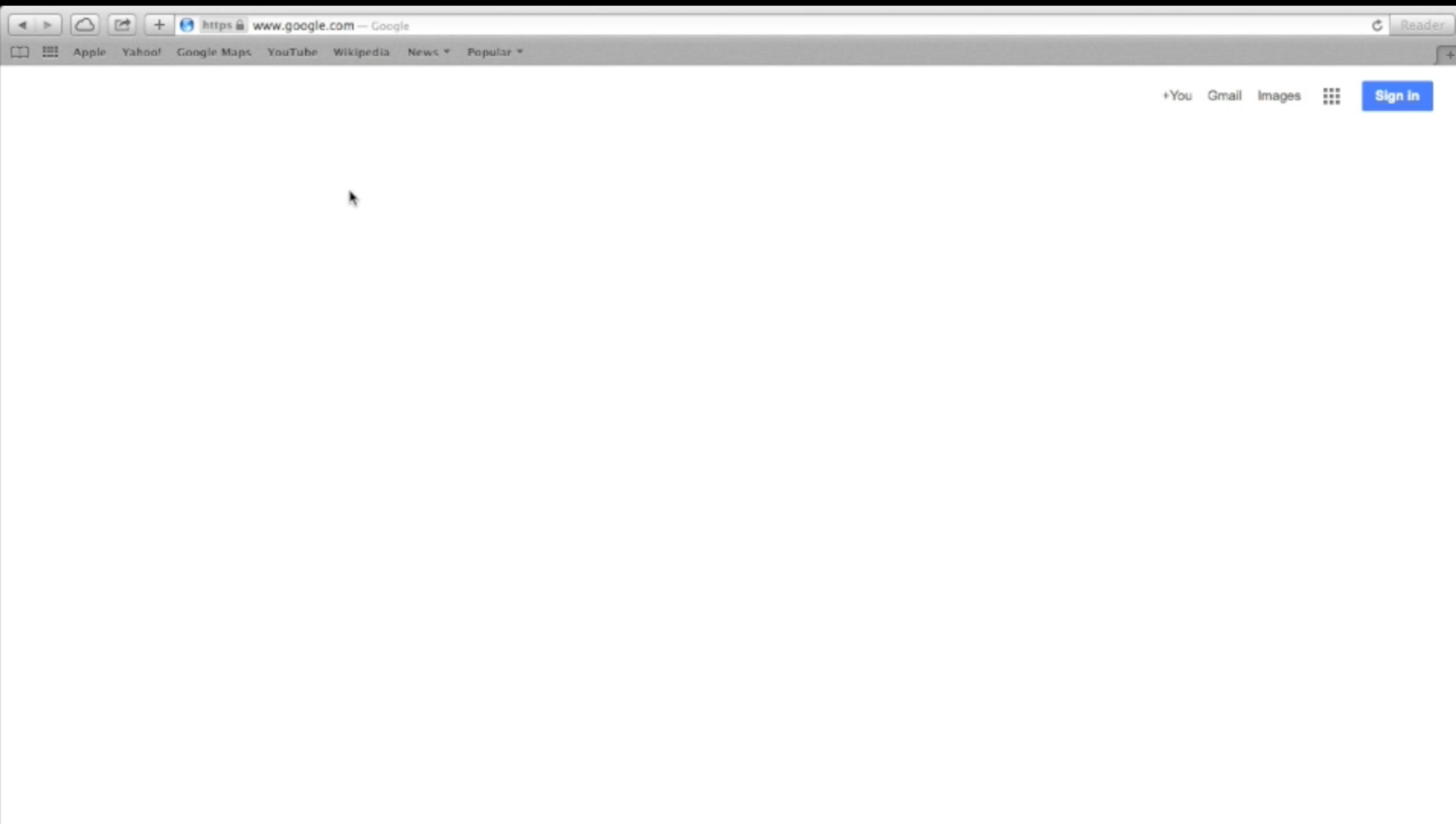
v chooses A if $p > q$
 $q = \frac{b}{a+b}$

Stanford University

Presentation slide

CS246, lecture 12
December 2, 2013

Online demo



<http://videosearch.stanford.edu>

Outline

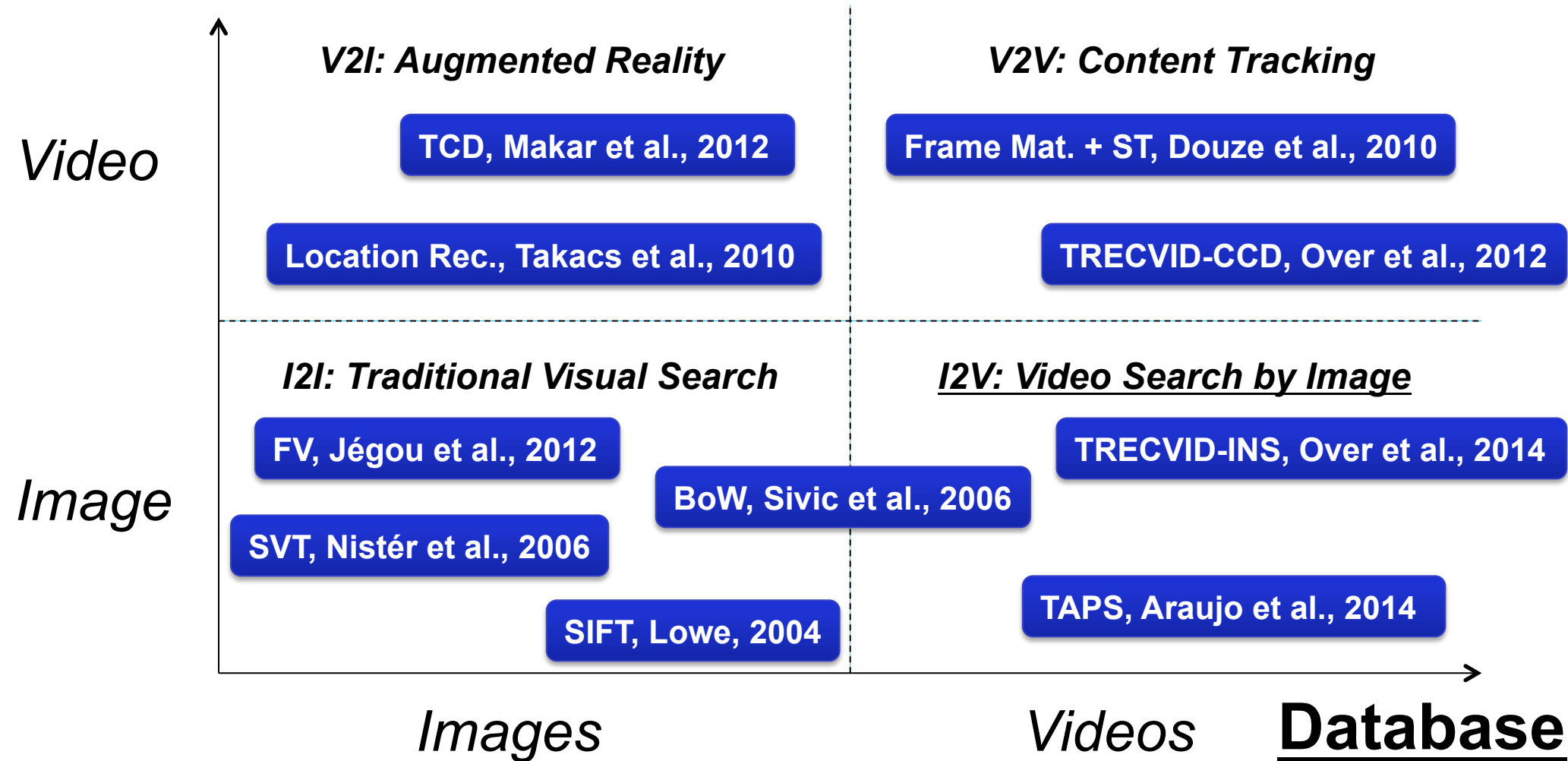
- Related Work
- Stanford I2V Dataset
- Dataset Construction
- Baseline Experiments

Outline

- Related Work
- Stanford I2V Dataset
- Dataset Construction
- Baseline Experiments

Related Work: Visual Search

Query



Related Work: Existing I2V Datasets

Dataset	Size	# Queries
Sivic et al., Video-Google, 2006	2h	164

Related Work: Existing I2V Datasets

Dataset	Size	# Queries
Sivic et al., Video-Google, 2006	2h	164
Over et al., TRECVID-INS, 2014	464h	30

Related Work: Existing I2V Datasets

Dataset	Size	# Queries
Sivic et al., Video-Google, 2006	2h	164
Over et al., TRECVID-INS, 2014	464h	30
Araujo et al., CNN2h, 2014	2h	139

Related Work: Existing I2V Datasets

Dataset	Size	# Queries
Sivic et al., Video-Google, 2006	2h	164
Over et al., TRECVID-INS, 2014	464h	30
Araujo et al., CNN2h, 2014	2h	139
Araujo et al., Stanford I2V, 2015	3,801h	229

Outline

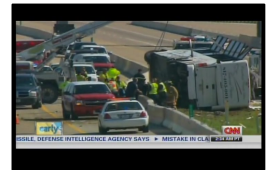
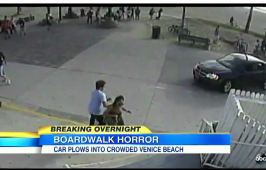
- Related Work
- **Stanford I2V Dataset**
- Dataset Construction
- Baseline Experiments

Stanford I2V Dataset

Query images



Database videos (selected frames)



Stanford I2V Dataset

Full version	Light version
3.8k hours	1k hours
84k video clips	23k video clips
229 query images	78 query images
14M keyframes@1fps	3.8M keyframes@1fps
2.7 minutes/clip	2.65 minutes/clip

Evaluation Procedure

Query



System



1st stage: Retrieval of Clips

1



2



3

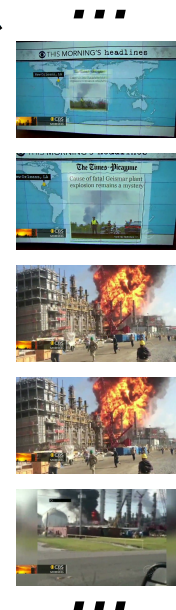


...

Ranked retrieval measures:

- Average Precision (AP)
- Precision at 1 (p@1)

2nd stage: Temporal Refinement



Unranked retrieval measure:

- Temporal Jaccard Index

Query/Annotation Viewer

blackhole1.stanford.edu

Query 1 (first one was 1), group 3 (first one was 1) (3 groups in total)
Query path: /home/ptvn/news_videos_dataset/final_dataset/201303/queries/001.jpg



Query image

Clip 1

Clip 2

Date: 20130302
Program: CNN_Newsroom_20130302_1400
Chunk: 01
Ground-truth segments:
Segment 1: 1:43 2:04

Date: 20130302
Program: Good_Morning_America_20130302_0400
Chunk: 02
Ground-truth segments:
Segment 1: 0:08 0:11
Segment 2: 0:46 0:51
Segment 3: 1:12 1:13

Outline

- Related Work
- Stanford I2V Dataset
- **Dataset Construction**
- Baseline Experiments

Dataset Construction: Video Collection

News Videos Recording

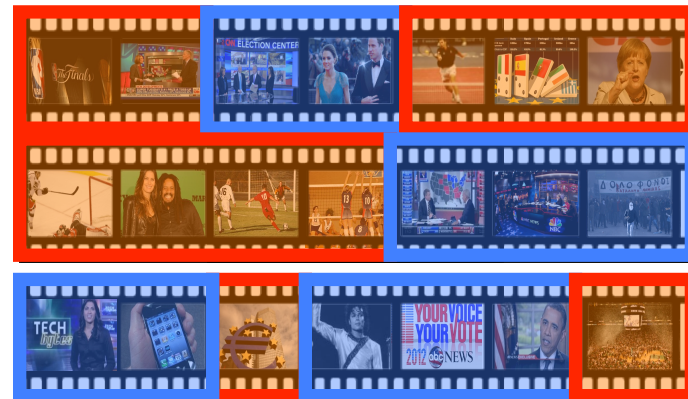
CNN



NBC NEWS

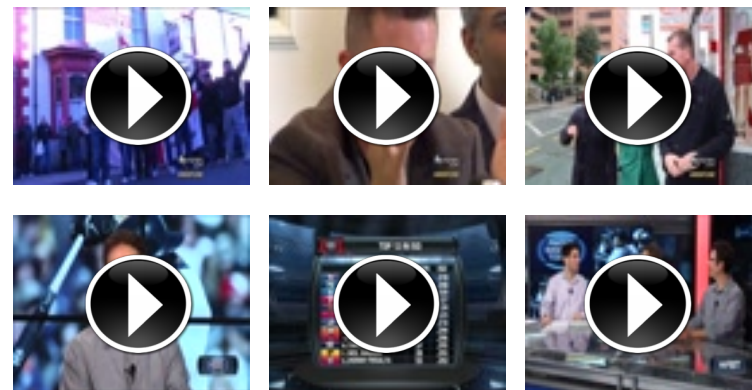


Story Segmentation



Video clips

Daneshi et al., 2013



Dataset Construction: Query Set Collection

- Collected images from news websites
 - Used the Internet Archive Wayback Machine
 - Collected 805 candidate images from dates between October 1st 2012 and September 30th 2013



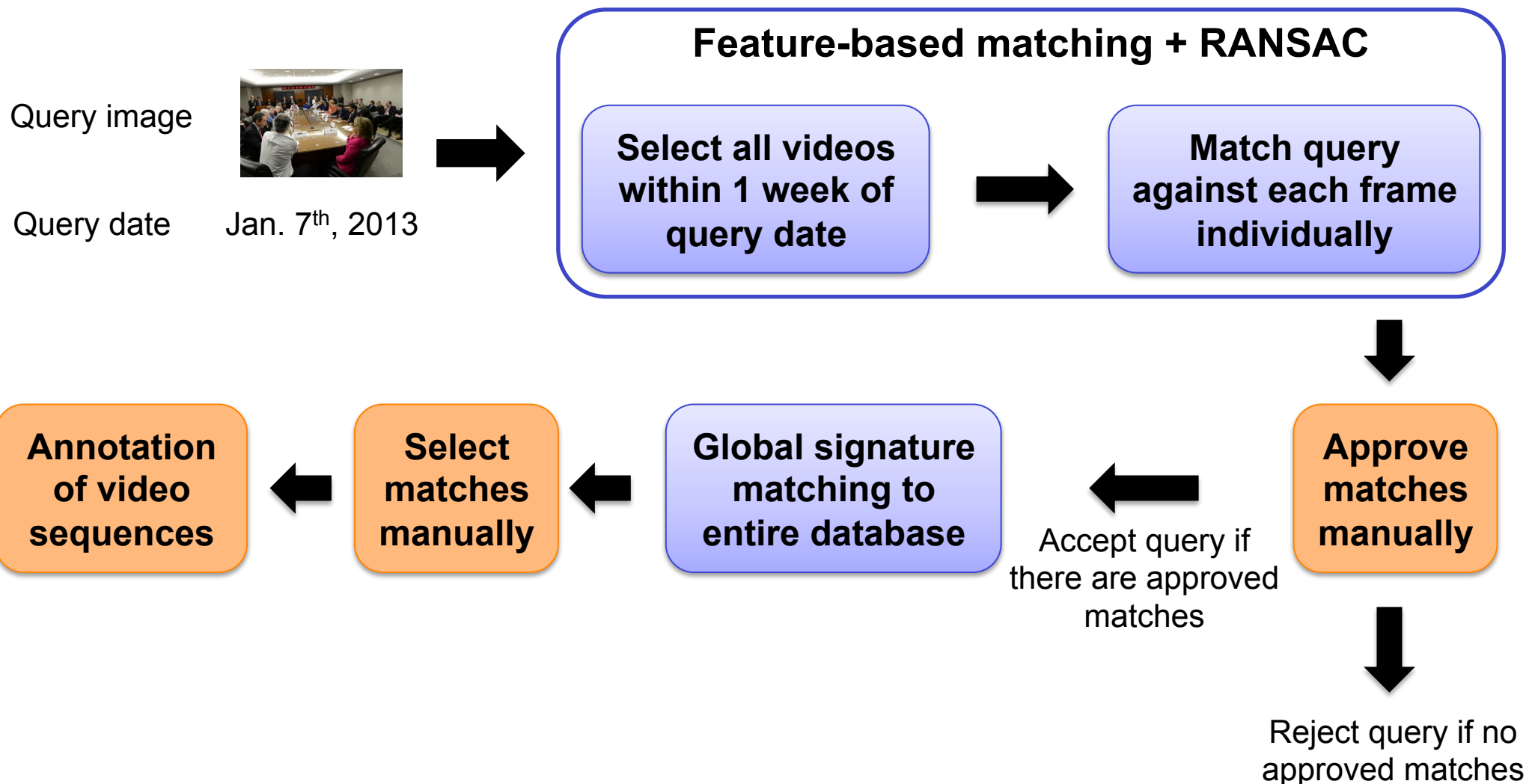
- Types of images:
 - Iconic images (events in the news)



- Magazine covers (Time, Economist)



Dataset Construction: Annotation



Outline

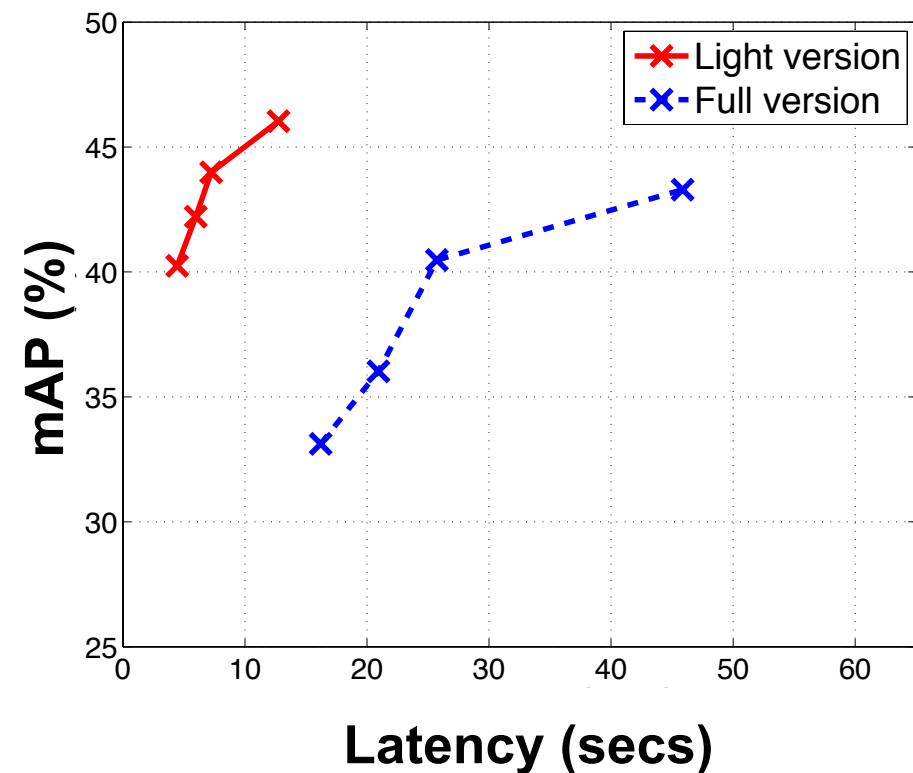
- Related Work
- Stanford I2V Dataset
- Dataset Construction
- **Baseline Experiments**

Example: Evaluation of Standard Technique

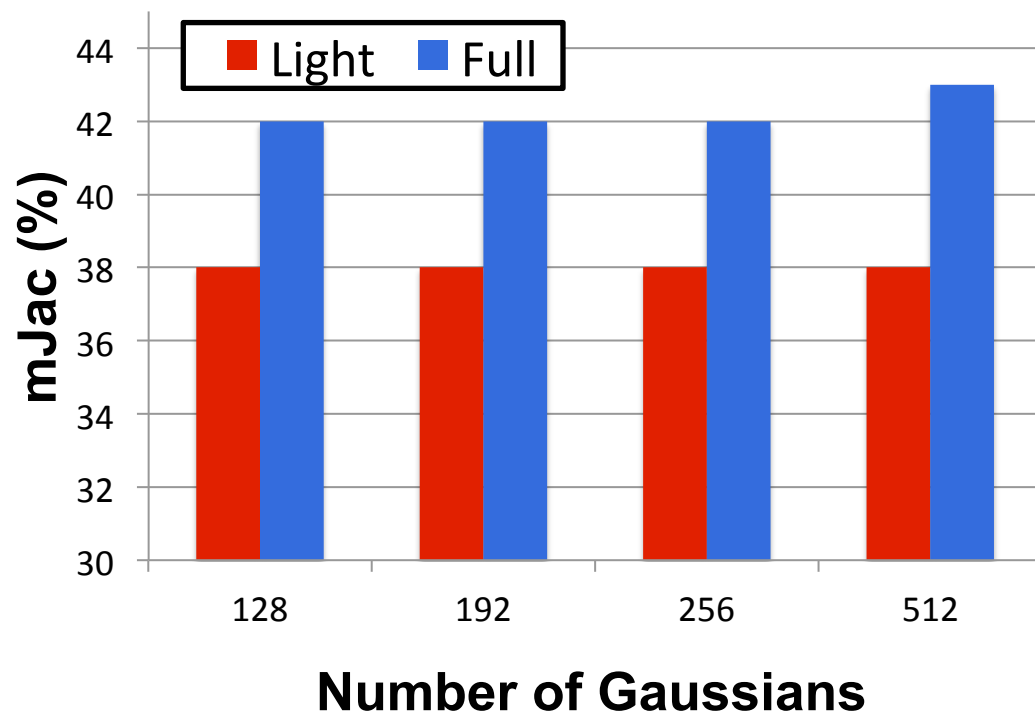
- SIFT descriptors + SCFV global signatures
[Lowe, 2004] [Duan et al., 2014]
- Retrieval of Clips evaluation:
 - Compare query signature to video frames' signatures (@1fps) from entire database
 - Evaluate performance over top 100 ranked clips
- Temporal Refinement evaluation:
 - Compare query signature to video frames' signatures (@1fps) from each correct matching video
 - Feature matching + RANSAC between query and top 50 frames (consider a match if at least 8 inliers are found)
 - Evaluate Jaccard index between matches and ground-truth segments

Example: Evaluation of Standard Technique

Retrieval of Clips: results



Temporal Refinement results



Summary

- Dataset for video retrieval using query images
- 3.8k hours of video and 229 queries – largest dataset yet
- First dataset to allow true large-scale experiments in this area
- Experiments using standard image retrieval technique were presented, serving as a baseline for future evaluations

Thank you! Questions?

Dataset webpage:

<http://blackhole1.stanford.edu/vidsearch/dataset/stanfordi2v.html>

Online demo:

<http://videosearch.stanford.edu>

André Araujo

<http://stanford.edu/~afaraujo>

afaraujo@stanford.edu