Digital Image Processing
(a modern approach)
(DIPAMA)

Zhu Li
Dept of CSEE, UMKC
Office: FH560E, Email: lizhu@umkc.edu, Ph: x 2346.
http://l.web.umkc.edu/lizhu
Outline

- Background
- Objective of the class
- Prerequisite
- Lecture Plan
- Course Project
- Q&A
An image is worth a thousand words….

- What we observe are pixels….

- The story:
  - The train wreck at La Gare Montparnasse, 1895

- What computer can do these days:
  - Figure out the building
  - The train
  - People walking around

- Still long way to go to figure out the semantics
  - Train crashes
  - It is an abnormal event (context)
Short Bio:

Research Interests:

- Immersive visual communication: light field, point cloud rendering, and low latency streaming
- What DL can do for compression
- What compression can do for DL
- Object Re-Identification, Content De-Duplication and ICN/CCN
- Signal processing and learning
- Image understanding
- Visual communication
- Mobile edge computing & communication
UMKC Media Computing & Communication Lab@FH262

- **Post-Doc**
  - **Li Li**, PhD USTC, Light Field, Point Cloud, 360 Video processing and compression
  - **Renlong Hang**, PhD NUIST, Deep learning in remote sensing problems.

- **PhD Students**:
  - **Zhaobin Zhang**: Deep learning in compression, Grassmann methods in transform optimization (now intern at Tencent Media Lab)
  - **Biren Kathariya**: Point cloud compression, deep learning model compression (now intern at Huawei Media Lab)
  - **Anique Akhatar**: Point cloud segmentation and classification, mobile edge computing for 3D map and auto-driving services (now intern at HERE)
  - **Yangfan Sun**: machine learning in video coding and rate-distortion optimization (now intern at Huawei Central Hardware Research Lab)
  - **Dewan Noor**: Super-resolution in biometrics.
  - **Raghunath Puttagunta**: Object recognition with deep learning and handcrafted features, aggregation schemes.
  - **Wei Jia**, PhD, Deep feature map and model compression, point cloud attributes coding, graph signal processing.
  - **Yue Li**, visiting PhD from USTC, deep learning in compression.
  - **Wenjie Zhu**, visiting PhD from SJTU, point cloud compression.

- **MS Student(s)**
  - **Paras Maharjan**: LSTM in genome data compression and point cloud entropy coding
DIP Related Research

2018
Media Computing & Communication Horizon

**Devices**

- Smart TV
- Mobile devices
- Wearable devices

**Networks**

- mmWave 5G
  - Frequency band
  - 4G frequencies
  - New higher frequencies
- FD-MIMO
  - Large scale Multi-antenna

**Applications**

- 4k/8k UHD Video
- Free Viewpoint TV
- Samsung VR/AR
- BIGDATA visual intelligence
Advances in Image Sensors: pixels and voxels

- **Hyperspectral Image Sensor**
  - $I(x,y)$ in $\mathbb{R}^D$, $D=48$, e.g.

- **3D/Depth Sensor: LiDAR, Stereo Capture**
  - $I(x,y,z)$ in $\mathbb{R}$

- **Panoramic Video Cameras**
  - $I(\alpha, \beta)$, $\alpha, \beta$ in $[0, 2\pi]$

- **Lightfield Capture**
  - Lenslet images
Embedded Deep Learning for Image Understanding

- **Automatic Image Tagging:**
  Mapping image pixel values to tags
- **Device based recognition:**
  Distilling the knowledge into a compact form
- **Compact CNN model for knowledge distilling**
  BIGDATA training set: >10k images per tag
  training set augmentation:
  - affine transforms, simulated lighting changes
  - Pruning filters from the CNN structure

NSF I/UCRC Center for Big Learning
Creating Intelligence

Z. Li, Digital Image Processing, Fall 2018.
Deep Learning in Video Compression

Deep Learning Chroma prediction (from Luma)

Results:
- BD rate:

<table>
<thead>
<tr>
<th></th>
<th>Y</th>
<th>U</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.2%</td>
<td>-5.7%</td>
<td>-0.6%</td>
</tr>
<tr>
<td></td>
<td>-0.5%</td>
<td>-4.3%</td>
<td>-2.1%</td>
</tr>
<tr>
<td></td>
<td>-0.2%</td>
<td>-2.8%</td>
<td>-1.0%</td>
</tr>
<tr>
<td></td>
<td>0.1%</td>
<td>-1.5%</td>
<td>-3.5%</td>
</tr>
<tr>
<td></td>
<td>0.1%</td>
<td>-0.1%</td>
<td>-3.1%</td>
</tr>
<tr>
<td></td>
<td>-0.2%</td>
<td>-3.1%</td>
<td>-2.0%</td>
</tr>
</tbody>
</table>
Subspace Indexing on Grassmann Manifold

- **Large Scale Face Identification/Image Understanding**
  - Improve the model Degree of Freedom (DoF) to accommodate the variations in the large training data set and large label space
  - A piece-wise linear local approximation of the underlying non-linearity, balance the DoF with the amount of discriminant information captured
  - Attempt to explain CNN as cascade of piece-wise linear projections.

- **Subspace Indexing on Grassmann Manifold**
  Develop a rich set of transforms that better captures local data characteristics, and
  Develop a hierarchical deep structure for subspaces on the Grassmann manifold.
  Applications: large subject set face recognition, speaker ID, and hierarchical transforms for image coding.
Mobile Visual Search: Object Re-Identification

Mobile Query-by-Capture

Object Recognition Pipeline
– KD: Keypoint Detection (Box Filtering)
– FS: Feature Selection (Affine transformed 2 way-matching)
– GD/LD: global and local descriptor generation (AKULA, Collision Optimized Hash)

Descriptor Extraction

Descriptor Encoding

Recall

Precision
Immersive Visual Communication

- Light Field Compression (6-DoF)
  - Sub-Aperture Image Based
  - Tensorial Display Decomposition Based

- Point Cloud Compression (6-DoF)
  - Video Based: Views + Depth
  - Binary Tree/Octree Geometry Coding + Graph Signal Compression

- 360 Video Compression (3-DoF)
  - Advanced Motion Model: Affine motion, Spherical motion models
  - Padding
Depth Sensing/SLAM

Key problems for auto driving cars
- Depth from Stereo Images
- Optical Flow
- Scene Flow
- 2D/3D data fusion and registration
- Image/3D features for SLAM
- Point Cloud Segmentation & Registration
Outline

- Background
- Objective of the class
- Prerequisite
- Lecture Plan
- Course Project
- Q&A
DIPama Objectives

- Understand basic image formation process, and its implications in pixel geometry and attributes
- Hands on experiences with point based operations, 2D convolution, frequency domain filtering, non-linear and deep convolution tools
- Applications:
  - Image Segmentation
  - Image Super Resolution
  - Image Enhancement (low light, denoise, ..., etc)
  - Basics of Image Compression
- Prepare students for more advanced topics in computer vision and video compression:
  - ECE 5578 Multimedia Communication:
  - ECE 5582 Image Analysis & Retrieval (Computer Vision)
Prerequisite & Textbook

- **Prerequisite**
  - For senior and graduate students in EE/CS
  - Taken Signal & System, or Digital Signal Processing or consent of the instructor
  - Programming experiences with Matlab
  - Will have different expectation for MS/PhD and undergrad students

- **Textbook:**
  - None required (saving $$) , will distribute relevant chapters, papers, and notes.

- **Key References:**
# Tentative Lecture Plan

<table>
<thead>
<tr>
<th>1. Image Formation</th>
<th>HW-1:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Geometry</td>
<td>color histogram</td>
</tr>
<tr>
<td>2. Color</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2. Image Sampling and Quantization</th>
<th>HW-2:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Sampling and aliasing</td>
<td>image sampling and quantization</td>
</tr>
<tr>
<td>2. Quantization and quantization error</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3. Image Filtering</th>
<th>HW-3:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Point based operations</td>
<td>image filtering</td>
</tr>
<tr>
<td>2. Linear Filtering</td>
<td></td>
</tr>
<tr>
<td>3. Non-Linear Filtering (Bilateral, Median)</td>
<td></td>
</tr>
<tr>
<td>4. Deep convolutional networks</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4. Applications</th>
<th>HW-4:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Super resolution</td>
<td>non-linear image filtering</td>
</tr>
<tr>
<td>2. Segmentation</td>
<td></td>
</tr>
<tr>
<td>3. Deep learning classification</td>
<td></td>
</tr>
<tr>
<td>4. Compression</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>HW-5:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>deep convolution networks</td>
</tr>
</tbody>
</table>

**Project:**
Choose from SR, Segmentation, Classification and Compression
Potential Course/MS thesis Project

- Resources from last year:

- Potential projects with 25% bonus points
  - Boosting based key point feature detection via box filtering for Hyperspectral images
  - Compact deep learning models for embedded vision
  - KTTI image + 3D point cloud benchmark
  - Image registration with point cloud
  - Rate-agnostic hash for video de-duplication in cache and networks.
Grading

- **Homeworks (50%)**
  - Color Histogram
  - Sampling & Quantization
  - Convolution & Freq Domain Filtering
  - Non-Linear Filters
  - Deep convolutional networks

- **2 Quizzes (20%)**: relax, quiz is actually on me, to see where you guys stand
  - Quiz-1: Sections 1.1 thru 3.3
  - Quiz-2: The remaining

- **Project (30%)**
  - Original work leads to publication, discuss with me by the mid of October. (15% bonus point)
  - Regular project: assign papers to read, implement certain aspect, and do a presentation.
Course Outcome

Upon completion of the DIPAMA course you will be able to:

- Understand the basic operations in image formation and its implications on pixel geometry and attributes
- Understand 2D signal sampling and quantization and its consequences in image quality
- Be able to perform point based operations on images
- Be able to perform 2D pixel domain and frequency domain filtering.
- Understand non-linear filtering on images
- Understand and perform deep convolutions on images
- Understand basic principles and issues in image segmentation, super resolution, classification, and compression.
- Can apply the knowledge an algorithms to solve real world image processing problems
- Well prepared for conducting advanced research and pursing career/PhD in this topic area. (PhD qualify required course)
Logistics

- **Office Hour:**
  - Tue/Thr 2:30-4:00pm, 560E FH
  - Or by appointment

- **TA:**
  - TBD
  - Lab Sessions are planned to cover certain software tools aspects.
  - Office Hour: TBA

- **Course Resources:**
  - Will share a box.com folder with slides, references, data set, and software
  - Additional reference, software, and data set will be announced.
Q&A