OTSU THRESHOLDING

Consider a histogram and its zeroth and first moments, p0 and p1:

A histogram is represented by its counts, c(i) for each graylevel from i=0 to 255 (for an 8-bit histogram).

If we normalize the counts by the total pixels in the image or count total of the histogram, CT, we have p(i)=c(i)/CT, where,

$$1 = \sum_{i=0}^{255} p(i)$$

Thus p(i) is a fraction and represents a “probability” whose sum is unity.

Now the zeroth and first histogram moments are:

$$p0 = \sum_{i=0}^{255} p(i) \quad \text{and} \quad p1 = \sum_{i=0}^{255} i * p(i)$$

In IM, the histograms have only non-empty bins, thus we change this to:

$$1 = \sum_{i=0}^{lastbin} p(i) \quad \text{and} \quad p0 = \sum_{i=0}^{lastbin} p(i) \quad \text{and} \quad p1 = \sum_{i=0}^{lastbin} p(i) * g(i)$$

where i = bin number from 0 to lastbin = (numbins - 1)

In the Otsu thresholding method, one seeks to maximize the Between Class Variance of the foreground (bright) data and background (dark) data. That is the data above and below some threshold T. The equation describing the Between Class Variance is:

(1) $BCV = Nb*(Mb - M)^2 + Na*(Ma - M)^2$

Where

Nb and Na are the cumulated normalized counts below the threshold and the cumulated normalized counts above the threshold, respectively.

MB and Ma are the mean graylevel values from the cumulated bins below the threshold and from the cumulated bin above the threshold, respectively.

M is the mean graylevel from the image or complete histogram.

Thus,

(2) $Nb = p0b = \sum_{i=0}^{T} p(i)$ and $Na = p0a = \sum_{i=T+1}^{lastbin} p(i)$

(3) $Gb = p1b = \sum_{i=0}^{T} p(i) * g(i)$ and $Ga = p1a = \sum_{i=T+1}^{lastbin} p(i) * g(i)$

where Gb and Ga are normalize cumulated graylevels below and above the threshold, respectively.

(4) $Mb = \frac{P1b}{P0b} = \frac{Gb}{Nb}$ and $Ma = \frac{P1a}{P0a} = \frac{Ga}{Na}$

The reason that we have to divide by Nb and Na to get Mb and Ma is that the means are computed by the sums divided by ONLY those pixels in each class. But the Nb, Na, Gb and Ga are sums divided by the total count. Thus to factor out the total counts and have result that are divide by only the counts in the class we need to divide Gb by Nb and Ga by Na.

The mean of the image (mean of complete histogram) can then be expressed as:

(5) $M = Nb*Mb + Na*Ma$

Also

(6) $1 = Nb + Na$

Now we would like to express BCV from equation (1) in a more efficient formula.

So if we substitute equation (5) into equation (1) and simplify (with a bit of effort), we get

(7) $BCV = Na*Nb*(Ma - Mb)^2$

This basically says that BCV is the difference between the means of the two classes and thus we are attempting to find the threshold that maximizes the separation of these two means.

Equations (7) and be made more efficient by substituting equation (4) to cast this using Gb and Gb instead of Ma and Mb, thus avoiding the extra divides. We therefore get

(8) $BCV = (Ga*Nb - Gb*Na)^2/(Na*Nb)$

However, we can alternately simplify equation (7) to completely avoid Mb and therefore the need for the second set of cumulative moments, Ga and Na. We do this by first expressing equation (5) as

(9) $Ma = (M - Nb*Mb)/Na$

Now, if we substitute equation (9) into equation (1) and simplify (with a bit of effort), we get

(10) $BCV = (M - Mb)^2 * Nb/(1-Nb)$

where we also have used Na = (1–Nb) from equation (6).

But we can take this even further. If we substitute into equation (10), Mb = Gb/Nb from equation (4) and simplify, we get

(11) $BCV = (Nb*M - Gb)^2/(Nb*(1-Nb))$

This is the most efficient way to express BCV as it only needs the one set of below threshold cumulative arrays, Nb and Gb along with the global mean.