2.3 Segmentation Using HMRF-EM Method

From the previous section, we find that although the DCT method can improve the segmentation result, identifying the region of interest is still necessary for almost every image. By applying the Hidden Markov Random Field and EM Algorithm Framework to our images, we can get rid of such inconvenience for most of the cases.

2.3.1 Hidden Markov Random Field and EM Framework

2.3.1.1 Hidden Markov Random Field Model

Mathematically, a HMRF model is characterized by an unobserved random field and an observable random field. Assume that S is a finite set of sites with a neighborhood system, and assume that a discrete Markov Random Field $X = \left\{X_i, i \in S\right\}$ is an underlying MRF taking values in a finite state space L. We have: for any $\boldsymbol{X}$,

- $P(X_i \mid \boldsymbol{X}_{S-\{i\}}) = P(X_i \mid X_j, j \in N(i))$,
- $P(\boldsymbol{X}) > 0$,

where $N(i)$ denotes the set of neighbours of i. According to the Hammersley-Clifford Theorem,

$$P(\boldsymbol{X}) = \frac{1}{Z} exp(-U(\boldsymbol{X})), \tag{2.8}$$

where Z is a normalizing factor and $U(\boldsymbol{X})$ is the energy function of $\boldsymbol{X}$, and

$$U(\boldsymbol{X}) = \sum_{c \in C} V_c(\boldsymbol{X}). \tag{2.9}$$

Here $V_c(\boldsymbol{X})$ is the clique potentials, which depends on a parameter Ψ. In general, this Ψ needs to be estimated. However, in convenience, we set the Ψ as a constant in this dissertation (See Equation 2.24). Celeux, Forbes and Peyrard [3] showed how to estimate Ψ in a general form.

Since the computation of the joint distribution $P(\boldsymbol{X})$ can be complicated [3], Besag [2] used a psedudo-likelihood:

$$P(\boldsymbol{X}) \approx \prod_{i \in S} P(X_i \mid \boldsymbol{X}_{N(i)}), \tag{2.10}$$

to approximate the equation 2.8. On the other hand, we consider an observable random field $Y = \{Y_i, i \in S\}$ as a random field taking values from a space D. Given an arbitrary configuration $X \in \chi$, Y_i follows the conditional probability distribution $P(Y_i \mid X_i, \theta)$, where θ is the involved parameters. By assuming that for any $X \in \chi$, the random variables Y_i are conditional independent, i.e.,

$$P(\boldsymbol{Y} \mid \boldsymbol{X}, \theta) = \prod_{i \in S} P(Y_i \mid X_i, \theta). \tag{2.11}$$

Based on the above, we can have the joint probability of (X, Y) as

$$\begin{aligned} P(\boldsymbol{Y}, \boldsymbol{X} \mid \theta,) =& P(\boldsymbol{X}) P(\boldsymbol{Y} \mid \boldsymbol{X}, \theta) \\ =& P(\boldsymbol{X}) \prod_{i \in S} P(Y_i \mid X_i, \theta) \\ \approx& \prod_{i \in S} P(X_i \mid \boldsymbol{X}_{N(i)}) \prod_{i \in S} P(Y_i \mid X_i, \theta). \end{aligned} \tag{2.12}$$

Thus, we have the marginal proability distribution of $\boldsymbol{Y}$ given parameters θ:

$$\begin{aligned} P(\boldsymbol{Y} \mid \theta) =& \sum_{\boldsymbol{X}} P(\boldsymbol{Y}, \boldsymbol{X} \mid \theta) P(\boldsymbol{X}) \\ \approx& \prod_{i \in S} \sum_{X_i} P(X_i \mid \boldsymbol{X}_{N(i)}) P(Y_i \mid X_i, \theta) \\ =& P(\boldsymbol{Y} \mid \theta, \boldsymbol{X}_{N(i)}). \end{aligned} \tag{2.13}$$

Such model is called the Hidden Markov Random Field (HMRF) Model. By assuming that the probability of Y_i given X_i follows a Gaussian Distribution with parameters $\theta = \{\mu_{X_i}, \sigma_{X_i}\}$, we have:

$$P(Y_i \mid X_i = l, \theta) = \frac{1}{\sqrt{2\pi}\sigma_l} exp\{-\frac{(y-\mu_l)^2}{2\sigma_l^2}\}. \tag{2.14}$$

We call this specific case as the Gaussian Hidden Markov Random Field (GHMRF) Model:

$$P(Y_i \mid \theta, \boldsymbol{X}_{N(i)}) \approx \prod_{i \in S} \sum_{X_i = l} P(X_i = l \mid \boldsymbol{X}_{N(i)}) \frac{1}{\sqrt{2\pi}\sigma_l} exp\{-\frac{(y-\mu_l)^2}{2\sigma_l^2}\}. \tag{2.15}$$

2.3.1.2 MAP Classification

In the image segmentation field, we need to recover the underlying MRF X, which is to maximize the $P(\boldsymbol{X} \mid \boldsymbol{Y})$, where Y refers to the observable intensity of a pixel of an image. This turns to find a maximizer of $P(\boldsymbol{X} \mid \boldsymbol{Y})P(\boldsymbol{X})$, that is, to seek a labeling $\hat{X}$ of an image, an estimate of the true labeling X, according to the MAP criterion

$$\hat{\boldsymbol{X}} = argmax_{\boldsymbol{X} \in \mathcal{X}} \big\{P(\boldsymbol{Y} \mid \boldsymbol{X}) P(\boldsymbol{X})\big\}. \tag{2.16}$$

By assuming that the pixel intensity Y_i follows a Gaussian distribution with parameters $\theta = \{\mu_{X_i}, \sigma_{X_i}\}$, and the conditional independence of Y, we have the joint likelihood probability of X and Y given parameter θ:

$$P(\boldsymbol{Y} \mid \boldsymbol{X}, \theta) = \prod_{i \in S} P(Y_i \mid X_i, \theta) = \prod_{i \in S} [\frac{1}{\sqrt{2\pi}} exp\{-\frac{(Y_i - \mu_{X_i})^2}{2(\sigma_{X_i})^2} - log(\sigma_{X_i})\}] \quad (2.17)$$

Therefore, incorporating with (2.8), we have:

$$P(\boldsymbol{Y} \mid \boldsymbol{X}, \theta) = \frac{1}{Z'} exp\{-U(\boldsymbol{Y} \mid \boldsymbol{X}, \theta)\}. \quad (2.18)$$

We then have:

$$U(\boldsymbol{Y} \mid \boldsymbol{X}, \theta) = \sum_{i \in S} U(Y_i \mid X_i, \theta) = \sum_{i \in S} [\frac{(Y_i - \mu_{X_i})^2}{2(\sigma_{X_i})^2} + log(\sigma_{X_i})] \quad (2.19)$$

and the constant normalization term $Z' = (2\pi)^{\frac{N}{2}}$, and

$$U(\boldsymbol{X} \mid \boldsymbol{Y}, \theta) = U(\boldsymbol{Y} \mid \boldsymbol{X}, \theta) + U(\boldsymbol{X}) + constant. \quad (2.20)$$

Therefore, the MAP estimation turns to become:

$$\hat{\boldsymbol{X}} = argmin_{\boldsymbol{X} \in \mathcal{X}} \{U(\boldsymbol{Y} \mid \boldsymbol{X}, \theta) + U(\boldsymbol{X})\}. \quad (2.21)$$

Define

$$V_c(X_i, X_j) = \frac{1}{\Psi}(1 - \delta_{X_i, X_j}), \quad (2.22)$$

where

$$\delta_{X_i, X_j} = \begin{cases} 0, X_i \neq X_j, \\ 1, X_i = X_j. \end{cases} \quad (2.23)$$

We get the full conditional distributions for binary MRF:

$$P(X_i = l \mid X_{-i}) = \frac{e^{\Psi m_{il}}}{e^{\Psi m_{i,0}} + e^{\Psi m_{i,1}}}, \tag{2.24}$$

where $l = 0, 1$, and $m_{i,1}$ & $m_{i,0}$ are the number of neighbors of s_i that are equal to 1, and 0, respectively. We take $\Psi = 2$ for convenience. Given a known estimate of X as $X^{(t)}$, we have:

$$X_i^{(t+1)} = argmin_{l \in L} \left\{ U(Y_i \mid l, \theta) + \sum_{j \in N(i)} V_c(l, X_j^{(t)}) \right\}. \tag{2.25}$$

In order to seek the MAP estimator, we need to know the estimates of the parameters of θ.

2.3.1.3 EM Algorithm

The EM Algorithm is one of the most widely used methods to solve such problems. The E-step is:

$$Q(\theta \mid \theta^{(t)}) = E[logP(\boldsymbol{Y}, \boldsymbol{X} \mid \theta) \mid \boldsymbol{Y} = \boldsymbol{y}], \tag{2.26}$$

and the M-step is:

$$\theta^{(t+1)} = argmax_\theta \{Q(\theta \mid \theta^{(t)})\}. \tag{2.27}$$

In our application, the EM Algorithm incorporated with MAP Estimation is stated as follows:

- Setting an initial value of θ, $\theta^{(0)}$, and an initial configuration of $\boldsymbol{X}$, $\tilde{\boldsymbol{X}}^{(0)}$, to start.

- Calculating the likelihood distribution $P^{(t)}(Y_i \mid X_i, \theta_{X_i})$.

- Estimating

$$\boldsymbol{X}^{(t)} = argmin_{\boldsymbol{X}\in\chi}\left\{U(\boldsymbol{Y} \mid \boldsymbol{X}, \theta^{(t)}) + U(\boldsymbol{X})\right\}. \tag{2.28}$$

  Such $\boldsymbol{X}^{(t)}$ can be found by using MAP Estimation.

- Calculating the posterior distribution:

$$\begin{aligned} P^{(t)}(\boldsymbol{X} = l \mid \boldsymbol{Y}, \theta^{(t)}) =& \prod_{i\in S}\{\frac{P(Y_i \mid X_i = l, \theta^{(t)})P(X_i = l)}{P^{(t)}(Y_i \mid \theta^{(t)})}\} \\ \approx& \prod_{i\in S}\{\frac{P(Y_i \mid X_i = l, \theta^{(t)})P(X_i = l \mid \tilde{\boldsymbol{X}}^{(t)}_{N(i)})}{\sum_{X_i=l} P(Y_i \mid X_i = l, \theta^{(t)})P(X_i = l \mid \tilde{\boldsymbol{X}}^{(t)}_{N(i)})}\} \\ =& P^{(t)}(\boldsymbol{X} = l \mid \boldsymbol{Y}, \tilde{\boldsymbol{X}}^{(t)}_{N(i)}, \theta^{(t)}). \end{aligned} \tag{2.29}$$

We have:

$$Q(\theta \mid \theta^{(t)}) \approx \sum_{i\in S}\sum_{X_i=l} P^{(t)}(X_i = l \mid Y_i, \boldsymbol{X}_{N(i)}, \theta^{(t)}) log P(Y_i \mid X_i = l, \theta). \tag{2.30}$$

Updating the parameters:

$$\mu_l^{(t+1)} = \frac{\sum_{i\in S} P^{(t)}(X_i = l \mid Y_i, \theta^{(t)})Y_i}{\sum_{i\in S} P^{(t)}(X_i = l \mid Y_i, \theta^{(t)})}, \tag{2.31}$$

$$(\sigma_l^{(t+1)})^2 = \frac{\sum_{i\in S} P^{(t)}(X_i = l \mid Y_i, \theta^{(t)})(Y_i - \mu_l^{(t+1)})^2}{\sum_{i\in S} P^{(t)}(X_i = l \mid Y_i, \theta^{(t)})}. \tag{2.32}$$

According to the above updating, we can obtain classified configuration, $\hat{X}$. Hence, we can obtain a segmented image based on the classification of each pixel.

We transform an RGB image to a weighted relative green ratio $(2G - R - B)/(R + G + B)$ as before. In order to use this framework, an initial classification is needed. It is set using k-means clustering, which is shown in Figure 2.7 (b). From the figure we can see that there are some noise around the plant and there are some gaps on its leaves, which cannot be used as a well segmentation result. Figure 2.7 (c) shows the segmentation result obtained by HMRF-EM method using Figure 2.7 (b) as an initial value. As we can see, most noise has been eliminated and the plant body still retains enough information. In order to remove the noise left in Figure 2.7 (c), we can use a Morphological closing followed by a Morphological opening with a 3-by-3 square structural element. Figure 2.7 (d) shows the final result of this procedure.

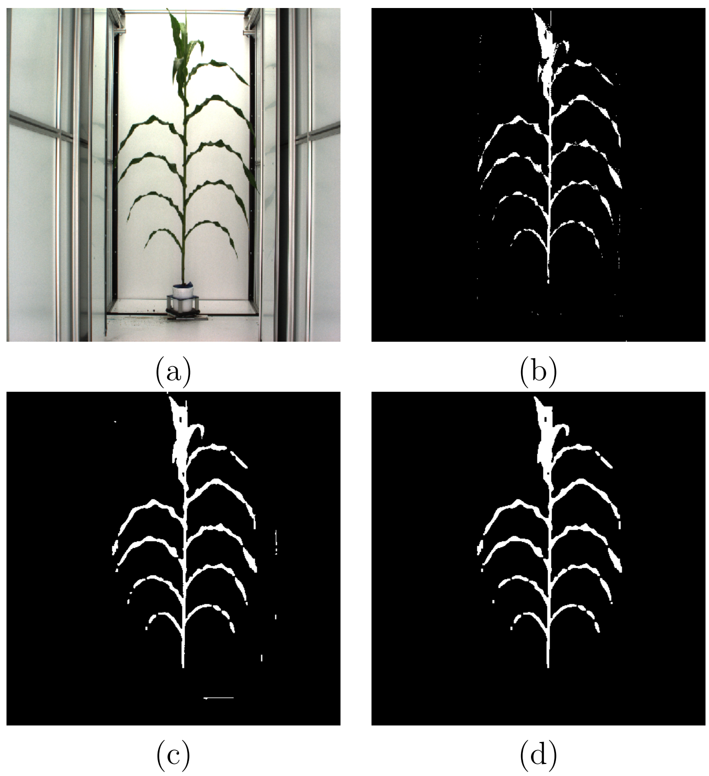

Figure 2.7: (a). Original image (b). Initial classification using k-means (c). Segmentation result using HMRF-EM (d). Applying Morphological closing and opening to (c)