Birla Institute of Technology & Science, Pilani
Hyderabad Campus

FIRST SEMESTER 2022-2023

Course Handout Part II

Date: 29-08-2022

In addition to part-I (General Handout for all courses appended to the time table) this portion gives further specific details regarding the course.

Course No.	***:***	***CS F441***
Course Title	***:***	***Selected Topics from Computer Science (Computer Vision)***
Instructor-in-Charge	***:***	***Dr. Aritra Mukherjee***

Scope and Objectives of the Course:

1. **Scope:**
   The scope of this course includes introduction to digital image format and camera model; point operations, linear and non linear filtering, Fourier transform, pyramids and wavelets, geometric transformation; Feature engineering: points and patches, edges and contours, contour tracking, lines and vanishing points, Segmentation; Image classification, object detection, semantic segmentation, video understanding; Tracking and registration, applications of registration and tracking: Image stitching and alignment, optical flow, stereo vision and depth estimation, structure from motion and SLAM (introductory); 3D scanning, surface representations; Deep learning for computer vision- supervised learning, unsupervised learning, deep neural networks, convolutional networks, more complex models

2. **Objectives:**
   - ➔ Provide the student with an understanding of the need for Computer Vision
   - ➔ To gain knowledge on important verticals of computer vision with OpenCV and Python/C++
   - ➔ To gain basic knowledge on advanced topics such as deep learning for CV, structure for motion, model generation, bio-vision etc.
   - ➔ To get a rough idea about designing autonomous systems using real time computer vision techniques and incorporation of other sensors to work in tandem with cameras.

Textbooks:

1. Computer Vision: Algorithms and Applications 2nd Edition, Richard Szeliski, Springer 2021
2. Computer Vision: A Modern Approach, second edition by David A. Forsyth and Jean Ponce, Pearson, 2012

Reference books

1. Deep Learning for Vision Systems by Mohamed Elgendy, Manning 2020

Course Plan:

Lecture No.	Learning objectives	Topics to be covered	Chapter in the Text Book
1	To introduce Computer Vision- what and why?	Brief history of the subject, different computer vision tasks and why they are needed	T1: chap 1
2-3	To get basics of image/video formation and basic transformations	Introduction to digital image format and camera model, Color space, pixel geometry, 2D and 3D transformations and rotations, Lighting and optics	T1: chap 2 T2: chap 1.1
4-8	To know about basics of Image Processing	Pixel and color transforms, histogram equalization	T1: chap 3.1 T2: chap 3.3
		Linear filtering on color spaces, bandpass and steerable filter. Nonlinear filtering- median and bilateral filtering. Binary image processing	T1: chap 3.2 and 3.3
		Fourier transform and applications in filtering. Pyramids and wavelets and need for multiresolution representations.	T1: chap 3.4 and 3.5 T2: chap 4.3
		Geometric transformations and mesh based warping and morphing	T1: chap 3.6
9-14	To get familiarized with feature engineering, patch, edge, corner and other features	Feature detector and descriptor, various features life SIFT, HOG, BRIEF etc. feature tracking and large scale matching	T1: chap 7.1 T2: chap 5.4
		Edge detection and contour detection, LoG, Sobel, Canny, edge kernels, snakes, splines, scissors, level sets	T1: chap 7.2, 7.3
		Hough transforms, RANSAC based lines and circles detection, using convex hulls, Vanishing point detection, perspective	T1: chap 7.4 T2: chap 10.1, 10.2
15-18	To know about segmentation- for image classification, object detection and video understanding	Graph based segmentations, superpixels, mean shifts, normalized cuts.	T1: chap 7.5 T2: chap 9.3, 9.4
		Semantic segmentation (classical approach), dense and sparse approaches, HOG and wavelet based face/object detection, blob based object detection	T1: chap 4.3, 6.3 T2: chap 16.1
19-22	To know about registration of landmarks: tracking, image stitching, alignment, optical flow	Pairwise alignment, 2D feature alignment, least square and RANSAC, 3D feature alignment	T1: chap 8.1 T2: chap 11.1
		Image stitching by rotational panorama, gap closing, projection on non cartesian coordinate systems	T1: chap 8.2
		Global alignment, bundle adjustment, parallax removal, composting, blending	T1: chap 8.3, 8.4 T2: chap 12.1
		Translational alignment by hierarchical motion estimation, spline based motion, video stabilization, optical flow (non deep approaches)	T1: chap 9.1, 9.2, 9.3 T2: chap 11.2, 11.3
23-28	To know about basics of Epipolar geometry, stereo vision, 3D	Epipolar geometry, sparse and dense correspondence, essential and	T1: chap 12.1, 12.2., 12.3, 12.4

	reconstruction, structure from motion, SLAM	fundamental matrices, subpixel estimation disparity map	T2: chap 7.1
		Multiview stereo, scene flow, volumetric and surface reconstruction	T1: chap 12.7 T2: chap 10.6
		Geometric intrinsic calibration, pose estimation by iterative nonlinear approach, two frame structure from motion (SFM) by multipoint algorithms, multiframe SFM by bundle adjustment	T1: chap 11.1, 11.2, 11.3, 11.4 T2: chap 8.1, 8.2
		Simultaneous Localization And Mapping for SFM, applications in augmented reality, area mapping and robot navigation	T1: chap 11.5 T2: chap 8.3
29-30	To know about 3D scanning and surface representations	Range data merging, various scanning techniques (Lidar, structured light etc.), surface interpolation, point and volumetric representation (point cloud and vector based)	T1: chap 13.2, 13.3, 13.4, 13.5 T2: chap 14.1, 14.2
31-40	To get the basics of Deep learning (in purview of computer vision): supervised and unsupervised approaches, CNN, different models of CNN for different tasks, spatio-temporal learning, semantic translation	Classical supervised learning: bayesian networks, nearest neighbor, logistic regression, support vector machines, decision trees and forests	T1: chap 5.1 T2: chap 15.2
		Classical unsupervised learning: Clustering- KMeans and Gaussian Mixture Models, Principal Component Analysis, Manifold learning	T1: chap 5.2 T2: chap 15.2
		Deep neural networks: weights, layers, regularization, normalization, activation functions, loss, gradient descent, backpropagation, training optimization	T1: chap 5.3 R1: chap 2
		Convolution neural networks: Kernels, pooling and unpooling, different types of layers (concat, FCN, dropout etc.) Network architectures, various pretrained models, dataset mechanisms, various performance metrics, adversarial examples, self supervised learning	T1: chap 5.4 R1: chap 3, 4, 5
		Complex models: 3D CNNs, UNets, RNN, Siamese, Transformer, BERT, GANs	T1: chap 5.5 Online resources
41-42	To get familiarized with applications in other domains of engineering	Applications in robot vision, biomedical vision, remote sensing, surveillance, agriculture, gaming and simulation etc.	Online resources

Evaluation Scheme:

Component	Duration	Weightage (%)	Date & Time	Nature of Component
Mid semester Exam*	90 mins	30%	05/11 3.30 - 5.00PM	Closed book
Comprehensive Exam*	180 mins	40%	31/12 FN	Closed book
Micro Project (10% of evaluation for Mid semester grading)		20%	To be notified	Open book
Coding Assignment	TBA	10%	To be notified	Open book

Note: minimum 40% of the evaluation to be completed by midsem grading.

Chamber Consultation Hour: To be discussed in the classroom/via email as per requirement.

Notices:
All notices pertaining to this course will be displayed on the Google Classroom/CS&IS Notice Board, as applicable.

Make-up Policy:
Make-up for Mid-semester test may be given for genuine cases with prior permission by IC. For the Comprehensive exam, make-up has to be approved by the IC and scheduled by AUGSD.

Academic Honesty and Integrity Policy:
Academic honesty and integrity are to be maintained by all the students throughout the semester and no type of academic dishonesty is acceptable.

INSTRUCTOR-IN-CHARGE