1985 Intl. Geoscience and Remote Sensing Symposium,
Amherst, MA, October, 1985

CONTEXT CLASSIFIER

Robert M. Haralick and Hyonam Joo

Machine Vision International

Ann Arbor, MI 48104

The context classifier is characterized by the fact that it classifies an unknown pixel using the entire context of the image or a substantially sized context neighboring the pixel. Basically, the effect of context is that a pixel can have certain properties, when it is viewed in isolation, which change when viewed in some context. One might expect that classification accuracy is higher if an unknown pixel is classified using context rather than when it is classified using only the measurement made on that pixel without context. This is true in most cases. For example, a single pixel is not likely to be classified as water if it is surrounded by the pixels classified as ground in a remotely sensed data. The classification result of the conventional context free classifier leaves many isolated pixels and many small groups of pixels not connected with the blob they belong to. Thus, in the last few years there has been a trend to increase the use of context in the labeling operations.

The most desirable kind of labeling process would give each pixel the highest probability label given the entire context of the image.

The next most desirable kind of labeling process would give each pixel the highest probability label given some substantially sized context neighboring the pixel. In this paper, a theory and an algorithm for such a context classifier is presented. The algorithm takes the form of a recursive neighborhood operator first applied in a top down scan of the image and then in a bottom up scan of the image. The algorithm itself is related to a forward dynamic programming algorithm put in a two dimensional mesh setting. To explain the meaning of what the algorithm produces, select any pixel in the image. Now consider all the row monotonically increasing paths which begin at any border pixel of the image above the selected pixel, go through the selected pixel, and end at some border pixel of the image below the selected pixel. Each such path represents a context for the pixel. Corresponding to each path and the observed pixel data on the path, there is an associated highest probability label for the given pixel. Among all the paths there is some best path whose associated highest probability label is higher than the highest probability label of every other path. In two scans of the image, the context algorithm is able to assign to each pixel of the image the highest probability label coming from its best path.

The theory for the algorithm requires two distinct ideas. The first idea produces a decomposition for the problem. Finding the highest probability label given the best path passing through the pixel can be accomplished by finding two probabilites, the probability for each possible label given the best path beginning above the pixel and terminating at the pixel and the probability for each possible label given the best path beginning at the pixel and terminating below the pixel. Finding these probabilities is what the algorithm accomplishes in top down scan and the bottom up scan. The decomposition tells how to combine these probabilities to determine the highest probability label given the context of the best path through the pixel. The second idea produces a recursive decomposition which tells how to determine the conditional probability for each label given the data on the pixels's best upper (or lower) path from this some kind of conditional probability of the pixel's neighbors which have already been processed. The decomposition bears a definite similarity to the one used in forward dynamic programming and as well bears some similarity to the iteration technique employed in some relaxation methods.

We now explain the assumptions and the probability function that we need to get the highest probability label. For pixel (i,j), let X_{ij} designate the measurements of the pixel (i,j). Let Q be any path and e_{rc} be the correct but unknown label at pixel (r,c). $f_{Z_{rc}}(e_{rc})$ is defined to be the probability that pixel (r,c) takes label e_{rc} and that the joint measurements on the best row monotonically increasing path

Q is $\{X_{ij}; (i,j) \in Q\}$.

Thus

$$f_{Z_{rc}}(e_{rc}) = \max_{Q \in Z_{rc}} P(e_{rc}, X_{ij}; (i,j) \in Q)$$

By Bayes formula,

$$P(e_{ij}, X_{ij}; (i,j) \in Q) = P(X_{ij}; (i,j) \in Q | e_{ij}; (i,j) \in Q) \cdot P(e_{ij}; (i,j) \in Q)$$

Assuming that the measurements at each pixel depend only on the true label at that pixel and measurement noise for one pixel does not influence the measurement noise for another pixel, we have

$$P(X_{ij}; (i,j) \in Q | e_{ij}; (i,j) \in Q) = \prod_{(i,j) \in Q} P(X_{ij} | e_{ij})$$

The probability $P(e_{ij}; (i,j) \in Q)$ is the joint prior probability of having the true labels for each pixel (i,j) on the path Q be e_{ij}. This probability encodes all the information we have about context. If for example, we had independence,

$$P(e_{ij};(i,j)\in Q) = \prod_{(i,j)\in Q} P(e_{ij})$$

Thus we would discover that the highest probability assignment we could make using the context is precisely the highest probability assignment we could make using only the local information.

The simplest assumption of higher order independence is a Markov like assumption in which the joint prior probability becomes a function expressible as the product of functions whose arguments are the label pairs for succesive pixels in the path Q. Letting R(Q) designate the set of all pairs of succesive pixels in the path Q, we have,

$$P(e_{ij};(i,j)\in Q) = \prod_{((i,j),(k,l))\in R(Q)} A(e_{ij},e_{kl})$$

Using the two assumptions described above, we could decompose $f_{Z_{rc}}(e_{rc})$ into two components as follows,

$$f_{Z_{rc}}(e_{rc}) = \frac{g_{U_{rc}}(e_{rc})g_{L_{rc}}(e_{rc})}{P(X_{rc}|e_{rc})}$$

where $g_{U_{rc}}(e_{rc})$ $(g_{L_{rc}}(e_{rc}))$ designates the probability of label e_{rc} and joint measurements $\{X_{ij};(i,j)\in Q\}$ arising from the best row monotonically increasing path in U_{rc} (L_{rc}). The set U_{rc} designates the set of all row monotonically increasing paths which begin at some border pixel of the image above or to the left of pixel (r,c) and terminate at pixel (r,c). The set L_{rc} designates the set of all row monotonically increasing paths which begin at (r,c) and terminate at some border pixel below or to the right of pixel (r,c).

Even though the context information is not used to its full extent due to the Markov like assumption, the improvement in overall classification accuracy was substantial. The overall classification accuracy is measured as the ratio of the number of correctly classified pixels to the number of total classified pixels. To compute the function A for the testing purpose in this paper, we assumed a Gaussian stationary two-dimensional process. Here stationary means that the correlation between pixels are position independent in the image. The function A is estimated for four directional pairs of pixels, horizontal, vertical, and two diagonals. It is approximated by the frequency distribution of each pairs of labels in all four directions from the known ground truth data. Assuming a multi-dimensional Gaussian normal distribution, the class conditional probability $P(X_{rc}|e_{rc})$ is computed from known or estimated class conditional covariance matrices and mean vectors, which in turn is used together with the function A to compute $f_{Z_{rc}}(e_{rc})$.

The recursive algorithm for the computation of $g_{U_{rc}}(e_{rc})$ is,

$$g_{U_{rc}}(e_{rc}) = P(X_{rc}|e_{rc}) \max\{ \sum_{e_{rc-1}} g_{U_{rc-1}}(e_{rc-1})A(e_{rc-1}, e_{rc}),$$

$$\sum_{e_{r-1c-1}} h_{U^*_{r-1c-1}}(e_{r-1c-1})A(e_{r-1c-1}, e_{rc}),$$

$$\sum_{e_{r-1c}} h_{U^*_{r-1c}}(e_{r-1c})A(e_{r-1c}, e_{rc}),$$

$$\sum_{e_{r-1c+1}} h_{U^*_{r-1c+1}}(e_{r-1c+1})A(e_{r-1c+1}, e_{rc})\}$$

where $h_{U^*_{rc}}(e_{rc})$ again designates the probability of label e_{rc} and joint measurements $\{X_{ij}; (i,j) \in Q\}$ arising from the best row monotonically increasing path in U^*_{rc}. The set U^*_{rc} designates the set of all row monotonically increasing paths which begin at some border pixel of the image at the same row or above pixel (r,c) and terminate at pixel (r,c), which differs from the set U_{rc} that the set U_{rc} does not include paths which contain any pixels on the same row r beyond column c while the set U^*_{rc} does.

$h_{U^*_{rc}}(e_{rc})$ can also be computed recursively as follows,

$$h_{U^*_{rc}}(e_{rc}) = \max\{g_{U_{rc}}(e_{rc}), P(X_{rc}|e_{rc})$$

$$\sum_{e_{rc+1}} h_{U^*_{rc+1}}(e_{rc+1})A(e_{rc+1}, e_{rc})\}$$

An obvious mirror image derivation applies to $g_{L_{rc}}(e_{rc})$.

First, a simulated image is used to examine the improvement in classification accuracy of the new context classifier as compared to the noncontext Bayes classifier under Gaussian distribution assumption, given that the class conditional covariance matrices and mean vectors were known. Then it is applied on real images to investigate its performance in more realistic case. The simulated image is generated from a real LANDSAT image. From the known ground truth data, the mean vectors and the covariance matrices are estimated for each class. Then a simulated image with the following characteristic is created. (1) each pixel in the simulated image represents the same class as in the ground truth data, (2) all classes have multi-dimensional Gaussian normal distribution having the means and covariance matrices estimated from the sample image, (3) all pixels are class-conditionally independent of adjacent pixels. Applied on this image, the context classifier shows better classification accuracy compared to the Bayes classifier (Overall classification accuracy for the Bayes classifier with equal priors, the Bayes classifier with correct priors, and the context classifier is 59.7%, 68.8%, and 80.0% respectively). On real images, it was observed that the context classifier gained 4-8% increase in overall classification accuracy over the context free Bayes classifier under Gaussian distribution assumption.