Computer Vision in MATLAB and Simulink

Steve Kuznicki - Pilot Engineering
IPCV - Image Processing Computer Vision
A look at new features in 14b/15a Releases

- Machine learning
  - Object recognition
  - Classification
  - Image retrieval

- Stereo vision
  - Depth estimation
  - 3D scene reconstruction

- 3D Point cloud processing
What is Image Processing?

Algorithms that enhance image quality or extract information from images

Measuring brain activity in MRI scans
What is Computer Vision?

Using images and video to detect, classify, and track objects or events in order to “understand” a real-world scene.
Object Detection

- **HOG feature extraction (13b)**
  - Useful to train pedestrian detection
  - New demo showing how to train a HOG-SVM detector

- **FREAK, BRISK (13b, 14a)**
  - Useful for feature-based object detection

- **Training Image Labeling App**
  - Simplifies training workflow (14a)
Machine Learning – What Problems Can You Solve?

- Clustering or Segmentation
- Object Recognition or Classification
- Object Detection
Example 1: Cluster or Group Similar Objects
Example 2: Clustering Images

http://www.robots.ox.ac.uk/~vgg/data/flowers/17/17flowers.tgz
Example: Recognize / Classify Objects

Known as **object classification or recognition**
Machine Learning

Machine learning uses **data** and produces a **program** to perform a **task**

**Task:** Image Category Recognition

If $\text{brightness} > 0.5$

then ‘hat’

If $\text{edge\_density} < 4$ and

$\text{major\_axis} > 5$

then “boat”

...
What is a Classifier?

Machine Learning

Training Data
Features

Classifier

Machine Learning

'hats'
'mugs'
'boats'

Classifier
Supervised Vs. Unsupervised Learning

- **Supervised learning**
  - Have *labeled* data (a.k.a. ground truth data)
  - Goal: make a prediction (categorical or numerical) given a new instance of data.

- **Unsupervised learning**
  - Have only data (no labels)
  - Goal: Find interesting patterns (connections) in the data

Data from MNIST digit dataset

Label = “2”  
Label = “1”
Machine Learning Workflow Using Images

- Training Data
- Feature Extraction
- Learning or Modelling
- Training
- Input Image
- Feature Extraction
- Classification
- ‘hat’
Challenges: Machine Learning Workflow Using Images

1. Training Data
2. Feature Extraction
3. Learning or Modelling

Challenge 1: Input Image

Challenge 2: Feature Extraction

Challenge 3: Classifier / Model

‘hat’
Common Challenges for Machine Learning with Images

- **Challenge 1:** Handling large sets of images
- **Challenge 2:** How to extract discriminative information from images
- **Challenge 3:** How to model problem using machine learning techniques
Managing Large Image Sets

- Acquiring images of your objects
  - Use publicly available image sets
  - Use image acquisition toolbox

- Labeling images
  - Put similar images into a named folder

- Using images based on the object label
  - Pull into MATLAB using the imageSet object

```matlab
imgSet = imageSet(imageLocation)

imgSetVector = imageSet(imgFolder,'recursive')
```
What is Feature Extraction?

Feature Extraction
- Representations often invariant to changes in scale, rotation, illumination
- More compact than storing pixel data
- Feature selection based on nature of problem

Sparse

Dense

Bag of Words

SURF

HOG

Image Pixels
Feature Extraction

- Choose the right features for the data set
  - Lots of edge information? Try HOG
  - Color images, can you leverage color info?

- Specific objects of interest,
  - use local features detectors/extractors

- Need to add spatial dependencies?
  - Append X,Y location to local features (Spatial augmentation)
  - Spatial pyramiding (beware – increases feature dimension)

- Whole scene important? use global image features
  - Dense SURF, HOG, color histograms
Bag of Words (Form of Features)

How do you tell if the content of these two books is similar?
Image Processing Toolbox

Perform image processing, analysis, and algorithm development

Image Processing Toolbox™ provides a comprehensive set of reference-standard algorithms, functions, and apps for image processing, analysis, visualization, and algorithm development. You can perform image analysis, image segmentation, image enhancement, noise reduction, geometric transformations, and image registration. Many toolbox functions support multicore processors, GPUs, and C-code generation.

Image Processing Toolbox supports a diverse set of image types, including high dynamic range, gigapixel resolution, embedded ICC profile, and tomographic. Visualization functions and apps let you explore images and videos, examine a region of pixels, adjust color and contrast, create contours or histograms, and manipulate regions of interest (ROIs). The toolbox supports workflows for processing, displaying, and navigating large images.
Bag of “Visual Words” (features)

Class / Label

Training Data

Vocabulary / Bag of Words

'mugs'
Image Classification with Bag of Words

Training Data

Bag = Visual Vocabulary

Input Image

Classifier

‘hat’
Machine Learning – a demo!
Differentiate between different Toy Vehicles

% Real-time Car Identification Using Image Data
% Copyright (c) 2015, MathWorks, Inc.

% Clear Workspace
clear all;

% Load image data
imset = imageSet('CarData', 'recursive');

% Display Class Names
imset.Description

% Display Sampling of Image Data
montage(imset(3).ImageLocation(1:5:end));

% Pre-process Training Data: *Feature Extraction using Bag Of Words*
% Create Visual Vocabulary
bag = bagOfFeatures(imset, 'VocabularySize', 250, 'PointSelection', 'Detector');
cordata = double(encode(bag, imset));

% Visualize Feature Vectors
img = read(imset(1), rand(imset.1.Count));
featureVector = encode(bag, img);
figure;
Using Machine Learning Functionality in MATLAB

```matlab
model = fit[c/r][model type] (X, Y)  

label = predict(model, input)
```

**Command Window**

```
>> knnClassifier = fitcknn(featureDataset,Y)
```

**Command Window**

```
>> bestGuess = predict(knnClassifier, featureDataset(1,:))
```
Many Options for Features and Machine Learning

Feature Extraction
- BRISK, FREAK, SURF
- Histogram of Oriented Gradients (HoG)
- Using box filters
  - integral images
- Bag of visual words
- Color-based features
- Frequency-domain features

Machine Learning
- SVM
- Decision trees
- AdaBoost
- Bagged trees
- k-NN
- Discriminant analysis
- Bayes classifiers

Bottom Line: Many permutations and combinations to fit the needs of your problem
Common Challenges for Machine Learning with Images

- **Challenge 1**: Handling large sets of images

- **Challenge 2**: How to extract discriminative information from images

- **Challenge 3**: How to model problem using machine learning techniques

- Easy to handle large sets of images
  - `imageSet`

- Bag of words for feature extraction
  - More available in Computer Vision System Toolbox
Examples of Object Recognition/Classification

- Automatic scene categorization
- Biometrics
  - Face recognition
  - IRIS recognition
  - Fingerprint recognition
- Part recognition for factory automation
- Autonomous robotics
Other Applications of Same Workflow

- Detection and Location
  - Is a coffee mug present?
  - Where is it?

- Anomaly Detection/Verification

‘yes’

‘no’

‘pass’

‘fail’
Recommended Next Steps

- Look into Parallel Computing Toolbox and MATLAB Distributed Computing Server
- Try Neural Network Toolbox
- Explore examples in Computer Vision System Toolbox
Stereo Vision (14a)

- Stereo calibration
- Semi-global disparity matching
  - Better results than block matching
- 3D scene reconstruction from disparity
- New demos available
  - Estimate real distance to objects in video
  - 3D scene reconstruction
  - Pedestrian detection (14b)
Recovering Scene Depth with Stereo Cameras
Epipolar Geometry

Left view

Right view
Fundamental Matrix

$X_L^T F X_R = 0$
Stereo Camera Calibration

- Simplifies and automates calibration process
Stereo Vision

- Process of extracting 3-D information from a pair of synchronized cameras
- Applications
  - ADAS
  - Autonomous Robotics
  - 3-D Mapping and Visualization
  - Gesture Recognition
Stereo Vision Workflow

- Calibration - App (14b)
- Rectification - Codegen (15a)
- Disparity Estimation - Block matching, semi-global matching (14b) - Codegen (14b)
- 3-D Reconstruction - Codegen (15a)
Point Clouds

- In computer vision and in computer graphics, point clouds are typically used to measure physical world surfaces.

- Point clouds can also represent other types of 3-D data:
  - Financial industry time series data (2-D points plus time).
Point Cloud Application – Robot Vision

- Robot Navigation

- Robot Perception
Point Cloud Application – Advanced Driver Assistance Systems (ADAS)

- Collision Detection

- Visual SLAM (Simultaneous localization and mapping) / Visual Odometry
Point Cloud Application – 3-D Vision

- Traditional 2-D CV applications extend to 3-D, i.e. Object Detection
- 3D Reconstruction / Digital Heritage

- Volumetric Data Processing
Point Cloud Application – Auto CAD

- CAD Re-engineering
- Industrial Inspection
Point Cloud Application – Remote Sensing

- Visualization

- Multi-modality registration
Real World Point Cloud Data

- What is a “point”?  
  - Coordinates, colors, normal vectors, high level descriptors…
- What is a “point cloud”?  
  - Collection of points potentially in the order of millions
- Organized vs. Unorganized point cloud  
  - Neighborhood considerations
- Measurement Noise, Sensor Calibration
Point Cloud Processing

New in R2015a

- I/O
  - File I/O
  - Acquire point clouds from Kinect

- Processing
  - Denoising
  - Downsampling
  - Merging
  - Rigid registration

- Visualization
  - Point cloud viewer
Point Cloud Registration

- **Rigid registration**
  - *pcregigid*: Fundamental operation across point cloud applications
  - ‘Iterative Closest Point’ Algorithm
  - Comparable to state-of-the-art c++ package on academic benchmarks
  - 3-D Point Cloud Registration and Stitching Featured Example

![Diagram of Point Cloud Registration Workflow]
Why Use MATLAB for Machine Learning?

- App driven workflow
  - Spend more time solving problems, less time coding

- Easily manage large sets of images

- Domain specific toolboxes
  - Image processing, signal processing, financial data, RADAR etc.

- Extensive documentation and examples to help you get started
Why Use MATLAB for Stereo Vision/Point Cloud Processing?

- **Stereo Calibration App**
  - Generate calibration data with ease and apply

- **Design applications requiring distance calculations**
  - SLAM: Simultaneous localization and mapping
  - SfM: Structure from Motion
  - Develop 3-D models of objects
  - Build 3-D world maps, Multi-Modal Registration

- And also again: more extensive documentation and examples to help you get started
Thank You!

Questions?