CS6350: Computer Vision

July-Nov Semester 2024
Slot; G
Slots: Monday: 12:00-12:50pm, Wednesday: 5:00-5:50pm
Thursday: 10:00-10:50am, Friday: 9:00-9:50am

Prof. Sukhendu Das, SSB 302; Phone: 4367
Email: *sdas@cse.iitm.ac.in*
TA(s:) Pooja Kumari (cs20d006@smail.iitm.ac.in),
Apu Sarkar (cs23d002@smail.iitm.ac.in),
S Meena Padnekar (cs21d015@smail.iitm.ac.in)

Updated on July 25, 2024

Note: The course webpage is **www.cse.iitm.ac.in/~vplab/computer_vision.html**. A google group will be formed for course related communications; please check the e-mail regularly.

1 Course Objectives

Computer Vision focuses on development of algorithms and techniques to analyze and interpret the visible world around us. This requires understanding of the fundamental concepts related to multi-dimensional signal processing, feature extraction, pattern analy- sis visual geometric modeling, stochastic optimization etc. Knowledge of these concepts is necessary in this field, to explore and contribute to research and further developments in the field of computer vision. Applications range from biometrics, medical diagnosis, document processing, mining of visual content, to surveillance, advanced rendering etc.

2 Learning Outcomes

- To apply mathematical modeling methods for low-, intermediate- and high- level image pro- cessing tasks
- To be able to design new algorithms to solve recent state of the art computer vision problems.
- To perform software experiments on computer vision problems and compare their performance with the state of the art.
- To develop a broad knowledge base so as to easily relate to the existing literature.
- To gather a basic understanding about the geometric relationships between 2D images and the 3D world.
- To build a complete system to solve a computer vision problem.

3 Course prerequisite(s)

COT

4 Classroom Mode

Traditional lectures (4 [50 mins. slots] per week) with slides. 5-6 tutorials are tentatively planned to be conducted within these slots.

5 Textbooks

- Richard Szeliski, Computer Vision: Algorithms and Applications, Springer-Verlag London Limited 2011.
- Richard Hartley and Andrew Zisserman, Multiple View Geometry in Computer Vision, 2ndEdition, Cambridge University Press, March 2004

6 Reference Books

- R. Bishop; Pattern Recognition and Machine Learning, Springer, 2006
- R.C. Gonzalez and R.E. Woods, Digital Image Processing, Addison- Wesley, 1992.
- K. Fukunaga; Introduction to Statistical Pat- tern Recognition, Second Edition, Academic Press,Morgan Kaufmann, 1990.

7 Course Requirements

You are required to attend all the lectures. If you miss any of them it is your responsibility to find out what went on during the classes and to collect any ma- terials that may be handed out. Class participation

is strongly encouraged to demonstrate an appropriate level of understanding of the material being discussed in the class. Regular feedback from the class regarding the lectures will be very much appreciated.

8 Planned Syllabus

- Digital Image Formation and low-level processing: Overview and State-of-the-art, *Fundamentals of Image Formation, Transformation: Orthogonal, Euclidean*, Affine, Projective, etc; *Fourier Transform*, Convolution and Filtering, Image Enhancement, Restoration, *Histogram Processing*

- Depth estimation and Multi-camera views: *Perspective, Binocular Stereopsis*: Camera and Epipolar Geometry; Homography, Rectification, DLT, RANSAC, 3-D reconstruction framework; Auto-calibration.

- Feature Extraction: Edges - Canny, LOG, DOG; Line detectors (Hough Transform), Corners - Harris and Hessian Affine, Orientation Histogram, SIFT, SURF, HOG, GLOH, Scale- Space Analysis- Image Pyramids and Gaussian derivative filters, Gabor Filters and DWT.

- Image Segmentation: *Region Growing, Edge Based approaches to segmentation*, Graph-Cut, Mean-Shift, MRFs, Texture Segmentation; Object detection.

- Pattern Analysis: *Basics of Probability and Statistics*, Clustering: K-Means, K-Medoids, Mixture of Gaussians, Classification: Discriminant Function, Supervised, Un-supervised, Semi-supervised; Classifiers: Bayes, KNN, ANN models; Dimensionality Reduction: PCA, LDA, ICA; Non-parametric methods.

- Motion Analysis: Background Subtraction and Modeling, Optical Flow, KLT, Spatio-Temporal Analysis, Dynamic Stereo; Motion parameter estimation.

- Shape from X: Light at Surfaces; Phong Model; Reflectance Map; Albedo estimation. Photometric Stereo; Use of Surface Smoothness Constraint; Shape from Texture, color, motion and edges.

The topics shown in *italic* font will be left for self-study.

9 Tentative Grading Policy

The following allocation of marks is tentative. The flexibility will be used to the benefit of average grades for the entire class.

Mid Sem (50 mins)	15-20
End Sem Exam (120-150 mins)	35-40
TPA*	30-35
Tuts(best n-1)	10-15
Total	100

10 Tentative Schedule (Dates)

Tentative dates of activities throughout the semester:

Finalizing the groups by students	TBA
Choice of TPA by students	TBA
TPA Assignment to each group	26/08/2024
TA Interaction for TPA	TBA
Interim TPA Review (optional)	TBA
Tutorial 1	16/08/2024,
Tutorial 2	05/09/2024
Mid Semester Exam	**19/09/2024**
Tutorial 3	03/10/2024
Tutorial 4	17/10/2024
Tutorial 5	30/10/2024
Tutorial 6	TBA (optional)
Extra Classes	TBA
End Semester Exam	**12/11/2024**
Final TPA Demo/Viva Slot 1	14/11/2024
Final TPA Demo/Viva Slot 2	23/11/2024

11 Academic Honesty

Academic honesty is expected from each student participating in the course. NO sharing (willing, unwill- ing, knowing, unknowing) of assignment code between students, submission of downloaded code (from the Internet, or anywhere else) is allowed.

Academic violations will be handled by IITM Sen- ate Discipline and Welfare (DISCO) Committee. Typically, the first violation instance will result in ZERO marks for the corresponding component of the Course Grade and a drop of one-penalty level in final course grade. The second instance of code copying will result in same as above for marks but drop of two levels of final course grades at end of semester. The DISCO Committee of IIT Madras will also be intimated of the matter.

Please protect your Moodle account password. Do not share it with ANYONE. Do not share your aca- demic disk drive space on the Campus LAN.

Each proxy in the attendance will be penalized by 5% of (absolute) marks. It becomes 2% each for the donor & beneficiary, if both accept the fault.

Tutorials are not exams. Occasional ex- change of technical ideas is permitted, but

not copying. No complaints of copying will be entertained after the tutorial. If you find anybody copying, immediately inform the TA's present. If guilty, their copies will be taken away and they will obtain 0 marks for that tutorial.Based on CSE-DCC approval, re-evaluation of End Sem exam papers will only be done in front. Only minor clarifications, to- taling errors, any missed corrections will be considered till that time. In case if classes are conducted in online mode, recording of any discussions on your digital device with- out taking permission of the teacher is disal- lowed. If any found to do that, will be re- ported to Deptt/IITM level ethical commit- tee. Recorded video lectures may not be dis- tributed to anyone else, other than TAs, registered students. Posting same on any platform online/digital media is illegal and will also be reported.

12 Basic Preparation and Background knowledge

The students are expected to have a basic knowledge of the following:

- Linear Algebra
- Euclidean Geometry (2D and 3D).
- Fundamentals of Probability and Statistics.
- Basics of DSP.
- Programming - Python/C/C++/OpenCV
- Basics of PRML(may give you an advantage) and Deep Learning(only for TPA).

Miscellaneous Instruction :

- Self-study topics of the syllabus, as announced, may be studied without delay.
- Advisable to learn & use OpenCV techniques for TPA.
- Clarification of topics covered in class can be asked over email to TA or Google form circu- lated.
- Students are expected to make their own arrangements- personal laptops, servers or desktops in their respective departments, or even use Google Colab. Only during the final demo, a GPU-based server will be kept ready for installation and use.