CS 3570
Final Project

5/24/2012

General Guideline

- 1-3 students form a group
- Select a topic for the final project.
- The final project consists of three stages
  - Project proposal
  - Project presentation in class
  - Project demo/report

Suggested Topics

- Image inpainting
- Image retargeting
- Image editing
- Non-local means image denoising
- Bilaterial image filtering
- Convolution reverb
- Speaker recognition
- Face recognition
- Audio retrieval
- Image retrieval
- background subtraction and substibution
- Video stabilization
- Surveillance video compression
- Video summarization

Image inpainting

- ***Inpainting*** is the process of reconstructing lost or deteriorated parts of images
- Ref: **http://ppt.cc/6hpy Image Inpainting**

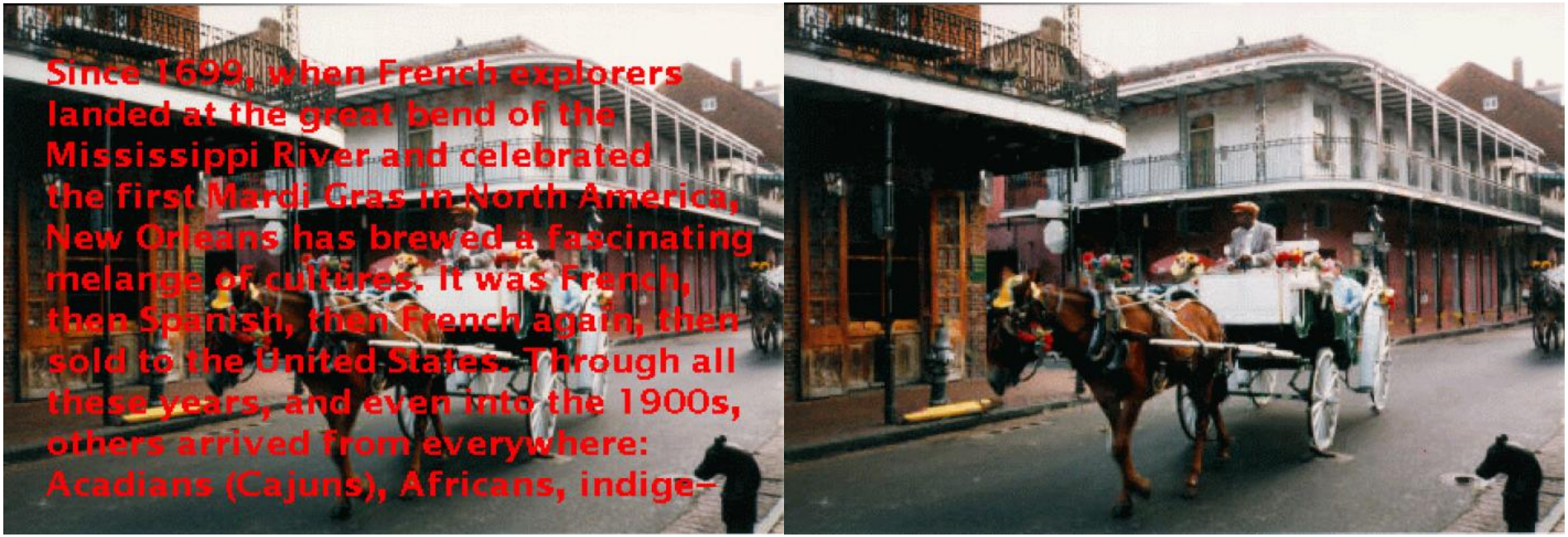


Image retargeting

- Resize images to different aspect ratios such that the actual photo's contents can be preserved as best as possible.
- The seam carving method is based on finding shortest paths.
- Seam Carving for Content-Aware Image Resizing, S. Avidan and A. Shamir, ACM SIGGRAPH, 2007 http://ppt.cc/SOsM

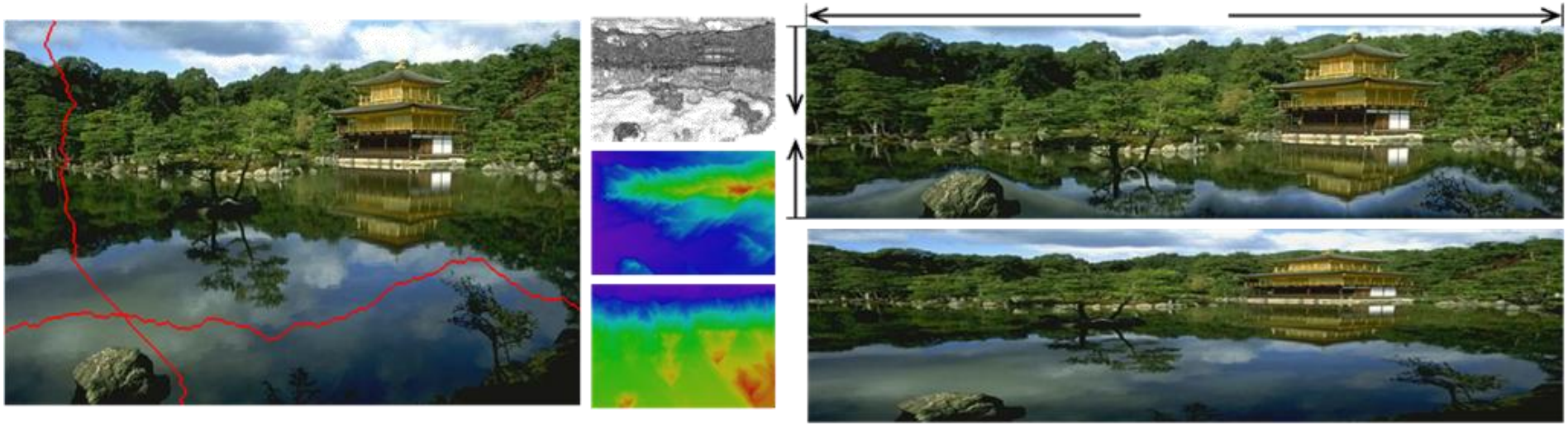
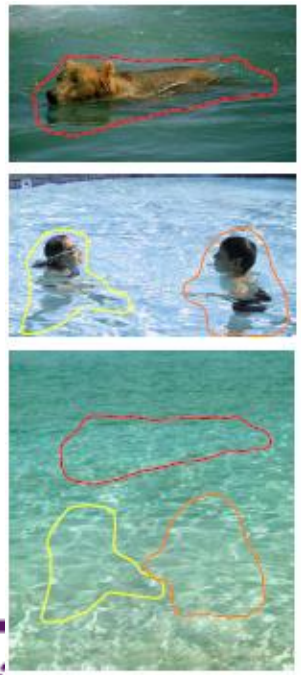


Image editing

- *Image editing* encompasses the processes of altering images and composing objects from different images
- Poisson Image Editing, P Pérez et al, ACM SIGGRAPH 2003 http://cs.engr.uky.edu/~jacobs/classes/2010_photo/readings/PoissonImageEditing.pdf



sources/destinations

cloning

seamless cloning

Non-local Means Image Denoising

- The goal of image denoising methods is to recover the original image from a noisy image for advanced image analysis
- A. Buades, B. Coll and J.M. Morel, “A non-local algorithm for image denoising”,IEEE Int. Conf. on Computer Vision and Pattern Recognition, 2005.

Bilaterial Image Filtering

$$g(\mathbf{x}) = \frac{1}{k(\mathbf{x})} \sum_{\xi} h(\mathbf{x},\xi)\, w(f(\xi) - f(\mathbf{x}))\, f(\xi)$$

From B. Weiss, Fast Median and Bilateral Filtering, SIGGRAPH'2006

C. Tomasi and R. Manduchi, "Bilateral filter for gray and color images", ICCV, 1998

Echo Cancellation

- In audio signal processing, echo cancellation denotes the process of removing echo from a voice communication in order to improve voice quality on a telephone call.
- Two sources of echo have primary relevance in telephony: **acoustic echo** and **hybrid echo**.
- Echo cancellation involves first recognizing the originally transmitted signal that re-appears, with some delay, in the transmitted or received signal. Once the echo is recognized, it can be removed by 'subtracting' it from the transmitted or received signal.
- http://dea.brunel.ac.uk/cmsp/Home_Saeed_Vaseghi/Chapter16-EchoCancellation.pdf

6. Speaker Recognition

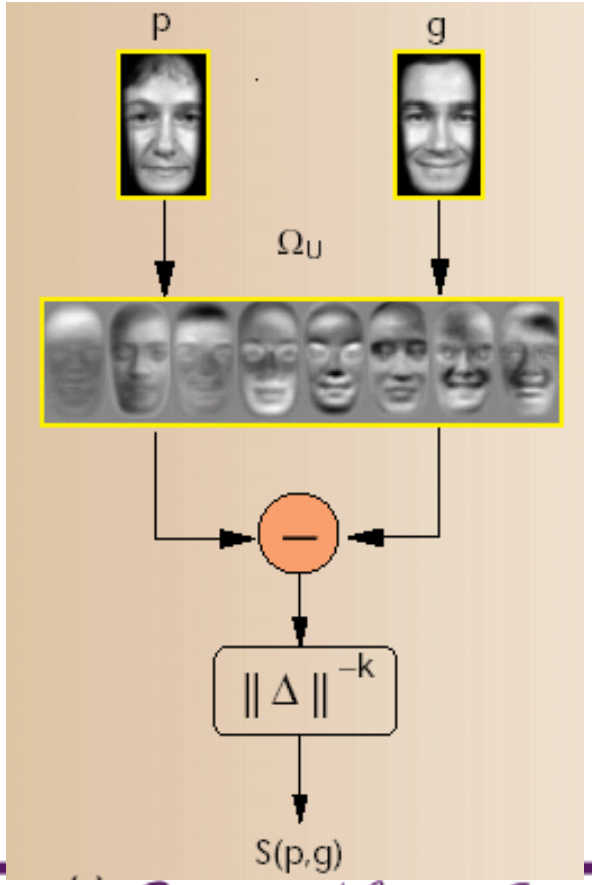
- Given a speech audio, we can recognize who makes this speech by the MFCC features extracted from the speech signal..
- This technique involves machine learning. We first determine the features of the signals, and use the training data to train Gaussian Mixture Models for speaker classification.

- GMM Based Speaker Recognition on Readily Available Databases, B R. Wildermoth & K. K. Paliwal, Nov. 2003.
- https://maxwell.ict.griffith.edu.au/spl/publications/papers/merc03_brett.pdf
- Speaker Verification Using Adapted Gaussian Mixture Models, D. A. Reynolds, Digital Signal Processing 10, 19–41 (2000).
- http://www.ll.mit.edu/mission/communications/ist/publications/000101_Reynolds.pdf
- Campbell J. P., ”Speaker recognition: a tutorial,”*Proc. IEEE, vol. 85, no.*9, pp. 1437-1462, 1997.

Face Recognition

- Recognition human faces from a database of face images.
- Eigenface (PCA) based recognition is a traditional method.

1. The system collects a database of face images.
2. It generates a set of *eigenfaces* by performing principal component analysis (PCA) on the face images. Approximately 100 eigenvectors are enough to code a large database of faces.
3. The system then represents each face image as a linear combination of the eigenfaces.
4. Given a test image, the system approximates it as a combination of eigenfaces. A distance measure indicates the similarity between two images.



Audio Retrieval

- An **audio retrieval** system is a computer system for searching and retrieving audios of similar contents from a large database.
- Compute representative features for audio (music) signals.
- Audio features: MFCC, FFT-based, LPC, etc.
- Music and Audio Retrieval Tools (MaART) is a set of software components used to investigate and implement retrieval and searching of music and audio.
- References:
  - *Audio Retrieval* by Rhythmic Similarity, by J. Foote et al., International Conference on Music Information Retrieval, 2002.
  - *www.fxpal.com/publications/FXPAL-PR-02-172.pdf*
  - MaART: http://maart.sourceforge.net/

Image Retrieval

- An **image retrieval** system is a computer system for searching and retrieving images of similar contents from a large image database.
- Compute representative features for images, such as color distribution, texture features, bag-of-words features, etc.
- User relevance feedback.
- A suitable database for image retrieval should be found.
- References:
  - http://savvash.blogspot.com/2009/10/image-retrieval-systems.html
  - Content-Based *Image Retrieval Systems*: A Survey
  - http://www.mathworks.com/matlabcentral/fileexchange/22030-image-retrieval-query-by-example-demo

Background Subtraction / Substitution

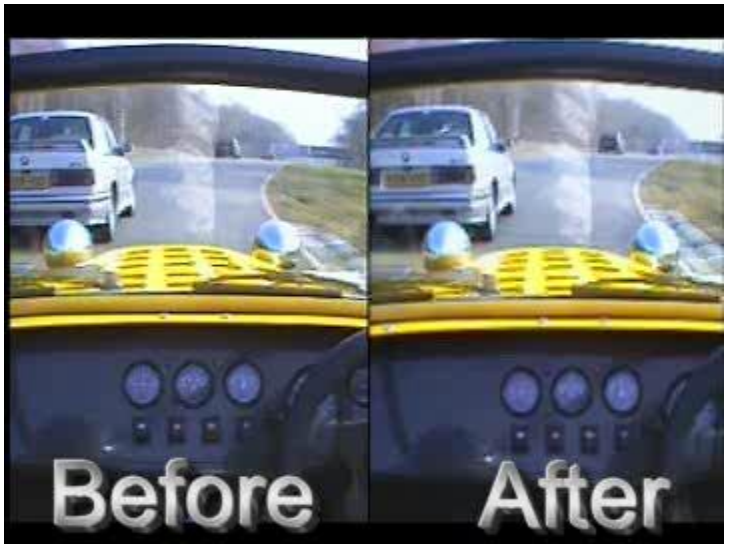
- Moving object detection in video sequences is one of the main tasks in many computer vision applications.
- Background subtraction is an common approach for this task. The idea is to compare the current image against the GMM background model
- C. Staufer, W.E.L. Grimson, "Adaptive background mixture models for real-time tracking," *CVPR*, Vol. 2, pp. 246-252, June 1999.

Background subtraction

synthesis

Video Stabilization

- Videos retrieved from video devices is affected by unwanted camera shakes and jitters, resulting in video quality loss.
- Video stabilization techniques are important to obtain high quality and stable video footages even in non-optimal conditions



Y. Matsushita, E. Ofek, X. Tang and H. Shum "Full-frame video stabilization", IEEE Int. Conf. on Computer Vision and Pattern Recognition, 2005

Surveillance Video Compression

- Video surveillance has been widely used in recent years to enhance public safety and privacy protection
- Surveillance video usually has constant background
- State-of-the-art video compression methods such as H.264/AVC often lead to high computational complexity
- The algorithm taught in class can be used to compress surveillance video and compare with other compression algorithms.

Video Summarization

- Video summarization methods attempt to abstract the main occurrences, scenes, or objects in a clip in order to provide an easily interpreted synopsis.
- Extract key frames from a video for summarization
- Shot boundary detection, i.e. temporal video segmentation
- Keyframe-based video summarization using Delaunay clustering, P. Mundur et al., International Journal on Digital Libraries, 2006.
- Temporal video segmentation: A survey, I. Koprinska and S. Carrato, Signal Processing: *Image Communication 16 (2001).*

Schedule of Final Project

- 5/31: Project proposal due in class (project title, project goal, team members, references)
- 6/11,14: Project presentation in class
  - 4/6/8 minutes for teams of 1/2/3 students
- 6/18(3pm-5pm), 6/21(2pm-4pm), 6/25 (2pm-6pm): Project presentation/demo and final project report due (Location to be announced)