EC4480: Image Processing and Recognition

By

Roberto Cristi, Professor

Dept of Electrical and Computer Engineering

Naval Postgraduate School

Monterey, CA 93943

rcristi@nps.edu

EC4480. Week1: *Color Representation*

Description.

Humans perceive light and colors through two classes of sensors on the retina: "rods", to perceive intensity only, and three classes of "cones" (Red, Green and Blue), to perceive colors. R,G,B are also the basis of video displays and by their combinations all perceivable colors can be displayed. Red, Green and Blue are called the "Primary Colors".

Another set of colors, called "Secondary Colors" are Yellow, Magenta and Cyan. These form the basis of displays based on absorption, such as ink on paper. These are also the ink cartridges you buy for your color printer.

Due to a number of reasons, including the need to be compatible with Black and White displays, rather than using R,G,B components for colors it is better to use Intensity (or Luminance) and then two parameters to describe the color information. This two dimensional color description leads to standard "Chromaticity Diagrams" where all the colors are displayed in a plane, which also shows the result of mixing different colors, as you might have seen when you buy paint in the hardware store.

This representation leads to a number of standards such YUV (used in JPEG), YIQ (used in the NTSC TV standard) and YCbCr. In all these standards the "Y" stands for Intensity (or Luminance) and the other two components describe the color. Different standards vary according to which company developed it.

Course Material:

- PPT Slides:
  - Color Representation
    http://faculty.nps.edu/rcristi/EC4480/slides/1-Color-Representation.ppt

- Videos:
  - Human Color Perception 25.28min
    http://faculty.nps.edu/rcristi/EC4480/videos/1-Color-Representation-Seg1.mp4
  - Chromaticity Diagrams 28.31min
    http://faculty.nps.edu/rcristi/EC4480/videos/1-Color-Representation-Seg2.mp4
  - Color Coding 20.19min
    http://faculty.nps.edu/rcristi/EC4480/videos/1-Color-Representation-Seg3.mp4

- Assignment:
  - Project 1 in Sakai (*description*)
    http://faculty.nps.edu/rcristi/EC4480/ComputerAssignments/ComputerAssignment1/EC4480-Assignment-1.docx

  - Project1 video 25.35min
    http://faculty.nps.edu/rcristi/EC4480/ComputerAssignments/ComputerAssignment1/Project1.mp4

EC4480. Weeks2-3: *2D Signals and Systems*

Description.

Signals and Basic Processing. A digital image is a matrix of pixels. In particular a color image is a matrix with three dimensions: Nr X Nc X 3, with Nr, Nc numbers of rows and columns and 3 the color dimension (say R,G,B or any other standard).
It is customary to separate Intensity from Color, so the intensity image (gray scale) is just a regular Nr X Nc matrix, with two indices. This defines a 2D (for "Two Dimensional") signal as a function of two independent variables, as opposed to a 1D signal, function of one independent variable, such as time.

Basic image processing algorithms are also introduced to detect local features, such as edges which, in turn, are used to identify geometric features such as lines.

Fourier Analysis of 2D Signals and Systems. The concept developed for 1D signals and systems (usually function of time) in previous courses, are extended to 2D signals and systems, where the independent variables are now vertical and horizontal spatial coordinates. In particular concepts like the Fourier Transform, Linear Shift Invariant filtering and Convolution are extended to the 2D domain.

Along similar lines as for 1D time signals, the Fourier Transform is used to identify features of the images with the purpose of classification and/or identification of disturbances. In addition, standard 1D techniques to design filters can be extended easily to the design of 2D filters by a number of 1D to 2D transformations, such as the McClellan Transformation.

Course Material:

- PPT Slides:
  - 2D Signals and Systems
    http://faculty.nps.edu/rcristi/EC4480/slides/2a-2D-Signals-and-Systems.ppt
  - 2D Fourier Analysis
    http://faculty.nps.edu/rcristi/EC4480/slides/2b-Fourier-Analysis-in-2D.ppt
- Videos:
  - Images as 2D Signals 33.44min
    http://faculty.nps.edu/rcristi/EC4480/videos/2a-2D-Signals-Systems-Seg1.mp4
  - Simple 2D Filters 30.20min
    http://faculty.nps.edu/rcristi/EC4480/videos/2a-2D-Signals-Systems-Seg2.mp4
  - 2D Fourier Transform 46.31min
    http://faculty.nps.edu/rcristi/EC4480/videos/2b-Fourier-Analysis-2D-seg1.mp4
  - 2D Sampling Theorem 16.27min
    http://faculty.nps.edu/rcristi/EC4480/videos/2b-Fourier-Analysis-2D-seg2.mp4

- 2D Frequency Response 20.56min
  http://faculty.nps.edu/rcristi/EC4480/videos/2b-Fourier-Analysis-2D-seg3.mp4
- 2D Discrete Fourier Transform (DFT) 24.28min
  http://faculty.nps.edu/rcristi/EC4480/videos/2b-Fourier-Analysis-2D-seg4.mp4
- 2D Filters 21.27min
  http://faculty.nps.edu/rcristi/EC4480/videos/2b-Fourier-Analysis-2D-seg5.mp4

- Assignment:
  - Project 2 in Sakai (*description*)
    http://faculty.nps.edu/rcristi/EC4480/ComputerAssignments/ComputerAssignment2/Project2.docx
  - Project 2 Video 44.22min
    http://faculty.nps.edu/rcristi/EC4480/ComputerAssignments/ComputerAssignment2/Project2.mp4
  - Project 2 Video 2
    http://faculty.nps.edu/rcristi/EC4480/ComputerAssignments/ComputerAssignment2/Project2-Filter.mp4

EC4480. Weeks4-5: *Image Compression by the Discrete Cosine Transform (DCT)*

Efficient Signal and Image Representation.. We see image (and also video) compression every day. Whenever we download an image from our digital camera or from the internet and we see it on our screen, we are still amazed at the excellent quality of the result. But if we look even further on the size of the file and we do simple back-of-the-envelope computations, we see that the size of the file is order of magnitudes smaller than what we expect by the 3 bytes per pixel needed to encode luminance and color. This efficiency is due to the application of basic signal representation techniques such as the Discrete Karhunen Loeve Transform (DKLT), and the Discrete Cosine Transform (the DCT, a "cousin" of the FFT).

Entropy as a Measure of Signal Information. The general problem of quantifying how much "information" is in a signal (whether it is a communication signal, or audio or an image) is addressed by using the concept of "entropy". This concept is introduced from basic concepts and well known formulas of "combinatorics". It yields a surprisingly simple expression of the information content of any source, in terms of "average number of bits per symbol", which can be used to assess how much a signal can be compressed without noticeable losses. A commonly used code (the Huffman code) for efficient representation is introduced.

Course Material:

- PPT Slides:
  - DKLT and DCT
    http://faculty.nps.edu/rcristi/EC4480/slides/3-DKLT-DCT.ppt
  - Entropy
    http://faculty.nps.edu/rcristi/EC4480/slides/4-Entropy.ppt
- Videos:
  - Discrete Karhunen-Loeve Transform (DKLT) 41.24min
    http://faculty.nps.edu/rcristi/EC4480/videos/3-DKLT-DCT-seg1.mp4
  - DKLT for 2D Signals 33.37min
    http://faculty.nps.edu/rcristi/EC4480/videos/3-DKLT-DCT-seg2.mp4
  - Discrete Cosine Transform (DCT) 17.08min
    http://faculty.nps.edu/rcristi/EC4480/videos/3-DKLT-DCT-seg3.mp4
  - Transmission of Information 16.21min
    http://faculty.nps.edu/rcristi/EC4480/videos/4-Entropy-seg1.mp4
  - Entropy of a Source of Information 30.32min
    http://faculty.nps.edu/rcristi/EC4480/videos/4-Entropy-seg2.mp4
  - Entropy Coding: Huffman Code 40.49min
    http://faculty.nps.edu/rcristi/EC4480/videos/4-Entropy-seg3.mp4

- Assignment:
  - Project 3 in Sakai (*description*)
    http://faculty.nps.edu/rcristi/EC4480/ComputerAssignments/ComputerAssignment3/Project3.docx
  - Project 3 Video 1 33.55min
    http://faculty.nps.edu/rcristi/EC4480/ComputerAssignments/ComputerAssignment3/Project3-seg1.mp4
  - Project 3 Video 2 41.58min
    http://faculty.nps.edu/rcristi/EC4480/ComputerAssignments/ComputerAssignment3/Project3-seg2.mp4

EC4480. Weeks6-7: *Multi Resolution Decomposition*

Description.

Multi Resolution Signal Decomposition. An alternative representation of signals (both 1D and 2D) is based on a Multiresolution Decomposition. By this approach an image is separated into low frequency approximations and high frequency details. This decomposition is obtained using a combination of multirate low pass and high pass filters, in a configuration called Quadrature Mirror Filters.

Quadrature Mirror Filters. It turns out that the Quadrature Mirror Filters (QMF) structure presented yields a very sparse decomposition of the signal, both in the 1D as well as in the 2D domains, with low entropy. Families of FIR filters for QMF design can be derived from basic principles, which can be effectively implemented. In particular two families of filters are derived: Daubechies and Biorthogonal.

Apart from being efficient in terms of data compression, QMF decomposition also lends itself to an algorithm which can be applied effectively to nonlinear filtering of disturbances while still preserving high frequency details of the image itself.

Course Material:

- PPT Slides:
  - Multi Resolution Decomposition 1
    http://faculty.nps.edu/rcristi/EC4480/slides/5-MultiResolution-1.ppt
  - Multi Resolution Decomposition 2
    http://faculty.nps.edu/rcristi/EC4480/slides/5-MultiResolution-2.ppt
- Videos:
  - General Approach to Multi Resolution Decomposition 23.30min
    http://faculty.nps.edu/rcristi/EC4480/videos/5-Multiresolution-1-seg1.mp4
  - Quadrature Mirror Filters: Ideal Case 14.24min
    http://faculty.nps.edu/rcristi/EC4480/videos/5-Multiresolution-1-seg2.mp4
  - Simple FIR Filter Example 44.14min
    http://faculty.nps.edu/rcristi/EC4480/videos/5-Multiresolution-1-seg3.mp4
  - Quadrature Mirror Filters with FIR Filters 27.10min
    http://faculty.nps.edu/rcristi/EC4480/videos/5-Multiresolution-2-seg1.mp4
  - Wavelet Decomposition of Images 47.46min
    http://faculty.nps.edu/rcristi/EC4480/videos/5-Multiresolution-2-seg2.mp4
  - Matlab Wavelet Tool 17.37min
    http://faculty.nps.edu/rcristi/EC4480/videos/5-Multiresolution-2-seg3.mp4

- Assignment:
  - Project 4 in Sakai (*description*)
    http://faculty.nps.edu/rcristi/EC4480/ComputerAssignments/ComputerAssignment4/Project4_AY16.doc

EC4480. Week8: ***Motion Estimation in Video Images***

Description.

A very important problem in processing of video signals is to detect moving objects, and estimate the motion vector. This is important not only in target tracking, but also in video compression, where the stationary background is treated separately from the moving foreground.

We introduce two important techniques in motion estimation: the Optical Flow, which is computed over the entire frames, and the Kanade Lucas Tomasi (KLT) technique which tracks only points of interests in the video images, such as corners.

Course Material.

- PPT Slides:
  - Optical Flow
    http://faculty.nps.edu/rcristi/EC4480/ slides/6-Optical-Flow.ppt
  - Morphological Operations
    http://faculty.nps.edu/rcristi/EC4480/slides/7-morphological.ppt
  - Kanade Lucas Tomasi (KLT)
    http://faculty.nps.edu/rcristi/EC4480/slides/8-KLT-Tracking.ppt
- Videos:
  - Motion Detection and Optical Flow 20.45min
    http://faculty.nps.edu/rcristi/EC4480/videos/6-OpticalFlow-seg1.mp4
  - Optical Flow in Simulink 16.32min
    http://faculty.nps.edu/rcristi/EC4480/videos/6-OpticalFlow-seg2.mp4
  - Optical Flow in Simulink 2 28.41min
    http://faculty.nps.edu/rcristi/EC4480/videos/7b-OpticalFlowSimulink.mp4
  - Morphological Operations 36.54min
    http://faculty.nps.edu/rcristi/EC4480/videos/7-Morphological.mp4
  - Detect Points of Interest by KLT 40.46min
    http://faculty.nps.edu/rcristi/EC4480/videos/8-KLT.mp4
- Assignment:
  - Project 5 in Sakai (*description*)
    http://faculty.nps.edu/rcristi/EC4480/ComputerAssignments/ComputerAssignment5/Project5.docx

- Project 5 Video 1: Optical Flow 10.20min
  http://faculty.nps.edu/rcristi/EC4480/ComputerAssignments/ComputerAssignment5/Project5-problem1.mp4
- Project 5 Video 2: KLT 27.59min
  http://faculty.nps.edu/rcristi/EC4480/ComputerAssignments/ComputerAssignment5/Project5-problem2.mp4
- Project 5 Video 3: track a car 14.13min
  http://faculty.nps.edu/rcristi/EC4480/ComputerAssignments/ComputerAssignment5/Project5-problem3.mp4

EC4480. Week9: *Video Surveillance*

Description.

An alternative way of detecting moving objects is by estimating the background from a number of video frames and then subtracting it from the videos themselves. This is based on the a-priori assumption that the background does not change between frames.

This problem can be extended also to video surveillance and detection of "abnormal behavior". For example, if you have a video of (say) the lobby of an airport, you see a lot of people moving or standing for a short amount of time, over a static background. However if someone leaves a piece of luggage for an amount of time, this luggage becomes part of the background which then changes. This change shows an "anomalous behavior" of the background which is clearly expected and would trigger a security alarm.

This can be extended to any situation in which, under normal circumstances, the intensities of the pixels exhibit given statistics (say a certain level of traffic on a street) and we want to detect "abnormal" cases, such increase or decrease in traffic.

Course Material:

- PPT Slides:
  - Background Estimation
    http://faculty.nps.edu/rcristi/EC4480/slides/9-background.ppt
  - Video Surveillance
    http://faculty.nps.edu/rcristi/EC4480/slides/12-Video-Surveillance.pptx
- Videos:
  - Background Estimation 14.12min
    http://faculty.nps.edu/rcristi/EC4480/videos/10-BackgroundEstimation.mp4
  - Video Surveillance 1 39.47min
    http://faculty.nps.edu/rcristi/EC4480/videos/11-VideoSurveillance-1.mp4
  - Video Surveillance 2 15.30min
    http://faculty.nps.edu/rcristi/EC4480/videos/11-VideoSurveillance-2.mp4
- Assignment:
  - Assignment 6 in Sakai (*description*)

http://faculty.nps.edu/rcristi/EC4480/ComputerAssignments/ComputerAssignment6/Project6.docx

- Project 6 Video 14.33min

  http://faculty.nps.edu/rcristi/EC4480/ComputerAssignments/ComputerAssignment6/Project6.mp4

EC4480. Weeks10-11: *Multiscale Image Representation*

Description.

A fundamental problem in Computer Vision is to recognize various objects in a given scene. The challenge is that the same object in two different images can be rotated, be further or closer to the viewer, have occlusions or different illumination. Although the human brain seems to overcome these difficulties in a very efficient way, for a computer this becomes a very challenging problem.

One of the approaches, which seem to mimic some of the workings of our vision system, is the Multiscale approach. This is similar to the Mulitiresolution representation we saw earlier in the course, but the goals are different. By the Multiscale approach a gray scale image is represented with an added dimension (thus becoming three dimensional). The added dimension is a scaling factor, which accounts for the fact that an image can be viewed at different scales (small or large as if it were viewed from further away or close up). In this way, Points of Interest (PoI) are identified by local maxima of gradients in the two dimensions (x and y) and also the third dimension (scale). Then, for every PoI a vector of features is computed, which varies according to different algorithms.

This is at the basis of the SIFT and SURF algorithms, well known techniques in Computer Vision.

Course Material:

- PPT Slides:
  - Multiscale Representation of Images
    http://faculty.nps.edu/rcristi/EC4480/slides/10-MultiScale-Image-Representation-1.pptx
  - SURF and SIFT algorithms
    http://faculty.nps.edu/rcristi/EC4480/slides/11-SIFT-SURF-algorithms.pptx
- Videos:
  - Introduction to Multi Scale Representation 23.25min
    http://faculty.nps.edu/rcristi/EC4480/videos/9-MultiScaleRepresentation-seg1.mp4
  - Gaussian Filtering and Laplacian-1 25.22min
    http://faculty.nps.edu/rcristi/EC4480/videos/9-MultiScaleRepresentation-seg2-1.mp4
  - Gaussian Filtering and Laplacian-2 27.44min
    http://faculty.nps.edu/rcristi/EC4480/videos/9-MultiScaleRepresentation-seg2-2.mp4

- SIFT and SURF algorithms 46.45min
  http://faculty.nps.edu/rcristi/EC4480/videos/9-MultiScaleRepresentation-seg3.mp4
- RANSAC algorithm 42.40min
  http://faculty.nps.edu/rcristi/EC4480/videos/9-MultiScaleRepresentation-seg4.mp4

- Assignment:
  - Assignment 7 in Sakai (*description*, *PPT slides*)
    http://faculty.nps.edu/rcristi/EC44800/ComputerAssignments/ComputerAssignment7/Project_7_description-new.docx
    http://faculty.nps.edu/rcristi/EC44800/ComputerAssignments/ComputerAssignment7/Project7Notes.pptx
  - Project 7 Video 16.16min
    http://faculty.nps.edu/rcristi/EC4480/ComputerAssignments/ComputerAssignment7/Project7SURF.mp4