

Μηχανική Μάθηση: Εργασία 2

Πρόδρομος Μαλακασιώτης

Η άσκηση αφορά την υλοποίηση του αλγορίθμου EM για τη μεγιστοποίηση της πιθανοφάνειας σε μία μίξη Gaussian κατανομών της μορφής:

$$p(\mathbf{x}) = \sum_{k=1}^K \pi_k \prod_{d=1}^D \frac{1}{\sqrt{2\pi\sigma_k^2}} e^{-\frac{1}{2\sigma_k^2}(x_d - \mu_{kd})^2} \quad (1)$$

Οι εξισώσεις και οι επαναλήψεις του EM περιγράφονται στην 9η διάλεξη. Η μόνη διαφορά είναι ότι, λόγω της μορφής των παραπάνω Gaussian συνιστωσών κατανομών που έχουν σφαιρικό πίνακα συμμεταβλητότητας, οι εξισώσεις ενημέρωσης των πινάκων σ_k απλοποιούνται σε:

$$[\sigma_k^2]^{\text{new}} = \frac{\sum_{n=1}^N \sum_{d=1}^D \gamma(z_{nk})(x_{nd} - \mu_{kd}^{\text{new}})^2}{D \sum_{n=1}^N \gamma(z_{nk})}, \quad k = 1, \dots, K \quad (2)$$

Ο υλοποιημένος αλγόριθμος στη συνέχεια θα χρησιμοποιηθεί για την κατάτμηση και συμπίεση εικόνων (βλ. Ενότητα 9.1.1 του βιβλίου). Ο αλγόριθμός θα εφαρμοστεί για διάφορες τιμές του K (π.χ., $K = 1, 2, 4, 8, 16, 32, 64$) στην εικόνα `im.jpg` που βρίσκεται στο `e-class`. Για κάθε τιμή του K θα οπτικοποιηθεί η κατατμημένη εικόνα (ομοίως με το Σχήμα 9.3 του βιβλίου) και θα υπολογιστεί το σφάλμα ανακατασκευής της εικόνας βάσει της σχέσης:

$$\text{error} = \frac{1}{N} \|\mathbf{x}_i^{\text{true}} - \mathbf{x}_i^{\text{r}}\|^2 \quad (3)$$

όπου \mathbf{x}_i^{r} είναι η τιμή του pixel που προβλέπεται από τη μίξη, δηλαδή η τιμή του μέσου $\boldsymbol{\mu}_k$ για το οποίο η αντίστοιχη εκ των υστέρων κατανομή γz_k είναι η μέγιστη, ενώ $\mathbf{x}_i^{\text{true}}$ είναι η πραγματική τιμή του pixel.