16-720 Computer Vision Fall 2026

Last updated August 7, 2026

TR | **3:30–4:50 PM** | **TEP 1403** | **Lecturer:** Kris Kitani | **TAs:** Ce Zhang, Wenli Xiao, Vihaan Misra, Aviral Chharia | **Canvas**

COURSE OVERVIEW

This course introduces the fundamental principles and algorithms of computer vision, focusing on how machines interpret, recognize, and reconstruct the visual world from images. Topics include filtering, feature detection and matching, image representations, classical and deep visual recognition, camera geometry, multi-view reconstruction, stereo vision, optical flow, and image alignment. The course concludes with modern generative models, including diffusion transformers, and neural 3D scene representations such as NeRF and Gaussian Splatting. Programming assignments use Python, NumPy, OpenCV, and PyTorch.

Prerequisites:
Linear Algebra, Multivariate Calculus, Probability Theory, Python Programming, and prior exposure to Machine Learning.

ASSESSMENT

Programming Assignment 1: Image Features	25%
Programming Assignment 2: Recognition	25%
Programming Assignment 3: Geometry and Motion	25%
Checkpoint 1	10%
Checkpoint 2	10%
Checkpoint 3	10%
Total:	**105%**

Grading:
A+ $\geq$ 100%, A 90–99%, B 80–89%, C 70–79%, D 60–69%, R $<$ 60%.

LEARNING OUTCOMES

Students should be able to:

1. Implement filtering, feature detection, description, and matching.
2. Build recognition systems using classical and deep learning methods.
3. Estimate image geometry and perform multi-view 3D reconstruction.
4. Analyze motion using optical flow and image alignment.
5. Explain modern generative and neural 3D representations.
6. Use Python, NumPy, OpenCV, and PyTorch for computer vision.

LATE DAYS AND COLLABORATION

Students receive five late days for the semester and may use at most two on any one programming assignment. No credit is given after all late days are exhausted. Regrade requests must be submitted within two weeks. Additional flexibility may be granted for circumstances outside a student's control, but not for planned activities such as conference deadlines or vacation travel. The three in-class checkpoints emphasize conceptual understanding and steady engagement. There is no final exam.

High-level discussion is encouraged, but submitted code and writing must be your own. Acknowledge collaborators, outside resources, and generative AI assistance. AI may support learning and debugging, but may not replace your understanding or authorship.

CLASS SCHEDULE

Date	No.	Topic	Assignments / Notes
Aug 25	L1	Introductions, Policies, Grading, Applications of Computer Vision	
Aug 27	L2	Filtering and Image Pyramids	
Sep 1	L3	Fourier Domain	
Sep 3	L4	Image Boundaries and Gradient Filters	
Sep 8	L5	Feature Detection	**HW1 Released: Image Features**
Sep 10	L6	Feature Descriptors	
Sep 15	L7	Image Representations: GIST, Bag of Visual Words	
Sep 17	–	**Checkpoint 1 (35 minutes)**, lecture continuation	
Sep 22	L8	Classification Methods: K-NN, Naive Bayes, SVM	
Sep 24	L9	Neural Networks: Perceptron, Gradient Descent, SGD	
Sep 29	L10	Initialization, Learning Rates, and Optimization Algorithms	**HW1 Due** **HW2 Released: Recognition**
Oct 1	L11	CNNs and Image Classification	
Oct 6	L12	Object Detection	
Oct 8	L13	Vision Transformers	
Oct 13	–	**NO CLASS — Fall Break**	
Oct 15	–	**NO CLASS — Fall Break**	
Oct 20	L14	2D Transformations: Euclidean, Similarity, Affine, Projective	
Oct 22	–	**Checkpoint 2 (60 minutes)**, lecture continuation	
Oct 27	L15	Homography Estimation	**HW2 Due** **HW3 Released: Geometry and Motion**
Oct 29	L16	Single-View Geometry	
Nov 3	–	**NO CLASS — Democracy Day**	
Nov 5	LXX	Special Topics / Buffer / Guest Lecture	
Nov 10	L17	Camera Models, Pose Estimation, and Triangulation	
Nov 12	L18	Epipolar Geometry	
Nov 17	L19	Structure from Motion	
Nov 19	–	**Checkpoint 3 (35 minutes)**, lecture continuation	
Nov 24	L20	Stereo, Optical Flow, and Image Alignment	**HW3 Due**
Nov 26	–	**NO CLASS — Thanksgiving**	
Dec 1	LXX	Generative Models (*e.g.*, GANs, Diffusion)	
Dec 3	LXX	3D Representations (*e.g.*, NeRF, Gaussian Splatting)	