Course Project (aka HW #5)

Requirements

- Thursday, **November 17**: Team members (3), tentative title, and abstract
- Thursday, **December 1**: Progress report
- **December 13** and **15**: Class presentations (5% of course grade)
- Tuesday, **December 20**: Final project report and web page (20% of course grade)

Project Ideas and Grading

- "Straightforward" approach: Pick a paper, implement it, extend it, and modify it in some ways, and perform experimental evaluation
- Pick a paper that's easy to understand and on a topic you're interested in
- Grading based on effort, initiative, ***creativity***, coolness, difficulty, focus, depth, implementation, quality of experimental results, originality, project report write-up

Project Ideas

- Best to pick a **narrower** topic and go **deeply** into it rather than pick a broad topic that is not very in-depth on any part

Sources for Finding Ideas

- Recent projects by researchers doing computational photography – see "Links" page on course web site
- Recent papers in computational photography, computer vision, or computer graphics conferences – see "HW" and "Links" pages
- Previous student projects in CS 534
- Other computational photography course projects and assignments
  - CMU, Illinois, Brown, Columbia, etc.
- Papers listed on "computational photography" page on Wikipedia
- ImageNet Challenges
  - http://image-net.org/challenges/LSVRC/2016/

Class Presentation

- December 13 and 15
- 5 minutes
- Conference-style "powerpoint" talk
- State problem, give motivation and example, background, description of method and main ideas of the approach, initial results, discussion of strengths and weaknesses of the method, possible future extensions

Project Report

- Due Tuesday, December 20 at 5 p.m.
- ~15 pages (pdf)
- Submit report, code and example results
  - Include how much code written; what work each person contributed
- Grade will be based on report and submitted materials
- Create web page with report and sample output
- Fill out Evaluation Report for each of your teammates

Project Policies

- 3-person project groups very strongly preferred
- Feel free to use code or data you find on the web, provided it does not make your project trivial
- Implementation does ***not*** need to be in Matlab
  - OpenCV is an alternative open source library with C++ interface
- All outside sources should be fully cited in the project report
- Feel free to talk to other people about the project, but do your own implementation
- Each person should have a clearly identifiable part that they are responsible for; describe in the project report

Sources of Image Data

- Lots of image datasets on the web!
- CV datasets on the web
- ImageNet
- Computer vision test images
- Images from Flickr, Twitter, Google, etc.

Some Topic Areas

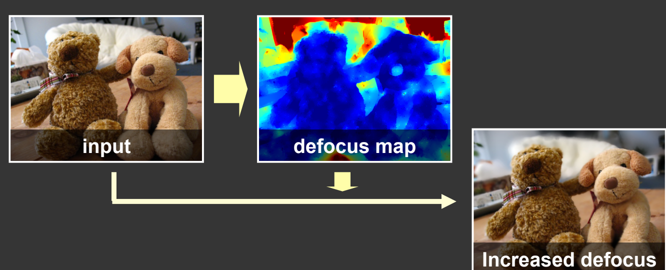
- Image quality improvement
- Photo composition
  - Panoramas, collages, matting, segmentation, cut-and-paste
- Internet vision
  - Using collections of images from web
  - Social photography
  - Image retrieval – see Google Image Swirl, for example
- Places
- People
- Beyond conventional cameras

Image Quality Improvement

- **Defocusing**
  - S. Bae and F. Durand, Defocus magnification, *Proc. Eurographics*, 2007
  - M. Levoy, SynthCam
    - Shallow depth of field is often desired
- **Denoising**
  - A. Buades et al., A non-local algorithm for image denoising, *Proc. CVPR*, 2005
    - One of the most effective denoising methods
  - C. Tomasi and R. Manduchi, Bilateral filtering for gray and color images, *Proc. ICCV*, 1998
- **Dehazing**
  - K. He et al., Single image haze removal using dark channel prior, *Proc. CVPR*, 2009
    - Uses matting

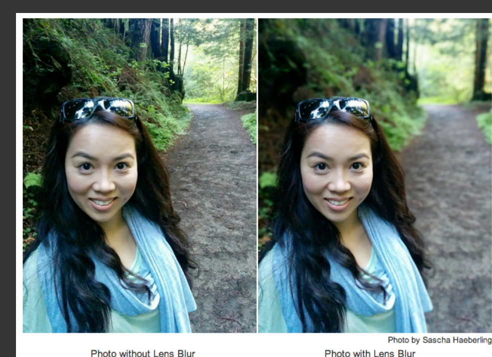
Defocus (Bae and Durand, 2007)

1. User provides a single input photograph
2. System automatically produces the defocus map
3. Uses Photoshop's lens blur to generate the defocus magnified result



Google Camera's Lens Blur App

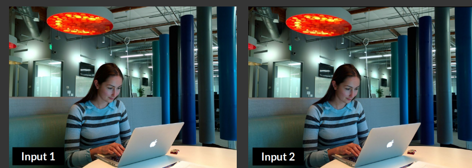
http://googleresearch.blogspot.com/2014/04/lens-blur-in-new-google-camera-app.html

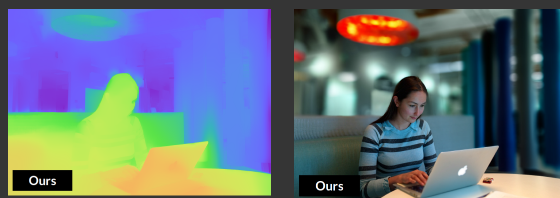


Original With Lens Blur

Defocus

J. Barron et al., Fast Bilateral-Space ***Stereo*** for Synthetic Defocus, *Computer Vision and Pattern Recognition Conf.*, 2015



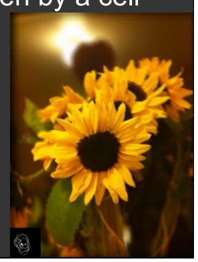


Tilt-Shift Photography

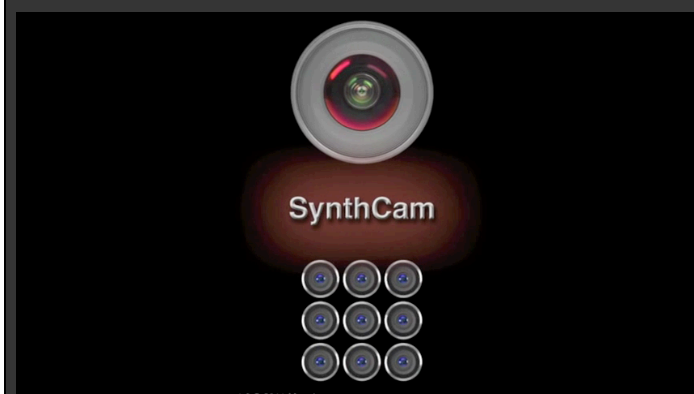
- **Miniature faking** is a process in which a photograph of a life-size location or object is made to look like a photograph of a miniature scale model
- Blurring parts of the photo to simulate a shallow depth of field normally encountered in close-up photography
- https://en.wikipedia.org/wiki/Miniature_faking

Changing the Depth of Field: Synthetic Aperture Photographs

- Phone cameras have small apertures (big f-number), giving a large depth of field, which may *not* be desirable
- Task: Synthesize a new image corresponding to a large aperture from a ***video*** taken by a cell phone
- Levoy's SynthCam app for iPhone
  - http://sites.google.com/site/marclevoy/



SynthCam



Dehazing

Goals of Haze Removal

depth

- Scene restoration
- Depth estimation

Image Quality Improvement

- **Tone Adjustment and Relighting**
  - D. Lischinski, et al., Interactive local adjustment of tonal values, *Proc. SIGGRAPH*, 2006
    - Easy to read and implement
  - S. Bae et al., Two-scale tone management for photographic look, *Proc. SIGGRAPH*, 2006
    - Easy to read; uses bilateral filtering
- **Shadow Editing**
  - T-P. Wu et al., Natural shadow matting, *ACM Trans. Graphics*, 2007
    - Uses matting; many useful application scenarios
- **Possible Application: Sky Editing and Enhancement**

Interactive Tone Adjustment

SIGGRAPH 2006

Interactive Local Adjustment of Tonal Values

Dani Lischinski
Zeev Farbman
Matt Uyttendaele
Richard Szeliski

Artifact Removal: Image De-Fencing

Y. Liu, T. Belkina, J. Hays, and R. Lublinerman, Image De-Fencing, *Proc. CVPR*, 2008

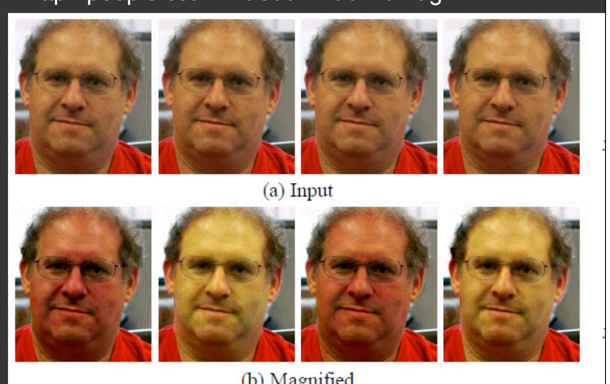
Super-Resolution

- From a single **photo** or a **video**
- D. Glasner et al.,, Super-resolution from a single image, Proc. Int. Conf. on Computer Vision, 2009

Eulerian Video Magnification

http://people.csail.mit.edu/mrub/vidmag/

Thanks Aaron Wurtinger-Knaack



Bottom row shows the subject's pulse signal amplified

Massachusetts Institute of Technology

CSAIL

Revealing Invisible Changes In The World

Created for the NSF International Science & Engineering Visualization Challenge 2012

Image Colorization

- R. Zhang et al., Colorful image colorization, *ECCV*, 2016
- http://richzhang.github.io/colorization/
- http://demos.algorithmia.com/colorize-photos/
- Uses deep learning

Image Style Transfer

- L. Gatys et al., Image style transfer using convolutional neural networks, *CVPR*, 2016
- G. Kogan http://www.genekogan.com/works/style-transfer.html
- C. Ham, Sketch-based image synthesis

Gene Kogan's Style Transfer

Deep Learning

- Unsupervised learning of a feature hierarchy
- Multiple layers work to build an improved feature space
  - 1st layer learns 1st-order features (e.g., edges)
  - 2nd layer learns higher-order features (combinations of first layer features)
  - Etc. for subsequent layers of features
- Each layer combines patches from previous layer using a set of convolution filters, followed by "pooling," which compresses and smooths the data

31

Deep Convolutional Neural Networks

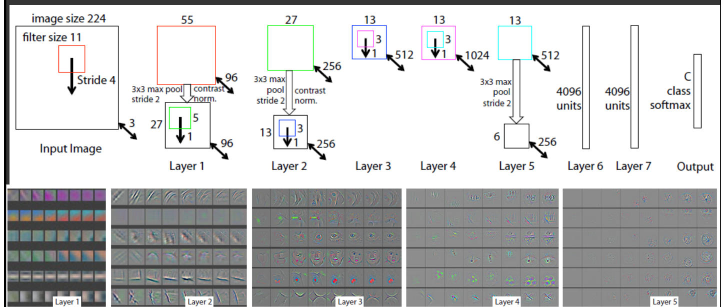


Figure 3. Evolution of model features through training. Each layer's features are displayed in a different block. Within each block, we show a randomly chosen subset of features at epochs [1,2,5,10,20,30,40,64]. The visualization shows the strongest activation (across all training examples) for a given feature map, projected down to pixel space using our deconvnet approach. Color contrast is artificially enhanced and the figure is best viewed in electronic form.

A. Krizhevsky et al., ImageNet classification with deep convolutional neural networks, *NIPS*, 2012

Feature Extraction

- Deep convolutional neural network
  - 7 feature layers, 650K neurons, 60M parameters, 630M connections
- Supervised learning used to train model on ImageNet (1.2 million images with 1,000 classes)
- Use the output of the 6th layer in the deep network as a feature vector (4,096-dimensional feature vector)

CNN Image Features

- https://github.com/rbgirshick/rcnn
- Downloadable, pre-computed R-CNN detectors ("regions with CNN features")
- Detectors trained on PASCAL VOC 2007 train+val, 2012 train, and ILSVRC13 train+val

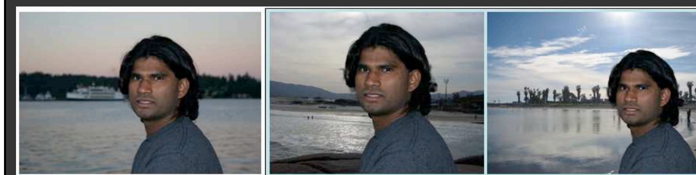
Image/Video Retargeting

- F. Liu and M. Gleicher. Automatic Image Retargeting with Fisheye-View Warping, *Proc. ACM UIST, 2005*
- F. Liu and M. Gleicher. Video Retargeting: Automating Pan-and-Scan, *ACM Multimedia, 2006*
- L. Wolf, M. Guttmann, D. Cohen-Or, Non-Homogeneous, Content-driven Video Retargeting, *ICCV*, 2007

Content-based Image Synthesis

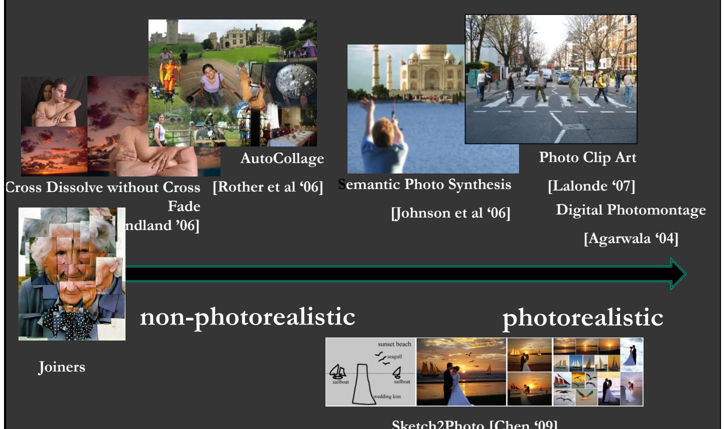
N. Diakopoulos *et al.*, *Conference on Image and Video Retrieval,* 2004

Background Replacement



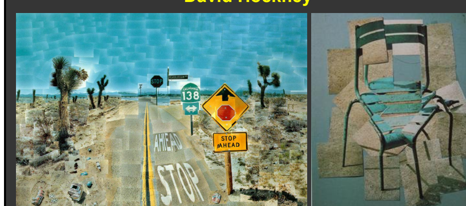
(a) input (b) result 1 result 2

Combining Multiple Images



Creating "Joiners"

David Hockney

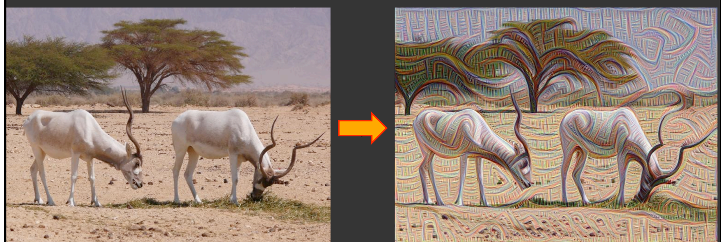


Flickr "Hockneyesque" pool

L. Zelnik-Manor and P. Perona, Automating Joiners, *Proc. 5th Int. Symp. Non-Photorealistic Animation and Rendering*, 2007

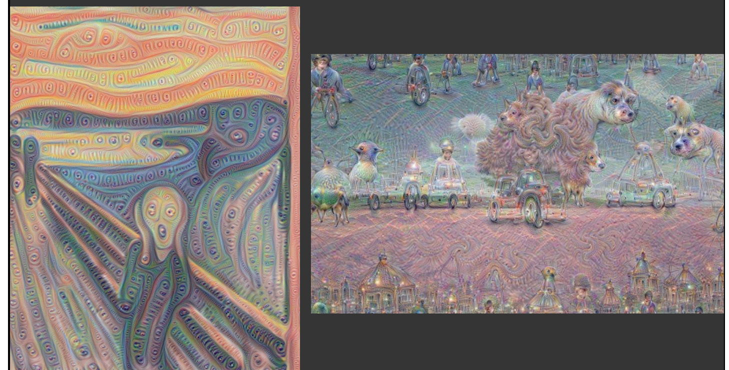
Deep Dreams / Inceptionism

Google project by A. Mordvintsev, C. Olah, and M. Tyka

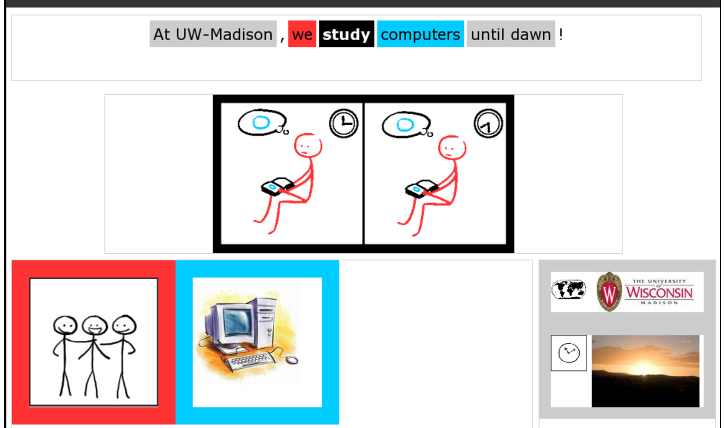


Produce results like these but without using a neural network approach

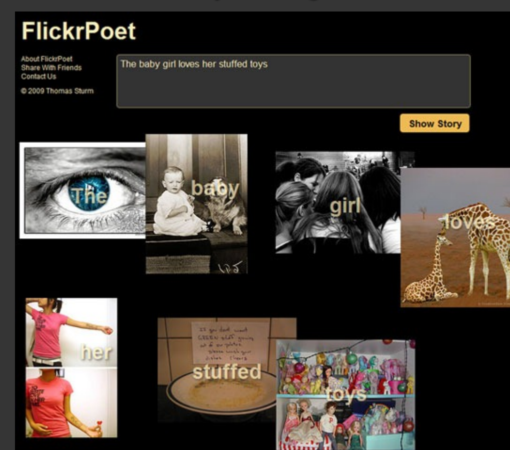
Thanks Aaron Wurtinger-Knaack



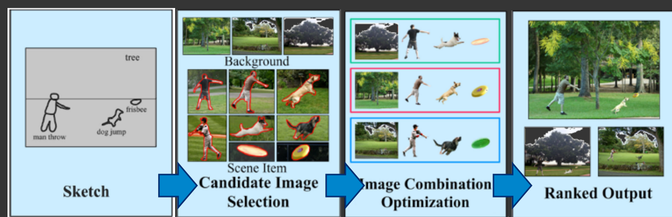
Visual Storytelling: Text-to-Picture



Visual Storytelling: FlickrPoet



Sketch-to-Photo



T. Chen *et al.*, Sketch2Photo, *Proc. SIGGRAPH Asia*, 2009

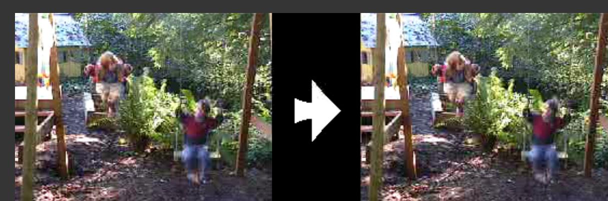
Very Long Panoramas

J. Sivic, B. Kaneva, A. Torralba, S. Avidan, and W. Freeman, Creating and Exploring a Large Photorealistic Virtual Space, *Proc. Internet Vision Workshop*, 2008

Video Textures

- A. Schodl, R. Szeliski, D. Salesin and I. Essa, Video textures, SIGGRAPH 2000
- A. Agarwala et al., Panoramic video textures, SIGGRAPH 2005
- Z. Liao, N. Joshi, N. Joshi, and H. Hoppe, Automated video looping with progressive dynamism, SIGGRAPH 2013

Video Textures



video clip **video texture**

Multi-Perspective Images

M. C. Escher, 1956

Rademacher and Bishop,1998

Images that depict *more* than can be seen from any *single* viewpoint, yet remain interpretable

Multi-Perspective Panoramas

Input Video

Space-time scene manifolds, Y. Wexler and D. Simakov, Proc. ICCV, 2005

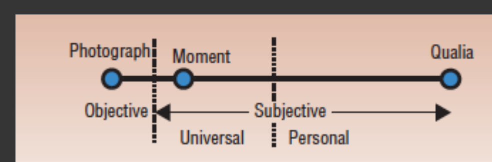
Multi-Perspective Panoramas

Space-time Scene Manifolds. Y. Wexler, D. Simakov in ICCV 2005

The Moment Camera

M. Cohen and R. Szeliski, The Moment Camera, *IEEE Computer Society magazine*, August 2006

“Future cameras will let us “capture the moment,” not just the instant when the shutter opens. The moment camera will gather significantly more data than is needed for a single image. This data, coupled with automated and user-assisted algorithms, will provide powerful new paradigms for image making.”

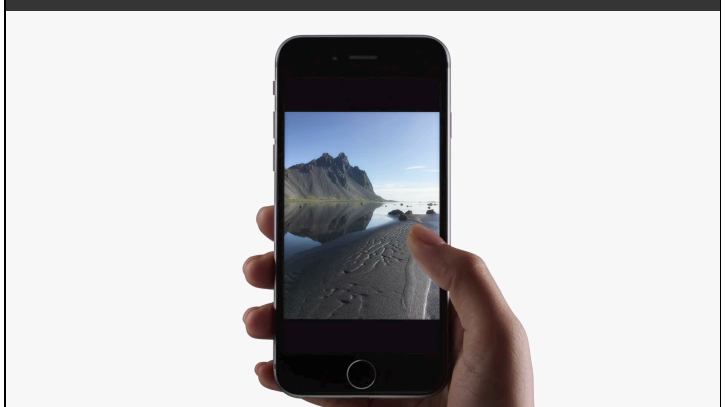


"Moment Camera" Video Clips

- Camera is *always* recording images using a finite round-robin buffer of 10s or 100s of frames, providing a short space-time video clip
- Instagram's Boomerang
  - 1 sec burst of 5 photos, played in a loop
- Apple's Live Photos
  - 1.5 sec buffer of frames before and after shutter pressed
- Google's Photos Assistant
  - Finds repeated photos and creates collages, animations, or panoramas
- Better animated GIFs

Instagram Boomerang

Apple Live Photos



Better Selfies

- Applied to still or video clips
- Snapchat's animated Lenses

- FaceTune
- Perfect365

Snapchat Lenses

Interactive Digital Photomontage

"Moment Camera" is *always* recording images using a finite round-robin buffer of perhaps 500 frames, or 5 seconds, resulting in a "space-time volume"

Interactive Digital Photomontage

set of originals

actual

output photomontage

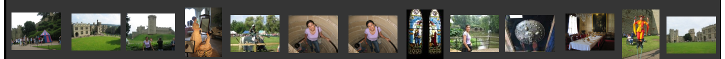
perceived

- Generalize to video
  - Combine short video clips of separate moving objects into a single composite video containing all moving objects in a single scene

AutoCollage

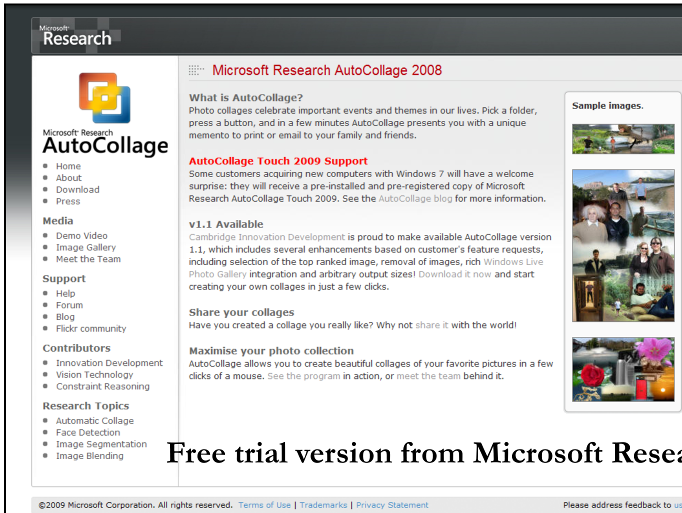
C. Rother et al., *SIGGRAPH 2006*

Microsoft Research AutoCollage



G

- Represe
- One coh in from
- Pack n appr (sky at top)
- Smooth image transitions

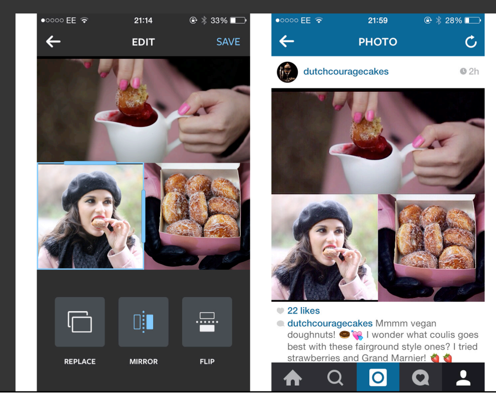


Free trial version from Microsoft Researc

Results

Other Collage Making

- Instagram Layout



Photomontages

Video Summarization

- Z. Lu and K. Grauman, Story-driven summarization for egocentric video, *CVPR*, 2013
- S. Uchihashi et al., Video manga: Generating semantically meaningful video summaries, *ACM Multimedia*, 1999
- B. Truong and S. Venkatesh, Video abstraction: A systematic review and classification, *ACM Trans. Multimedia Computing, Communications and Applications*, 2007

Time-Lapse and Hyper-Lapse Photography

Video stabilization and frame selection

- N. Joshi et al., Real-time hyperlapse creation via optimal frame selection, *SIGGRAPH 2015*
- J. Kopf, M. Cohen and R. Szeliski, First-person hyperlapse videos, *SIGGRAPH 2014*
- R. Martin-Brualla, D. Gallup and S. Seitz, Time-lapse mining from Internet photos, *SIGGRAPH 2015*
- E. Bennett and L. McMillan, Computational time-lapse video, *SIGGRAPH 2007*
- Instagram Hyperlapse
- Microsoft Hyperlapse

Microsoft Hyperlapse

Stereoscopic and 3D Photography

Use of stereo and 3D cameras, and stereo displays (e.g., Oculus Rift, Microsoft HoloLens, and Google Cardboard)

- F. Zhang and F. Liu, Casual stereoscopic panorama stitching, *CVPR*, 2015
- F. Zhang and F. Liu, Parallax-tolerant image stitching, *CVPR*, 2014
- F. Liu, Y. Niu, and H. Jin, Joint subspace stabilization for stereoscopic video, *ICCV*, 2013
- Microsoft Kinect 2 available to use

Using Large Photo Collections

- Photo Tourism / Photosynth
  - Snavely et al., Proc. SIGGRAPH, 2006
- Internet stereo
  - Goesele et al., Proc. ICCV, 2007
- Image completion
  - Hays et al., Proc. SIGGRAPH, 2007
- Photo clipart
  - Lalonde et al., Proc. SIGGRAPH, 2007
- Object recognition
  - Torralba et al., IEEE Trans PAMI, 2008
  - Dataset available containing 1.5 million images of size 32 x 32
- Scene summarization
  - Simon et al., Proc. ICCV, 2007
- Duplicate image discovery
  - Wang et al., CVPR workshop, 2013

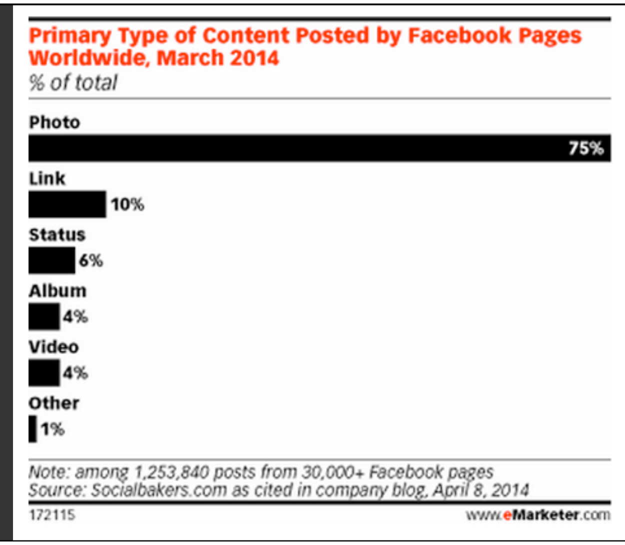
Time-Lapse Mining

Social Photography

- Mobile social media provides near-real-time data about intentional or unintentional communities of users, which can be used for tasks such as surveillance and monitoring
- CNN/Photosynth "The Moment" containing images of Obama's presidential inauguration
  - http://www.cnn.com/SPECIALS/2009/44.president/inauguration/themoment/
- "A Moment in Time" photos taken around the world on the same day at the same time (May 2, 2010, 15:00 UTC)
  - http://www.nytimes.com/interactive/2010/05/03/blogs/a-moment-in-time.html
- How can images (+ text) be used for enhanced communication?

Social Media Users as Sensors

- Social media collects spatio-temporal data of our environment at a vast scale
  - 500 million tweets per day on Twitter
  - 100 million messages per day on Sina Weibo (China)
  - 4.75 billion pieces of content shared daily on Facebook
- Visual content is growing rapidly
  - 350 million photo uploads per day on Facebook
  - 58 million photos shared on Twitter in Dec 2011
  - 60 million photos shared per day on Instagram



Challenges using Social Media Data

- Text often ambiguous due to language and brevity
- Unstructured, diverse images/videos that contain complex content and poor quality
- Social media users can't be controlled
- Distribution of posts depends on many factors, including population density and time of day
- Location and time stamps associated with social media posts may be erroneous or missing

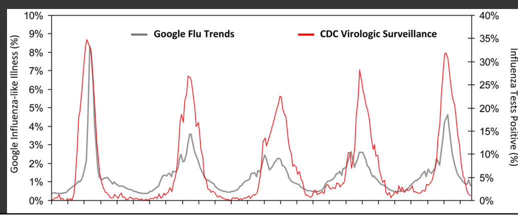
- Beyond "in the wild" and into the "Wild, Wild West" of image (and text) data

Advantages of using Social Media Data

- Lots of data, including multiple modalities (text, images, video, audio)
- Often groups of images taken at a time by users
- Data available over many locations and times
- Many tasks involve measuring spatiotemporal signals, e.g., when, where, how much
- While user's primary intention for a post may be one (unknown) thing, there is often unintended, serendipitous information available

Public Health Surveillance

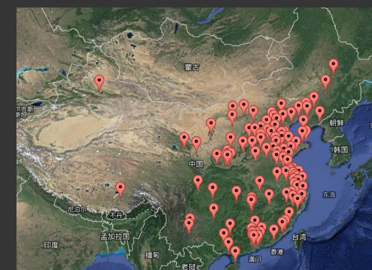
- **Google Flu Trends: Uses aggregated Google search data to estimate flu activity**
- **CDC "Predict the Influenza Season Challenge" (2014)**
- **Most methods use a fixed set of manually-specified text keywords**



Inferring Air Pollution from Social Media

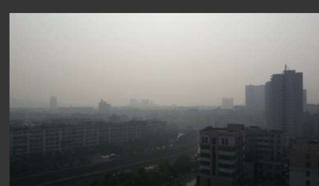
Can we use social media (text and images) as a data source for estimating the Air Quality Index (AQI)?

S. Mei, H. Li, J. Fan, X. Zhu and C. Dyer, *IEEE/ACM Int. Conf. Advances in Social Networks Analysis and Mining*, 2014



Examples of Cities Without a Monitoring Station

Anqing



Pingxiang

Photo Analytics

Analyze images on social media such as Twitter, Tumblr and Instagram to find logos and other brand information

Photo Forensics

- Determine if a photo has been digitally modified
  - http://www.getghiro.org/
  - http://www.imageforensic.org/
  - http://web.cs.dartmouth.edu/research/projects/digital-image-forensics

Projects on Places

Using Images of Places

- Where am I?
- Im2GPS
- Reconstructing building interiors
- Landmark recognition
- Auto-annotation of photo collections
  - "Annotating personal albums via web mining"
- Organizing geo-tagged photo collections
- Make3D

Google Maps' Photo Tours

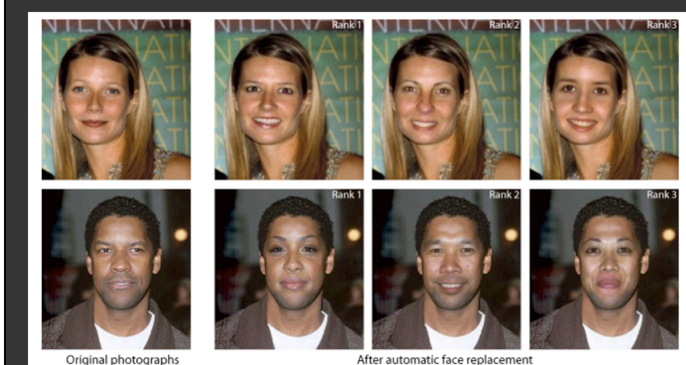
- Photo tours are available for more than 15,000 sites around the world
- http://google-latlong.blogspot.com/2012/04/visit-global-landmarks-with-photo-tours.html

Watch photo tours
in Google Maps

Google

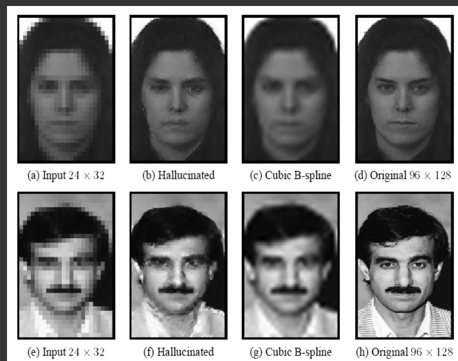
Projects on People

Face Transfer



D. Bitouk, N. Kumar, S. Dhillon, P. Belhumeur, S. K. Nayar, Face Swapping: Automatically Replacing Faces in Photographs, *Proc. SIGGRAPH*, 2008

Hallucinating Faces



S. Baker and T. Kanade, Hallucinating Faces, *Proc. 4th Int. Conf. Automatic Face and Gesture Recognition*, 2000

Hallucinating Faces

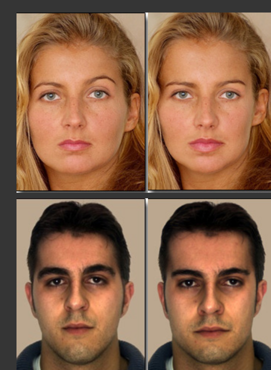
C. Liu, H. Shum, and W. Freeman, Face Hallucination: Theory and Practice, *IJCV*, 2007

Face Shaving

synthesize →

M. Nguyen et al., Image-based Shaving, *Proc. Eurographics*, 2008

Face Beautification



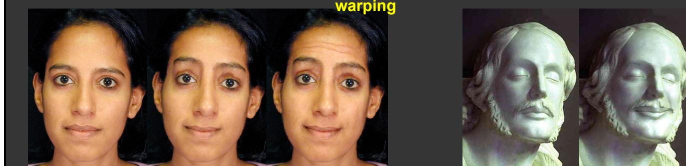
T. Leyvand, D. Cohen-Or, G. Dror and D. Lischinski, Data-Driven Enhancement of Facial Attractiveness, SIGGRAPH 2008

Example-Based Cosmetic Transfer

W.-S. Tong, C.-K. Tang, M. Brown, Y.-Q. Xu
Proc. Pacific Graphics, 2007

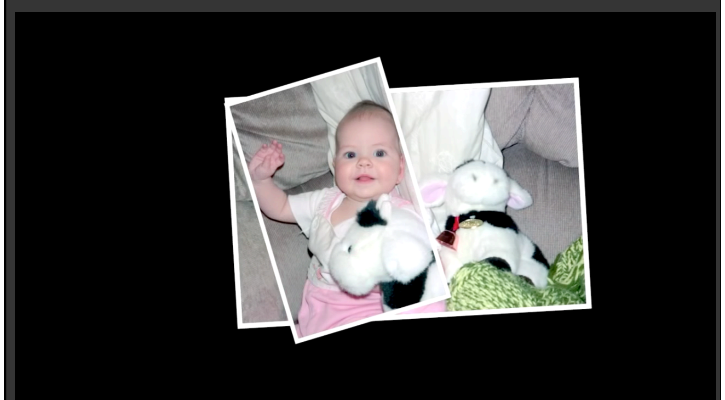
Facial Expression Transfer

Original image | Expression transfer by warping | Proposed method



Z. Liu, Y. Shan, Z. Zhang,
Expressive Expression Mapping with Ratio Images, SIGGRAPH 2001

Picasa Face Movies



Anonymizing Photographs or Video

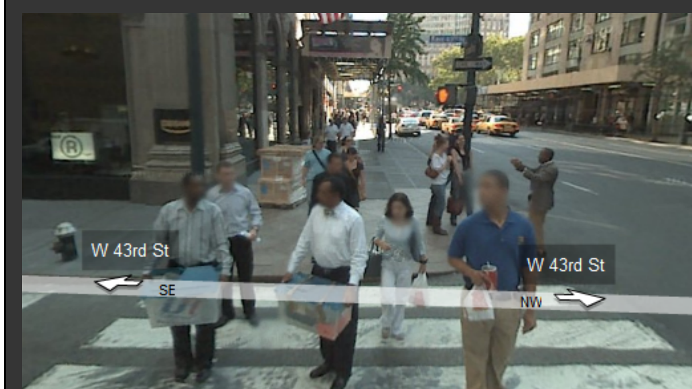


Image Search: SkyFinder

Attribute-based search based on learned sky attributes such as category, layout, richness, horizon. Example query: "Whole blue sky with white clouds"