Edge Detection

A.L. Yuille (UCLA)

We introduce low-level vision by edge detection.
Why do we care about edges?

• A Line Drawing is simple representation of the image (far fewer bits than a normal image).
• They are often sufficient to interpret the entire image (caveats).
• The Line Drawing is composed of edges – an “edge map”.
What Does the Line Drawing Represent?

- They represent the boundaries of objects.
- They represent interior edges of objects.
- They represent texture edges.

- Mooney images – a caveat: edges alone are not sufficient. Black and white helps.
But how can we find edges?

- Edge detection applies local operations (filtering) to images in order to detect edges. Local – low-level vision.
- Typically look at a local image patch – 3x3 pixels, or 8 x8 pixels – and decide if there is an edge or not.
- This is binary classification task – machine learning/statistics.

\[
\begin{bmatrix}
120 & 118 & 110 & 115 & 116 & 120 \\
115 & 21 & 20 & 16 & 19 & 121 \\
112 & 19 & 17 & 18 & 20 & 117 \\
119 & 118 & 121 & 117 & 116 & 112
\end{bmatrix}
\]

(a) Intensity matrix of an image

(b) Image of the matrix (150×100)
Typically edges occur at places where the intensity gradient changes.

- Idealized edges: Images \( I(x) \). Derivative of Image \( dI(x)/dx \)
But Images are much more complex than this simple picture

• The intensity gradient can be very small at the boundaries of an object. This is surprising to humans since we use context to interpret images. This context can be non-local or high-level (e.g., recognition), later in this course.

• Fox. Local edges – threshold gradient. Steeple Image (ambiguous)
It is impossible to detect edges locally.

• So what can we do?

• Use local cues to determine “edge candidates” or “edge hypotheses”. (In later lectures we will discuss how context can be used to exploit these edges for other visual tasks).

• How to use local cues?
• Classic approach – define an ideal model of an edge and obtain an optimal edge detector (Canny 1986).
• Modern approach – statistical/machine learning – treat as a classification problem. Obtain a dataset with groundtruth (positions of edge specified), train a classifier. (Konishi, Yuille, Coughlan, Zhu 2003), (Malik et al. 2004), (Dollar et al. 2014).
Sowerby Dataset

• Konishi: “There have been a thousand PhD theses on edge detection. Why should I do another?”

• Yuille: “Trust me. Nobody has done edge detection by learning”
P-on and P-off: Statistics versus Canny
Comparing Edge Cues:

- Different Filters. Different scales. Color.
- Left: Sowerby. Right: South Florida.

Figure 9: Chernoffs for Sowerby and South Florida. The edge detector operators are labelled by stars for \((N_1, N_2)\), crosses for \(N_2\), triangles for \(\|\nabla\|\), and diamonds for \(\nabla^2\). The three leftmost panels plot the Chernoff Information for Sowerby for full colour, greyscale, and chrominance respectively. The far right panel plots Chernoff for South Florida for greyscale. The horizontal axis shows the filter scale \((\sigma = 1, 2, 4)\). Decision trees are not needed.
Combining Edge Cues

- Combining cues give stronger edge detectors.
Edges at Multiple Scales

- Hard (Sowerby) and Easy (S. Florida) Datasets.

Sowerby consists of outdoor images. Much texture and vegetation.

South Florida consist of indoor images. Very little texture.
Local Edge Structure

- Konishi et al. classified pixels are edges or non-edges.
- But edges have local structure such as orientation.
- Malik et al. (2004). Trained classifiers to detect edge segments at eight different orientations. Eight orientations – classify by comparing statistics of image features on either side of segment (red and blue).
Local Edge Structure

- Pb at different orientations.
Learn local edge structure. Sketch tokens

• From a dataset of ground truth edges.
• Cluster the local edges structures.
• Obtain a dictionary of edges structures – each local edge configuration is similar to an element of this dictionary.
• Sketch token. Shape epitomes.
Figure 1. Illustration of our proposed dictionary of Shape Epitomes in the context of semantic labeling. Segmentation templates are generated from the shape epitomes, by specifying the values of the hidden variables. Image labels are assigned to the regions within the templates, and thus the local relationship between object classes is explicitly modeled. Note the shift-invariance illustrated in the third shape epitome. Those two generated segmentation templates are shifted version of each other.
Classify using sketch tokens

- Dollar et al. 2014.
- Obtain dictionary of sketch tokens.
- Classify each part of the image dataset by the sketch tokens – includes a default sketch token with no edge.
- Train a sketch token classifier – using random forests.
- This classifies the center of the sketch token as an edge (or not).

Dangers of Datasets

• Datasets are great because: (i) they enable us to learn algorithms, (ii) they allow us to benchmark algorithms.

• But they are also problematic. They can lead to a research culture which aids only at small incremental improvements in performance without adding any understanding.

• Also there are three other dangers/issues:

  (I) Failure to generalization within datasets – classic machine learning/statistics.

  (II) Failure to generalize across datasets – dataset bias.

  (III) Labeling Bias.
Dangers of Datasets (I)

• (I) Failure to Generalize within datasets.
• Machine Learning/Statistical Methods divide the dataset into two set – training and testing (sometimes also validation).
• The classifier is trained on the Training Dataset. It is evaluated on the Testing Dataset.
• This is to ensure that the classifier “generalizes” from Training to Testing datasets. Otherwise it may be simply fitting the training data – or memorizing it. Learning requires the classifier to perform well on data that it has not seen yet (but comes from the same source).
• Machine Learning researchers have studied these issues carefully and written books and thousands of papers about them.
Dangers of Datasets (II)

• (II) Failure to generalize across datasets. Dataset bias.
• Is the dataset representative of the class of images that you want the algorithm to work on?
• Classifiers trained on the South Florida dataset work poorly on the Sowerby dataset (more complex, much texture). Even if the classifiers pass all the Machine Learning/Statistical tests for generalization.
• We need generalization across datasets (not just generalization within a dataset). Medical image example.
• Ideally image datasets should be representative of the set of natural images. This is a tough challenge – only now do computer vision researchers have the ability to make enormous datasets (Pascal, ImageNet, Coco). We need generalization ac

• Or you can restrict yourself to a specific visual environment (e.g., circuit board images).
Dangers of Datasets (III)

• Labeler Bias
• X. Hou, et al. CVPR 2013

An example image and the corresponding labels from BSDS 300. Top figure shows the original image overlapping with all 6 boundary maps from labelers. There is a clear difference among different labelers. Red circle gives an example boundary segment that is labeled by only one out of 6 labelers (labeler 4). Boundary segment in the orange circle is labeled by two labelers (labeler 3 and 4). The boundary segment in green circle is unanimously labeled by all 6 labelers.
Extensions: What else can you classify locally?

• What about “sky”, “vegetation”, “water”?

• These typically require larger context.

• But some surprisingly good results can be found for some “region classes”.
Label Regions Sowerby

• Konishi and Yuille.
Label Regions: San Francisco