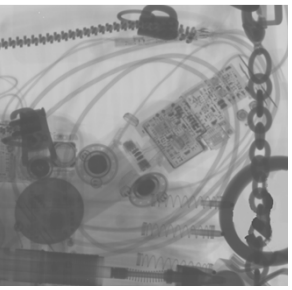
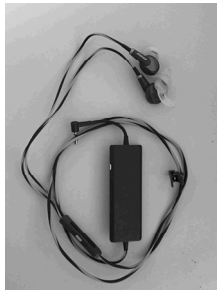
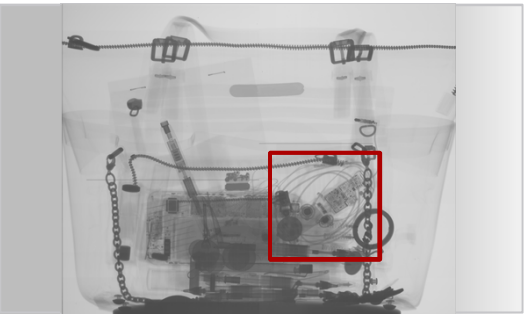
SAND2015-7124C

X-ray radiographs and photo acquired at Sandia National Laboratories

Exploring the Feasibility of Traditional Image Querying Tasks for Industrial Radiographs

Iliana Bray, Stephany Tsai, and Edward Jimenez
Sandia National Laboratories – Software Systems R&D
SPIE Optics and Photonics
August 14th, 2015

Sandia National Laboratories

Exceptional service in the national interest

U.S. DEPARTMENT OF ENERGY

NNSA National Nuclear Security Administration

INFORMATION SOLUTIONS

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. SAND NO. 2011-XXXXP

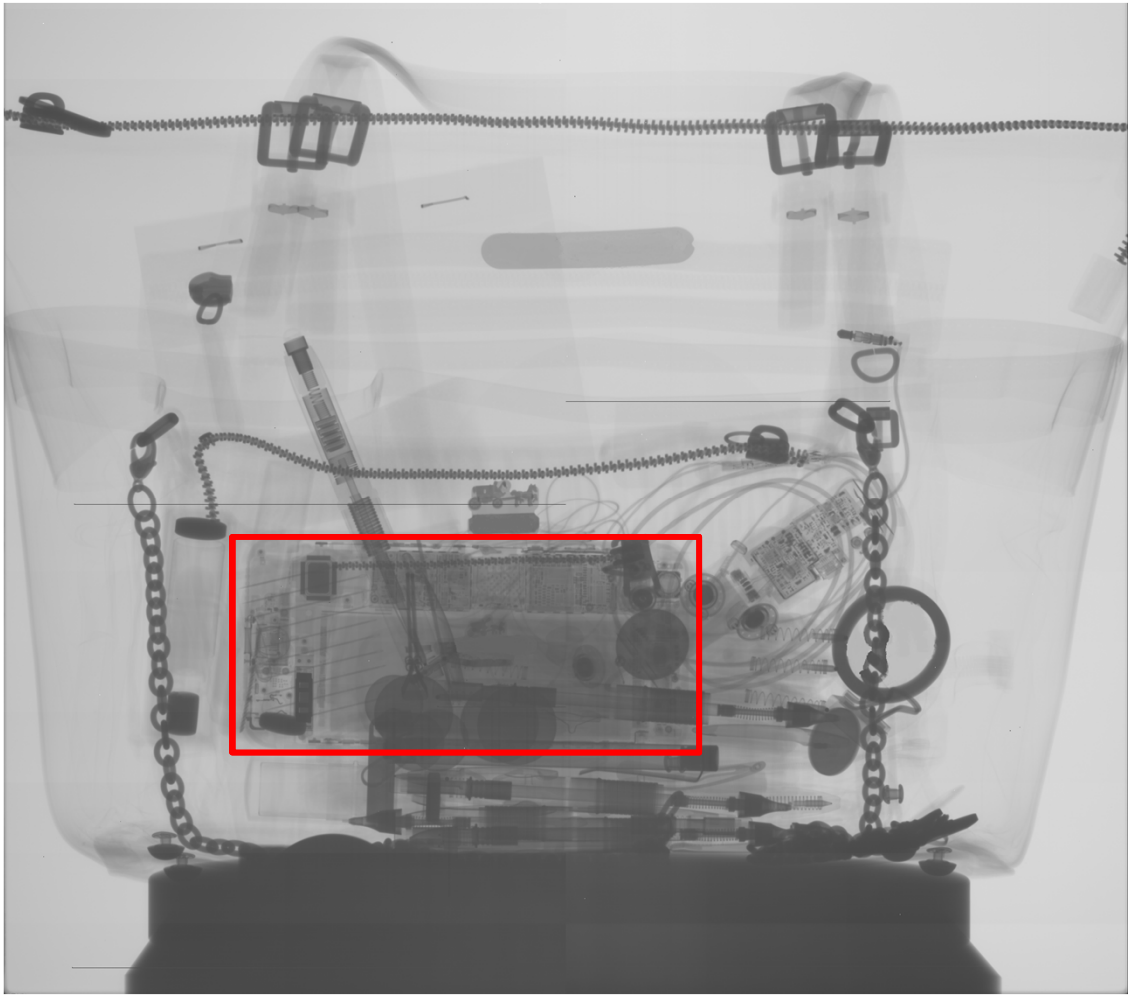
Introduction

- Objective:

  Identify objects in a region of interest using an existing database of X-ray radiographs

- Status:
  - Created clean exemplar images from database
  - Developed pre-processing algorithm for radiographs
  - Developing sorting and matching algorithm

- Importance:
  - Inspection applications
  - National Security



X-ray radiograph acquired at Sandia National Laboratories

Database Characteristics

Sandia National Laboratories

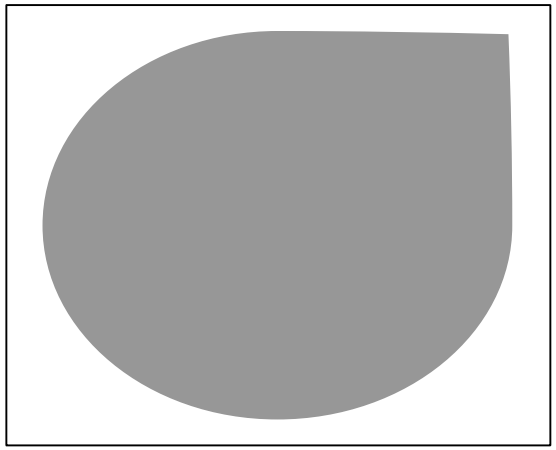
- Compiled for a previous effort
- Over 50,000 images (80.1 GB)
- Combination of photographs and radiographs
- JPEG compression
- Duplicate copies of images
- Extraneous objects in some images
- Limited perspectives of each object (1-7 images/object)

Cleaning:

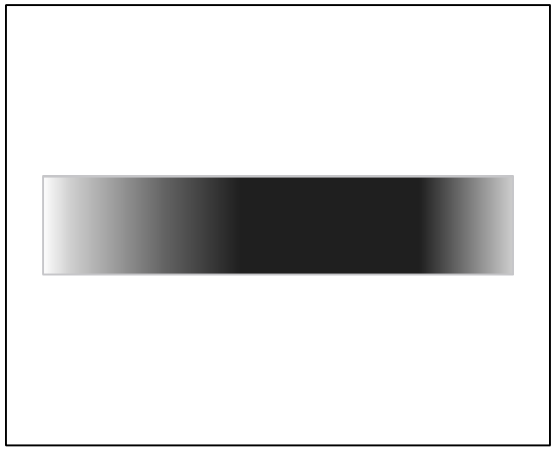
1. Delete duplicate images
2. Remove photographs from the database (leaving only radiographs)
   - Based on pixel values
3. Crop radiographs to only contain object of interest
   - Targeted common appearance of extraneous objects

Example of Cleaned Images

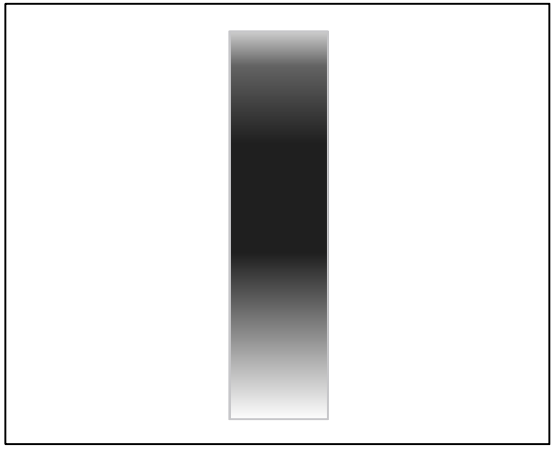
Imaginary Object



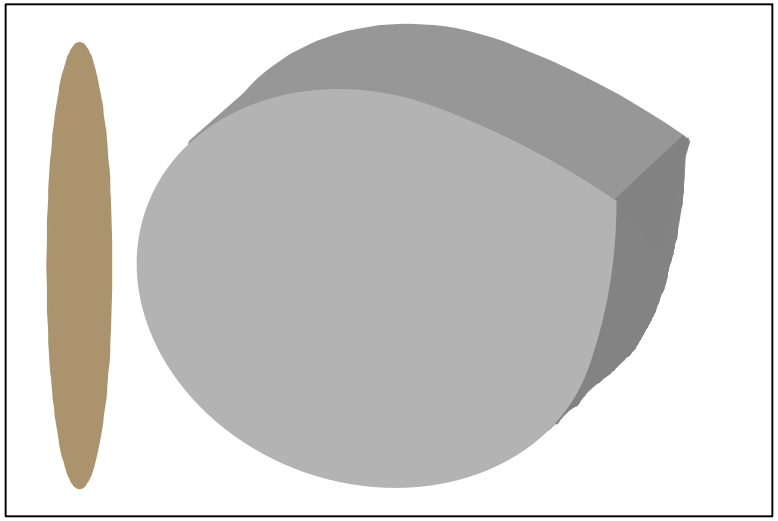
Front



Top



Side



Removed Image

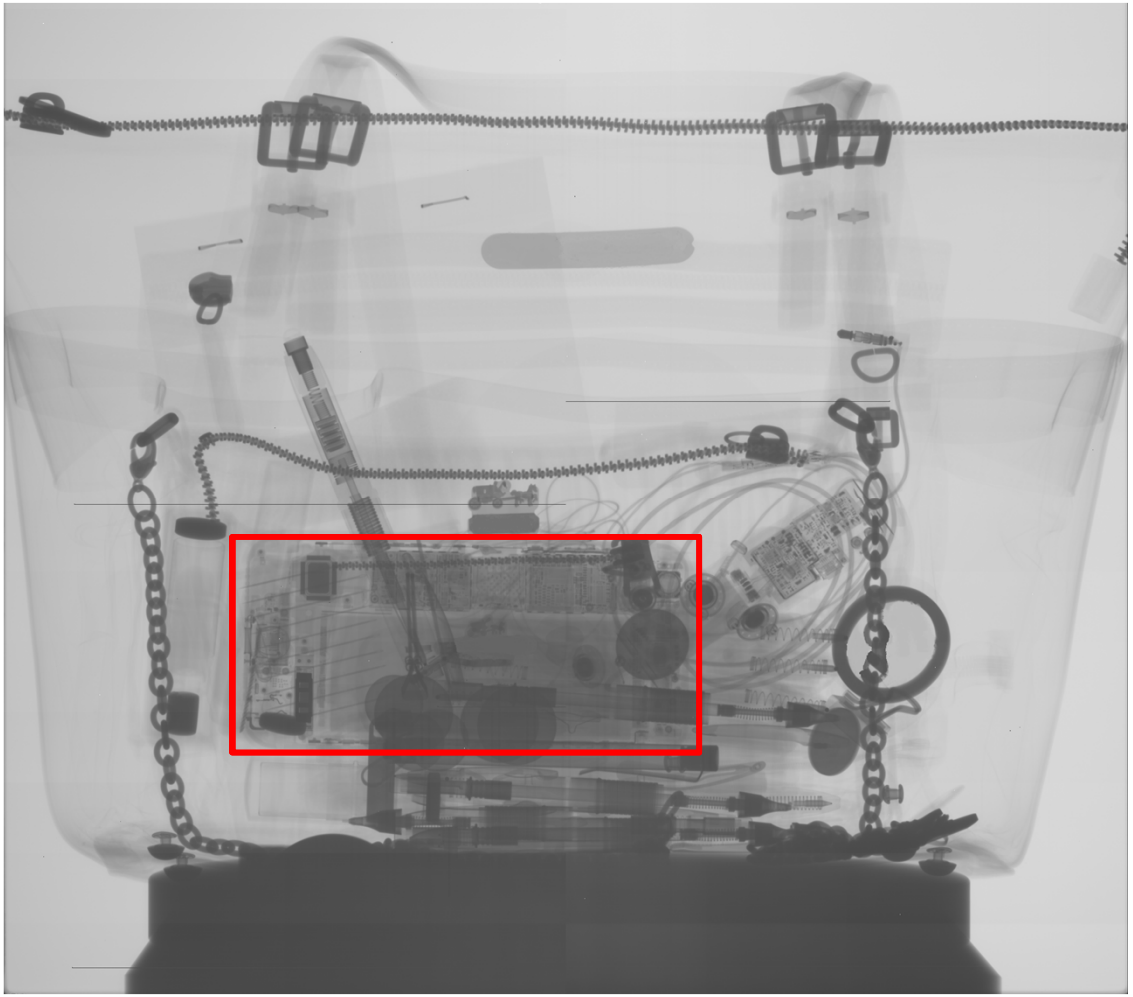
Introduction

- Objective:

  Identify objects in a region of interest using an existing database of X-ray radiographs

- Status:
  - Created clean exemplar images from database
  - Developed pre-processing algorithm for radiographs
  - Developing sorting and matching algorithm

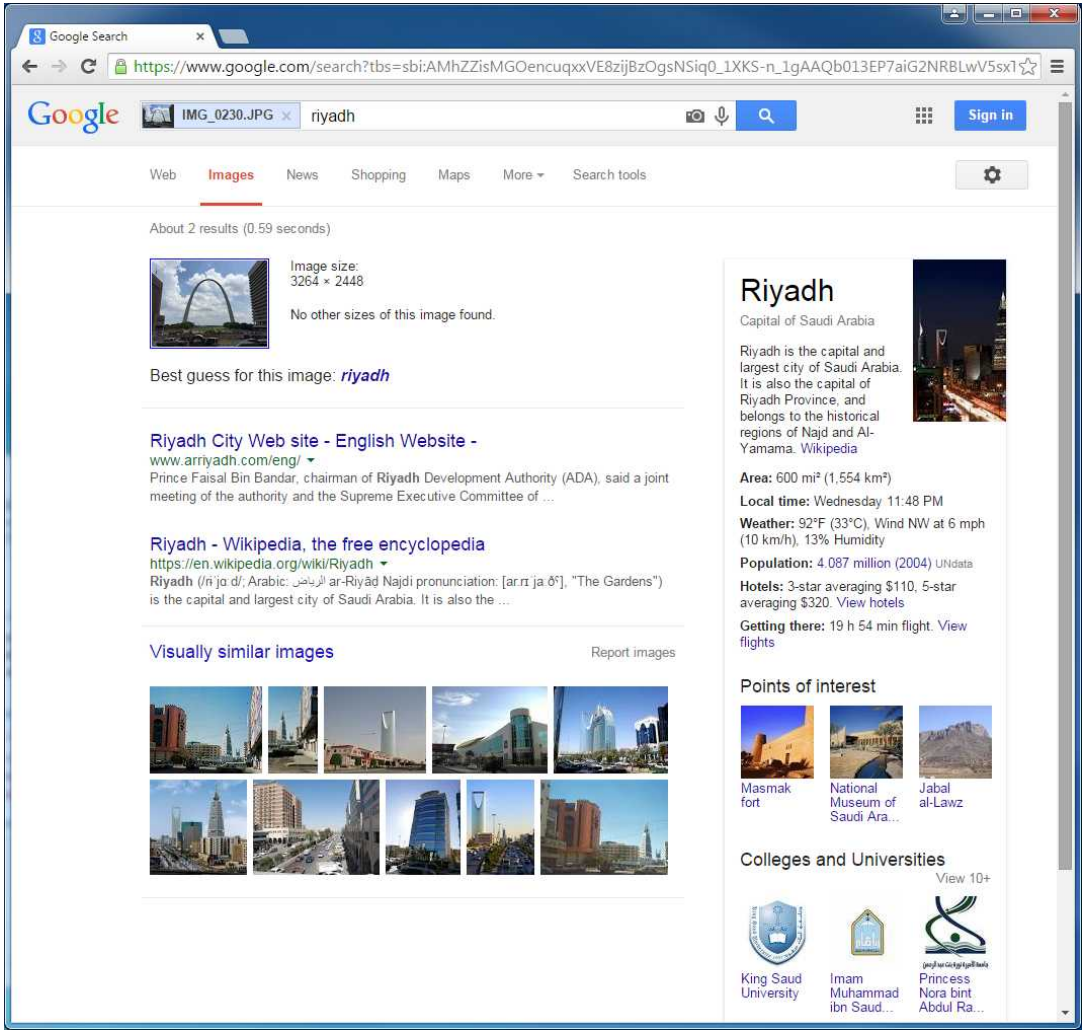
- Importance:
  - Inspection applications
  - National Security



X-ray radiograph acquired at Sandia National Laboratories

Background

Personal picture of St. Louis Arch



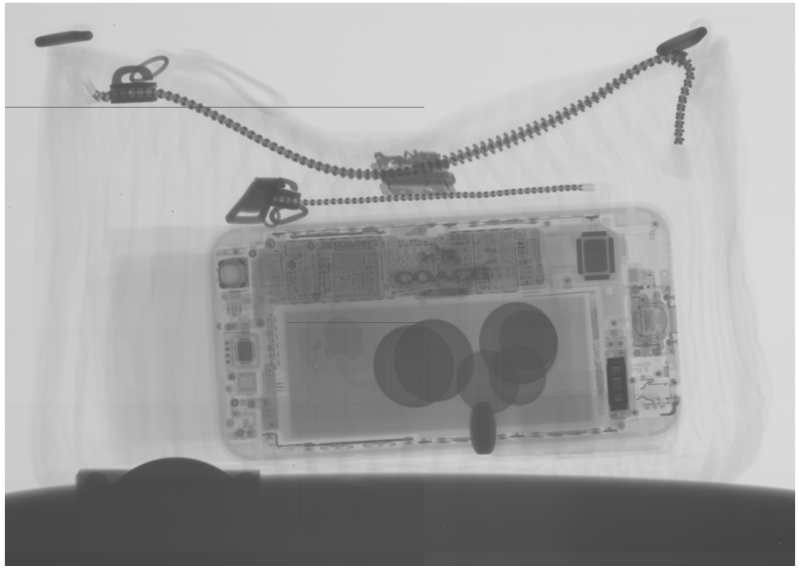
Screenshot generated using Google reverse image search

Radiographs vs Photographs

Sandia National Laboratories

- Transmission
- Grayscale with little texture
- Transmission overlap

- Reflection
- Color with texture



X-ray radiograph acquired at Sandia National Laboratories

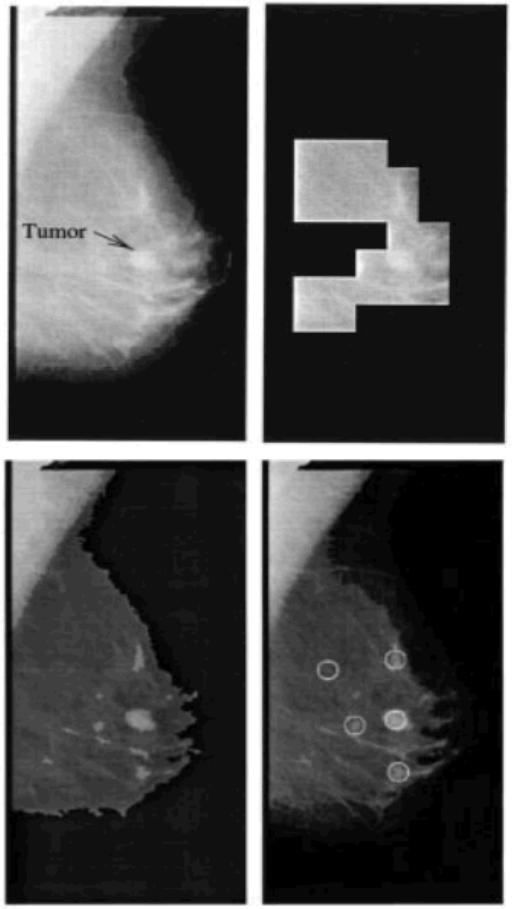
Photograph of the purse in the X-ray image

→ Must develop a unique object recognition technique for radiographs

Other Object Recognition Methods

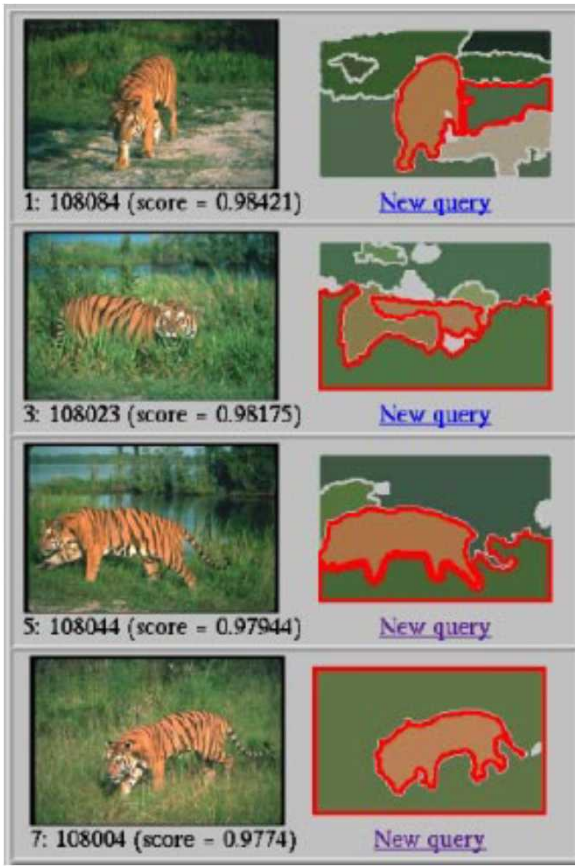
Sandia National Laboratories

Tumor Detection



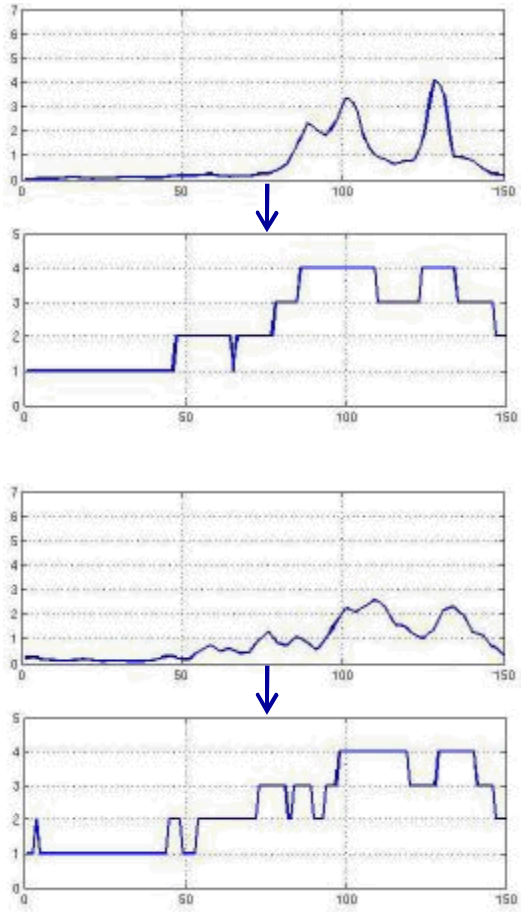
An Artificial Intelligent Algorithm for Tumor Detection in Screening Mammogram
Zheng and Chan, 2001

Blobworld



Blobworld: A System for Region-Based Image Indexing and Retrieval
Carson et al., 1999

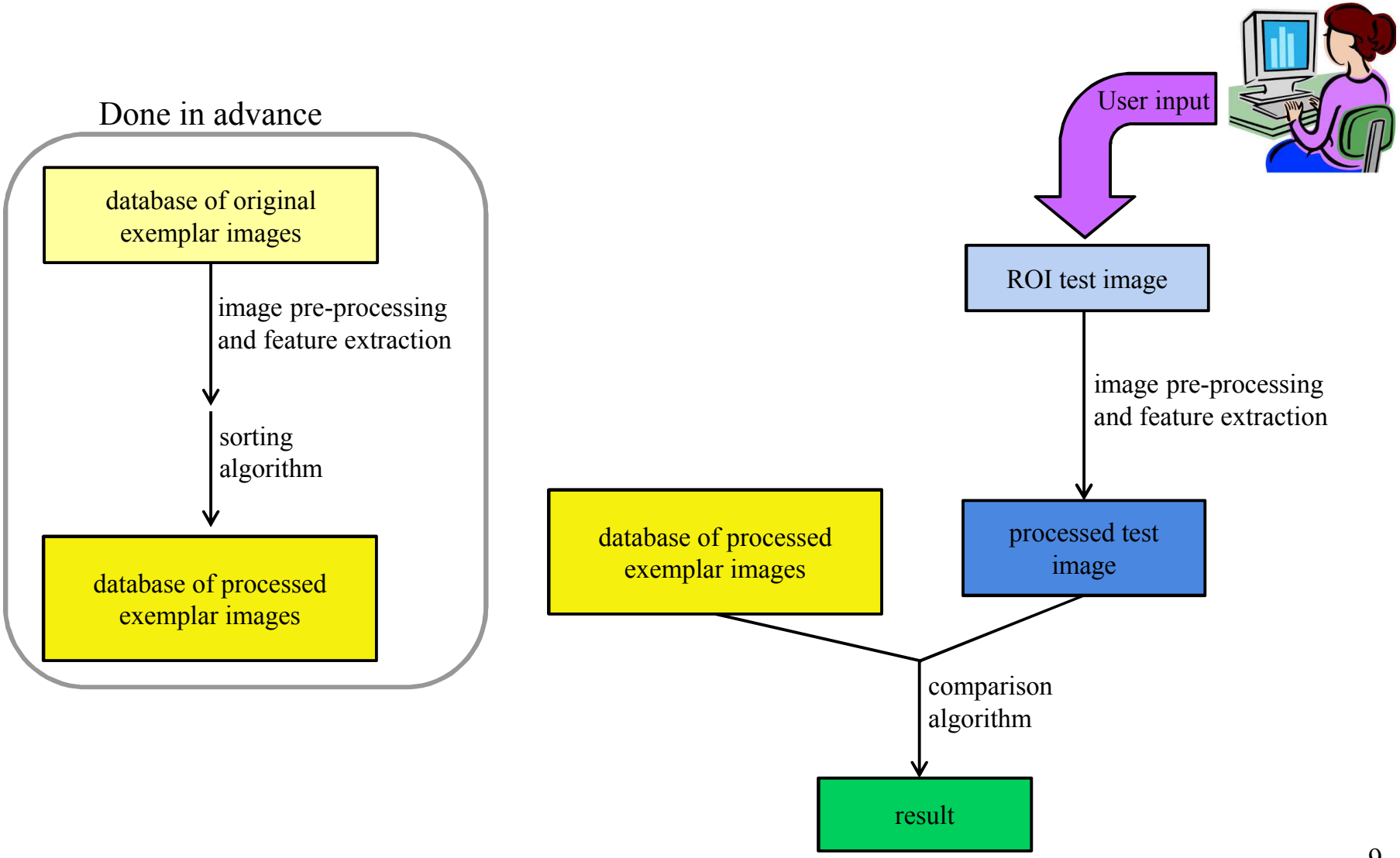
Multinomial Pattern Matching (MPM)



Multinomial Pattern Matching for High Range Resolution Radar Profiles
Koudelka et al., 2007

Hybrid Object Recognition (OR) System

Sandia National Laboratories

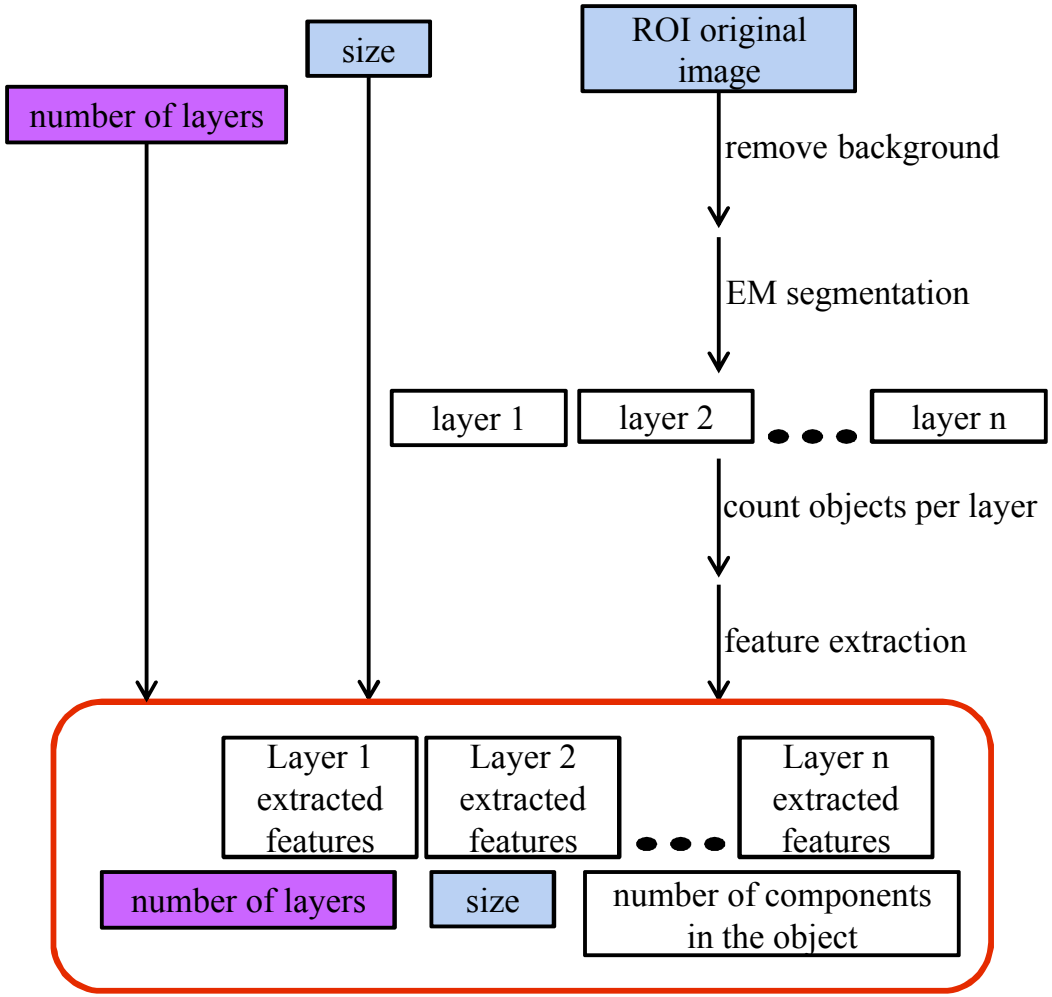


Assumptions

Sandia National Laboratories

- User selects region of interest to be identified from input image
- Perspective Concerns:
  - Database only contains orthogonal perspectives
  - Ignore perspective invariance (match to a specific perspective)
  - Perspectives of the same object are stored as individual images (not grouped together)

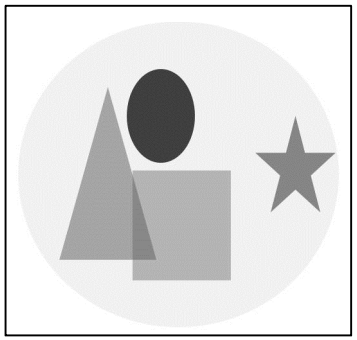
OR Pre-Processing



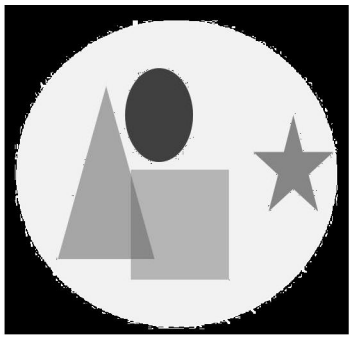
for sorting and matching process

- Determine properties and extract features for sorting and matching process
- Divide into layers
  - Layer corresponds to brightness quantile
- Number of quantiles based on user input
  - Typically for our dataset, 3-6 quantiles
- Extracted features qualify shapes in the layer

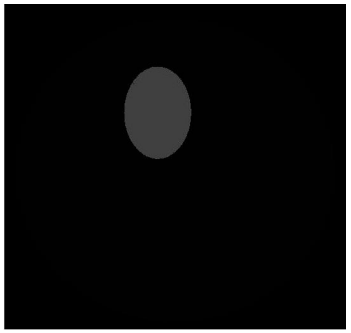
Example of OR Pre-Processing



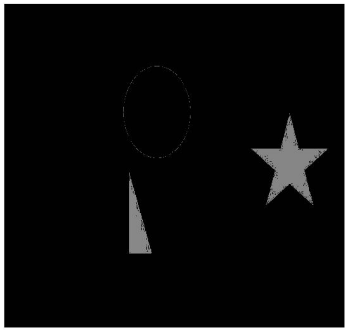
Original Image



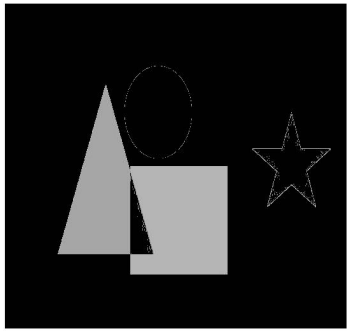
After Background Removal



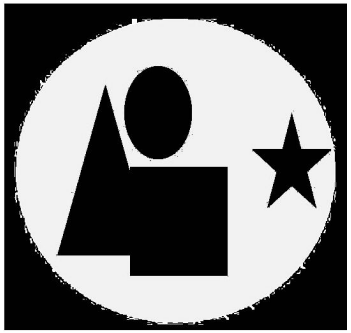
Layer 1
1 Object



Layer 2
2 Objects



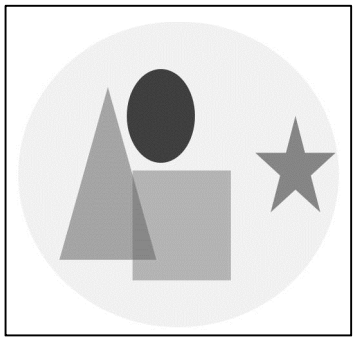
Layer 3
2 Objects



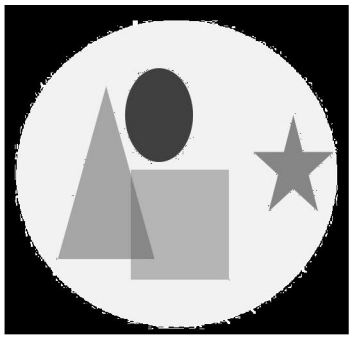
Layer 4
1 Object

Example of OR Pre-Processing

Mathworks detectBRISKFeatures



Original Image

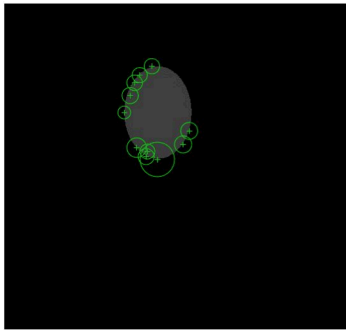


After Background Removal

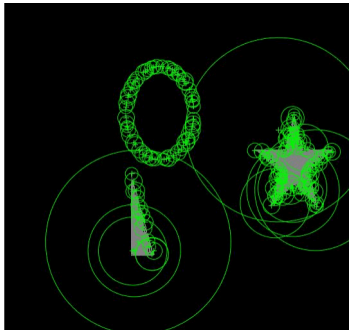
size: small

number of layers: 4

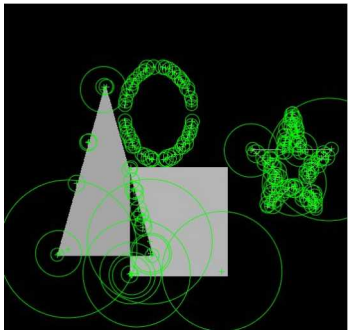
number of components: 6



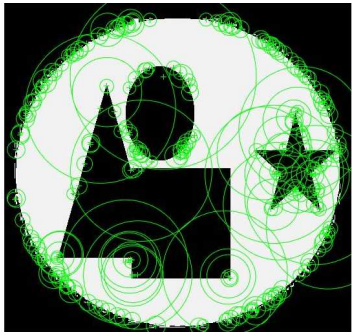
Layer 1
1 Object



Layer 2
2 Objects



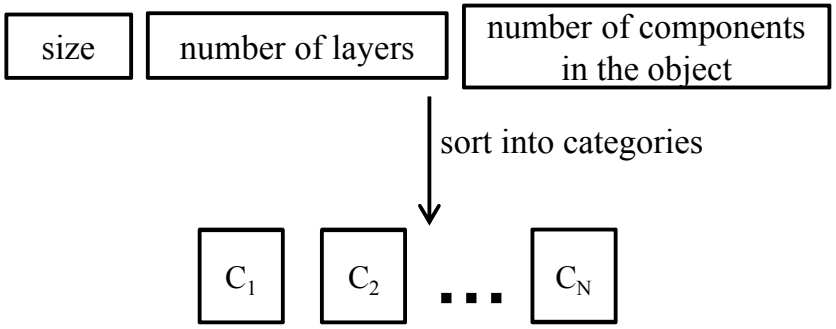
Layer 3
2 Objects



Layer 4
1 Object

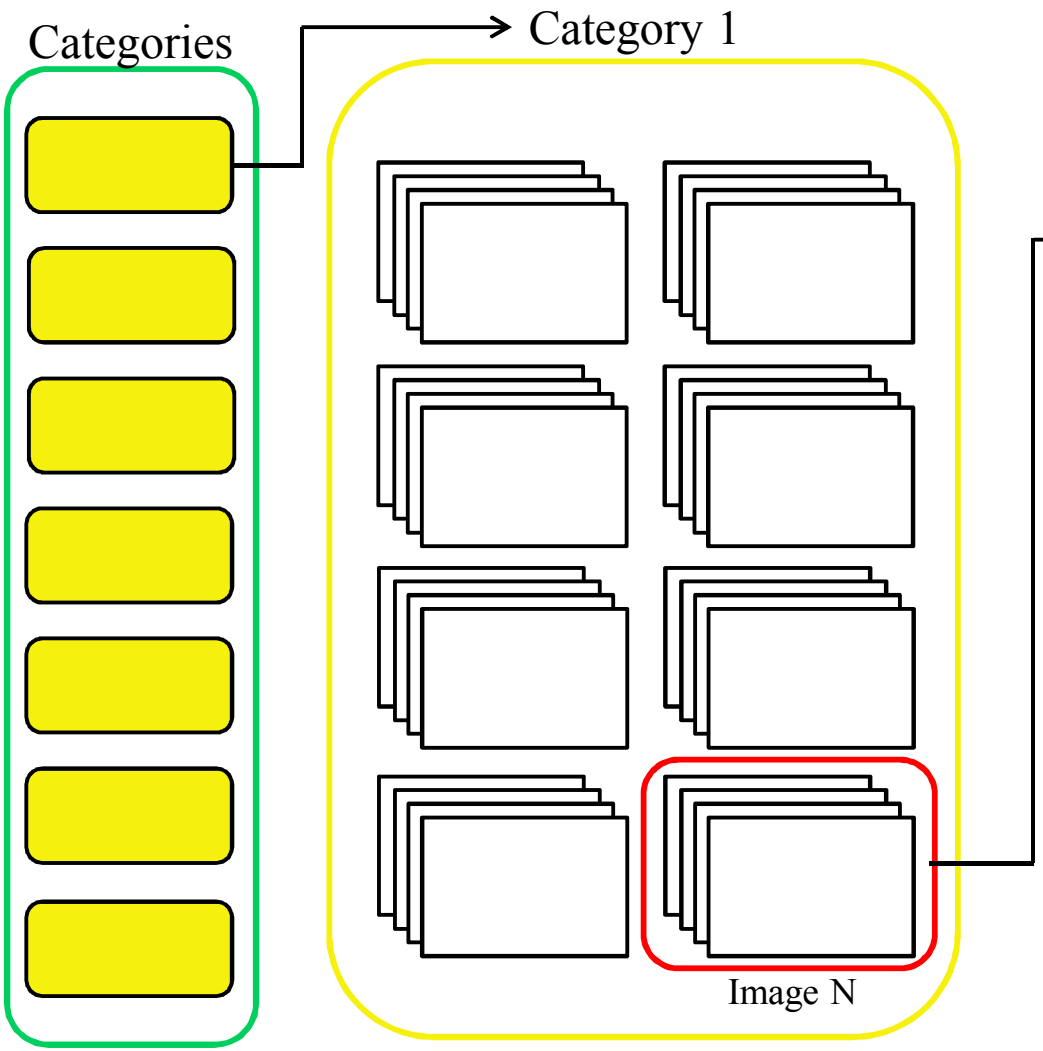
OR Sorting

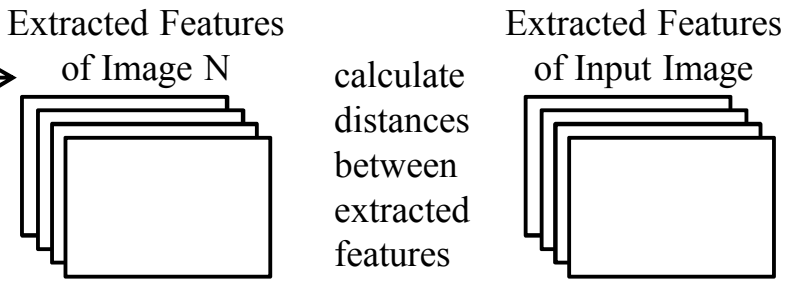
- Exemplar database is very large. For speed in the comparison step, we bin the database into smaller groups based on object size, number of layers and number of number of components in the object
  - Call these smaller groups of exemplar images "categories"



Searching over multiple close categories helps address variation in the parameters due to non calibration, object overlap, imprecise size, etc.

OR Comparison





Return to user :

- Top ten closest matching images
- Metric of how close each match is (how small are the distances)

Methods

Sandia National Laboratories

- MATLAB 2015a with
  - Image Processing Toolbox
  - Computer Vision Toolbox
  - Parallel Computing Toolbox

- Subset of Database Used for Testing
  - 42 objects
  - 151 radiographs (3-6 radiographs/object or only 1)
  - Manually removed photographs
    and cropped out extraneous objects

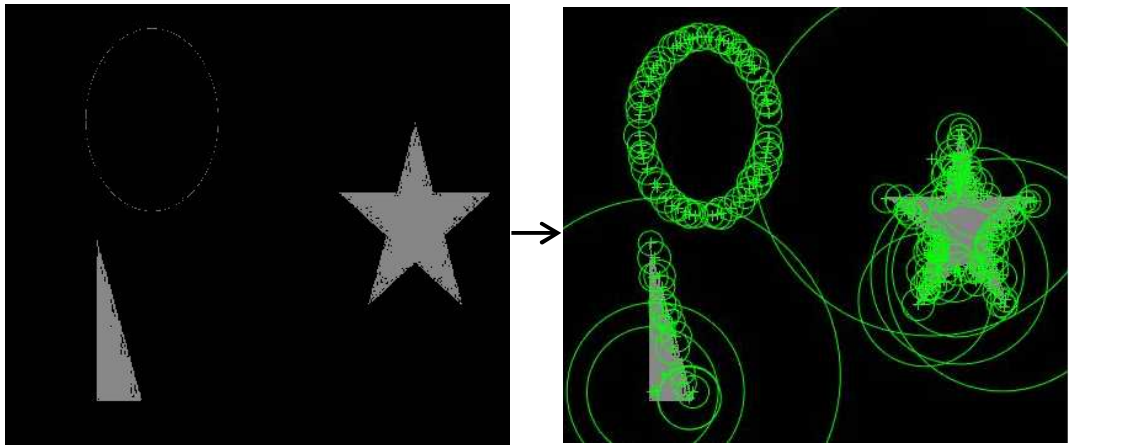
Results

Sandia National Laboratories

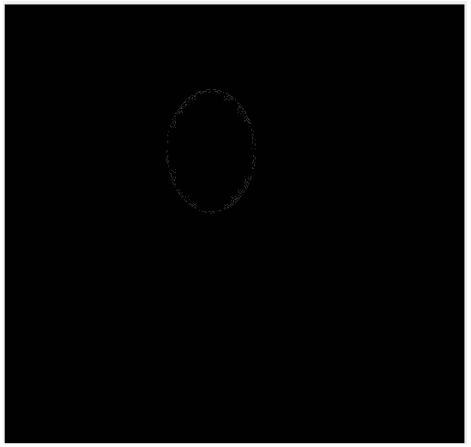
- Data Set Cleaning
  - Successfully deleted duplicates and extracted X-ray images
  - Still verifying results of image cropping
    (checking for images exhibiting any cropping errors)

Results

- OR Pre-processing
  - Quality of Segmentation
    - Speckling in pixel placement leads to extra features that aren't very repeatable
    - Deficiency on images with many components (difficult to separate them out properly)
    - Increasing number of layers can lead to meaningless "shadow" layers



zoomed layers to show speckling and resulting features



additional "shadow" layer generated when input 5 instead of 4

Results

- OR Pre-processing
  - Non-parallelized run time:
    - 151 images → 1398 seconds
  - Parallelization to increase speed
    - 151 images, 12 threads → 138 seconds
    - Overhead for starting and shutting down parallel pool
- OR Sorting
  - Successful file transfer
  - Experimenting with other data storage methods (large structure instead of many data files)
  - Non-parallelized run time:
    - 151 images → 5 seconds

Conclusion

Sandia National Laboratories

- Preliminary work holds promise for performing OR on radiographs
  - Algorithms implemented thus far have been successful
- Need to Improve:
  - Quality of segmentation
  - Overall performance
- What's next:
  - Developing the OR Comparison Algorithm
  - Test system with full database
- Potential applications of this work:
  - Inspection applications
  - National Security