

Schedule for CS4241: Semester 2, 20011/12

Lecturer: CHUA, Tat-Seng **Co-Lecturer:** ZHA, Zhengjun/Wang, Meng
Tutor: (to be added later) **Web Site:** <http://ivle.nus.edu.sg/>
Lecture: Thu, 9-11 (SR@I3) **Tutorial:** Thu 11-12 (SR@I3) (2%)
Midterm: 8 Mar (Thu) 11-12 (20%) **Final Exam:** 28 Apr (Sat), afternoon (30%)
Assignments: 3 Programming Assignments (14+16+18) = 48%

Wk	Date	Lecture/Tutorial Topics	Remarks
1.	12 Jan	L1: Intro: Why MMIR	
2.	19 Jan	L2: Most MMIR is Text-based -1	Outline of Assignments
3.	26 Jan	L3: Most MMIR is Text-based -2 T1: Text-based MMIR	
4.	2 Feb	L4: Color-based Retrieval Models	Briefing A1: Content-based MMIR
5.	9 Feb	L5: Other Visual Features: Texture & Critical Point-based Features T2: Color-based and other visual features	
6.	16 Feb	L6: Indexing of Media Contents T3: Critical point-based features	
	23 Feb	Recess Week	
7.	1 Mar	L7: Visual Concept Annotation -1 T4: Indexing Issues in MMIR	Briefing A2: Interactive Concept-based MMIR Submission of A1 (1 Mar) & Grading (2 Mar)
8.	8 Mar	L8 Visual Concept Annotation -2 Mid-Term Test (8 Mar, 11-12pm)	
9.	15 Mar	L9: Interactive MMIR T5: Visual Concept Annotation	Briefing A3: MM + UGCs
10.	22 Mar	L10: Ranking for MMIR	Submission of A2 (22 Mar) + Grading (23 Mar)
11.	29 Mar	L11: Location and Mobile MMIR & Applies T6: Ranking & Interactive MMIR	
12.	5 Apr	L12: Multimedia Question-Answering T7: Location and Mobile MMIR & Apps	
13.	12 Apr	L13: Integration of MM & Real World Data	Submission of A3 (12 Apr) +Grading (13 Apr)
-	19 Apr	Reading Week	
-	28 Apr	Final Exam (28 Apr, Sat, Afternoon)	

DETAILS AND REFERENCE MATERIALS

Aim: To design a multimedia information retrieval (MMIR) course targeting at the senior undergraduate level. It covers the key concepts, techniques and system design for MMIR.

Target Audience: This course is designed to be a second course in multimedia for senior undergraduate class. Students should have gone through a first course in multimedia that covers the basic image, video and audio compression technologies and standards, as well as basics of hypermedia, WWW and Web 2.0. Students are also expected to have the knowledge equivalent to a first course in Artificial Intelligence and possess good knowledge of machine learning.

MODULE OUTLINE

Week 1: Why is MMIR Important?

- Basics; History; MMIR and Society; Use of MMIR technologies; Overview of Retrieval Process; Key Components of Retrieval Systems; Impact on our Life
- Applications; Commercial Activities and Systems; Mobile Searches.
- Overview of Module; Objectives; Workloads and Assignments.

REFERENCE:

- [1] M. S. Lew, N. Sebe, C. Djeraba, and R Jain, "Content-Based Multimedia Information Retrieval: State-Of-The-Art and Challenges," ACM Transactions on Multimedia Computing, Communications, and Applications, 2006.
- [2] M. S. Kankanhalli and Y. Rui, "Application Potential of Multimedia Information Retrieval," Proceedings of IEEE, vol. 96, no. 4, 2008.
- [3] J. Magalhaes and S. Ruger, "Semantic Multimedia Information Analysis for Retrieval Applications," chapter XVI, Idea Group Inc., 2007.

Weeks 2-3: Most MMIR Systems are Text-based

- Text-based IR techniques; Sources of Text (Surrounding Text, Annotation and ASR etc); Text Content Representation; Similarity Measures; Relevance Feedback
- Text-based MMIR; Relevance feedback; Example Systems; Limitations and Effectiveness
- Applications to Web Media Search; Commercial Systems; State-of-the-Arts

REFERENCE:

- [1] CD Manning, P Raghavan, and H Schutze. "An Introduction to Information Retrieval", Cambridge University Press, England, 2009. Chapters 1, 6, and 9.
- [2] R Datta, D Joshi, J Li, and JZ Wang. Image Retrieval: Ideas, Influences, and Trends of the New Age, ACM Computing Survey, 2005
- [3] S-Y Neo and T-S Chua. News Video Retrieval using Multi-modal Query-dependent Model and Parallel Text Corpus. Proceedings of 'Multimedia Information Retrieval' Workshop, in SIGIR'05, Brazil, Aug 2005.
- [4] H. Xu and T-S Chua. Fusion of AV Features and external information sources for event detection in team sports video. ACM Transactions on Multimedia Computing, Communications and Applications (TOMCCAP), Volume 2 (1), Feb 2006, 44-67.
- [5] R Price, T-S Chua and S Al-Hawamdeh. Applying relevance feedback to a photo archival system. Journal of Information Science, Vol 18(3), 203-215, 1992.

Weeks 4: Media Contents are Important: Let's Start with Colors

- Importance of Inherent Contents; Basics of Multimedia; Types of Features
 - The Color Features: Different Color Models; Color Histogram; CCV; Color Correlegram
 - Basic Similarity Measures and Retrieval; Retrieval Examples
-

REFERENCE:

- [1] F Long, H Zhang, and DD Feng. "Fundamentals of Content-Based Image Retrieval," Chapter 1, Multimedia Information Retrieval and Management, Springer.
- [2] O Penatti and R Torres. "Color Descriptors for Web Image Retrieval: A Comparative Study," SIBGRAPI, 2008.
- [3] YB Mahdy, KM Shaaban and A El-Rahim. "Image Retrieval Based on Content," GVIP Journal, vol. 6, no. 1, 2006.

Week 5: Other Visual Features: Feature Extraction and Media Representation

- The Texture Features: Edge; Direction; Garbor Filter; Wavelet
- Critical Point-based Features: SIFT and SURF; Visual Keywords; Higher-level Representations

REFERENCE:

- [1] T Tuytelaars and K Mikolajczyk. "Local Invariant Feature Doctors: A Survey," Foundations and Trends in Computer Graphics and Vision, 2008.
- [2] J Sivic and A Zisserman. "Video Google: A Text Retrieval Approach to Object Matching in Videos", International Conference on Computer Vision, 2003.
- [3] Y-T Zheng, M Zhao, S-Y Neo, T-S Chua, Q Tian. "Visual Synset: towards a Higher-level Visual Representation", IEEE International Conference on Computer Vision and Pattern Recognition (CVPR) 2008, Anchorage, Alaska, U.S., June 24-26, 2008.

Week 6: Indexing of Media Contents

- Need to Scale up to Huge Media Collections
- Inverted File vs. Hash-based Indexing; Different Hash-based Indexing Approaches

REFERENCE:

- [1] J Sivic and A Zisserman. "Video Google: A Text Retrieval Approach to Object Matching in Videos", International Conference on Computer Vision, 2003.
- [2] A Torralba, R Fergus, and Y Weiss. "Small Codes and Large Image Databases for Recognition," International Conference on Computer Vision and Pattern Recognition, 2008.
- [3] A Gionis, P Indyk, and R Motwani. "Similarity Search in High Dimensions via Hashing," Proceedings of VLDB Conference, 1999.

Weeks 7-8: Visual Concept Annotation on Image and Video: Towards Concept-based IR

- Visual Concept Annotation: Introduction; Why; and Approaches
- Concept-based MMIR
- User Social Tagging; Tag Refinement; Leveraging on User Tags and High Quality Results of Commercial Systems
- Video Concept Detection and Search: Introduction to Scene Detection, Shot Detection, and Key-frame extraction; Introduction to Motion Feature; TRECVID Experience

REFERENCE:

- [1] TRECVID online Proceedings: <http://www-nlpir.nist.gov/projects/typubs/tv.pubs.org.html>
- [2] GM Snoek and M Worring. "Concept-Based Video Retrieval," Foundations and Trends in Information Retrieval, 2009.
- [3] M Wang and H Zhang. "Video Content Structuring," Scholarpedia, 2009. http://www.scholarpedia.org/article/Video_Content_Structuring

Week 9: Put Users in the Loop- Interactive MMIR

- Overview of Relevance Feedback Techniques; Active Learning Approaches
 - UI Design; Human Computation Model
 - Maximizing Human Efforts: Image-based vs. Cluster-based Feedback
 - TRECVID: Experience and Lessons Learnt
-

REFERENCE:

- [1] Y Rui, TS Huang and MMS Ortega. "Relevance Feedback: A Power Tool for Interactive Content-Based Image Retrieval," IEEE Transactions on Circuits and Systems for Video Technology, 1998.
- [2] M Wang, B Ni, X-S Hua and T-S Chua, "Assistive multimedia tagging: a survey of multimedia tagging with human-computer joint exploration," ACM Computing Surveys. (Available at <https://sites.google.com/site/mengwangsite/>)
- [3] H Luan, S-Y Neo, H-K Goh, S-X Lin and T-S Chua. Segregated Feedback with Performance-based Adaptive Sampling for Interactive News Video Retrieval. Proceedings of ACM Multimedia 2007, Short Paper. Augsburg Germany, September 2007. 293-296
- [4] Video Olympics. <http://www.videolympics.org/>

Week 10: Ranking for Multimedia IR

- Similarity Measures and Fusion Models
- Reranking for Text-Based Image Retrieval
- Ranking and Reranking for Multimodal Search

REFERENCE:

- [1] C Beecks, MS Uysal and T Seidl. "A Comparative Study of Similarity Measures for Content-Based Multimedia Retrieval," International Conference on Multimedia & Expo, 2010.
- [2] X Tian, L Yang, J Wang, Y Yang, X Wu and X-S Hua. "Bayesian Video Search Reranking," ACM Multimedia, 2008.
- [3] L Kennedy, S-F Chang and A Natsev. "Query-Adaptive Fusion for Multimodal Search," Proceedings of the IEEE, 2008.

Weeks 11: Location and Mobile MMIR & Applications

- Location Multimedia and Applications
- Mobile Multimedia and Applications

REFERENCE:

- [1] Location-based services: Foursquare <https://foursquare.com/>; Gowalla <http://gowalla.com/>; Facebook Places <http://www.facebook.com/about/location/>; and Google Places <http://www.google.com/places/>.
- [2] Y-T Zheng, Z-J Zha and T-S Chua. "Research and applications on geo-referenced multimedia: a survey," Multimedia Tools and Application, 2011.
- [3] Y-T Zheng, M Zhao, Y Song, H Adam, H Adam, U Buddemeier, A Bissacco, F Brucher, T-S Chua and H Neven. "Tour the World: building a web-scale landmark recognition engine". In Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), Miami, Florida, U.S., June 20-25, 2009.
- [4] X Xie, L Lu, M Jia, H Li, F Seide and W-Y Ma. "Mobile Search with Multimodal Queries," Proceedings of the IEEE, 2008.

Week 12: Multimedia Question-Answering

- Question analysis and answer media selection
- Media answer verification
- Multimedia QA

REFERENCE:

- [1] H Yang, L Chaisorn, Y Zhao, S-Y Neo and T-S Chua. VideoQA: Question Answering in News Video. ACM Multimedia'03, Berkeley, Nov 2003. 632-641.
- [2] G Li, H Li, Z Ming, R Hong, S Tang and T-S Chua. Question Answering over Community-Contributed Web Videos. IEEE MultiMedia 17(4): 46-57 (2010).
- [3] L Nie, M Wang, Z-J Zha, G Li and T-S Chua: Multimedia answering: enriching text QA with media information. SIGIR 2011: 695-704
- [4] G Li, M Wang, Z Lu and T-S Chua. In-Video Product Annotation with Web Information Mining. ACM Transactions on Multimedia Computing, Communications and Applications (TOMCAP), 2011

Week 13: Integration of MM with Real World Data

- Introduction to real world data: user generated contents (UGC's)
 - Check-in venues and user generated videos/ images
 - Integration of multimedia with real world data
-

- Towards the generation of social graph of cities

REFERENCE:

- [1] VK Singh, M Gao and R Jain. "Social Pixels: Genesis and Evaluation", ACM Multimedia' 2010, Oct, Florence, Italy.
- [2] R Jain and D Sonnen. Social Life Networks. IEEE IT Professional, Sept 2011.
- [3] T-S Chua (Keynote Talk). Extracting Knowledge form Community QA (cQA) Sites. COLING Workshop (People's Web Meets NLP), COLING'10 (23rd International Conference on Computational Linguistics 2010), Beijing, China, 23-28 August 2010. (<http://lms.comp.nus.edu.sg/slides/chua/2010-chua-coling.pdf>)

COURSE ASSIGNMNETS

ASSIGNMENT 1: A content-based image retrieval system

Submission date: 1 Mar (Thu) by 1200Hrs; Grading date: 2 Mar (Fri)

ASSIGNMENT 2: An interactive concept-based image/video retrieval system

Submission date: 22 Mar (Thu) by 1200Hrs; Grading date: 23 Mar (Fri)

ASSIGNMENT 3: Integration of MM and user-generated contents for real-world applications

Submission date: 12 Apr (Thu) by 1200Hrs; Grading date: 13 Apr (Fri)

Assignment Grouping: This is a group assignment. Each group comprises up to 2 members. Each group needs to make a 20-minute presentation cum demo for each assignment during evaluation.

Later Submission Policy: The following penalties are imposed for late submissions.

- (a) Late but within 24 hours: 25% reduction in grades.
 - (b) Later but within 5 days: 50% reduction in grades.
 - (c) After 5 days: zero mark.
-