University of Florida

Department of Electrical and Computer Engineering

EEE 6512, Section 012A

Image Processing and Computer Vision

Fall 2021

Course Description

This is a 3-credit course.

This course introduces fundamental concepts and techniques for image processing and computer vision. We will address 1) how to efficiently represent and process image/video signals, and 2) how to deliver image/video signals over networks. Topics to be covered include: image acquisition and display using digital devices, properties of human visual perception, sampling and quantization, image enhancement, image restoration, two-dimensional Fourier transforms, linear and nonlinear filtering, morphological operations, noise removal, image deblurring, edge detection, image registration and geometric transformation, image/video compression, video communication standards, video transport over the Internet and wireless networks, object recognition and image understanding.

Course Prerequisites

- EEL 3135 (Discrete-Time Signals and Systems) or undergraduate-level signals and systems
- EEL 4516 (Noise in Devices and communication Systems) or undergraduate-level probability theory/stochastic processes

- Some exposure to MATLAB or Python programming language
- Knowledge of basic matrix theory (linear algebra) would be helpful, but not necessary

Required Textbook

- Rafael C. Gonzalez, Richard E. Woods, ``Digital Image Processing,'' 3rd Edition, Prentice Hall; ISBN: 013168728X; August 2007.

or

- Rafael C. Gonzalez, Richard E. Woods, ``Digital Image Processing,'' 4th Edition, Pearson; ISBN-13: 9780133356724; 2018.

Recommended Readings

- George Siogkas, "Visual Media Processing Using Matlab Beginner's Guide," Packt Publishing, 2013. ISBN-10: 1849697205|ISBN-13: 978-1849697200
- Oge Marques, "Practical Image and Video Processing Using MATLAB," Wiley, New York, NY, 2011. ISBN-10: 0470048158 | ISBN-13: 978-0470048153
- Rafael C. Gonzalez, Richard E. Woods, and S. L. Eddins, ``Digital Image Processing Using MATLAB,'' Prentice Hall, 2004. ISBN 0130085197.
- Anil K. Jain, ``Fundamentals of digital image processing,'' Englewood Cliffs, NJ : Prentice Hall, 1989.
- Y. Wang, J. Ostermann, and Y.Q.Zhang, "Video Processing and Communications," 1st ed., Prentice Hall, 2002. ISBN: 0130175471.
- D. Taubman and M. Marcellin, "JPEG2000: Image Compression Fundamentals, Standards, and Practice," Kluwer, 2001. ISBN: 079237519X.
- David A. Forsyth, Jean Ponce, "Computer Vision: A Modern Approach," Prentice Hall; 1st edition (August 14, 2002), ISBN: 0130851981.
- Richard Hartley, Andrew Zisserman, "Multiple View Geometry in Computer Vision," Paperback: 672 pages; Publisher: Cambridge University Press; 2 edition (March 25, 2004) ISBN: 0521540518
- Yi Ma, Stefano Soatto, Jana Kosecka, S. Shankar Sastry, "An Invitation to 3-D Vision," Hardcover: 526 pages ; Publisher: Springer-Verlag; (November 14, 2003) ISBN: 0387008934
- A. Ardeshir Goshtasby, "2-D and 3-D Image Registration," Wiley Press, April. 2005. [ebook on NetLibrary]
- John W. Woods, "Multidimensional Signal, Image, and Video Processing and Coding," Academic Press; (March 13, 2006), ISBN-10: 0120885166, ISBN-13: 978-0120885169.
- Linda G. Shapiro and George C. Stockman, "Computer Vision," Prentice-Hall, Inc., Upper Saddle River, New Jersey, 2001 (ISBN 0-13-030796-3).

- Emanuele Trucco and Alessandro Verri, "Introductory Techniques for 3-D Computer Vision," Prentice-Hall, Inc., Upper Saddle River, New Jersey, 1998 (ISBN 0-13-261108-2).
- Iain E G Richardson, "H.264 and MPEG-4 Video Compression," John Wiley & Sons, September 2003, ISBN 0-470-84837-5
- M. E. Al-Mualla, C. N. Canagarajah and D. R. Bull, "Video Coding for Mobile Communications: Efficiency, Complexity and Resilience", Elsevier Science, Academic Press, 2002. ISBN: 0120530791
- A. Gersho, and R. Gray. Vector Quantization and Signal Compression. Boston: Kluwer Academic Publishers, 1992.

Course Information

Instructor:

Dr. Dapeng Wu
Office: NEB 431
Email: dpwu@ufl.edu

TA:

1) Heng Qiao
Email: hengqiao@ufl.edu

2) Haotian Jiang
Email: haotian.jiang@ufl.edu

3) Tianqi Liu

Email: tianqi.liu1@ufl.edu

4) Nidish Vashistha

Email: nidish@ufl.edu

Course website: http://www.wu.ece.ufl.edu/courses/eee6512f21

Meeting Time

Monday, Wednesday, Friday, period 8 (3 pm - 3:50 pm)

Meeting Room

NEB 100

Office Hours

- Dr. Wu: Monday, Wednesday, period 7 (1:55 pm - 2:45 pm), and by appointment via email.

Structure of the Course

The course consists of lectures, 6 homework assignments, a quiz, and 1 project.

This course is primarily a lecture course. I cover all important material in lectures. Since EEL 3135 and EEL 4516 are prerequisites, I assume some previous knowledge about DSP, probability theory and stochastic processes, and hence I will cover some material very quickly. Thus, depending on what and how much you recall from earlier study, varying amounts of reading in introductory books on DSP, probability theory and stochastic processes (other than the course textbook) may be necessary; these readings are up to the student. I will only give reading assignments from the course textbook.

Attending lecture is quite important as I may cover material not available in any book easily accessible to you. I use Powerpoint presentation during lecture. Lecture notes will be posted on the course website before the class. The lecture is to engage the students in independent thinking, critical thinking, and creative thinking, help the students organize the knowledge around essential concepts and fundamental principles, and develop conditionalized knowledge which tells them when, where and why a certain method is applicable to solving the problem they encounter.

I do not intend for the WWW material to be a substitute for attending lecture since engaging the students in active thinking, making logical connections between the old knowledge and the new knowledge, and providing insights are the objectives of my lecture. The lecture notes are posted on the web so that you can miss an occasional lecture and still catch up, and it makes taking notes easier.

Course Outline

- Overview of image processing systems, Image formation and perception, Continuous and digital image representation
- Image quantization: uniform and nonuniform, visual quantization (dithering).
- Image contrast enhancement: linear and non-linear stretching, histogram equalization.
- Continuous and discrete-time Fourier Transforms in 2D; and linear convolution in 2D.
- Image smoothing and image sharpening by spatial domain linear filtering; Edge detection.
- Discrete Fourier transform in 1D and 2D, and image filtering in the DFT domain.
- Median filtering and Morphological filtering.
- Color representation and display; true and pseudo color image processing.

- Image sampling and sampling rate conversion (resize).
- Lossless image compression: The concept of entropy and Huffman coding; Runlength coding for bi-level images; CCITT facsimile compression standards.
- Lossy image compression: Image quantization revisited; Predictive coding; Transform coding; JPEG image compression standard.
- Imaging Geometry; Coordinate transformation and geometric warping for image registration.
- Object recognition

Course Objectives

Upon the completion of the course, the student should be able to

- know the fundamental techniques for image processing, video processing, and computer vision
- understand the basics of analog and digital video: video representation and transmission
- acquire the basic skill of designing image/video compression
- familiarize himself/herself with image/video compression standards

Handouts

Please find handouts here.

Course Policies

- **Attendance:**
  - Perfect class attendance is not required, but regular attendance is expected.
  - It is the student's responsibility to independently obtain any missed material (including handouts) from lecture.
- **During lecture, cell phones should be turned off.**
- **No late submissions of your homework solution, and project proposal/report, are allowed unless U.F. approved reasons are supplied and advance permission is granted by the instructor.** Excused late submissions must be consistent with university policies in the Graduate Catalog (https://catalog.ufl.edu/graduate/regulations) and require appropriate

documentation. Additional information can be found here: https://catalog.ufl.edu/graduate/regulations/

- **Software use**
  - All faculty, staff and student of the University are required and expected to obey the laws and legal agreements governing software use. Failure to do so can lead to monetary damages and/or criminal penalties for the individual violator. Because such violations are also against University policies and rules, disciplinary action will be taken as appropriate. We, the members of the University of Florida community, pledge to uphold ourselves and our peers to the highest standards of honesty and integrity.
- **Announcements:**
  - All students are responsible for announcements made in lecture, on the student access website, or via the class email list.
  - It is expected that you will check your email several times per week for possible course announcements.
- **Students Requiring Accommodations**
  - Students with disabilities who experience learning barriers and would like to request academic accommodations should connect with the disability Resource Center by visiting https://disability.ufl.edu/students/get-started/. It is important for students to share their accommodation letter with their instructor and discuss their access needs, as early as possible in the semester.
- **University Honesty Policy**

UF students are bound by The Honor Pledge which states, "We, the members of the University of Florida community, pledge to hold ourselves and our peers to the highest standards of honor and integrity by abiding by the Honor Code. On all work submitted for credit by students at the University of Florida, the following pledge is either required or implied: "On my honor, I have neither given nor received unauthorized aid in doing this assignment." The Honor Code (https://www.dso.ufl.edu/sccr/process/student-conduct-honor-code/) specifies a number of behaviors that are in violation of this code and the possible sanctions. Furthermore, you are obligated to report any condition that facilitates academic misconduct to appropriate personnel. If you have any questions or concerns, please consult with the instructor or TAs in this class.

Students are encouraged to discuss class material in order to better understand concepts. All homework answers must be the author's own work. However, students are encouraged to discuss homework to promote better understanding. What this means in practice is that students are welcome to discuss problems and solution approaches, and in fact can communally work solutions at a board. However, the material handed in must be prepared starting with a clean sheet of paper (and the author's recollection of any solution session), but not refer to any written notes or existing code from other students during the writing of the solution. In other words, writing the homework report shall be an exercise in demonstrating the student understands the materials on his/her own, whether or not help was provided in attaining that understanding.

All work submitted in this course must be your own and produced exclusively for this course. The use of sources (ideas, quotations, paraphrases) must be properly

acknowledged and documented. For the copy of the UF Honor Code and consequences of academic dishonesty, please refer to http://www.dso.ufl.edu/sccr/honorcodes/honorcode.php. Violations will be taken seriously and are noted on student disciplinary records. If you are in doubt regarding the requirements, please consult with the instructor before you complete any requirement of the course.

Course Evaluation

Students are expected to provide professional and respectful feedback on the quality of instruction in this course by completing course evaluations online via GatorEvals. Guidance on how to give feedback in a professional and respectful manner is available at https://gatorevals.aa.ufl.edu/students/. Students will be notified when the evaluation period opens, and can complete evaluations through the email they receive from GatorEvals, in their Canvas course menu under GatorEvals, or via https://ufl.bluera.com/ufl/. Summaries of course evaluation results are available to students at https://gatorevals.aa.ufl.edu/public-results/.

In-Class Recording

Students are allowed to record video or audio of class lectures. However, the purposes for which these recordings may be used are strictly controlled. The only allowable purposes are (1) for personal educational use, (2) in connection with a complaint to the university, or (3) as evidence in, or in preparation for, a criminal or civil proceeding. All other purposes are prohibited. Specifically, students may not publish recorded lectures without the written consent of the instructor.

A "class lecture" is an educational presentation intended to inform or teach enrolled students about a particular subject, including any instructor-led discussions that form part of the presentation, and delivered by any instructor hired or appointed by the University, or by a guest instructor, as part of a University of Florida course. A class lecture does not include lab sessions, student presentations, clinical presentations such as patient history, academic exercises involving solely student participation, assessments (quizzes, tests, exams), field trips, private conversations between students in the class or between a student and the faculty or lecturer during a class session.

Publication without permission of the instructor is prohibited. To "publish" means to share, transmit, circulate, distribute, or provide access to a recording, regardless of format or medium, to another person (or persons), including but not limited to another student within the same class section. Additionally, a recording, or transcript of a recording, is considered published if it is posted on or uploaded to, in whole or in part, any media platform, including but not limited to social media, book, magazine, newspaper, leaflet, or third party note/tutoring services. A student who publishes a recording without written consent may be subject to a civil cause of action instituted by a person injured by the publication and/or discipline under UF Regulation 4.040 Student Honor Code and Student Conduct Code.

Commitment to a Safe and Inclusive Learning Environment

The Herbert Wertheim College of Engineering values broad diversity within our community and is committed to individual and group empowerment, inclusion, and the

elimination of discrimination. It is expected that every person in this class will treat one another with dignity and respect regardless of gender, sexuality, disability, age, socioeconomic status, ethnicity, race, and culture.

If you feel like your performance in class is being impacted by discrimination or harassment of any kind, please contact your instructor or any of the following:

• Your academic advisor or Graduate Program Coordinator

• Jennifer Nappo, Director of Human Resources, 352-392-0904, jpennacc@ufl.edu

• Curtis Taylor, Associate Dean of Student Affairs, 352-392-2177, taylor@eng.ufl.edu

• Toshikazu Nishida, Associate Dean of Academic Affairs, 352-392-0943, nishida@eng.ufl.edu

Software Use

All faculty, staff, and students of the University are required and expected to obey the laws and legal agreements governing software use. Failure to do so can lead to monetary damages and/or criminal penalties for the individual violator. Because such violations are also against University policies and rules, disciplinary action will be taken as appropriate. We, the members of the University of Florida community, pledge to uphold ourselves and our peers to the highest standards of honesty and integrity.

Student Privacy

There are federal laws protecting your privacy with regards to grades earned in courses and on individual assignments. For more information, please see: http://registrar.ufl.edu/catalog0910/policies/regulationferpa.html

Campus Resources:

Health and Wellness

Covid-19 Protocols:

- You are expected to wear approved face coverings at all times during class and within buildings even if you are vaccinated. Please continue to follow healthy habits, including best practices like frequent hand washing. Following these practices is our responsibility as Gators.

- If you are sick, stay home and self-quarantine. Please visit the UF Health Screen, Test & Protect website about next steps, retake the questionnaire and schedule your test for no sooner than 24 hours after your symptoms began. Please call your primary care provider if you are ill and need immediate care or the UF Student Health Care Center at 352-392-1161 (or email covid@shcc.ufl.edu) to be evaluated for testing and to receive further instructions about

returning to campus. UF Health Screen, Test & Protect offers guidance when you are sick, have been exposed to someone who has tested positive or have tested positive yourself. Visit the UF Health Screen, Test & Protect website for more information.

U Matter, We Care:

Your well-being is important to the University of Florida. The U Matter, We Care initiative is committed to creating a culture of care on our campus by encouraging members of our community to look out for one another and to reach out for help if a member of our community is in need. If you or a friend is in distress, please contact umatter@ufl.edu so that the U Matter, We Care Team can reach out to the student in distress. A nighttime and weekend crisis counselor is available by phone at 352-392-1575. The U Matter, We Care Team can help connect students to the many other helping resources available including, but not limited to, Victim Advocates, Housing staff, and the Counseling and Wellness Center. Please remember that asking for help is a sign of strength. In case of emergency, call 9-1-1.

Counseling and Wellness Center: https://counseling.ufl.edu, and 392-1575; and the University Police Department: 392-1111 or 9-1-1 for emergencies.

Sexual Discrimination, Harassment, Assault, or Violence

If you or a friend has been subjected to sexual discrimination, sexual harassment, sexual assault, or violence contact the **Office of Title IX Compliance**, located at Yon Hall Room 427, 1908 Stadium Road, (352) 273-1094, title-ix@ufl.edu

Sexual Assault Recovery Services (SARS)

Student Health Care Center, 392-1161.

University Police Department at 392-1111 (or 9-1-1 for emergencies), or http://www.police.ufl.edu/.

Academic Resources

E-learning technical support, 352-392-4357 (select option 2) or e-mail to Learning-support@ufl.edu. https://lss.at.ufl.edu/help.shtml.

Career Resource Center, Reitz Union, 392-1601. Career assistance and counseling. https://www.crc.ufl.edu/.

Library Support, http://cms.uflib.ufl.edu/ask. Various ways to receive assistance with respect to using the libraries or finding resources.

Teaching Center, Broward Hall, 392-2010 or 392-6420. General study skills and tutoring. https://teachingcenter.ufl.edu/.

Writing Studio, 302 Tigert Hall, 846-1138. Help brainstorming, formatting, and writing papers. https://writing.ufl.edu/writing-studio/.

Student Complaints Campus: https://www.dso.ufl.edu/documents/UF_Complaints_policy.pdf.

On-Line Students Complaints: http://www.distance.ufl.edu/student-complaint-process.

TOP

Grading

Grading:

Grades	Percentage	Due Dates
Homework	30%	See the course calendar
Project proposal	10%	4pm, October 29
Quiz	10%	December 8
Project report	50%	4pm, December 15

The project report consists of

1. (50%) A written report for your project
2. (25%) Computer programs that you develop for your project
3. (10%) Powerpoint file of your presentation
4. (15%) Your presentation/demo video on YouTube

Grading scale:

Top 25% students will receive A. Average score will be at least B+.

More information on UF grading policy may be found at:
https://catalog.ufl.edu/ugrad/current/regulations/info/grades.aspx

Homework:

- Due dates of assignments are specified in the course calendar.
- **No late submissions** are allowed unless U.F. approved reasons are supplied and advance permission is granted by the instructor. Excused late submissions must be consistent with university policies in the Graduate Catalog (https://catalog.ufl.edu/graduate/regulations)

and require appropriate documentation. Additional information can be found here: https://catalog.ufl.edu/graduate/regulations/

- If you wish to dispute a homework grade, you must return the assignment along with a succinct written argument within one week after the graded materials have been returned to the class. Simple arithmetic errors in adding up grade totals are an exception, and can normally be handled verbally on-the-spot during office hours of the TA. For all other disputes, the entire homework may be (non-maliciously) re-graded, which may result in increase or decrease of points.

Class Project:

The class project will be done individually (that is, teaming with other students is not allowed). Each project requires a proposal and a final report. The final report is expected to be in the format of a conference paper plus computer programs, a Powerpoint file, and a video. On Oct. 29, the project proposal (up to 2 pages) is due. On Dec. 15, the final report (up to 10 pages) is due. For details about the project, please read here.

Suggested topics for projects are listed here.

TOP

Calendar

Course calendar can be found here.

TOP

Links Study Guides

Related courses in other schools:

George Mason University, Computer Vision

Johns Hopkins University, Image Compression and Packet Video

Polytechnic University, Video Processing

Purdue University, Digital Video Systems

Stanford University, Digital Video Processing

University of California, Berkeley, Multimedia Signal Processing, Communications and Networking

University of Maryland, College Park, Digital Image Processing

University of Maryland, College Park, Multimedia Communication & Information Security: A Signal Processing Perspective

---

Useful links

- Anaconda: Anaconda is the leading open data science platform powered by Python.
- Theano: Theano is a Python library that lets you to define, optimize, and evaluate mathematical expressions, especially ones with multi-dimensional arrays (numpy.ndarray).
- TensorFlow: TensorFlow is an open source software library for numerical computation using data flow graphs. Nodes in the graph represent mathematical operations, while the graph edges represent the multidimensional data arrays (tensors) communicated between them. The flexible architecture allows you to deploy computation to one or more CPUs or GPUs in a desktop, server, or mobile device with a single API.
- Keras: Keras is a minimalist, highly modular neural networks library, written in Python and capable of running on top of either TensorFlow or Theano. It was developed with a focus on enabling fast experimentation. Being able to go from idea to result with the least possible delay is key to doing good research.
- PyTorch: PyTorch is a deep learning framework for fast, flexible experimentation.
- A curated list of resources dedicated to recurrent neural networks
- Source code in Python for handwritten digit recognition, using deep neural networks: [another link]
- Source code in PyTorch for handwritten digit recognition, using deep neural networks
- Source code in Python for TF-mRNN: a TensorFlow library for image captioning
- Source code in Python for the following work on image captioning:
    - Oriol Vinyals, Alexander Toshev, Samy Bengio, Dumitru Erhan, Show and Tell: A Neural Image Caption Generator, CVPR 2015
        - Implementation
- Image captioning:
    - Zhe Gan, et. al, Semantic Compositional Networks for Visual Captioning, CVPR 2017
        - Implementation Source code in Python (Theano)
    - Bottom-Up and Top-Down Attention for Image Captioning and Visual Question Answering: source codes (Caffe) and source codes (PyTorch)
- Microsoft COCO datasets

- Visual Question Answering:
  - Bottom-Up and Top-Down Attention for Image Captioning and Visual Question Answering: source codes (Caffe) and VQA source code(PyTorch)
- Semantic Propositional Image Caption Evaluation (SPICE)
  - Source code in JAVA to calculate SPICE
- Region-based Convolutional Neural Networks (R-CNN)
  - References:
    - Ren, Shaoqing, Kaiming He, Ross Girshick, and Jian Sun. "Faster R-CNN: Towards real-time object detection with region proposal networks." In Advances in neural information processing systems, pp. 91-99. 2015. [pdf]
    - Dai, Jifeng, Yi Li, Kaiming He, and Jian Sun. "R-FCN: Object detection via region-based fully convolutional networks." In Advances in neural information processing systems, pp. 379-387. 2016. [pdf] [source code]
    - Huang, Jonathan, Vivek Rathod, Chen Sun, Menglong Zhu, Anoop Korattikara, Alireza Fathi, Ian Fischer et al. "Speed/accuracy trade-offs for modern convolutional object detectors." arXiv preprint arXiv:1611.10012 (2016). [pdf] (E.g., for Inception V3, extract features from the “Mixed 6e” layer whose stride size is 16 pixels. Feature maps are cropped and resized to 17x17.)
  - Source codes:
    - A Faster Pytorch Implementation of Faster R-CNN (PyTorch)
    - Bottom-Up and Top-Down Attention for Image Captioning and Visual Question Answering: source codes (Caffe)
- Source code in Python for end-to-end training of LSTM
  - Implementation
- Bidirectional Encoder Representations from Transformers (BERT)
  - Implementation in TensorFlow
  - Implementation in PyTorch
- Source code in Python for sequence-to-sequence learning (language translation, chatbot)
  - TensorFlow seq2seq library
  - Implementation 1 on Tensorflow with separable encoder and decoder
  - Implementation 2 on Keras
- Visual Storytelling Dataset (VIST)
  - Visual storytelling algorithms:
    - No Metrics Are Perfect: Adversarial REward Learning for Visual Storytelling: source codes (TensorFlow)
- Visual Genome is a dataset, a knowledge base, an ongoing effort to connect structured image concepts to language.
- MPII Movie & Description dataset for automatic video description, video summary, video storytelling
- Bidirectional recurrent neural networks (B-RNN):

    - Graves, Alan, Navdeep Jaitly, and Abdel-rahman Mohamed. "Hybrid speech recognition with deep bidirectional LSTM." IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU), 2013. [pdf]
- Deep reinforcement learning
    - UCL Course on reinforcement learning: [ppt] [video]
    - References:
        - Mnih, Volodymyr, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller. "Playing atari with deep reinforcement learning." *arXiv preprint arXiv:1312.5602* (2013).
        - Mnih, Volodymyr, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves et al. "Human-level control through deep reinforcement learning." *Nature* 518, no. 7540 (2015): 529-533. [source code]
        - How to Study Reinforcement Learning
    - Source codes:
        - Implementation of Reinforcement Learning Algorithms. Python, OpenAI Gym, Tensorflow. Exercises and Solutions to accompany Sutton's Book and David Silver's course. [link]
- Generative Adversarial Network (GAN)
    - References:
        - Goodfellow, Ian, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. "Generative adversarial nets." In *Advances in neural information processing systems*, pp. 2672-2680. 2014.
        - Radford, Alec, Luke Metz, and Soumith Chintala. "Unsupervised representation learning with deep convolutional generative adversarial networks." *arXiv preprint arXiv:1511.06434* (2015).
        - Arjovsky, Martin, Soumith Chintala, and Léon Bottou. "Wasserstein GAN." *arXiv preprint arXiv:1701.07875* (2017).
    - Types of GAN
        1. Vanilla GAN
        2. Conditional GAN
        3. InfoGAN
        4. Wasserstein GAN
        5. Mode Regularized GAN
        6. Coupled GAN
        7. Auxiliary Classifier GAN
        8. Least Squares GAN
        9. Boundary Seeking GAN
        10. Energy Based GAN
        11. f-GAN
        12. Generative Adversarial Parallelization
        13. DiscoGAN
        14. Adversarial Feature Learning & Adversarially Learned Inference
        15. Boundary Equilibrium GAN

16. Improved Training for Wasserstein GAN
17. DualGAN
18. MAGAN: Margin Adaptation for GAN
19. Softmax GAN

- Source codes:
  - A Tensorflow Implementation of "Deep Convolutional Generative Adversarial Networks": python code
  - Collection of generative models, e.g. GAN, VAE in Pytorch and Tensorflow: python code

- Sequential Generative Adversarial Network (GAN)
  - References:
    - Yu, Lantao, Weinan Zhang, Jun Wang, and Yong Yu. "SeqGAN: Sequence Generative Adversarial Nets with Policy Gradient." In *AAAI*, pp. 2852-2858. 2017.
    - Mogren, Olof. "C-RNN-GAN: Continuous recurrent neural networks with adversarial training." *arXiv preprint arXiv:1611.09904* (2016).
    - Im, Daniel Jiwoong, Chris Dongjoo Kim, Hui Jiang, and Roland Memisevic. "Generating images with recurrent adversarial networks." *arXiv preprint arXiv:1602.05110* (2016).
    - Press, Ofir, Amir Bar, Ben Bogin, Jonathan Berant, and Lior Wolf. "Language Generation with Recurrent Generative Adversarial Networks without Pre-training." *arXiv preprint arXiv:1706.01399* (2017).
  - Source codes:
    - Implementation of C-RNN-GAN
    - Tensorflow Implementation of GAN modeling for sequential data

- Subjective evaluation for content aware video processing techniques
- **MATLAB Tutorial**
- **MATLAB Central**
- **Matlab Primer**, **Matlab Manuals**, **Image Processing Toolbox**
- **Matlab implementation of image/video compression algorithms**
- **Introduction to Matarix Algebra** **(free book by Autar K Kaw, Professor, University of South Florida).**
- **Matrix Reference Manual**
- **HIPR2: a WWW-based Image Processing Teaching Materials with J**
- **LIDAR**
- **Learning by simulations**
- **OpenCV**
- **OpenGL**
- Download the following free (open source) program to record video with screen capture: http://www.nchsoftware.com/capture/index.html?gclid=CNadwsW6-6wCFSVjTAodbjzTSg
- SD and HD video sequences for evaluating coding performance of video codec: http://media.xiph.org/video/derf/
- WebRTC: WebRTC is a free, open-source project that enables web browsers with Real-Time Communications (RTC) capabilities via simple JavaScript APIs.

The Missing Semester of Your CS Education

Standards:

- H.264 tutorial
- H.263
- MPEG4 overview can be found at http://www.chiariglione.org/mpeg/standards/mpeg-4/mpeg-4.htm
- JPEG XR
- KTA (contender for future H.265)

ATSC (Advanced Television Systems Committee) & HDTV (High Definition Television):

ATSC.org

HDTV

SMPTE.org

MPEG (Moving Picture Experts Group):

MPEG.org

MPEG standards committee

MPEG TV

MP3

- MPEG Audio Layer-3

Software:

- Video codec
- Virtual Dub: VirtualDub is a video capture/processing utility for 32-bit Windows platforms (95/98/ME/NT4/2000/XP), licensed under the GNU General Public License (GPL).
- XnView: is an efficient multimedia viewer, browser and converter.
- ImageJ: Read and write GIF, JPEG, and ASCII. Read BMP, DICOM, and FITS. [Open Source, Public Domain]

- Open source for image processing tasks: http://octave.sourceforge.net/doc/image.html
- Photosynth: you can access gigabytes of photos in seconds, view a scene from nearly any angle, find similar photos with a single click, and zoom in to make the smallest detail as big as your monitor.
    - Refer to: http://labs.live.com/photosynth/
    - A demo video: http://www.ted.com/index.php/talks/view/id/129
- Video filtering and compression, by the Video Group, Moscow State University
- MSU Lossless Video Codec, by the Video Group, Moscow State University

HSI color model

Compression link: http://cchen1.et.ntust.edu.tw/compression/compression.htm

---

JOURNALS

Elsevier

- Computer Vision and Image Understanding
- Digital Signal Processing: A Review Journal
- Graphical Models and Image Processing
- Journal of Visual Commuication and Image Representation
- Real-Time Imaging
- Computers & Graphics
- Data & Knowledge Engineering
- Image and Vision Computing
- Pattern Recognition
- Pattern Recognition Letters
- Signal Processing
- Signal Processing: Image Communication

IEEE

- IEEE Transactions on Circuits and Systems for Video Technology
- IEEE Transactions on Multimedia
- IEEE Transactions on Image Processing
- IEEE Transactions on Medical Imaging
- IEEE Transactions on PAMI

Kluwer

- International Journal of Computer Vision
- Journal of Intelligent Information System
- Multidimensional Systems and Signal Processing

SPIE

- Journal of Electronic Imaging

Digital Video and Multimedia Standards Pages

- MPEG Pointers and Resources
- JPEG Tutorial
- JPEG FAQ
- Compression FAQ
- VRML Homepage
- Internet Engineering Taskforce Homepage

Digital TV and DVD

- Worldwide TV Standards
- More on Digital TV
- DVD FAQ

Overview of the AVI format

Signal Processing Information Base (SPIB)

Computer Vision

- *Computer Vision Homepage at CMU*
- Annotated Computer Vision Bibliography from USC IRIS
- *CVonline: The Evolving, Distributed, Non-Proprietary, On-Line Compendium of Computer Vision*
- **3-D for Everyone**

- Red-blue glasses or anaglyph for 3D viewing: http://www.best3dglasses.com/anaglyph.html
- Shutter glasses for 3D viewing: http://www.stereo3d.com/shutter.htm
- 3D cameras: http://www.ptgrey.com/index.asp
- 3D photos at http://www.jessemazer.com/3Dphotos.html
- 3D video sequences can be downloaded at: http://research.microsoft.com/vision/InteractiveVisualMediaGroup/3DVideoDownload/

Public Domain Image Databases

CMU Database

Patent licensing

As with MPEG-2 Parts 1 and 2 and MPEG-4 Part 2 amongst others, the vendors of H.264/AVC products and services are expected to pay patent licensing royalties for the patented technology that their products use. The primary source of licenses for patents applying to this standard is a private organization known as MPEG-LA, LLC (which is not affiliated in any way with the MPEG standardization organization, but which also administers patent pools for MPEG-2 Part 1 Systems, MPEG-2 Part 2 Video, MPEG-4 Part 2 Video, and other technologies).

To search patents, visit free patent searching site: www.FreePatentsOnline.com.