Boston University, College of Engineering
Department of Electrical and Computer Engineering

ENG EC720 A1 – Fall 2021

Digital Video Processing

Instructor: Prof. Janusz Konrad (office: PHO 443, e-mail: `jkonrad@bu.edu`, tel.: 353-1246)

Lectures: Tue-Thu 3:30–5:15/PHO 201, **Office hours:** TBD

Objectives: This course extends EC520 ("Digital image processing and communication") to time-varying imagery, such as digital video, and to other types of image sequences (multi-view, multi-modal, etc.). The goal is to provide the understanding of the theory behind various video processing tasks as well as the practical experience in implementing them in a high-level computer language. The material covered in the course will extend numerous concepts from still images (2-D, i.e., $x-y$) to image sequences (3-D, i.e., $x-y-t$) and multi-view sequences (4-D, i.e., $x-y-t-\alpha$), but will also introduce new concepts unique to spatio-temporal data (timeline, motion, motion-induced occlusions, etc.) and spatio-angular data (viewing angle, angle-dependent perspective, angle-induced occlusions, etc.) The course format will be a combination of regular lectures and homework assignments, and of compulsory readings followed by in-class discussions. A very important aspect of the course will be a practical project. Students will select a topic, find suitable literature (with instructor's guidance) and carry out a simulation in *Matlab/Python/C/C++*. Upon the completion of this course students will have acquired in-depth knowledge to carry out research in image-sequence-related areas and also an understanding of modern applications of video processing (e.g., video analytics, advanced video compression and streaming, stereo vision for autonomous navigation).

Content: The course will start with an accelerated review of EC520 material, including multi-dimensional (M-D) sampling theory, M-D filtering, human visual system properties and digital representation of image data. The distinguishing feature of image sequences compared to still images is the ability to capture scene dynamics. Therefore, a substantial part of the course will be devoted to motion analysis in image sequences, and in particular to motion detection and motion estimation. A related concept of video segmentation, i.e., of video-domain partitioning into disjoint sets of pixels corresponding to "events" in the observed scene, will be presented as well. Although motion in an image sequence is necessarily a space-time concept, its description in the frequency or space/frequency domain has proven beneficial to many video analysis and compression tasks, and will be covered in detail. This will be followed by a discussion of video enhancement methods (noise reduction, super-resolution, error concealment) and recent advances in video compression. The most successful approach to video compression to date, namely the motion-compensated hybrid DCT/DPCM coding, will be introduced first. Then, salient features of the H.26X and MPEG-X families of video compression standards will be highlighted including the most recent AVC and HEVC standards. The course will conclude with a discussion of error resilience issues (due to IP or wireless transmission) and means of addressing them. Video processing has been traditionally **model-based** due to its high computational complexity, however **learning-based methods** are slowly emerging. Throughout the course, examples of such methods, primarily based on **neural networks**, will be provided.

Prerequisites: EC520 or EC516 or equivalents, and experience with *Matlab*.

Tentative outline:

- *Introduction*: scope of the course, historic background on moving images, overview of applications, present and future challenges, basic definitions
- *Review*: spatio-temporal (3-D) sampling and filtering of signals, digital representation of image sequences, human visual system (motion perception)
- *Motion analysis*: motion detection, 2-D motion estimation, 3-D motion estimation
- *Image sequence segmentation*: spatial segmentation (frame-by-frame), temporal segmentation (scene cut detection), spatio-temporal segmentation
- *Spectral analysis of image sequences*: Fourier-, DCT-, and wavelet-domain analysis
- *Video enhancement:* noise reduction, super-resolution, error concealment
- *Video compression:* motion-compensated hybrid DCT/DPCM coding, H.26X and MPEG-X families of compression standards, error resilience in video coding, introduction to motion-compensated wavelet-domain video compression (MCTF)

Grading:

50% Assignments	5 assignments covering theory and practice (*Matlab*); penalty for late submission, and no assignment accepted after one week.
50% Project	Project involving experimental work; report and in-class presentation mandatory at the end of the course.

Homework: The homework must be the result of your individual work. You may discuss the contents and general approach to a problem with your classmates but not the detailed solution. You are expected to formulate your approach and write the solutions by yourself. You must clearly reference any sources you used in your work: books, Internet, and your collaborators! Copying the solution and/or answer from another student or source is considered academic misconduct. Procedures that will be followed should academic misconduct be discovered are defined in Student Handbook.

Course material: The course is primarily based on my lecture notes and various handouts, but I will also use the following book: M. Tekalp, *Digital Video Processing*. Signal Processing Series, Prentice Hall, 2015 (2nd edition only). Additionally, I will make reading material (journal and conference papers) available on the course website.

Other references helpful in the course:

- J.W. Woods, *Multidimensional Signal, Image and Video Processing and Coding*. Academic Press, 2nd edition - 2012 (alternatively, 1st edition - 2006).
- Y. Wang, J. Ostermann and Y.-Q. Zhang, *Video Processing and Communications*. Signal Proc. Series, Prentice Hall, 2002.
- A. Bovik (Ed.), *The Essential Guide to Video Processing*. Academic Press, 2009.
- A. Jain, *Fundamentals of Digital Image Processing*. Information and System Sciences Series, Prentice Hall, 1989.
- E. Dubois, "The sampling and reconstruction of time-varying imagery with application in video systems," *Proc. IEEE*, vol. 73, pp. 502–522, Apr. 1985.
- J.S. Lim, *Two-dimensional Signal and Image Processing*. Prentice Hall, 1990.