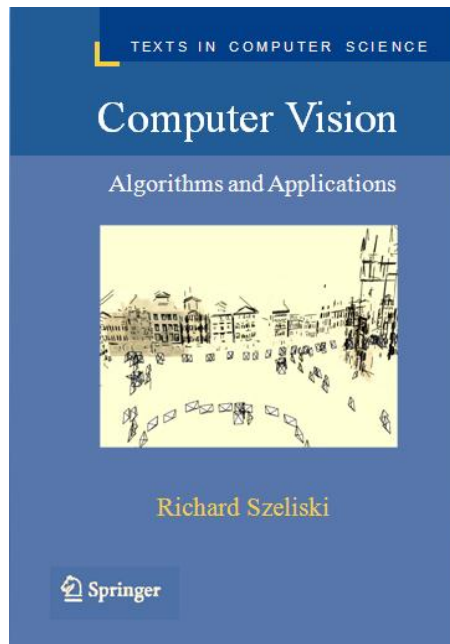


Vision-based Natural User Interfaces

Richard Szeliski
Principal Researcher
Microsoft Research

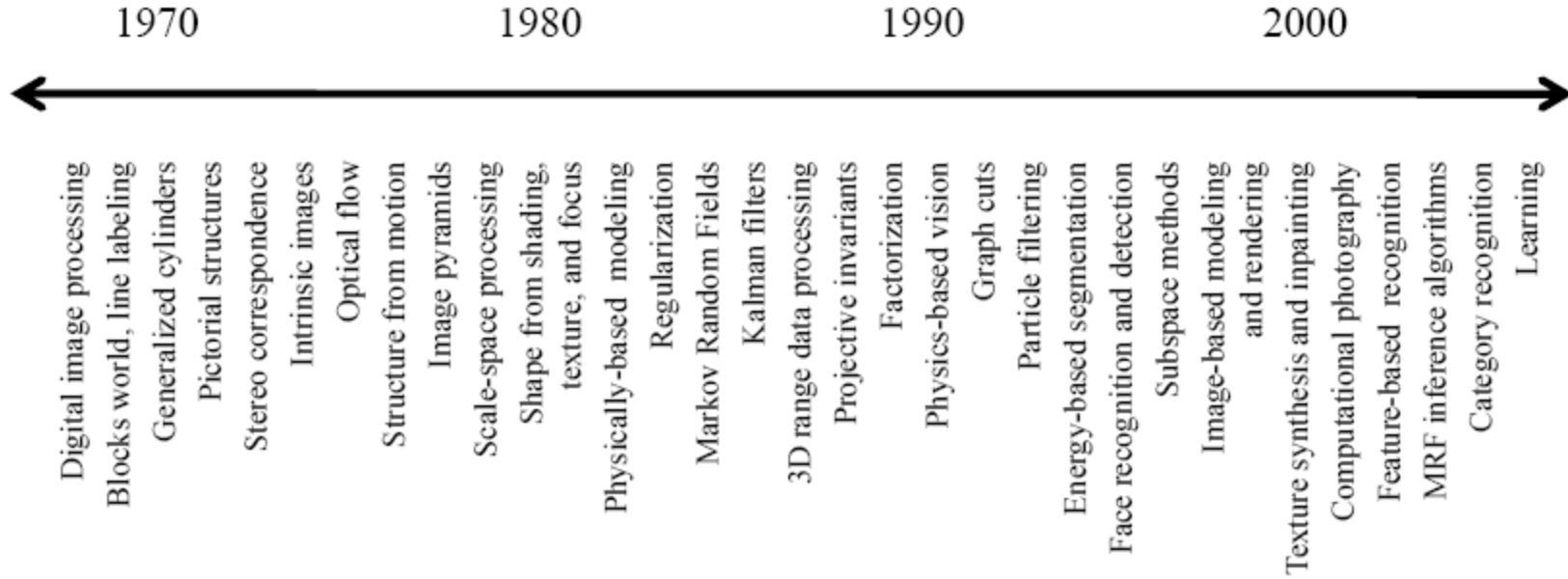
A Brief History of Computer Vision



A Brief History of Computer Vision

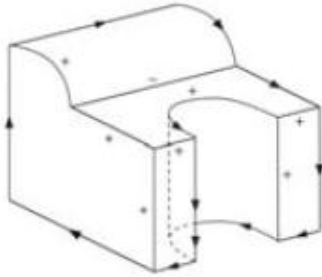
1970s. When computer vision first started out in the early 1970s, it was viewed as the visual perception component of an ambitious agenda to mimic human intelligence and to endow robots with intelligent behavior. At the time, it was believed by some of the early pioneers of artificial intelligence and robotics (at places such as MIT, Stanford, and CMU) that solving the “visual input” problem would be an easy step along the path to solving more difficult problems such as higher-level reasoning and planning. According to one well-known story, in 1966, Marvin Minsky at MIT asked his undergraduate student Gerald Jay Sussman to “spend the summer linking a camera to a computer and getting the computer to describe what it saw” (Boden 2006, p. 781).⁵ We now know that the problem is slightly more difficult than that.⁶

A Brief History of Computer Vision

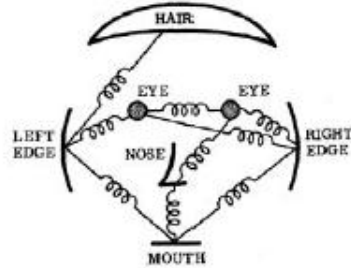


A Brief History of Computer Vision

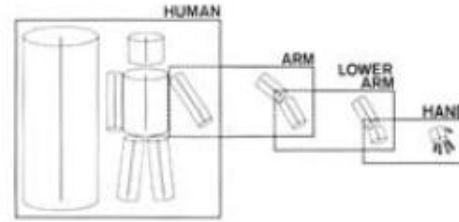
1970s



(a)



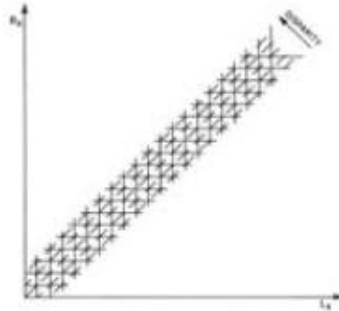
(b)



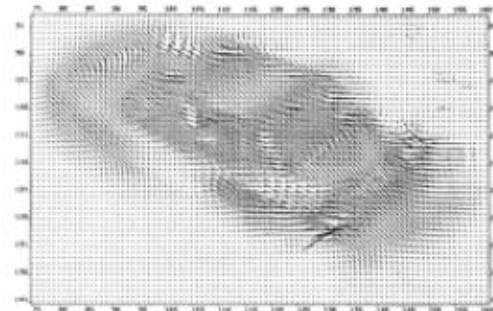
(c)



(d)



(e)



(f)

A Brief History of Computer Vision

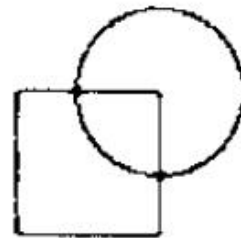
1980s



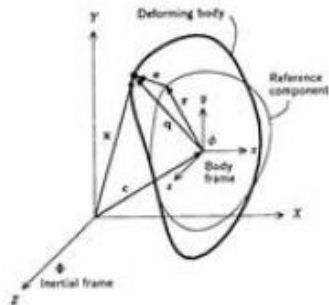
(a)



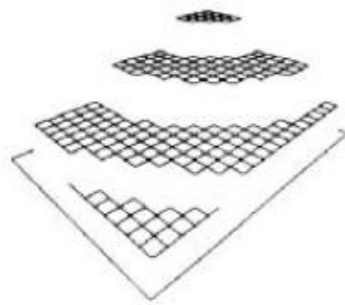
(b)



(c)



(d)



(e)



(f)

A Brief History of Computer Vision

1990s



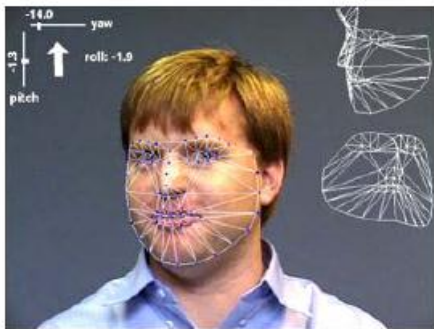
(a)



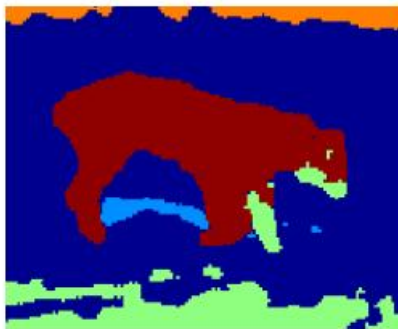
(b)



(c)



(d)



(e)



(f)

A Brief History of Computer Vision

2000s



(a)



Original photograph with marked edges

Recovered model

Model edges projected onto photograph

(b)



(c)



(d)

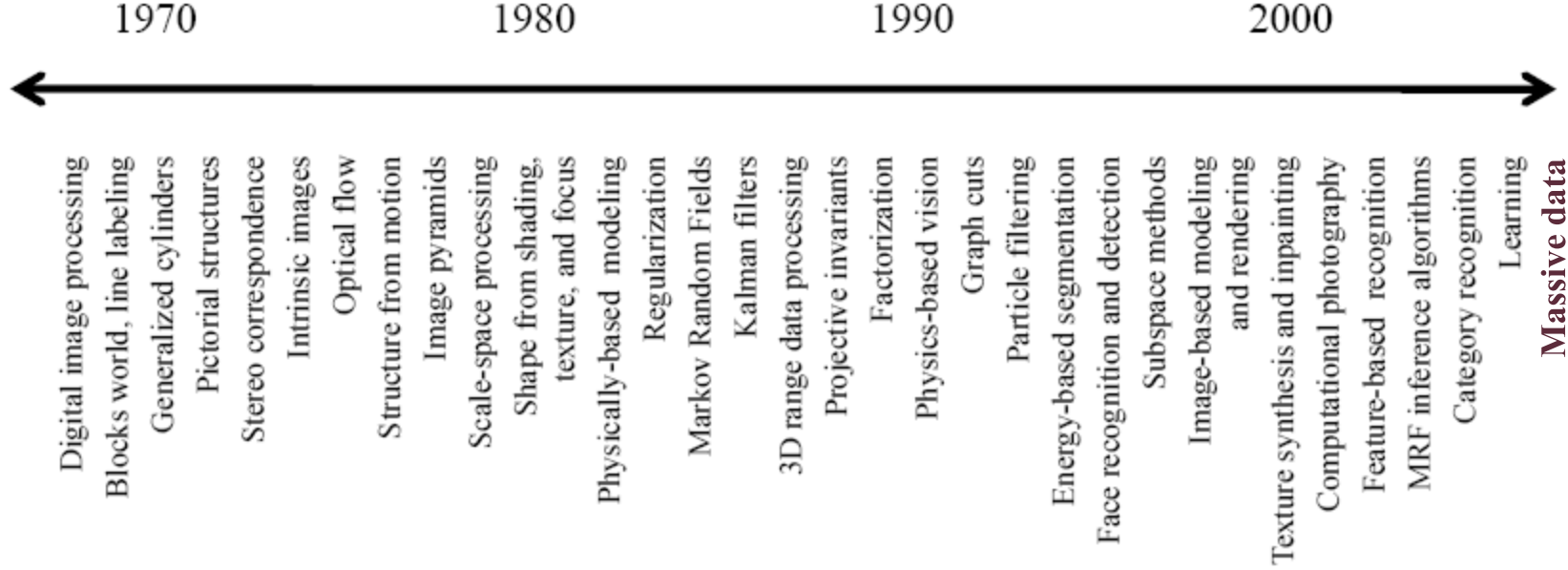


(e)



(f)

A Brief History of Computer Vision



Outline

Computer vision and machine learning techniques are maturing and having major impact:

1. 3D body tracking [Kinect]
2. Medical image segmentation [Amalga]
3. Object (product) recognition [Bing Vision]
4. Multi-image matching and navigation [Photosynth]

Massive (Internet) data is playing a key role

1. Body part recognition for Kinect

Jamie Shotton
Microsoft Research Cambridge
FG 2011 & CVPR 2011

2. Medical Image Segmentation

Antonio Criminisi
Microsoft Research Cambridge

To find out more...

| | | |
|------------|---|---------|
| 9:00–10:30 | Breakout Sessions | |
| | <p>Session: Medical Visualization Medical Imaging on the Microsoft Platform Session Chair: Rick Benge, Microsoft Research</p> <p>Presentations:</p> <ul style="list-style-type: none">■ Advanced Medical Imaging Research at Microsoft and its Applications on Product Groups—Khan Siddiqui, Microsoft■ Inner Eye: Toward a Computational Platform for Imaging Metadata—Steve White, Microsoft■ Applications of Advanced Semantic Tagging in Clinical Settings—David Haynor, University of Washington <p>Analysis and metadata extraction and from medical image data represent significant computational challenges, but current open source efforts in the field of medical imaging focus on sharing code rather than sharing information. A common platform enabling researchers to benchmark and integrate very different analysis techniques in a common environment, and exchange both data and analyses on the web, would greatly accelerate research in this area. In this session, the speakers will present three different aspects of how Microsoft and its partners are addressing these challenges in terms of research, development, and real-world deployment.</p> | Rainier |

3. (Mobile) Object Recognition

Larry Zitnick
Microsoft Research Redmond

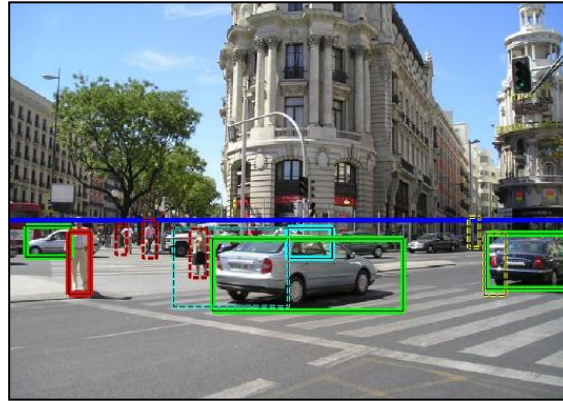
David Nister
Bing Vision

Object Recognition

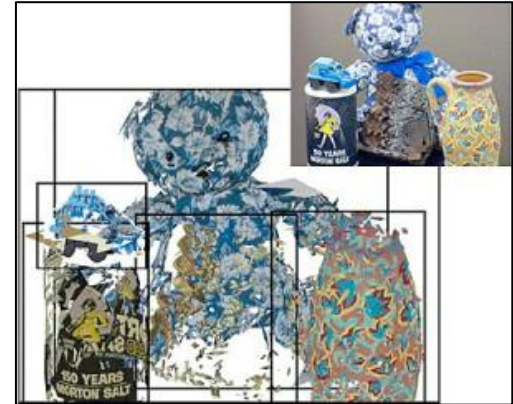
- Detecting and localizing objects in images



Lowe IJCV 04

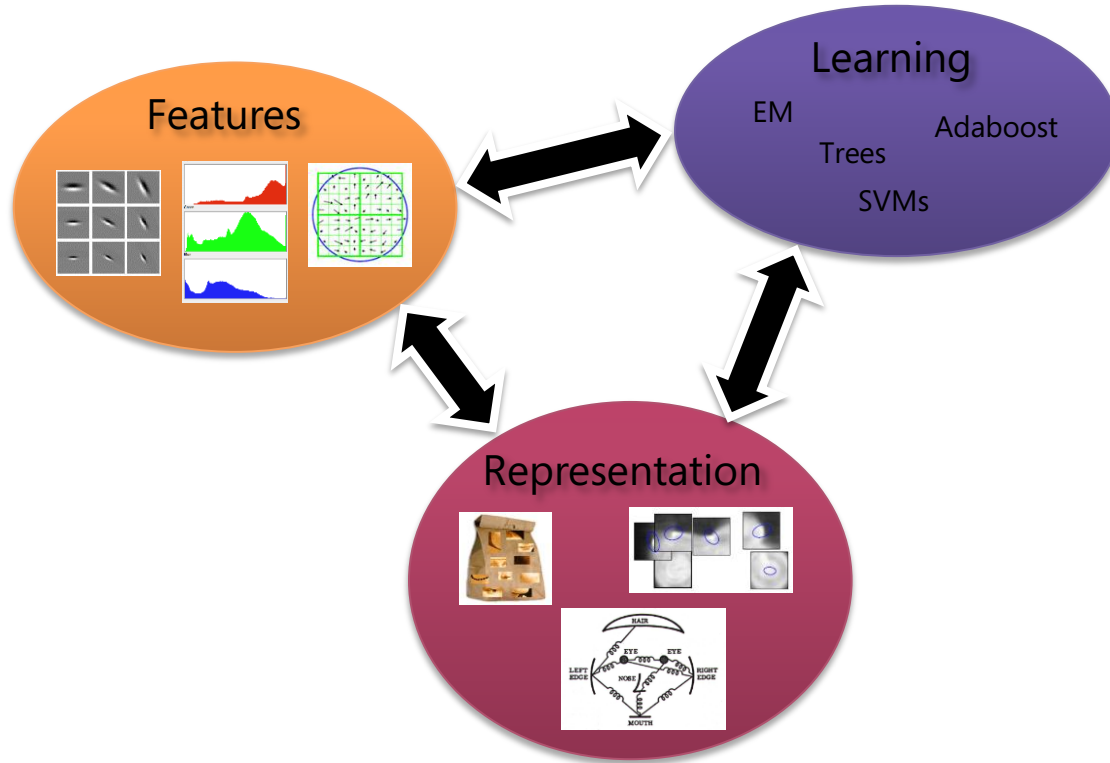


Hoiem et al. CVPR 06

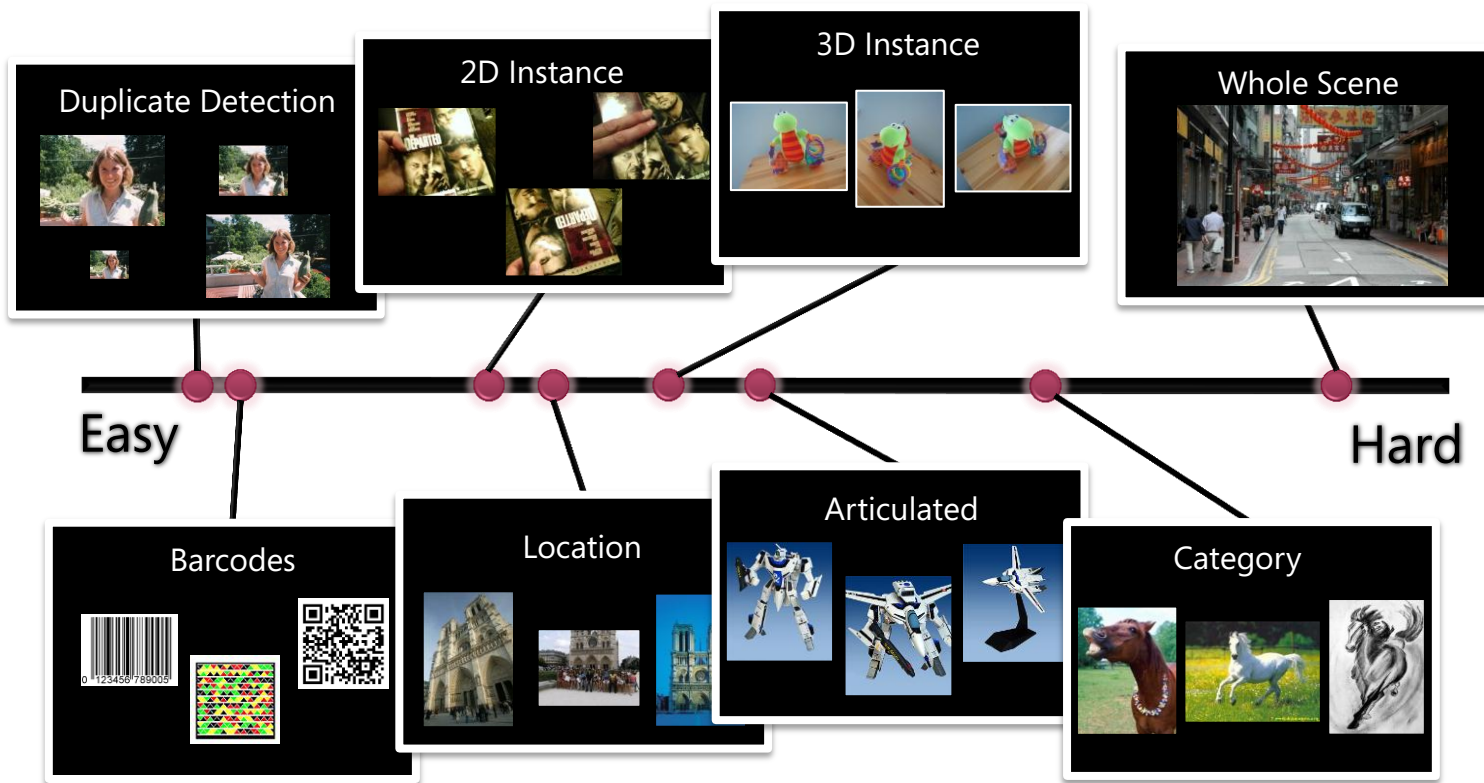


Rothganger et al. IJCV 06

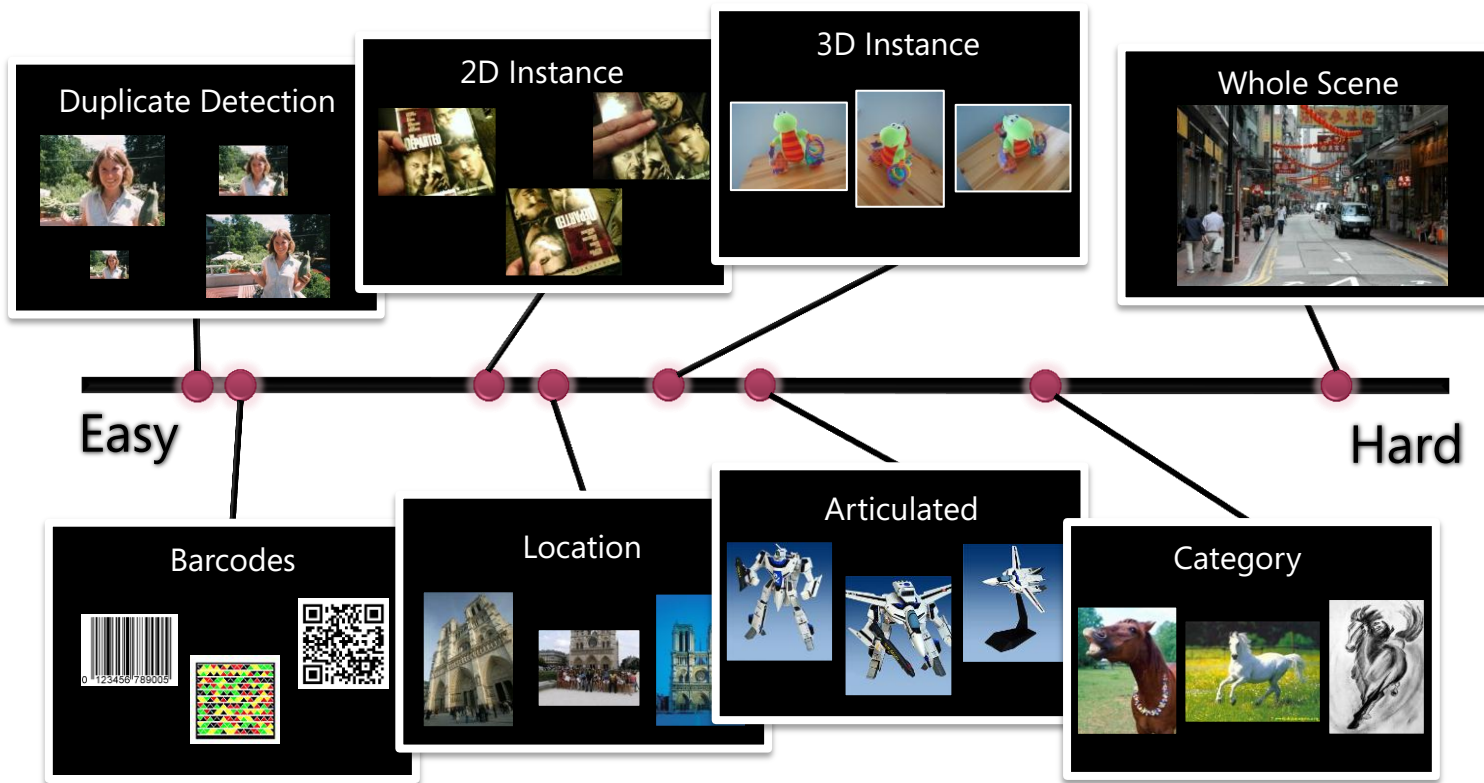
Problems



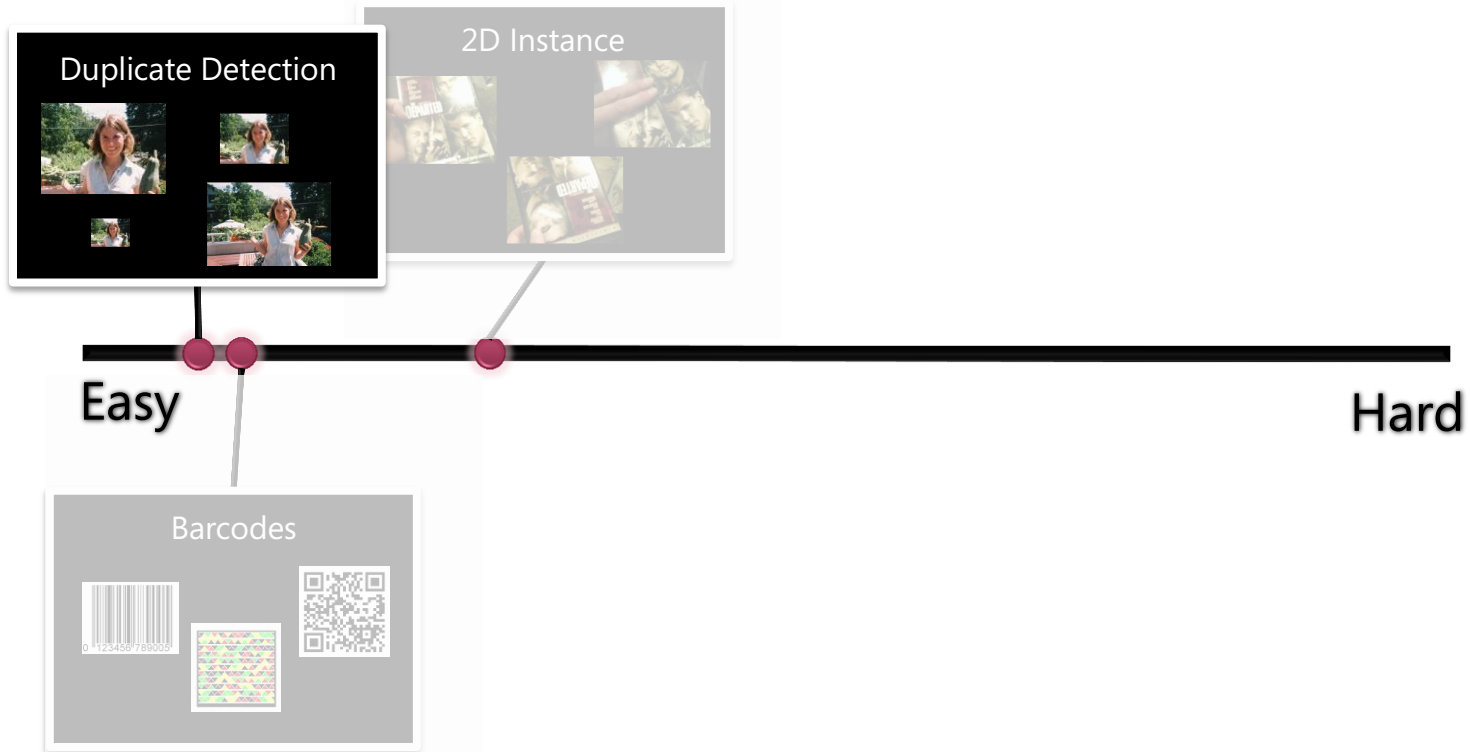
Spatial Complexity



Lincoln (Live Labs → Bing Mobile)



Near Duplicate Detection



Near Duplicate Detection

Are these images the same?



Near Duplicate Detection

Why is duplicate detection important?

- Increase search relevance
- Copyright search
- Remove illegal images



PhotoDNA

In 2009, Microsoft, working with Dartmouth College, developed PhotoDNA, a technology that aids in finding and removing some of the “worst of the worst” images of child sexual exploitations from the Internet. Microsoft donated the PhotoDNA technology to the National Center for Missing & Exploited Children (NCMEC), who established a PhotoDNA-based program for online service providers to help disrupt the spread of child pornography online. Over the next year, Microsoft, working with NCMEC, implemented a gradual rollout of PhotoDNA on Bing, SkyDrive and Hotmail services. In early 2011, Facebook joined Microsoft in sublicensing the technology for use on its network. It is our hope that other online service providers will follow Microsoft and Facebook’s lead in adopting this game-changing technology.


Facebook Implements Microsoft’s PhotoDNA Technology


May 19, 2011

Facebook adopts PhotoDNA and joins Microsoft and The National Center for Missing & Exploited Children to disrupt the proliferation of online child exploitation.

 **Blog:** 500 Million Friends Against Child Exploitation

 **Blog:** Facebook To Use Microsoft’s PhotoDNA Technology to Combat Child Exploitation

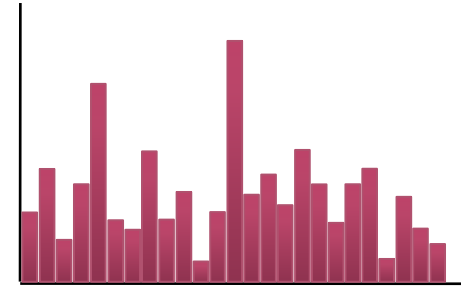
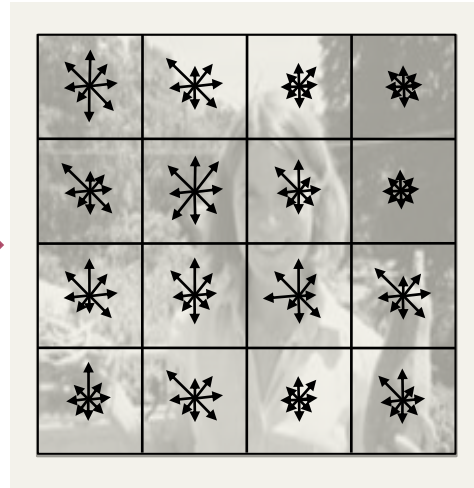
 **Video:** The Next Chapter in Protecting Children Online

 **Interactive:** Join the Facebook live event, May 20, 3 pm EDT



Approach

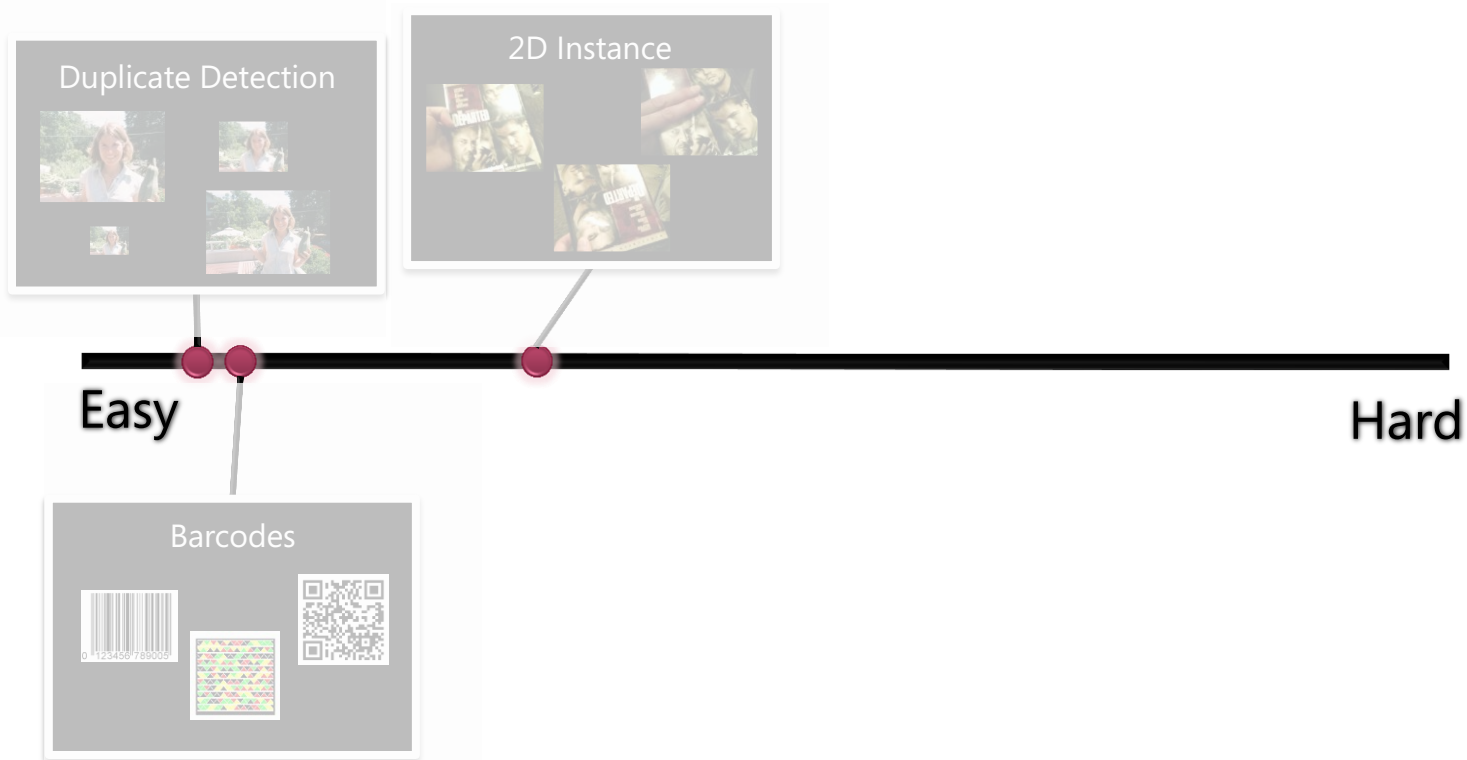
- Speed is most important



Fuzzy Hash

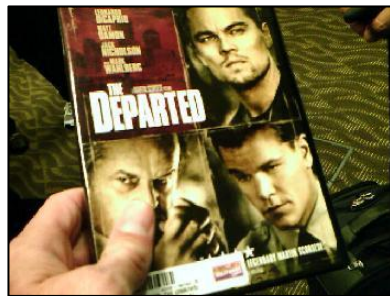
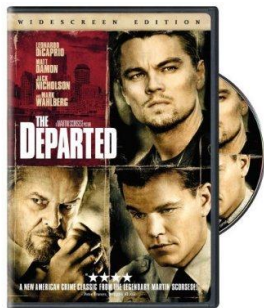
95% found
1 in 100 million false positives

2D Object Instance Recognition



2D Object Instance Recognition

Is this the same planar object?



2D Object Instance Recognition

Why is this interesting?

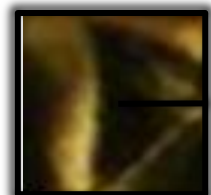
- Recognize real world objects
- Search using images
- Add metadata to images



Approach

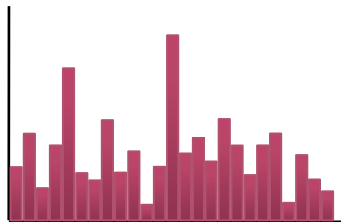


Find interest points

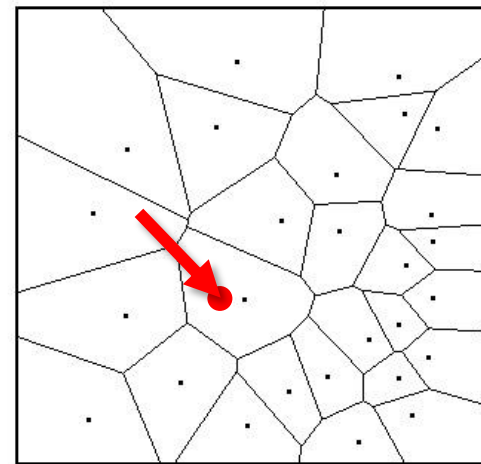


Extract patches

Approach



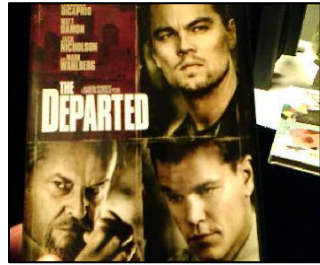
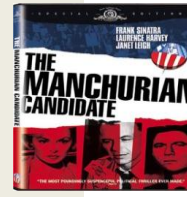
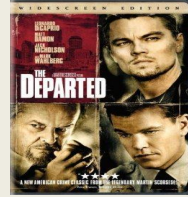
Compute descriptors













Quantize


kd-tree
vocabulary tree

Approach

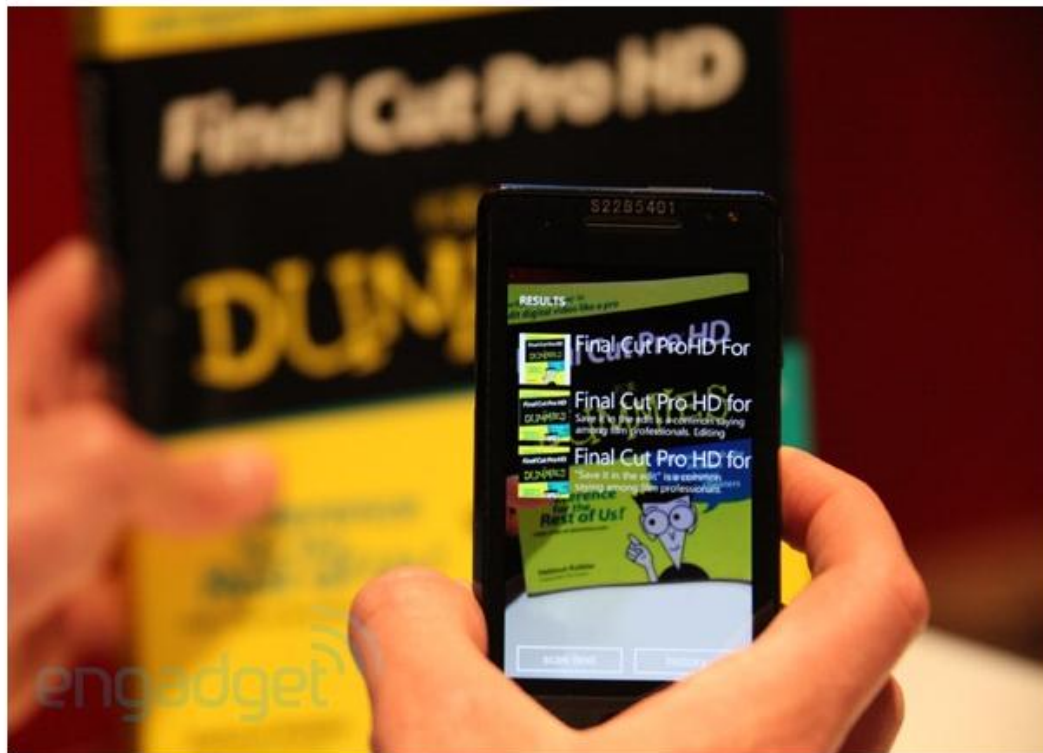


| | | | | |
|---|---|--|---|---|
|  | |  | | |
|  | | | | |
|  | |  |  | |
|  | | | |  |
|  | |  | | |
| | 0 | 3 | 0 | 1 |

Windows Phone Mango and Bing Vision hands-on

By Tim Stevens  posted May 24th 2011 12:43PM

HANDS-ON



Object Recognition Landscape

| Yes | Maybe | No |
|--|--|---|
|  <p>Magazines/ Books</p> |  <p>DVDs/CDs</p> |  <p>Places</p> |
|  <p>Products</p> |  <p>Storefronts</p> |  <p>People</p> |
|  <p>Posters/ Advertisements</p> |  <p>Toys/ 3D objects</p> |  <p>Pets</p> |
|  <p>Beer mats</p> |  <p>Scenery</p> |  <p>Plants</p>  <p>Clothes</p>  <p>Cars</p>  <p>Shoes</p> |

Microsoft
Research

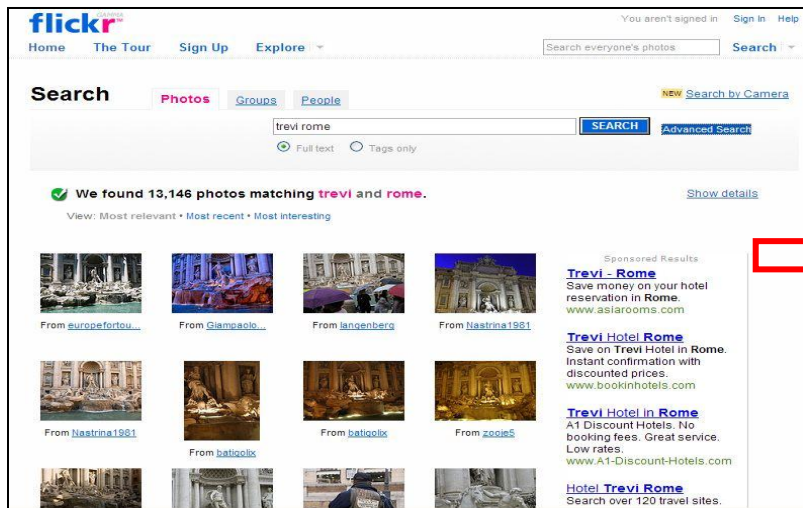
4. Internet Images

Noah Snavely
Cornell

+ ...

Photo Tourism

[Snavely, Seitz, Szeliski, SIGGRAPH 2006]



Images on the Internet

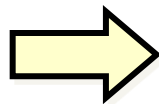


Computed 3D structure

Photo Tourism overview



Input photographs



Scene
reconstruction



Relative camera positions
and orientations

Point cloud

Sparse correspondence

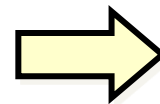


Photo Explorer

Incremental structure from motion



Navigation: Prague Old Town Square



Annotations: Notre Dame



Microsoft Photosynth

The screenshot shows the Microsoft Photosynth website interface. At the top, there is a navigation bar with the Microsoft Photosynth logo, links for Home, Explore, About, and My Photosynths, a search bar, and buttons for Sign Out and Upload. The main content area features a large 3D model of the Pyramid of Menkaure, which is a 3D reconstruction of the pyramid. The model is shown in a perspective view, with a white wireframe overlay indicating the 3D structure. The background of the model is a photograph of the pyramid in a desert setting. Below the model is a control panel with various icons for navigation and interaction. To the right of the model, there is a sidebar with the text "Use your camera to stitch the world." and a list of categories for 3D results, including Towers, Collections, Museums, National Parks, Markets, Forests, Insects, Archaeology, Galleries, Aerial Views, and Bridges. Below the list, there is a section for "Browse the best Photosynths uploaded in the last 7 days, or of all time." and a link to "Bing Maps." At the bottom of the page, there is a navigation bar with five icons and their corresponding labels: "Create your Synth" (camera icon), "About Photosynth" (leaf icon), "Explore Synths" (3D model icon), "Latest Synth News" (document icon), and "Discussion Forum" (leaf icon).

Microsoft Photosynth Home | Explore | About | My Photosynths Search Sign Out Upload

Pyramid of Menkaure 47 100%
NationalGeographic 9/29/2008 2062 Views PHOTOS SYNTH 0 2

Use your camera to stitch the world.

See the amazing 3D results for:

- Towers
- Collections
- Museums
- National Parks
- Markets
- Forests
- Insects
- Archaeology
- Galleries
- Aerial Views
- Bridges

Browse the best Photosynths uploaded in the last 7 days, or of all time.

You can also explore the world of Photosynth on Bing Maps.

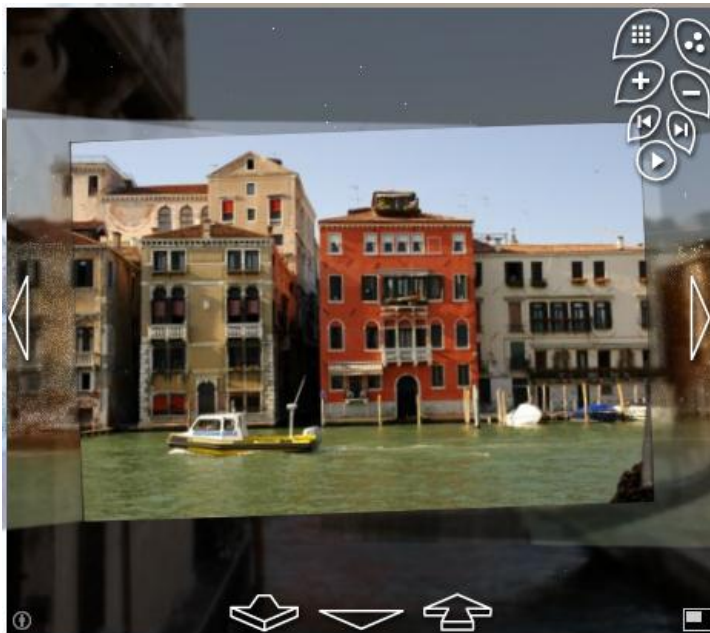
Create your Synth About Photosynth Explore Synths Latest Synth News Discussion Forum

<http://photosynth.net/>

Microsoft Photosynth

- 3D reconstruction
- Multi-resolution streaming & zooming
- Quad-based exploration
- Community photo sharing

<http://photosynth.net/>



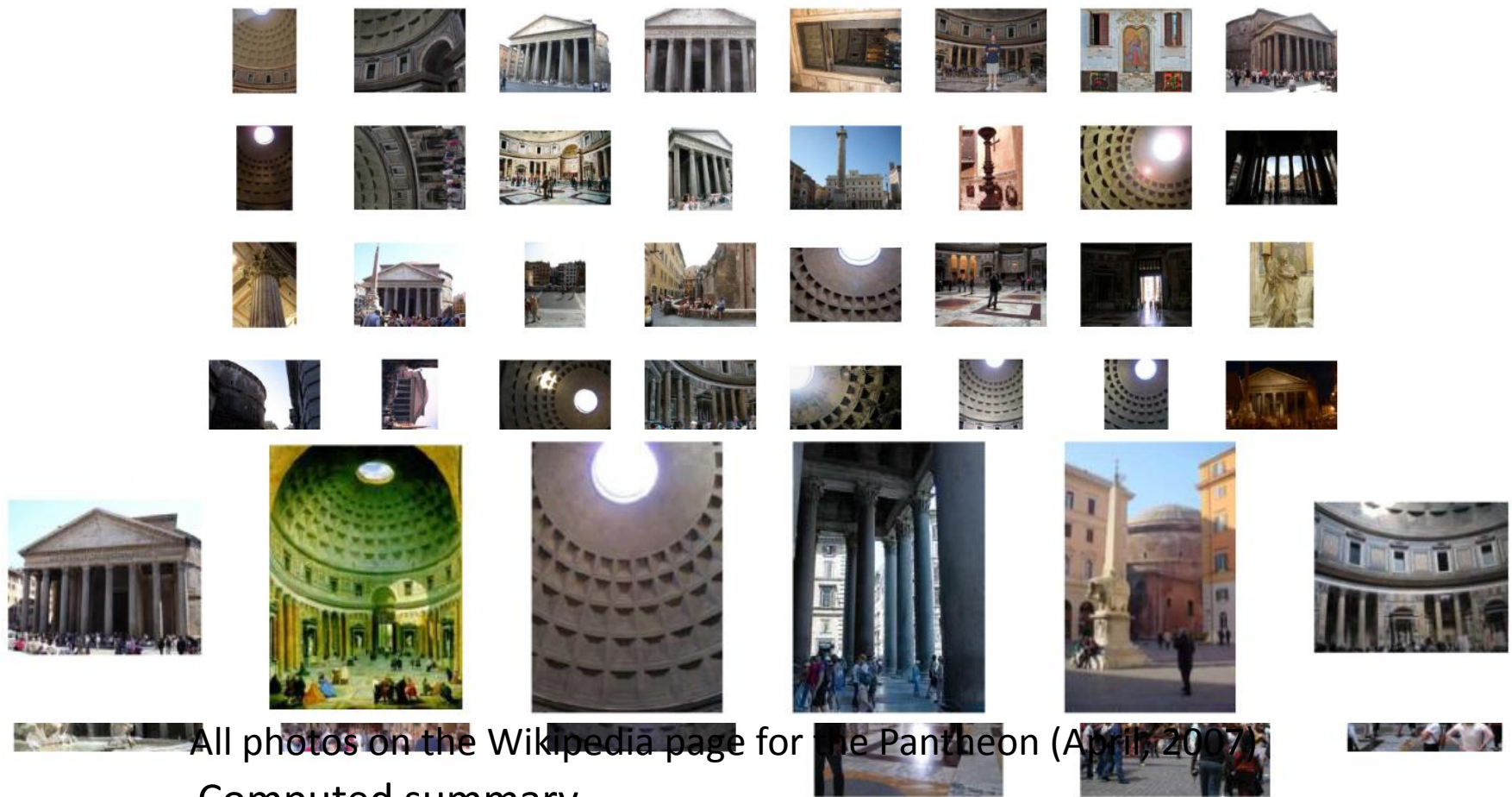
How well does this (Internet) scale?

Scene summarization for online photo collections

[Simon, Snavely, Seitz, ICCV 2007]

The screenshot shows the Flickr website interface. At the top, it says "Signed in as Jimantha" with links for "Help" and "Sign Out". Navigation links include "Home", "You", "Organize", "Contacts", "Groups", and "Explore". A search bar contains the text "Search everyone's photos" and a "Search" button. Below the search bar, there are tabs for "Photos", "Groups", and "People". The search results section shows "Everyone's Photos" selected, with the search term "pantheon rome" entered. There are options for "Full text" and "Tags only". A message states "We found 29,504 results for photos matching pantheon and rome." with a "View as slideshow" link. Below this, there are view options: "Most relevant", "Most recent", "Most interesting", "Details", and "Thumbnails". The main content area displays a grid of photo thumbnails. The first row shows four photos: three from various users and one sponsored result. The second row shows four photos from different users. The third row shows four photos, including one from "batiqolix" and another from "batiqolix". The sponsored results on the right include "Pantheon Hotels In Rome - Save 75%", "Pantheon - Free Download", "Hotel Pantheon - Rome", and "Hotel Pantheon - Rome".

flickr
Signed in as Jimantha Help Sign Out
Home You Organize Contacts Groups Explore Search everyone's photos Search
Search Photos Groups People
Everyone's Photos pantheon rome SEARCH Advanced Search Search by Camera
Full text Tags only
We found 29,504 results for photos matching pantheon and rome. View as slideshow (#)
View: Most relevant • Most recent • Most interesting Show: Details • Thumbnails
From georgeogoodm...
From georgeogoodm...
From marklsto
From Piry
From batiqolix
From batiqolix
From batiqolix
From batiqolix
From BRUNO MENDEZ...
From d_cherubini
From Andrei S
From zoojie5
Sponsored Results
Pantheon Hotels In Rome - Save 75%
Book Pantheon Hotels Online. Save Up To 75% Off Standard Rates.
www.Priceline-Europe.com/pantheon
Pantheon - Free Download
Download Pantheon free when you try GamePass.
www.rearcade.com/download
Hotel Pantheon - Rome
Pay 75% Less if you book Online. No reservation costs. Pay at Hotel.
www.caupona.net/Rome
Hotel Pantheon - Rome
Save up to 75% when you book Hotel Pantheon in



All photos on the Wikipedia page for the Pantheon (April, 2007)

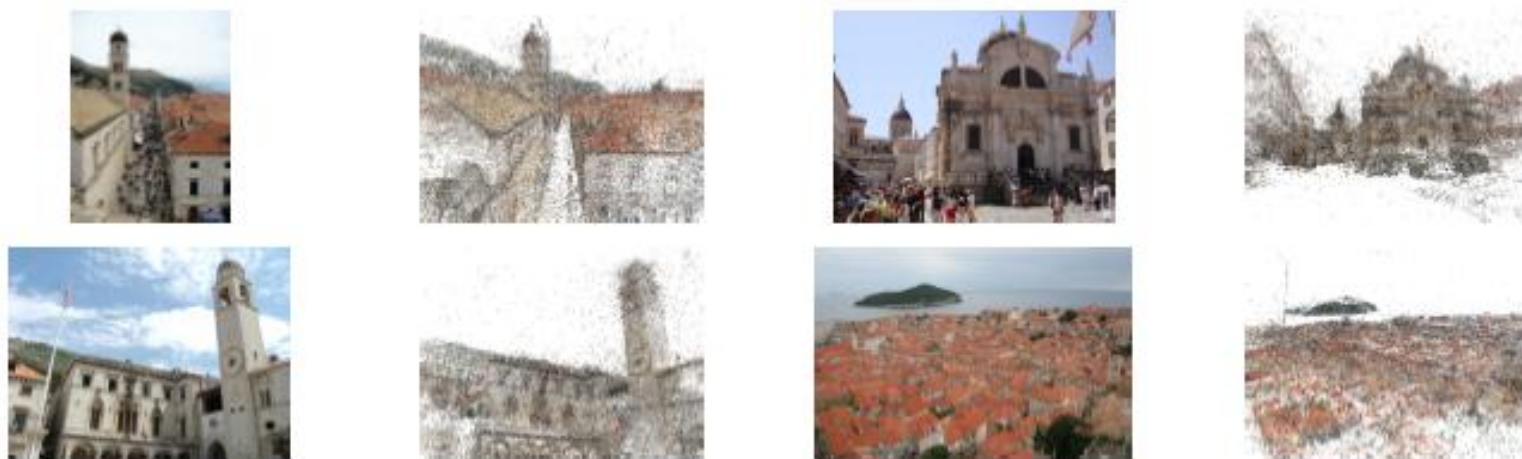
Computed summary

Building Rome in a Day

Sameer Agarwal, Noah Snavely,
Ian Simon, Steven M. Seitz,
Richard Szeliski
ICCV'2009



Results: Dubrovnik



(a) Dubrovnik: Four different views and associated images from the largest connected component. Note that the component captures the entire old city, with both street-level and roof-top detail. The reconstruction consists of 4,585 images and 2,662,981 3D points with 11,839,682 observed features.

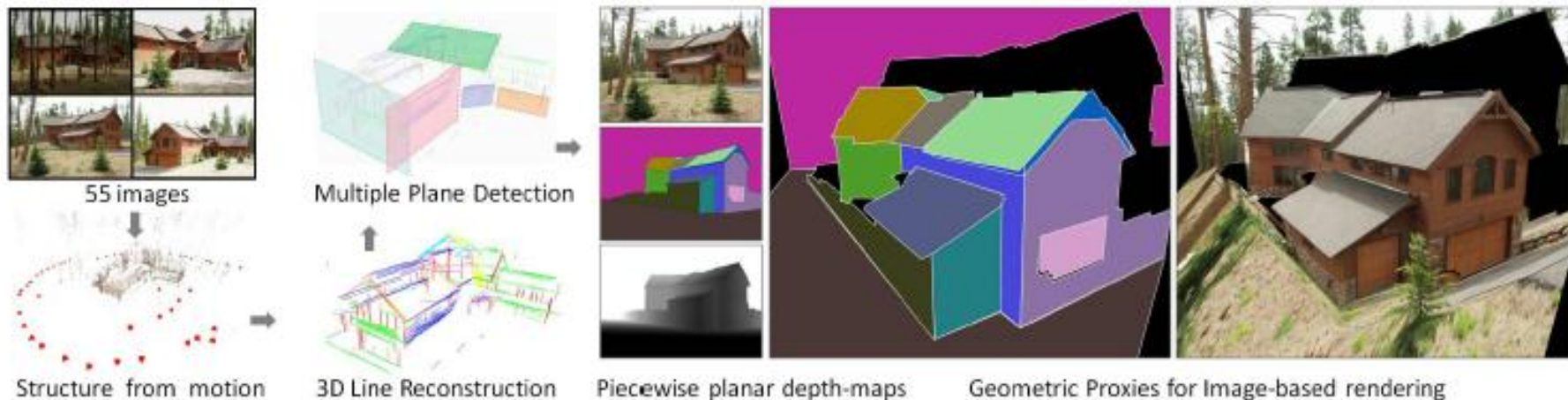


What about “real” 3D?

Piecewise Planar Stereo for Image-based Rendering

[Sinha, Steedly, and Szeliski. ICCV 2009]

Per-image piecewise-planar proxies



View Interpolation

Examples



Blogs

TechFest 2011: 3D Scanning with a regular camera or phone!

Posted: Mar 09, 2011 at 6:39 PM

By: [Laura Foy](#)

★★★★★ (3) | 44,866 Views | 14 Comments

Avg Rating: 5

Tweet 57

Share 73



< > embed

Download ?

Right click "Save as..."

High Quality WMV

(PC, XBox, MCE)

MP3

(Audio only)

WMA

(Audio only)

Mid Quality WMV

(Lo-band, Mobile)

High Quality MP4

(iPad, WP7)

MP4

(iPod, Zune HD)

Related posts



Expert to Expert: Helen Wang and Alex Moshchuk - ...



Project JSMeter: JavaScript Performance Analysis...



E2E: Erik Meijer and Leslie Lamport - Mathematical.

3-D television is creating a huge buzz in the consumer space, but the generation of 3-D content remains a largely professional endeavor. Our research demonstrates an easy-to-use system for creating photorealistic, 3-D-image-based models simply by walking around

What about “regular” Internet Images?

ImageNet is an image database organized according to the **WordNet** hierarchy (currently only the nouns), in which each node of the hierarchy is depicted by hundreds and thousands of images. Currently we have an average of over five hundred images per node. We hope ImageNet will become a useful resource for researchers, educators, students and all of you who share our passion for pictures. [Click here](#) to learn more about ImageNet, [Click here](#) to join the ImageNet mailing list.

SEARCH



What do these images have in common? *Find out!*

Internet Computer Vision



Computer Vision and the Internet (09w5126)

Arriving Sunday, August 30 and departing Friday September 4, 2009



COMP 790-096: Con

Fall 2007, Tuesdays 3:30-4:30, S

Instructor: [Svetlana Lazebnik](#) (C

Quick links: [presentation sched](#)



Proceedings OF THE IEEE

AUGUST 2010 / VOL. 98 / NO. 8

CONTENTS

Second IEEE Workshop on Intern CVPR 2009)

Program

[IV Home Page](#)

[Workshop Home](#)

[Areas of Interest](#)

[Program](#)

Workshop Home

General Chairs

Thomas S. Huang, *UIUC*
Harry Shum, *MSR*

Program Chairs

Shai Avidan, *Adobe Research*
Simon Baker, *MSR*
Ying Shan, *Microsoft AdCenter Labs*

SPECIAL ISSUE

INTERNET VISION

Edited by S. Avidan, S. Baker, and Y. Shan

- 1370 Scene Reconstruction and Visualization From Community Photo Collections**
By N. Snavely, I. Simon, M. Goesele, R. Szeliski, and S. M. Seitz
| INVITED PAPER | Recent progress is described in digitizing and visualizing the world from data captured by people taking photos and uploading them to the web.
- 1391 Infinite Images: Creating and Exploring a Large Photorealistic Virtual Space**
By B. Kaneva, J. Sivic, A. Torralba, S. Avidan, and W. T. Freeman
| INVITED PAPER | This proposed system uses 3-D-based navigation to browse large

DEPARTMENTS

- 1363 POINT OF VIEW**
Cyber-Physical
Systems: Close
Encounters Between
Two Parallel Worlds
By R. Poovendran
- 1367 SCANNING THE
ISSUE**
Internet Vision
*By S. Avidan, S. Baker,
and Y. Shan*

What *else* can we do with these photos?

ShadowDraw: Real-Time User Guidance for Freehand Drawing

Yong Jae Lee, Larry Zitnick, and Michael Cohen

[SIGGRAPH 2011](#)

Computer Assisted Drawing

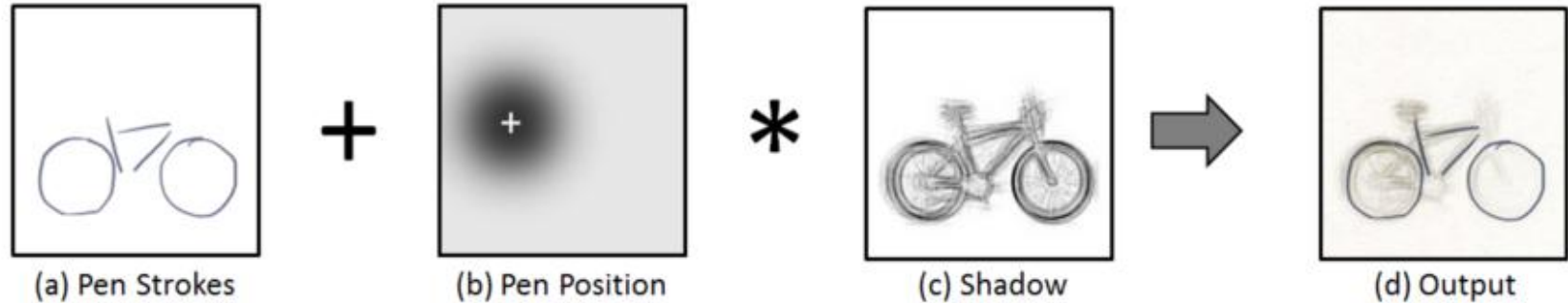
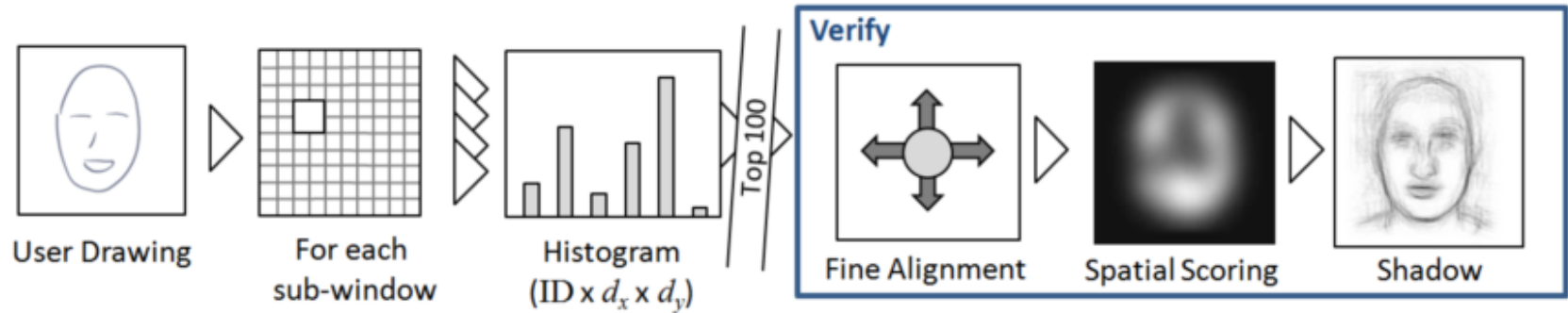
Main idea:

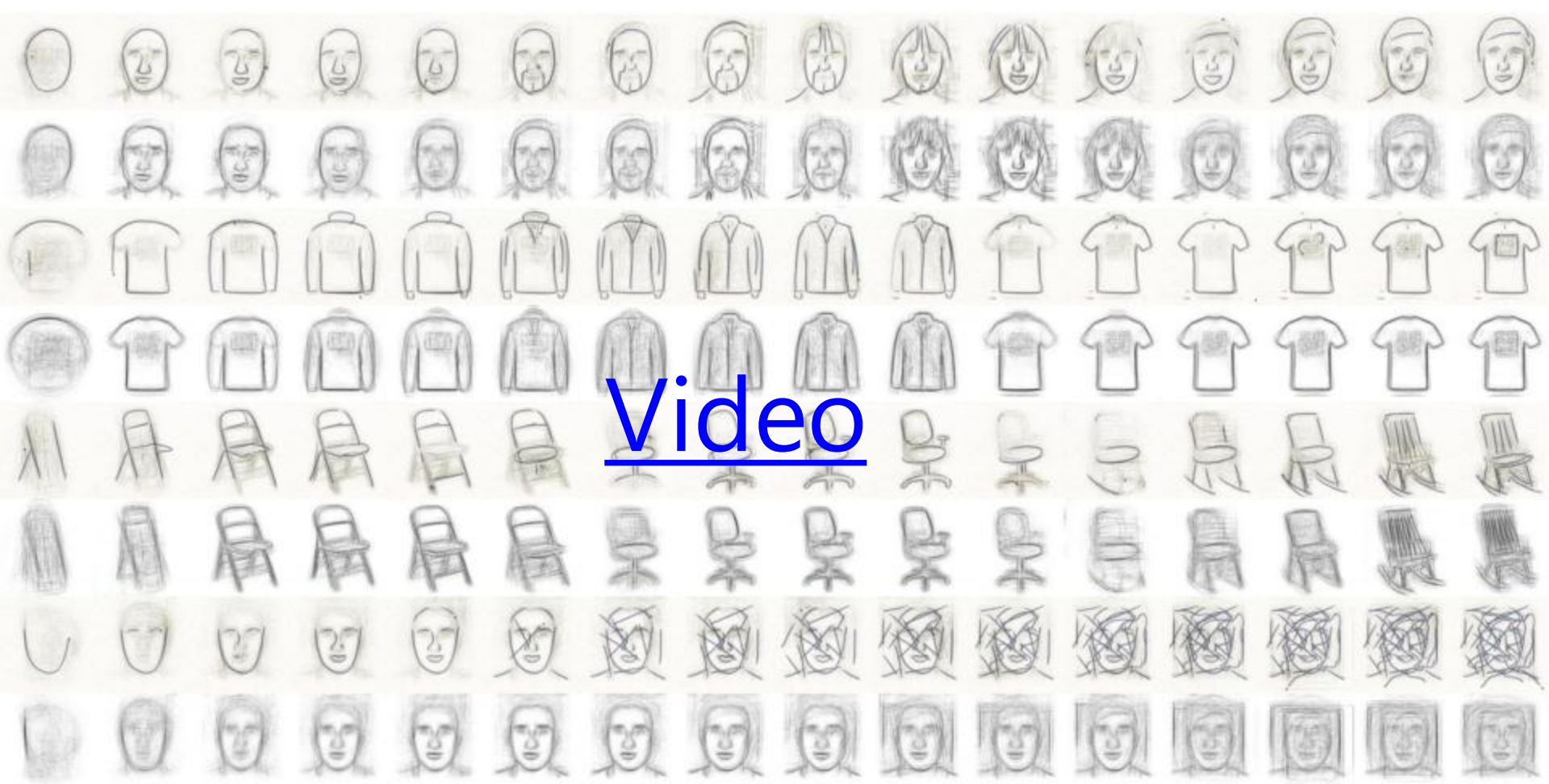
- Use a large database of images to suggest good contours for your drawing



Examples of database images and corresponding edge maps

Computer Assisted Drawing





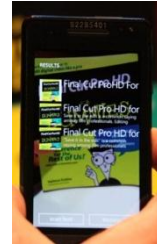
Video

Wrapping up

- Exciting time to be working in computer vision
- Machine learning and massive (Internet) data are having a huge impact
 - Body tracking
 - Medical imaging
 - Object recognition
 - Internet-scale reconstruction
- General image recognition is still an open problem



Question?



URLs for Web content

- <http://www.engadget.com/2011/05/24/windows-phone-mango-and-bing-vision-hands-on/>
- <http://channel9.msdn.com/posts/TechFest-2011-3D-Scanning-with-a-regular-camera-or-phone>
- <http://www.microsoft.com/presspass/presskits/photodna/>
- <https://webpace.utexas.edu/yl3663/~ylee/shadowdraw/shadowdraw.html>

Microsoft Research

Faculty Summit



FUTURE WORLD

2011 ← → 2031