Multimedia 664/864 Project List (subject to additions and edits)

Note: All peer-reviewed papers should be accessible through the SFSU library (if unavailable directly). Teams are welcome to add to the advanced learning goals to maximize your project score.

1. ***Project Title:*** *Determining Motion Fields Using Optic Flow*

Description and basic learning goals: Optic flow is a technique to determine and characterize the apparent motion of objects in a visual scene. In this project, you are expected to learn about the basic optic flow algorithm and code & apply it to quantify the motion in videos that will be given to you.

Advanced learning goals: The basic optic flow technique is limited by computational complexity. Research, implement and test a more efficient method.

Data: Zebrafish greyscale and binary data will be posted on iLearn. Contact TA for questions on data.

References:

1. C. Marcellino, J. Gut, K. C. Lim, R. Singh, J. McKerrow, J. Sakanari, "WormAssay: A Novel Computer Application for Whole-Plate Screening of Macroscopic Parasites", *PLoS Neglected Tropical Diseases,* Vol. 6(1):e1494, 2012

2. B. Kitt, B. Ranft, and H. Lategahn, "Block-matching based optical flow estimation with reduced search space based on geometric constraints," *13th International IEEE Conference on Intelligent Transportation Systems*, 2010.

2. ***Project Title:*** *Visualization of large numbers of 3D points using LPCA embedding*

Description and basic learning goals: Most data sets today are large, complex, and/or high-dimensional. Data visualization and exploration techniques are essential for sensemaking. In this project, you are expected to learn about and develop visualization and visual data exploration techniques for large data sets using the local PCA method [2]. This method "locally" groups the data before applying linear dimensionality reduction (PCA). At the minimum, your implementation must support operations such as "orbit" (rotation of a camera about a point), "pan" (lateral movement of a camera), zoom, and display of associated data about each point when clicked (*e.g.* each point may have an associated name or could be an image). Implementations should support visualization/interactions with thousands of points. Implementations that can run in a browser are preferred but not required. You may find the three.js library useful.

Advanced learning goals: High dimensional data (*i.e.* data that has more than 3 attributes) is impossible to visualize directly. In such cases the data has to be embedded in a lower dimensional space (2D or 3D) to support visualization and interactions. Two common linear embedding techniques that are used include Principal component analysis (PCA) and Multi-dimensional scaling (MDS). How would the results vary if you apply the MDS method instead of PCA? Justify your findings. How would you design a visualization that can account for time-dependencies in the data (assume you are given time coordinates for the data).

Data: The basic data set(s) consisting of points from complex data distributions will be posted on

iLearn. This data will describe COVID-19 variants from two different countries. Contact TA for questions on data.

References:
1. threejs.org

2. N. Kambhatla, T. K. Leen, Dimension Reduction by Local Principal Component Analysis, *Neural computation*, 1997-07-01, Vol.9 (7), p.1493-1516

3. R. Singh and R. Jain, "From Information-Centric to Experiential Environments", in *Interactive Computation: The New Paradigm, D. Goldin, S. Smolka, and P. Wegner, eds.,* Springer Verlag, pp. 323 – 351, 2006 (Long read. Initially focus on the properties of experiential interfaces and implement them in your prototype)

3. ***Project Title*****:** *Comparative visualization and analysis of large numbers of high dimensional points using linear and non-linear embedding*

Description and basic learning goals: See project 2. The distinction of this project from project 2 is the following: (1) You should use 3-4 linear and non-linear embedding techniques to analyze the data and compare which embedding technique is better. You can employ library implementations of these techniques. (2) You are expected to implement interactive visualizations as described in project 2. (3) Significant emphasis should be placed on detailed comparisons and analyses of the results obtained using the different dimensionality reduction methods.

Advanced learning goals (required for 864 students): The results from the comparisons should be analyzed by taking into account the theory underlying PCA, MDS, and any one non-linear technique of your choice. How would you design a visualization that can account for time-dependencies in the data (assume you are given time coordinates for the data).

Data: Data will be posted on iLearn. For questions about the data, contact TA

Reference:
1. C. J. C. Burges, Dimension Reduction: A Guided Tour, *MSR Tech Report MSR-TR-2009-2013*

2. Look into the following non-linear dimensionality reduction techniques: Locally linear Embedding, Isomap, Laplacian Eigenmaps, t-SNE (t-distributed stochastic neighbor embedding)

4. ***Project Title:*** *Video Tracking Using Distance Transforms*

Description and basic learning goals: Tracking objects in a video constitutes an essential task in Multimedia and Computer Vision. This project will track an animal called zebrafish in video using a technique called distance transforms. The largest part of a fish is usually its head. Solutions would take the distance transform of the masks and identify points where the gradient is zero. These points likely correspond to the center point of fish heads. Identified head points should be filtered based on the value of the distance transform at that point.

Advanced learning goals: To distinguish points corresponding to true head centers from those that are false, a threshold distance transform value would need to be determined. Research, implement, and compare at least two automatic thresholding methods for this purpose.

Data: Zebrafish greyscale and binary data will be posted on iLearn. Contact TA for questions on data.

References:
1. H. Breu, J. Gil, D. Kirkpatrick, and M. Werman, "Linear time Euclidean distance transform algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 5, pp. 529–533, May 1995.

5. ***Project Title:*** *Image analysis for biometric feature identification.*

Description and basic learning goals: The data consists of still frames of zebrafish larvae in a 2D image. Your solution should identify points in a frame that correspond the center point of a fish's head by analyzing the intensity characteristics of the larvae. Your solution should not depend on threshold values that are manually determined. Investigate the accuracy of your determinations.

Advanced learning goals: Investigate if the problem can be solved by taking into account the geometry of the larvae and combining this information with the results from intensity-based analysis.

Data: Zebrafish greyscale and binary data will be posted on iLearn. Contact TA for questions on data.
References: see lectures on image analysis

6. ***Project Title:*** *Tracking objects in a video by blob matching*

Description and basic learning goals: Object tracking can be thought of as establishing correspondences between segmented (binary) regions across frames. Due to a number of factors, a semantic object may not correspond to a single binary region. In this project you will study and implement a tracking method based on region (blob) matching. In particular, your should implement and test the method listed in the reference (sections I-II) to the data set provided to you and assess it experimentally.

Advanced learning goals: Following section III of the paper, build a zebrafish movement model.

Data: Zebrafish greyscale and binary data will be posted on iLearn. Contact TA for questions on data.

References: O. Masoud and N.P. Papanikolopoulos, A Novel Method for Tracking and Counting Pedestrians in Real-Time Using a Single Camera, IEEE transactions on vehicular technology, 2001-09, Vol.50 (5), p.1267-1278

7. ***Project Title:*** *Foreground-background separation in video*

Description and basic learning goals: Identifying the foreground and background in the video of a scene is crucial for any analysis. In this project you shall study, implement and experimentally evaluate a codebook-based technique that considers changes in values at each

pixel to identify foreground and background instances.

Advanced learning goals: Identify the critical parameters in the algorithm; one such example is the parameter λ which defines the duration during which a codeword has not recurred in the learning phase of the method. Investigate how this/these parameter(s) could be determined automatically.

Data: Videos of adult *S. mansoni* parasites will be posted on iLearn. Contact the TA for questions about the data.

References: K. Kim, T. H Chalidabhongse, D. Harwood, and L. Davis, "Real-time foreground–background segmentation using codebook model", *Real-time imaging*, 11(3) 2005

8. ***Project Title:*** *Unknotting objects in images/video*

Description and basic learning goals: Many articulated objects can self-overlap or self-cross during movement. This leads the images of their shape to form a knot. Examples include cursive letters, pedestrians, snakes etc. The goal of this project is to take images/videos of specific objects (parasites) that self-cross and unknot them – that is, you will take feature points on the object and linearize their ordering based on local feature characteristics potentially occurring across the video. Initially you may compute a shape skeleton and apply the method from Reference 1 to see if it helps to linearize the object shape.

Data: Videos of adult *S. mansoni* parasites that overlap will be made available on iLearn. See TA for questions regarding the data.

References

1. P. M. Magwene, P. Lizardi, and J. Kim, "Reconstructing the temporal ordering of biological samples using microarray data," *Bioinformatics*, vol. 19, no. 7, pp. 842–850, 2003

2. R. Eshleman and R. Singh, "Reconstructing the Temporal Progression of Biological Data Using Cluster Spanning Trees", *IEEE Transactions on NanoBioscience*, Vol. 16, No. 2, pp. 140-147, 2017

9. ***Project Title:*** *Image retrieval by image appearance using histogram intersection*

Description and basic learning goals: Given a set of images with object masks (i.e. the objects have been pre-segmented for you), extract the objects and create a feature set that describes them in terms of their geometry, grey-scale, texture, and color (if relevant) characteristics. Store these images and their features in a simple database (desirable) or file system and build an interface that allows users to use any of these images as a query to retrieve similar images from the database. The retrieval should use the histogram intersection method from the reference. Note that the project requires you to extend the method to geometric features. Analyze the accuracy of your implementation using precision, recall, and F1 scores.

Advanced learning goals (required for 864): How can you ensure that all the relevant images have been retrieved without an exhaustive comparison? Such a goal can be achieved by indexing. Research and implement a relevant indexing algorithm and analyze its impact on your system.

Data: A collection of parasite images. Contact the TA for questions about the data.

Reference:
1. Swain, M. J., & Ballard, D. H. (1991). Color indexing. *International journal of computer vision*, 7(1), 11-32.
2. R. Mehta, and K. Egiazarian, Dominant Rotated Local Binary Patterns (DRLBP) for Texture Classification. *Pattern Recognition Letters*, Vol. 71 (1), pp. 16-22, 2016

10. ***Project Title:*** *Image retrieval by contour matching*

Description and basic learning goals: Given a set of images with object masks (i.e. the objects have been pre-segmented for you), extract the objects and characterize the geometry of their shape contours. Build an interface that allows users to use any of these images as a query to retrieve similar images from the set. The retrieval should implement the shape context method from the reference. Analyze the performance of your implementation.

Advanced learning goals: Can you ensure that all the relevant images have been retrieved without an exhaustive comparison? Such a goal can be achieved by indexing. Research and implement any relevant indexing algorithm and analyze its impact on your system.

Data: A collection of parasite images. Contact the TA for questions about the data.

Reference: S. Belongie, J. Malik and J. Puzicha, "Shape matching and object recognition using shape contexts," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 4, pp. 509-522, 2002

11. ***Project Title:*** *Automated tuberculosis interpretation using CNNs*

Description and learning goals: Tuberculosis (TB) is an infectious disease usually caused by Mycobacterium tuberculosis (MTB) bacteria. Most infections show no symptoms and it is also time consuming to manually interpret each image to decide whether the patient is affected or not. Therefore, automation in interpretation and detection of this disease is important. In this project, you will apply a deep learning technique (convolution neural networks) to detect and classify patterns in such images

Data: https://www.kaggle.com/saife245/tuberculosis-image-datasets

References
1. K. P. Smith, A. D. Kang, and J. E. Kirby, Automated Interpretation of Blood Culture Gram Stains by Use of a Deep Convolutional Neural Network, *Journal of Clinical Microbiology* , 2018, 56 (3) e01521-17.

2. C. Szegedy, W. Liu, J. Yangqing, P. Sermanet, *et al*. Going deeper with convolutions. *The IEEE Conference on Computer Vision and Pattern Recognition* (CVPR), pp. 1-9, 2015.

3. R. Girshick, J. Donahue, T. Darrell and J. Malik, "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation," *IEEE Conference on Computer Vision and*

Pattern Recognition (CVPR), pp. 580-587, 2014

12. ***Project Title:*** *Automated tuberculosis interpretation using tensor voting*

Description and learning goals: Same as above. In this project, you will apply a pattern recognition technique (tensor voting, from reference 2) that allows you to automatically detect and classify patterns in such images. Evaluate the performance of the method on the data set.

Data: https://www.kaggle.com/saife245/tuberculosis-image-datasets

References

1. Tensor voting tutorial slides by P. Mordohai
https://mordohai.github.io/public/TensorVotingTutorial_2007.pdf

2. L. Loss, G. Bebis, and B. Parvin, Iterative Tensor Voting for Perceptual Grouping of Ill-Defined Curvilinear Structures, *IEEE Transaction on Medical Imaging*, Vol. 30 (8), pp. 1503-1513, 2011

13. ***Project Title:*** *Automatic event-based grouping and interaction with personal multimedia information*.

Description and learning goals: As covered in the lectures, "events" are information/media generating occurrences. Consequently, one can use the notion of an "event" to organize potentially different types of media. In this project, using characteristics of an "event" such as location, time, participants, you will design techniques to organize large personal photo collections (> 1000 items such as photos, videos, documents etc.). You will design a basic UI to interface with the data based on using the notion of an "event" as the primary organization principle.

Advanced learning goals: Using the references, design and test a high-quality event-based personal information management system.

Data: Collect on your own

References:

1. R. Singh and J. C. Pinzon, "Study and Analysis of User Behavior and Usage Patterns in a Unified Personal Multimedia Information Environment", *Proceeding of the IEEE International Conference on Multimedia and Expo (ICME),* pp.1031 – 1034, 2007
2. J. C. Pinzon, R. Singh, W. Taube, J. Galan, "Designing Interactions In Event-Based Unified Management Of Personal Multimedia Information", *Proceeding of the IEEE International Conference on Multimedia and Expo (ICME)*, pp. 337 – 340, 2006
3. R. Singh, R. L. Knickmeyer, P. Gupta, and R. Jain "Designing Experiential Environments for Management of Personal Multimedia", *Proceeding of the ACM Multimedia Conference pp. 496 - 499, 2004*
4. M. Zhao, Y. W. Teo, S. Liu, T-S. Chua, and R. Jain, Automatic Person Annotation of Family Photo Album, *International Conference on Image and Video Retrieval* (CIVR) 2006, pp. 163-172

14. ***Project Title:*** *Capturing and analyzing variability in high-level image semantics by combining image annotations with linguistic analysis* **(664 only)**

Description and learning goals: An image can be interpreted differently by different people (or even by the same person). Dealing with such variability in semantics associated with an image represents the principal challenge to image-based querying. In this project, you will learn how to organize a collection of images based on varying annotations and use linguistic analysis to identify related images. You will begin by collecting a large image collection and having a group of your friends annotate some (or all images) in the collection as they see fit. You will then apply basic text analysis to identify key terms in the annotations for each image and create a bipartite graph where vertices at one level represent a collection of key words (or phrases) while vertices at the next level represent images such that each image is connected to the key words (or phrases) associated with it. You will then quantitatively analyze the connectivity patterns of the resultant bipartite graph to characterize the variability in semantics.

Advanced learning goals: Support querying of the resultant data structure by using linguistic relationships (see reference)

Data: Collect on your own

Reference: T. Lam and R. Singh, "Semantically Relevant Image Retrieval by Combining Image and Linguistic Analysis", *Proceeding of the International Symposium on Visual Computing (ISVC), Lecture Notes in Computer Science Vol. 4292,* pp. 1686 - 1695, Springer Verlag, 2006

15. ***Project Title:*** *Web search with linguistic expansions from term-frequency histograms*

Description and learning goals: Query definition is crucial to the success of web search. In this project, you will explore how starting from a given query, alternative query formulations can be devised that can help improve information retrieval on the web. The basic idea of this project will involve the use of WordBars (see reference). WordBars is a simple notion where the results of a web query by the most frequently appearing terms within the titles and snippets of the top search results. The potential query refinement terms can then be selected from the top search results returned by the underlying Web search engine. In this project you will implement the WordBars methodology and identify the high frequency terms. You will then use WordNet to explore how well different senses of these terms can help refine the query formulation.

Advanced learning goals: (1) create a basic UI to support the idea, and (2) conduct detailed analysis to see how effective the methodology is as compared to basic web search and WordBars (see Ref. 2 for study design ideas)

Data: Collect on your own

References:

1. O. Hoeber and X. D. Yang, Evaluating the effectiveness of term frequency histograms for supporting interactive Web search tasks, In Proceedings of the ACM Conference on Designing Interactive Systems, pp. 360-368, 2008

2. R. Singh, Ya-Wen Hsu, and N. Moon, "Multiple-Perspective Interactive Search: A Paradigm for Exploratory Search and Information Retrieval on the Web", *Journal of Multimedia Tools and*

Applications, Vol. 62, pp. 507-543, 2013

16. ***Project Title:*** *Analyze Twitter data to automatically categorize and subsequently analyze health attitudes, beliefs, and behaviors related to non-medical opioid use.*

Description and learning goals: Social media is a real-time source of opioid-use information that is not subject to disclosure concerns. In this research, you will extract tweets (> 2000) related to opioid use (through keyword based search based on the methodology from reference [1]). You will next read about a software called DUI (Drug Use Insights) developed in my lab to analyze these tweets [2]. Implement a parser using the substance-use term list identified in the DUI paper (supplemental materials) and analyze the data you have collected based on the DUI categories. Finally, following the methodology in Ref [1] you will compare the results of your analysis with the one in Ref [1] and analyze/explain any differences.

Data: Collect from Twitter following the methodology from Reference [1].

References:

1. Chan B, Lopez A, Sarkar U (2015) The Canary in the Coal Mine Tweets: Social Media Reveals Public Perceptions of Non-Medical Use of Opioids. PLoS ONE 10(8): e013507

2. Z. Prince, D. Jha, and R. Singh, "DUI: the drug use insights web server", *Bioinformatics*, Vol. 37 (24), pp. 4895-4897, 2021

17. ***Project Title:*** *Assigning scores to movie reviews by sentiment analysis*

Description and learning goals: Given an annotated dataset of imdb movie reviews, your task is to score the sentiment associated with a review by utilizing its text. That is, you want to map the review to a score (*e.g.* 1 – 10). Analyze the method in terms of accuracy and consistency.

Advanced learning goals: design experiments that demonstrate your approach to be both consistent and explainable

Data: Collect yourself. Look at: https://ai.stanford.edu/~amaas/data/sentiment/

Reference:

1. Martineau, J.C. and Finin, T., 2009, March. Delta tfidf: An improved feature space for sentiment analysis. 3rd International AAAI conference on weblogs and social media.

2. Tripathy, A., Agrawal, A. and Rath, S.K., 2016. Classification of sentiment reviews using n-gram machine learning approach. Expert Systems with Applications, 57, pp.117-126.

18. Propose a project/problem you find interesting. Must get approval from the professor. Email the professor a description of the project that has been written up in a manner similar to the above projects and be prepared to meet and discuss.