Tile EM Model

April 8, 2017

1 Problem Formulation

$\mathcal{T} = \{t_k\}$ is the set of all non-overlapping tiles for an object i. T is the ground truth tile set. T' is some combination of tiles chosen from $\mathcal{T}$.

For our problem, we consider only finding the tile regions that could be constructed from worker bounding boxes. In other words, our objective is to find the tile combination T' such that we maximize the probability that it would be ground truth p(T'=T), given a set of worker qualities Q_j and tile indicator labels l_{kj}. The indicator label l_{kj} is one when worker j votes on the tile t_k (i.e. the bounding box that he draws contains t_k), and zero otherwise. The indicator matrix consisting of tile indicator for all workers is denoted as $\mathbf{l_{kj}}$. Given these definitions, the objective function of the segmentation problem is formulated as:

$$T = \arg\max_{T' \subseteq \mathcal{T}} P(T = T' | \mathbf{l_{kj}}, Q_j) \tag{1}$$

Using Bayes rule we can rewrite this in terms of the posterior probability of the tile-based values($\mathbf{l_{kj}}$) or worker-based values(Q_j), which we can use for the E and M step equations respectively.

2 Inference

For the E step, we assume T' is ground truth and estimate the Q_j parameters. We can rewrite Eq.1 as:

$$p(T'|Q_j, \mathbf{l_{kj}}) = \frac{p(Q_j|\mathbf{l_{kj}}, T')\cancel{P(T')}}{\cancel{p(Q_j)}} \tag{2}$$

Our optimization function then becomes:

$$\hat{Q}j = \arg\max p(Q_j|\mathbf{l_{kj}}, T') \tag{3}$$

We use the binary random variable w to indicate whether the worker makes a correct vote (w=1) or an incorrect vote(w=0) for a tile. Our worker error model follows a Bernoulli distribution by describing worker quality $Q_j = q_j$ as the probability that the worker makes a correct vote, otherwise $Q_j = 1 - q_j$ for an incorrect guess. We can write the worker quality probability as the product of the probabilities that they would assume these two independent states (correct/incorrect).

$$p(Q_j) = \prod_j p_j(w=1) \cdot p_j(w=0) \tag{4}$$

$$= q_j^{p_j(w=1)} \cdot [1 - q_j]^{p(w=0)} \tag{5}$$

There is a closed form of the maximum likelihood solution for the Bernoulli distribution.

$$\mathcal{L} = \ln p(Q_j) = p(w=1)\ln q_j + p(w=0)\ln(1-q_j) \tag{6}$$

$$\frac{\partial \mathcal{L}}{\partial q_j} = \frac{p(w=1)}{\hat{q}_j} - \frac{p(w=0)}{1-\hat{q}_j} = 0 \tag{7}$$

$$\tag{8}$$

Solving for $\hat{q}_j$:

$$\hat{q}_j = \frac{p(w=1)}{p(w=0)+p(w=1)} = \frac{\frac{n_{w=1}}{\cancel{n_{total}}}}{\frac{n_{w=0}}{\cancel{n_{total}}} + \frac{n_{w=1}}{\cancel{n_{total}}}} \tag{9}$$

$$\boxed{\hat{q}_j = \frac{n_{correct}}{n_{total}}} \tag{10}$$

For the M step, we maximize the likelihood of the tile combination T' for a fixed set of worker qualities, $\{Q_j\}$. Following Eq.1,

$$p(T'|Q_j, \mathbf{l_{kj}}) = \frac{p(\mathbf{l_{kj}}|Q_j, l_k)\cancel{P(T')}}{\cancel{p(\mathbf{l_{kj}})}} = p(\mathbf{l_{kj}}|Q_j, l_k) \tag{11}$$

Our optimization function is written as:

$$\hat{T'} = \arg\max \prod_j p(\mathbf{l_{kj}}|Q_j, l_k) \tag{12}$$

The product over T' can be further decomposed into its tile components:

$$= \prod_j \left[\prod_{t_k \in T'} p(t_k \in \mathrm{T}|Q_j, l_k) \prod_{t_k \notin T'} p(t_k \notin \mathrm{T}|Q_j, l_k) \right] \tag{13}$$

We can evaluate these quantities by the observed indicator labels. The worker makes a correct response in the scenario where they vote for a tile that actually lies within T, $t_k \in$T, or they don't vote for a tile that is excluded from T. In those cases, the probability of the tile being in T is the probability that the worker had made the correct response: $p(l_{kj}) = q_j$. Likewise, a worker makes an incorrect response when their vote contradicts with the inclusion of the tile in T ($\{t_k \in \mathrm{T} \quad \& \quad l_{kj} = 0\}, \{t_k \notin \mathrm{T} \quad \& \quad l_{kj} = 1\}$), in which case $p(l_{kj}) = 1 - q_j$. We compute these probabilities and take their product over all tiles and workers.