

RANDOM FORESTS

SUPPLEMENTAL MATERIAL TO RANDOM FORESTS LECTURE

CSE 455/555

PROF. J CORSO

MATERIALS ARE TAKEN FROM AMIT AND GEMAN, “SHAPE QUANTIZATION AND RECOGNITION WITH RANDOMIZED TREES,” NEURAL COMPUTATION, 9(7):1545—1588, 1997.

THERE ARE A NUMBER OF PAPERS ON RANDOM TREE/FORREST CLASSIFIERS; THIS IS AN EARLY (THE FIRST?) ONE WITH THE BENEFIT OF PRINCIPLED MATHEMATICAL DERIVATION AND AN INTERESTING APPLICATION.

GENERATING TRAINING DATA BY PERTURBATION



SHAPE QUERIES

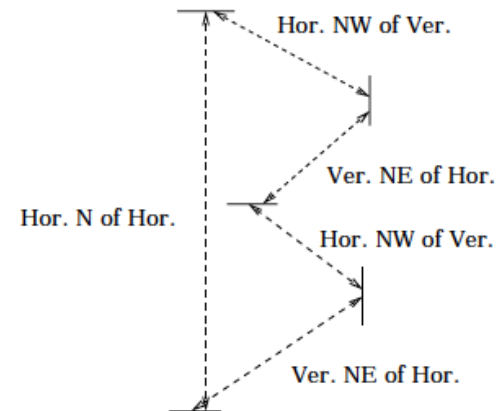
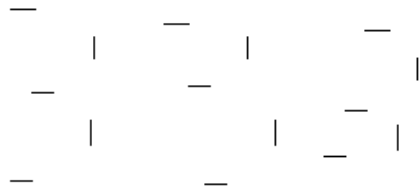


Figure 3: Left: Three curves corresponding to the digit '3'. Middle: Three tangent configurations determining these shapes via spline interpolation. Right: Graphical description of relations between locations of derivatives consistent with all three configurations.

EXAMPLE OF THE TAG CODE FOR SHAPE QUANTIZATION

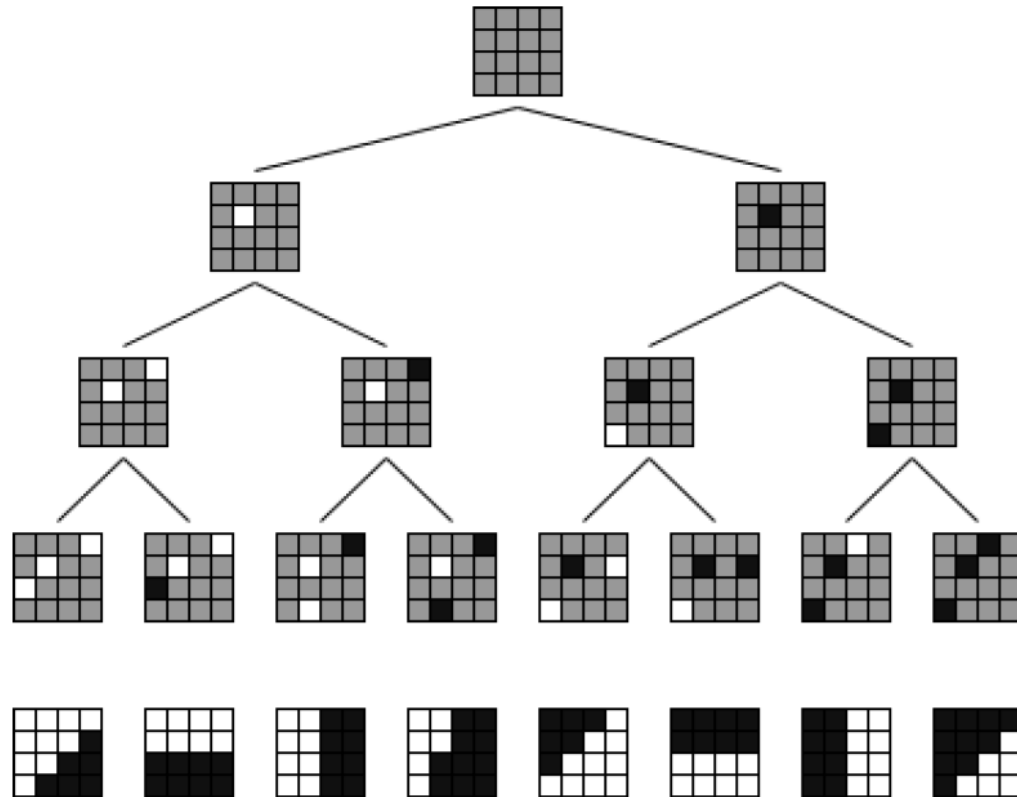


Figure 4: First three tag levels with most common configurations.

TAGS ARE ENCODED IN GEOMETRIC ARRANGMENTS

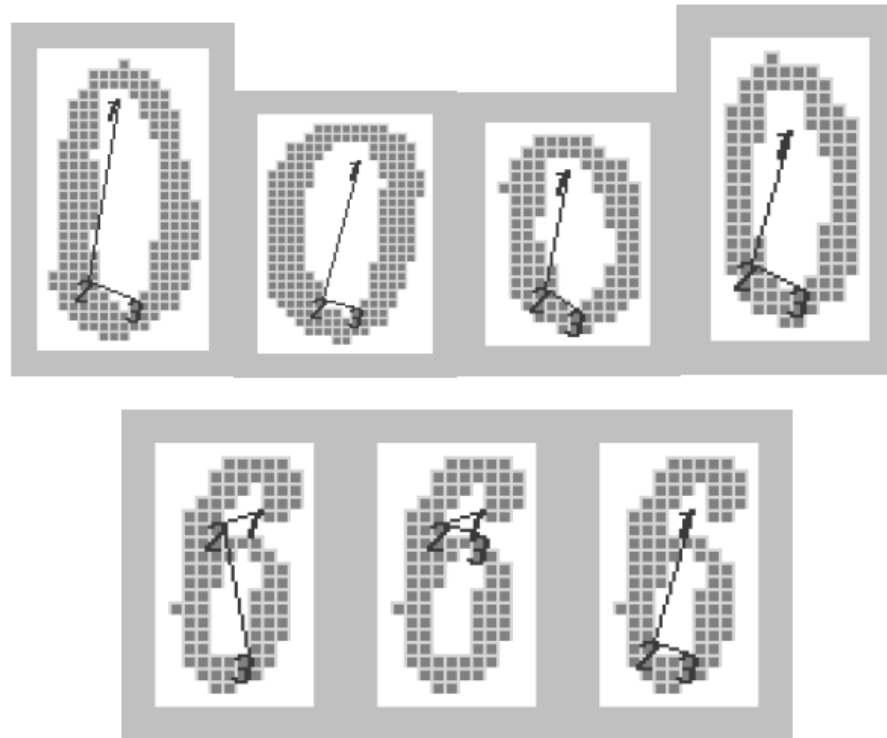


Figure 6: Top: Instances of a geometric arrangement in several 0's. Bottom: Several instances of the geometric arrangement in one 6

EXAMPLE STRUCTURE GRAPHS FROM MULTIPLE RANDOMIZED TREES

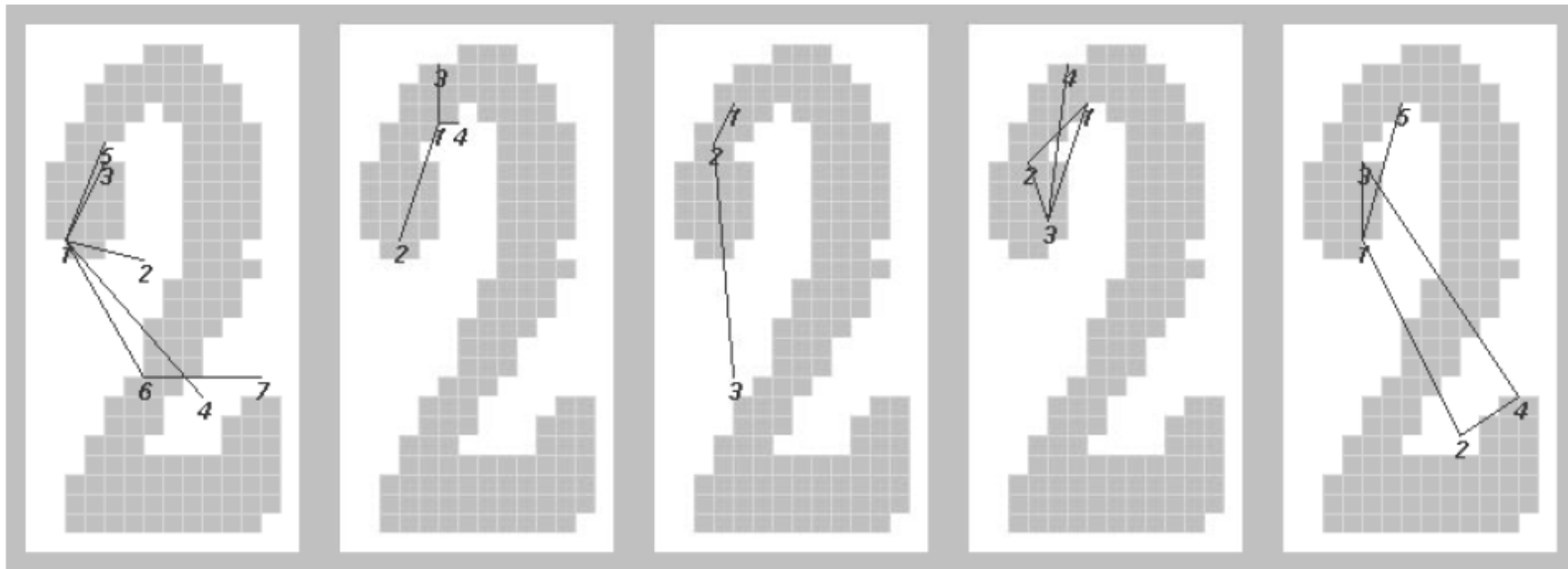


Figure 9: Graphs found in an image at terminal nodes of five different trees

CLASSIFICATION RATE

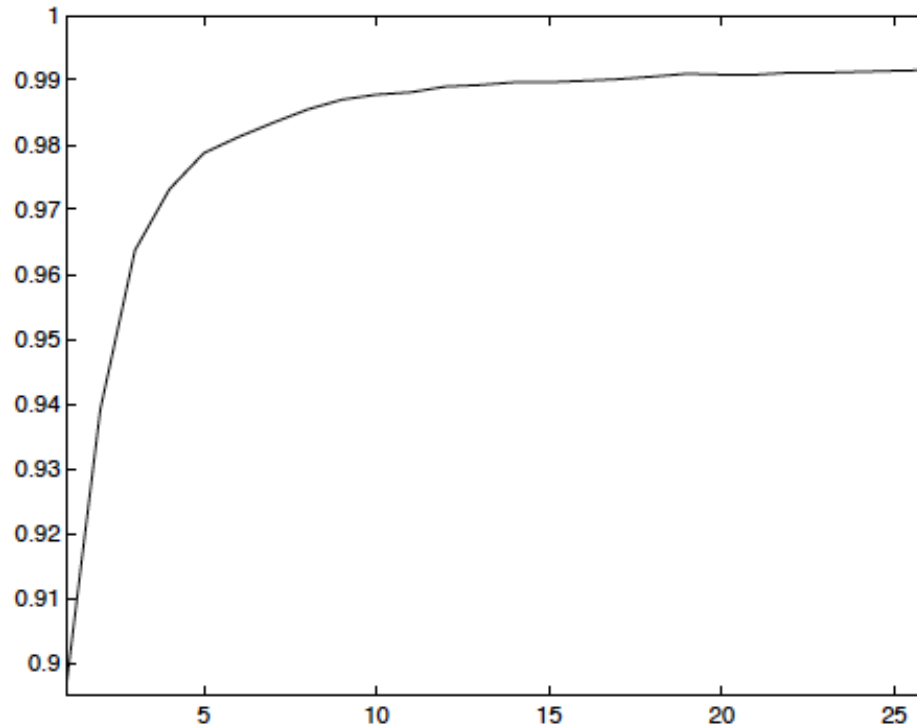


Figure 13: Classification rate vs. Number of trees.