

The background of the slide is a faded image of a large, ornate building with a central gable and a cross on top, likely a Stanford University building. The building has multiple levels of arches and windows.

The Stanford Mobile Visual Search Data Set

**Vijay Chandrasekhar, David Chen, Sam Tsai, Ngai-Man Cheung,
Huizhong Chen, Gabriel Takacs, Bernd Girod**
Stanford University

Yuriy Reznik
Qualcomm, San Diego

Ramakrishna Vedantham, Radek Grzeszczuk
Nokia Research Center, Palo Alto

Jeff Bach
Navteq, Chicago

Mobile Visual Search Applications



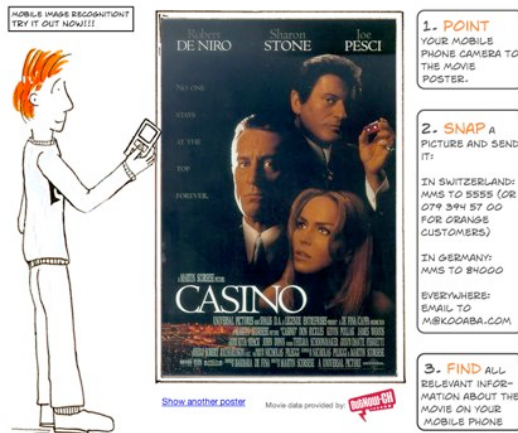
Ads/Catalogs



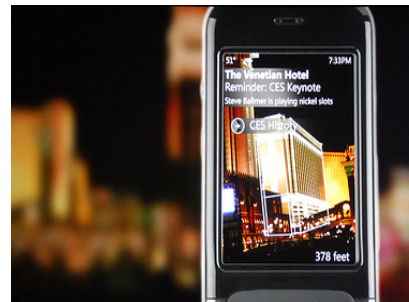
Wine Labels



Museum Guide



Movie Posters



Landmarks



Media covers
(CD, DVD, Books)

Landmark Recognition



Media Cover Recognition



Outline

- Survey of computer vision data sets
- Limitations of existing data sets
- Stanford Mobile Visual Search Data Set
- Pairwise Image Matching Results

Survey of existing data sets

- **Zurich Buildings**
 - 201 buildings
- **Oxford Buildings**
 - 11 landmarks
- **INRIA Holidays**
 - 500 outdoor locations
- **University of Kentucky**
 - 2550 different objects
- **Image Net**
 - > 1 million categories

Zurich Buildings



Mild illumination changes

Mild perspective changes

Simple approaches like color histograms work

Easy Data Set with ~100% accuracy

Oxford Buildings



Very few landmarks (11)

Different facades of the same building carry the same label

INRIA Holidays



Mild illumination changes

Non-rigid objects in outdoor scenes (fire, sunsets, animals)

University of Kentucky



Mild illumination changes

No foreground / background clutter

Image Nets



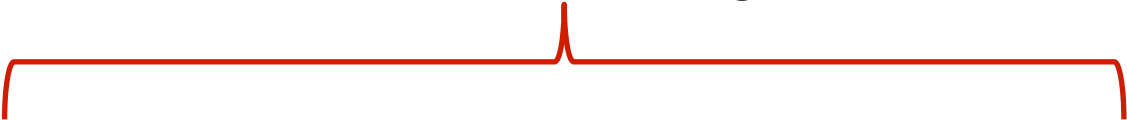
Many object categories

Non-rigid objects (animals)

More suitable for classification experiments

Comparison of Data Sets

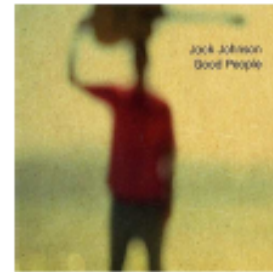
Desirable traits for testing
mobile visual search algorithms



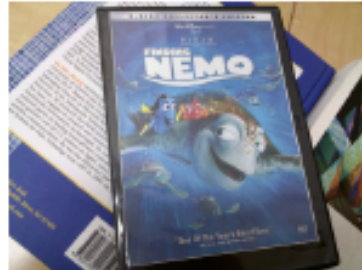
Data Set	Database (#)	Query (#)	Classes (#)	Rigid	Lighting	Clutter	Perspective	Camera Phone
ZuBuD	1005	115	200	✓	—	✓	✓	—
Oxford	5062	55	17	✓	✓	✓	✓	—
INRIA	1491	500	500	—	—	✓	✓	—
UKY	10200	2550	2550	✓	—	—	✓	—
ImageNet	11M	15K	15K	—	✓	✓	✓	—
SMVS	1200	3300	1200	✓	✓	✓	✓	✓

SMVS Examples

CDs



DVDs



Books

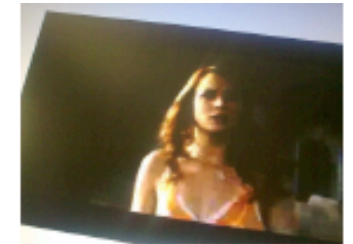
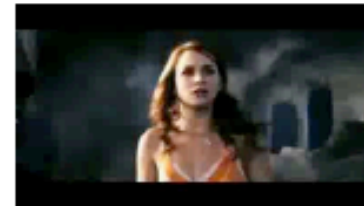
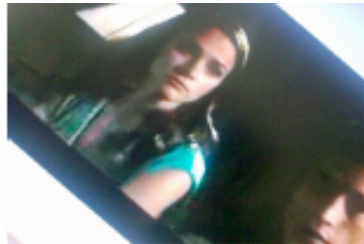
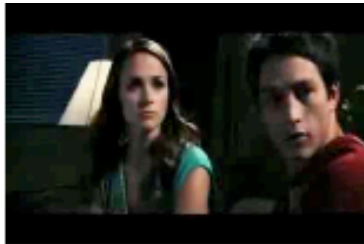


Landmarks



SMVS Examples

Video Clips



Cards



Print



Paintings



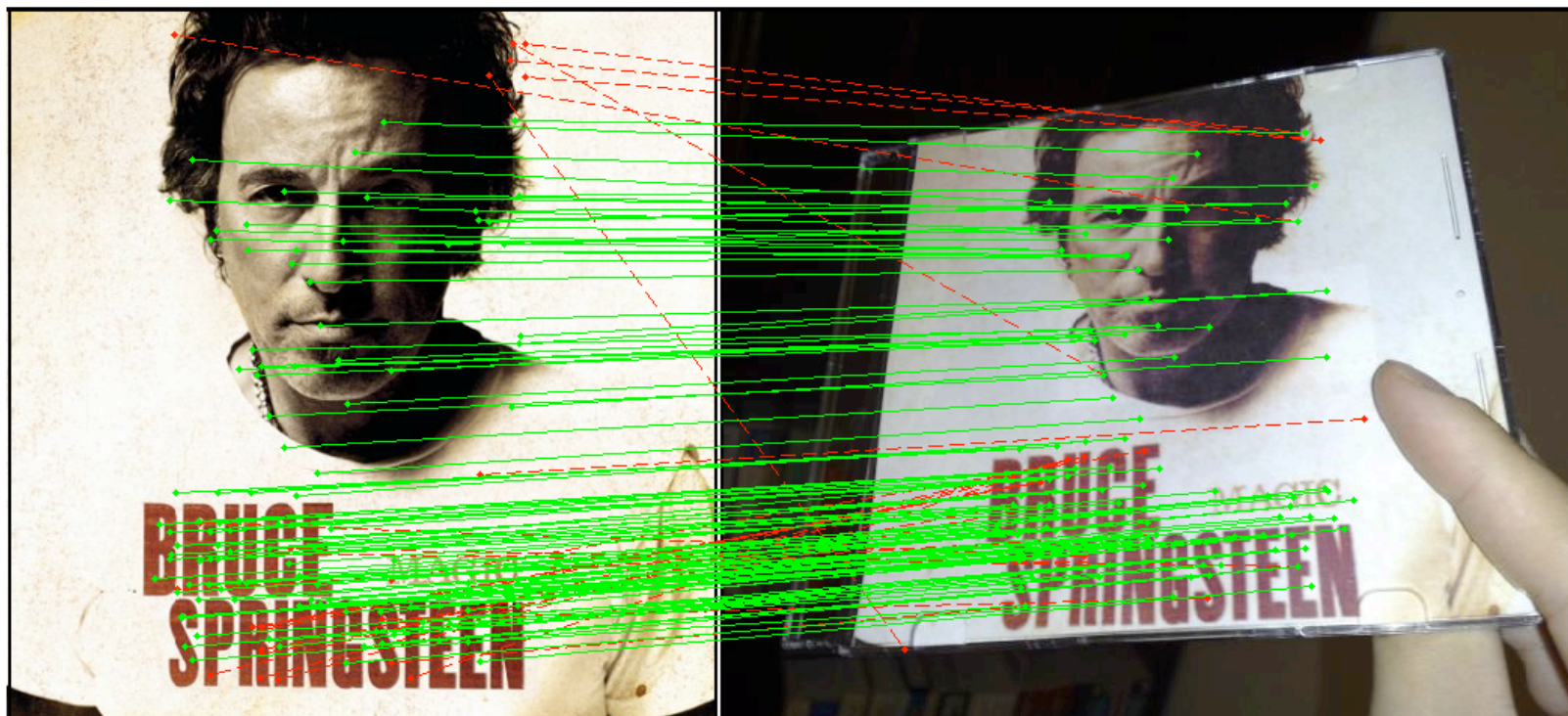
Summary of SMVS Images

Category	Database	Query
CD	100	400
DVD	100	400
Books	100	400
Video Clips	100	400
Landmarks	500	500
Business Cards	100	400
Text documents	100	400
Paintings	100	400

Camera phones used to capture query images:

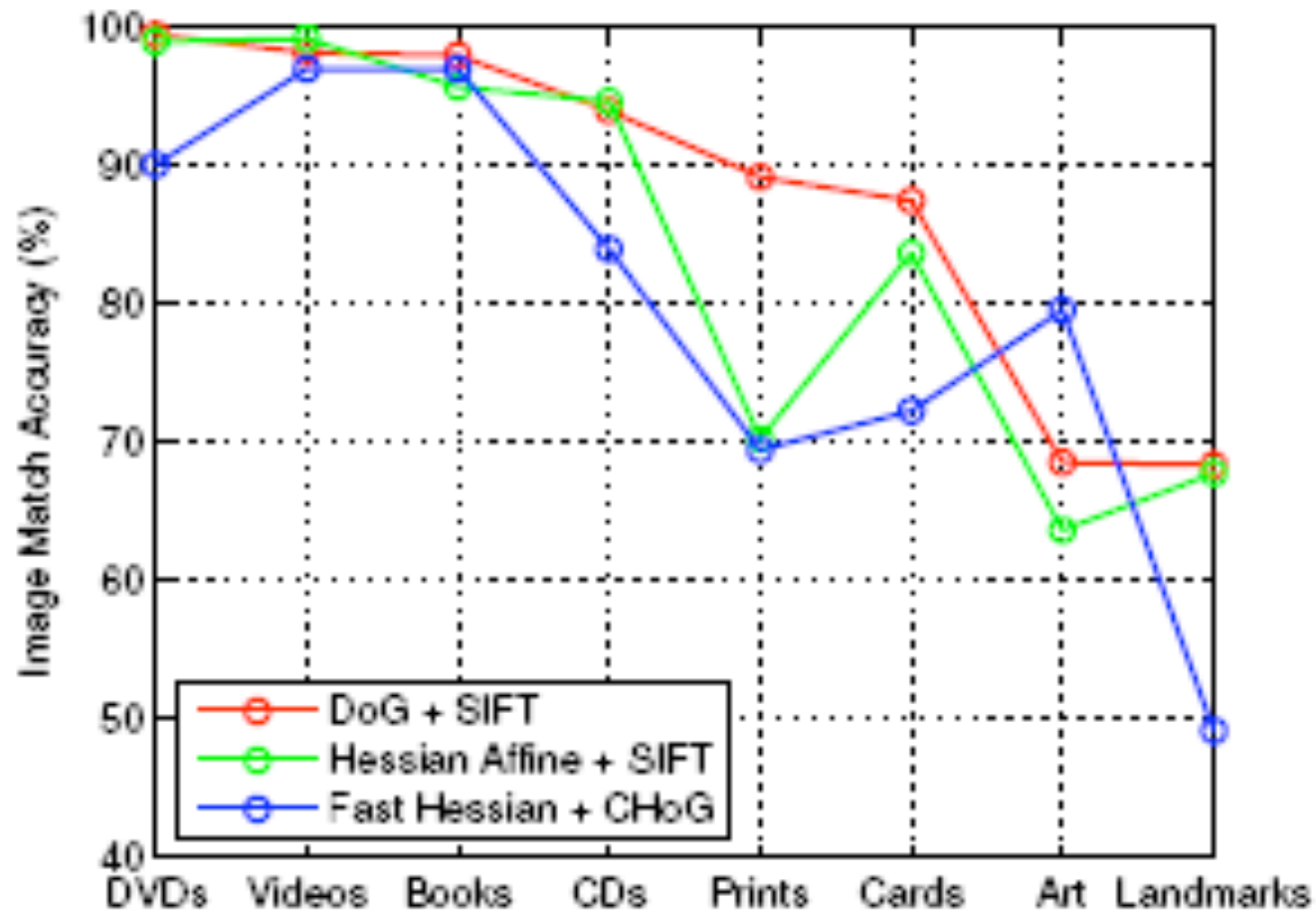
Nokia (N900, 5800, N97, E63, N86), Apple iPhone 4, Palm Pre, Motorola Droid, LG 300

Pairwise Image Matching



Robust local features
Distance ratio test
RANSAC with affine model

Pairwise Image Matching Results



Conclusions

- Survey of popular computer vision data sets
- Existing data sets are limited
- New SMVS data set
 - Specific to mobile visual search applications
 - 8 different categories of objects
 - Many labeled database and query pairs
 - Query images taken by different camera phones
- Stimulate further research in MVS

Questions

The image shows the front facade of Mission San Juan Bautista, a historic Spanish Colonial building. The central feature is a large triangular pediment containing a vibrant fresco depicting the Last Judgment, with figures in various states of resurrection and damnation. Below the pediment is a large arched window. The base of the facade consists of three large, ornate arched doorways. The entire building is constructed of light-colored stone or adobe with intricate carvings and painted details. The word "Questions" is superimposed in a large, bold, orange font across the center of the image.

SMVS Data Set

- Good ground truth images
- Query images with a wide range of camera phones
- Widely varying lighting conditions
- Captures typical perspective variations, foreground/background clutter
- Eight different categories (buildings, books, CDs, DVDs, text documents)
- Rigid objects
- Bounding box information in query images