Διαδικασια

Σκοπός της άσκησης είναι η χρήση των τεχνικών:

- Αποσύνθεσης Ιδιαζουσών Τιμών (Singular Value Decomposition (SVD)) και
- Ανάλυσης Κύριων Συνιστωσών (Principal Component Analysis (PCA)).

που είναι πολύ γνωστές τεχνικές αποσύνθεσης μητρώων από το μάθημα της Γραμμικής Άλγεβρας και χρησιμοποιούνται στην επεξεργασία σημάτων.

Στην συγκεκριμένη άσκηση οι παραπάνω τεχνικές θα εφαρμοστούν για την εξαγωγή του φόντου από τα πλαίσια (frames) μίας ακολουθίας εικόνων (video). Συγκεκριμένα θα δοθούν τρία αρχεία mat (αρχεία που περιέχουν δεδομένα για να επεξεργαστείτε στην Matlab) που περιέχουν τρεις (3) διαφορετικές σκηνές που διαπραγματεύονται διαφορετικά πρακτικά προβλήματα.

Για την εφαρμογή των παραπάνω τεχνικών θα πρέπει να αλλάξει η αναπαράσταση των δεδομένων ώστε να περιγράφονται από ένα μητρώο που θα έχει ως γραμμές του τα εικονοστοιχεία ενός πλαισίου της ακολουθίας. Έστω $\mathbf{X}$ το μητρώο διαστάσεων $T \times N$ που αναπαριστά ολόκληρη την ακολουθία εικόνων, όπου N είναι ο αριθμός των εικονοστοιχείων ενός πλαισίου και T ο αριθμός των πλαισίων που θα επεξεργαστούμε.

1 Η SVD του μητρώου $\mathbf{X}^T$ εκφράζεται ως:

$$\mathbf{X}^T = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$$

όπου $\mathbf{U}$ είναι ένα $N \times N$ μητρώο, το $\mathbf{\Sigma}$ είναι ένα $N \times T$ θετικό διαγώνιο μητρώο που περιέχει τις ιδιάζουσες τιμές του μητρώου και το $\mathbf{V}$ είναι ένα $T \times T$ μητρώο.

2 Εφαρμόστε SVD στα πρώτα 800 πλαίσια των τριών σκηνών και κρατώντας τις πρώτες 10 ιδιάζουσες τιμές μηδενίζοντας τις υπόλοιπες ανακτήστε την ακολουθία.

   1. Τι παρατηρείτε ότι έχει αλλάξει στη σκηνή;
   2. Η συμπεριφορά στα τρία βίντεο είναι η ίδια; Αν όχι καταγράψτε τις διαφορές.

3 Επαναλάβατε το προηγούμενο βήμα και κρατήστε μόνο τις πρώτες 3 ιδιάζουσες τιμές.

4 Επαναλάβατε τα προηγούμενα δύο βήματα αλλά αυτή τη φορά να γίνετε η επεξεργασία κατά ορμαθούς των 100 πλαισίων. Καταγράψετε όλες τις διαφορές που παρατηρείτε.

5 Για να μπορέσετε να χρησιμοποιήσετε την τεχνική της ανάλυσης κύριων συνιστωσών, αφαιρέστε την αριθμητική μέση τιμή από κάθε στήλη του μητρώου $\mathbf{X}$. Ας συμβολίσουμε με $\mathbf{X}_o$ το μητρώο που προέκυψε μετά την εφαρμογή της παραπάνω διαδικασίας.

6 Για να υπολογίσουμε την Ανάλυση Κύριων Συνιστωσών του μητρώου, δημιουργούμε το μητρώο $\mathbf{X}_o^T\mathbf{X}_o$ μεγέθους $N \times N$ και κάνουμε Αποσύνθεση Ιδιαζουσών Τιμών.

7 Ακολουθώντας τα παραπάνω, επαναλάβετε όλα τα παραπάνω ερωτήματα, εφαρμόζοντας τώρα αντί της SVD PCA και καταγράψτε τα συμπεράσματά σας.

Για τον υπολογισμό της SVD να χρησιμοποιήσετε την ομώνυμη συνάρτηση της Matlab.

Ιδιοπροσωπα-(Eigenfaces)

Έστω $\mathbf{c}$ ένα διάνυσμα διάστασης $NM \times 1$ που έχει προκύψει από μια $N \times M$ εικόνα I. Θέλουμε να αναπαραστήσουμε το διάνυσμα $\mathbf{c}$ σε χώρο μικρότερης διάστασης. Συγκεκριμένα αν:

$$\mathbf{v} = \mathbf{c} - \mathbf{m}_c$$

με $\mathbf{m}_c$ να παριστά την "*μέση εικόνα*" τότε, το παραπάνω διάνυσμα αποκλίσεων μπορεί να εκφραστεί ως ακολούθως:

$$\mathbf{v} = \sum_{i=1}^{k} \alpha_i \mathbf{q}_i \; k \ll NM$$

όπου $\alpha_i, \; \mathbf{q}_i, \; i = 1, 2, \cdots, k$ κατάλληλες βαθμωτές και κατάλληλα ορθοκανονικά διανύσματα τα οποία θα ορίσουμε στη συνέχεια.

Διαδικασια Εκπαιδευσης

1. Έστω $I_1, I_2, ..., I_K$ οι εικόνες που θα χρησιμοποιήσουμε στη φάση εκπαίδευσης του συστήματος. Οι εικόνες αυτές θα πρέπει να έχουν κοινό κέντρο και το ίδιο μέγεθος.

2. Αναπαριστούμε την κάθε εικόνα ως ένα διάνυσμα $\mathbf{c}_i$

3. Υπολογίζουμε τη μέση εικόνα

$$\mathbf{c}_m = \frac{1}{K} \sum_{i=1}^{K} \mathbf{c}_i$$

4. Αφαιρούμε από κάθε διάνυσμα τη μέση εικόνα:

$$\mathbf{v}_i = \mathbf{c}_i - \mathbf{c}_m.$$

5. Υπολογίζουμε το ακόλουθο $NM \times NM$ μητρώο συνδιασπορών:

$$C = \frac{1}{K} \sum_{i=1}^{K} \mathbf{v}_i \mathbf{v}_i^t = VV^T$$

   όπου $V = [\mathbf{v}_1 \, \mathbf{v}_2 \, ... \, \mathbf{v}_K]$.

6. Θέλουμε τώρα να υπολογίσουμε τα ιδιοδιανύσματα $\mathbf{q}_i \, i = 1, 2, ..., NM$ του C. Όμως ο πίνακας C είναι πολύ μεγάλος και ο απευθείας υπολογισμός των ιδιοδιανυσμάτων του έχει πολύ μεγάλη πολυπλοκότητα.
   Αντί αυτού, θεωρούμε το $K \times K$ μητρώο $\bar{C} = V^T V$ και υπολογίζουμε τα ιδιοδιανύσματά του $\mathbf{p}_i$. Εύκολα μπορούμε να αποδείξουμε (δείξτε το) ότι ο C και ο

C_M έχουν τις ίδιες ιδιοτιμές, ενώ οποιοδήποτε από τα K ιδιοδιανύσματα $\mathbf{q}_i,\ i = 1, 2, \cdots, K$ του C προκύπτει από το αντίστοιχο ιδιοδιάνυσμα $\mathbf{p}_i,\ i = 1, 2, \cdots, K$ του $\bar{C}$ ως:

$$\mathbf{q}_i = \frac{1}{\lambda_i} V \mathbf{p}_i,\ i = 1, 2, \cdots, K.$$

7. Διατηρούμε τα k ($k \leq K$) ιδιοδιανύσματα που αντιστοιχούν στις k μεγαλύτερες ιδιοτιμές, τα οποία ονομάζονται ιδιοπρόσωπα, και δημιουργούμε το ακόλουθο $NM \times k$ μητρώο:

$$Q_k = [\mathbf{q}_1\ \mathbf{q}_2 \cdots \mathbf{q}_k],\ k \leq K.$$

8. Χρησιμοποιώντας αυτό το μητρώο, μπορούμε τώρα να εκφράσουμε σε συμπιεσμένη μορφή κάθε διάνυσμα αποκλίσεων $\mathbf{v}_i$, που καθένα σχετίζεται με την αντίστοιχη εικόνα $\mathbf{c}_i$, ως:

$$\mathbf{w}_i = Q_k^T \mathbf{v}_i,\ i = 1, 2, \cdots, K$$

όπου $\mathbf{w}_i$ ένα διάνυσμα μήκους k που κάθε στοιχείο του εκφράζει (ισούται με) την προβολή της απόκλισης της εικόνας στο αντίστοιχο ιδιοδιάνυσμα του υπόχωρου $\mathbb{R}^k$ του $\mathbb{R}^{NM}$.

Ερωτηση: Αν θεωρήσουμε ότι κάθε εικόνα είναι μεγέθους 512×512 και το πλήθος των εικόνων που χρησιμοποιούμε κατά την διαδικασία της εκπάιδευσης είναι 64, δώστε μία εκτίμηση του συντελεστή συμπίεσης των δεδομένων σας.

Αναγνωριση με χρηση Ιδιοπροσωπων

Για μια εικόνα $\mathbf{c}$, σε διανυσματική μορφή, που έχει κοινό κέντρο με τις εικόνες του συνόλου εκπαίδευσης, η διαδικασία αναγνώρισης είναι η ακόλουθη:

1. Υπολογίζουμε το $\mathbf{v} = \mathbf{c} - \mathbf{c}_m$

2. Το διάνυσμα αυτό το προβάλουμε στον υπόχωρο $\mathbb{R}^k$, δηλαδή:

$$\mathbf{w} = Q_k^T \mathbf{v}$$

Είναι φανερό ότι το διάνυσμα:

$$\hat{\mathbf{v}} = Q_k \mathbf{w}$$

εκφράζει την ανακατασκευή (προσέγγιση) του διανύσματος $\mathbf{v}$.

3. Υπολογίζουμε την απόσταση:

$$e_{l^\star} = \min_{l \in \{1,2,\cdots,K\}} \|\mathbf{w}_l - \mathbf{w}\|_p \tag{8}$$

όπου με $\|\cdot\|_p$ συμβολίζουμε την l_p στάθμη του ορίσματος.

Στους ταξινομητές *ελάχιστης απόστασης* σημαντικό ρόλο έχει η επιλογή της συνάρτησης απόστασης που θα χρησιμοποιήσουμε. Στις παραπάνω συναρτήσεις απόστασης η Ευκλείδεια αντιστοιχεί στην επιλογή του $p = 2$.

Η εικόνα $\mathbf{c}$ ταξινομείται ως η εικόνα $\mathbf{c}_{l^\star}$ από το σύνολο εκπαίδευσης.

4. Υποθέστε ότι κανονικοποιούμε τα διανύσματα $\mathbf{w}_i,\ i = 1, 2, \cdots, k$ [4], όπως και το προς ταξινόμηση διάνυσμα $\mathbf{w}$, και στη συνέχεια λύνουμε το πρόβλημα βελτιστοποίησης (8). Αποδείξτε ότι αν δημιουργήσουμε το $K \times k$ μητρώο:

$$W = \begin{bmatrix} \bar{\mathbf{w}}_1^t \\ \bar{\mathbf{w}}_2^t \\ \vdots \\ \bar{\mathbf{w}}_K^t \end{bmatrix}$$

το παραπάνω πρόβλημα βελτιστοποίησης, είναι ισοδύναμο με το ακόλουθο:

$$\rho_{l^\star} = \max_{l \in \{1,2,\cdots,K\}} \mathbf{p} \tag{9}$$

όπου το μήκους K διάνυσμα $\mathbf{p}$ ορίζεται ως ακολούθως:

$$\mathbf{p} = W\bar{\mathbf{w}}.$$

Διαδικασια

Στο πλαίσιο της άσκησης αυτής θα χρησιμοποιήσετε μία βάση εικόνων που θα την βρείτε στο συμπιεσμένο αρχείο eigenfaces.zip.

1. Αποσυμπιέστε το αρχείο.

2. Χρησιμοποιείστε τις εικόνες που υπάρχουν στον φάκελο training και ακολουθείστε την διαδικασία εκπαίδευσης που περιγράψαμε σε προηγούμενη ενότητα, για την εκπαίδευση του συστήματος.

3. Χρησιμοποιείστε τις εικόνες που υπάρχουν στον φάκελο testing και ακολουθείστε την διαδικασία αναγνώρισης που περιγράψαμε στην προηγούμενη ενότητα.

4. Χρησιμοποιείστε κατάλληλες ποσότητες από αυτές που ορίζονται στο σύνδεσμο $https://en.wikipedia.org/wiki/Sensitivity_and_specificity$, για να καταγράψετε την απόδοση των τεχνικών που εξετάσατε.

[4]Το κανονικοποιημένο διάνυσμα $\mathbf{x}$ θα το συμβολίζουμε με $\bar{\mathbf{x}}$.