Recent years have witnessed an explosive increase in electronic multimedia data. Facebook is currently managing more than 200 billion of images. Over 6 billion videos are watched per month on YouTube. The number of smartphone around the world has topped 1 billion in 2012 and that number is projected to double by 2015. Such big multimedia data has presents a number of compelling computational challenges for all aspects of multimedia computing.

The goal of this course is to introduce new multimedia algorithms, novel multimedia systems, as well as intriguing multimedia applications, with a focus on the acquisition, generation, transmission, storage, processing, and retrieval of large-scale, heterogeneous, and semantic-rich multimedia information. We will also discuss future challenges in multimedia information processing and related software systems.

List of Course Topics (Tentative):

- New multimedia signal processing algorithms using ubiquitous computing devices, such as smartphone, smart watch, Kinect for Xbox, body sensors, brain–computer interface, and etc.
- Novel multimedia communication and stream systems, such as 3D video streaming, HD video streaming, mobile audio/video streaming, multimedia sensor networks
- New paradigm of multimedia feature engineering, such as multimodal deep learning, deep learning-based 3D video understanding
- Large scale multimedia indexing algorithms, such as MapReduce-based petabyte image/video indexing algorithms, Locality-sensitive hashing for massive multimedia indexing
- Scalable multimedia retrieval and mining algorithms, such as ultra-fast approximate nearest neighbor algorithm for massive image/video retrieval, online semantics-preserving algorithms for enhancing Bag-of-words (BoW) model
- Distributed and parallel computing platform for massive multimedia process, such as new computing system based on MapReduce and Key/value pair expansion
- Distributed machine learning algorithms for multimedia semantic learning, such as learning to rank boosted decision tree for large scale multimodal learning, applications of Apache Mahout machine learning library or Berkeley MLBase to scalable multimedia learning
- GPU-based multimedia architectures and systems, such as GPU-based real-time video tracking system for high speed (5000 frames per second) cameras
- Intriguing applications of multimedia computing in healthcare, biology, social network, gaming, virtual and augmented reality, and etc.
Format:

The class will include the following three main activities

- Class lectures (by instructor)
- Homework Assignments (by students), include but not limited to paper reading, class presentations, paper writing, and in-class discussion
- Course projects (by students and being supervised by instructor)

One interesting feature of this class is called “project early intervention”. Specifically, in the beginning of the semester, the instructor will introduce potential projects (include background, objectives, challenges, on-going research efforts, possible research activities, expected outcomes, etc.) to the students. Based on students’ own interests, their own paper reading, and consulting with instructor, the students will pick an appealing project. Then, the instructor will assign more research papers to each student and the students will also select papers by themselves with the approval from the instructor. Based on the project picked by the students, the instructor may group them into different teams, each team for one research and/or development area. We hope after the first few weeks, each student will have a clear idea and vision on his/her course project.

Class Schedule: Wednesday; 5:30 PM to 8:20 PM; Room: Olsen Hall 401

Office Hour: Wednesday, 10:00 AM to 12:00 PM; 1:00 PM to 2:00 PM, Room: Olsen Hall 213

Prerequisites: Programming experiences, understanding of data structures, algorithms, database, networking, and computer system. Other desirable (but not mandatory) qualifications include prior exposure to: multimedia computing, data mining and machine learning, decision support systems, information retrieval, networking, sensor design and development, distributed and parallel computing or other research experiences. Please contact instructor (Dr. Yu Cao, ycao@cs.uml.edu) for permission to enroll.

Textbook: There is no required textbook in this course. Lecture notes and papers will be distributed in class and/or Blackboard. Reference books are listed as follows:

Fundamentals of Multimedia
by Ze-Nian Li and Mark S. Drew

Multimedia Analysis, Processing and Communications