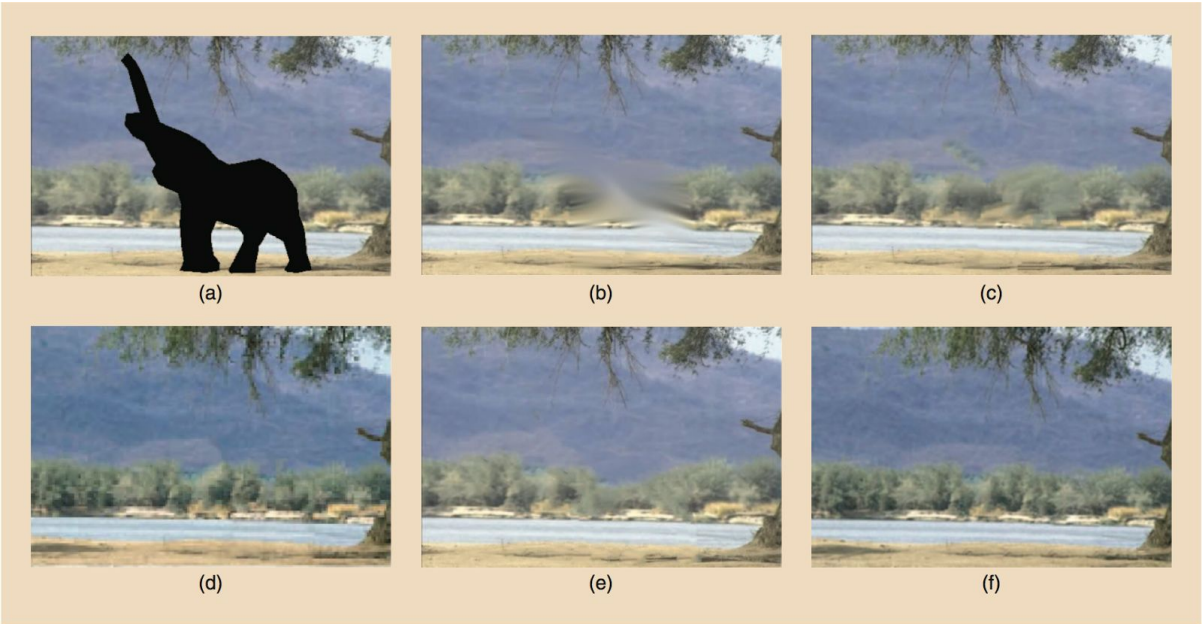
CS 532 Project Proposal

Tentative Title: Dictionary Learning on Image/Audio Inpainting
Team Members: Xiaomin Zhang, Jinman Zhao, Jinyu Xia

Brief Overview of Topic and Motivation:

1. Inpainting is an interesting task. An example is presented as following:



[FIG11] **Object removal application (a) mask and inpainting results with methods from different categories, (b) anisotropic diffusion (courtesy of [11]), (c) examplar-based with LLE (courtesy of [66]), (d) patch sparse representation (courtesy of [32]), (e) hybrid with one global energy minimization (courtesy of [21]), and (f) patch offsets (courtesy of [39]). [(a) courtesy of www.magazinehive.com.]**

Magic!!!

(Fig. 11 from "Image inpainting: Overview and recent advances")

2. Image inpainting is a widely investigated task and it would be interesting to see how it is related to pattern recognition techniques mentioned in the course and compare our results with the state-of-arts. Hopefully we can extend the method to audio inpainting which is generally under-investigated. Or even use learned dictionry as an indicator for phonemes in spoken language data.
3. Dictionary learning along with sparse coding is an efficient method for processing large data set. "Online Dictionary Learning for Sparse Coding" shows that it can help restore text embedding images.

   The method is related to the course as it is equivalent to minimize

   $$e_D(b) \equiv \min_w ||Dw - b||_2^2 + \lambda ||w||_0^2 .$$

   Matrix D is a learned dictionary with a much smaller size campared to the dataset, b is the element in the data set that to be reconstructed by the dictionary, w is a sparse

coding in the sense that it prefers few non-zero entries. (1-norm is Lasso, 2-norm is Tikhonov, here $||\cdot||_0$ means the number of non-zero elements).
A good dictionary is the one that minimize the total error for all data with its limited size.

$$\min_D \sum_{b \in B} e_D(b)$$

4. Similar with scale-invariant which is used as the basis of extracting features, we may involve with shift-invariant dictionary by taking all the possible shifts of some patterns. And we could use extending SVD to do this. This really arouses our interest and we will apply SVD method to study it.

Have your initial investigations led to any interesting questions or challenges?

1. Figuring out how to implement the online dictionary is a challenging topic. We need to amend the alogrithm as we test the learning result. This is owing to unsupervised learning.
2. The problem with audio inpainting may lies in the processing with sound signals. Examples are extracted and recover signals between sampled digit sound files and feature vectors that are suitable for learning.
3. We may involve with shift-invariant dictionary by taking all the possible shifts of certain patterns into consideration.

What sort of machine learning or data analysis tasks do you think will arise in your project?

1. We will make it an automatic process to find a good dictionary and coding for a given dataset so that thay can well reconstruct all orignal data. This is an unsupervised dictionary learning task.
2. We will create images or audios by adding noise on some “good” pictures or sound. The contrived examples could facilitate us to check the performance of our models, by comparing with the original “good” data. This is a task of (image/audio) data restoration or signal denoising.
3. Specifically, we will add some text on images and then try to remove them. These are opposite operations of text embedding and image inpainting.
4. We will also find some images or audios online with “noise”. For example, an image may contain watermark or some stains.

Core Concepts:

What concepts and tools may be involved in your project/lab?

Concepts:

1. **Dictionary Learning**
2. **Online Optimization**
3. **Sparse coding**
4. **Image inpainting**
5. **Audio inpainting**
6. **Image processing**
7. **Signal processing**

Involved softwares:

1. **Matlab**
2. **TBD**

Techniques used that have cover in the course.

1. **Least Squares**
2. **Lasso and Ridge regression**
3. **Extending SVD**

Related Papers, Datasets, or Resources:

List any that you have found so far (titles, urls, etc).

1. **Dictionary Learning and Sparse Coding**
   Online Dictionary Learning for Sparse Coding
   http://www.di.ens.fr/willow/pdfs/icml09.pdf
   SHIFT-INVARIANT DICTIONARY LEARNING FOR SPARSE REPRESENTATIONS: EXTENDING K-SVD
   http://www.irisa.fr/metiss/gribonval/Conf/2008/eusipco_shiftinvariantksvd_final.pdf
   K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation
   http://www.cs.technion.ac.il/~elad/publications/journals/2004/32_KSVD_IEEE_TSP.pdf

2. **On image restoration and inpainting**
   Sparse learned representations for image restoration
   http://www.di.ens.fr/willow/pdfs/IASC08.pdf
   Image inpainting: Overview and recent advances
   http://www.researchgate.net/profile/Christine_Guillemot/publication/260722824_Image_Inpainting__Overview_and_Recent_Advances/links/54d08c590cf298d6566707c1.pdf

3. **On audio restoration and inpainting**
   Dictionary Learning for Audio Inpainting
   http://dumas.ccsd.cnrs.fr/dumas-00725263/file/Guichaoua.pdf

Find some real-world data that are a good fit for your project (not necessary at this stage, but hopefully you will have a good dataset for your final project).

1. Image inpainting

   The TUM-Image Inpainting Database
   http://www.mmk.ei.tum.de/tumiid/

2. Image

   http://www.multitel.be/cantata/
   http://www.cis.upenn.edu/~jshi/ped_html/

3. Audio

   (some speech data)