



ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ

**Αναγνώριση Προτύπων και Ειδικά Θέματα Όρασης Υπολογιστών**

**Σύστημα Ανάκτησης Εικόνων**

**2η Εργασία**  
**Εμμανουήλ Παπαδημητρίου**  
**ΑΜ: mcse19021**

*Μεταπτυχιακό Πρόγραμμα "Επιστήμη και Τεχνολογία της Πληροφορικής και των Υπολογιστών"*

Αρχικά να αναφερθεί, ότι υπάρχουν όλα τα **\*.mat** αρχεία που έχουν γίνει generated για τις τρεις περιπτώσεις του kmeans για κάθε μέθοδο εξαγωγής των χαρακτηριστικών

Ειδικά για τις περιπτώσεις που το **kmeans** έχει τιμή 50 και 100, έκανε πάρα πολύ ώρα να εκτελεσθεί και εφόσον απαιτήθηκαν δοκιμές μόνο το τρέξιμο του αλγορίθμου αυτό κατέστησε αναγκαία την αποθήκευσή τους.

Θα βρείτε λοιπόν τους φακέλους:

- HOG
  - Για kmeans = 20 στον φάκελο **hog/kmeans20**
  - Για kmeans = 50 στον φάκελο **hog/kmeans50**
  - Για kmeans = 100 στον φάκελο **hog/kmeans100**
- SIFT
  - Για kmeans = 20 στον φάκελο **sift/kmeans20**
  - Για kmeans = 50 στον φάκελο **sift/kmeans50**
  - Για kmeans = 100 στον φάκελο **sift/kmeans100**

Όμως, στον τελικό φάκελο **ergasia2**, έχουν τροποποιηθεί τα ονόματα των αρχείων **\*.mat**, για να μπορείτε να τα εκτελέσετε τα αρχεία χωρίς αλλαγές. Συγκεκριμένα, το αρχείο **all\_classes.mat** που αντιστοιχεί σε όλα τα αρχεία, έχει ονομαστεί σύμφωνα με τα αρχεία και τις εκτελέσεις.

Έχει γίνει δυναμική υλοποίηση που προσαρμόζεται στις δύο μεθόδους εξαγωγής χαρακτηριστικών (HOG και SIFT).

## Πειραματικά Αποτελέσματα

i)

Υπολογισμός Average Precision (AP) για την query εικόνα **s10\_01.pgm** (1η εικόνα 10ου προσώπου)

	HOG Patches No reduction Cosine similarity	HOG Patches No reduction Voting	SIFT features No reduction Cosine similarity	SIFT features No reduction Voting
KMeans K=20	0.1561	0.1746	0.1807	0.1337
KMeans K=100	0.3762	0.2041	0.1887	0.1281

## Ερώτηση 1

Παρουσίαση TOP 5 αποτελεσμάτων για κάθε περίπτωση. Τα αποτελέσματα είναι αποθηκευμένα στον πίνακα top5 διαστάσεων 5x2, με την πρώτη στήλη να περιέχει το score για κάθε περίπτωση ανάκτησης (cosine και voting) και η δεύτερη στήλη περιέχει την θέση που βρίσκεται στον πίνακα, δηλαδή στην ουσία την εικόνα.

- HOG, Kmeans = 20, Cosine

1	91
0.8292	360
0.8240	356
0.8206	278
0.8204	309

- HOG, Kmeans = 20, Voting

14004	91
12918	276
12768	288
12679	257
12620	356

- SIFT, Kmeans = 20, Cosine

1	91
0.8597	249
0.8539	373
0.8417	121
0.8411	22

- SIFT, Kmeans = 20, Voting

9056	91
8581	142
8578	130
8403	358
8280	179

ii)

Υπολογισμός Mean Average Precision (MAP) για όλες τις εικόνες της βάσης

	HOG Patches No reduction Cosine similarity	HOG Patches No reduction Voting	SIFT features No reduction Cosine similarity	SIFT features No reduction Voting
KMeans K=20	0.2969	0.2020	0.1820	0.1624
KMeans K=100	0.2002	0.1633	0.1775	0.1572

## Ερώτηση 2

Έχει ζητηθεί το αρχείο **final\_k20\_hog\_cosine.m** αλλά, έχει δημιουργηθεί και το αντίστοιχο αρχείο με τον αντίστοιχο υπολογισμό του Mean Average Precision για την Voting περίπτωση με όνομα **final\_k20\_hog\_voting.m**. Για σαφήνεια και παρουσίαση της δημιουργίας και να είναι ξεκάθαρος ο τρόπος και τα αποτελέσματα.

Τα συμπεράσματα που προκύπτουν είναι ότι η τεχνική HOG, δίνει καλύτερα αποτελέσματα από το SIFT και το Cosine δίνει καλύτερα από το Voting.

Αρχικά παρατηρούμε κοινή συμπεριφορά και του HOG και του SIFT όταν αυξάνεται το πλήθος K των κλάσεων, καθώς και στις δυο περιπτώσεις μειώνεται το MAP από kmeans = 20 σε kmeans = 100.

Παρατηρήθηκε, ότι για το query μας (s10\_01) η αύξηση του kmeans είχε καλύτερα αποτελέσματα (εκτός του SIFT και Voting που είχε μείωση).

Φαίνεται από τα δεδομένα, να ωφελείται η τεχνική του SIFT καθώς έχει λιγότερο ποσοστό μείωσης του MAP από το HOG που έχει αρκετά περισσότερο.

### Ερώτηση 3

Το καλύτερο MAP το πετυχαίνουμε με την τεχνική εξαγωγή χαρακτηριστικών HOG, για τιμή του kmeans ίσον με 20 και με τεχνική υπολογισμού ομοιότητας Cosine.

### Ερώτηση 4

Η καλύτερη περίπτωση είναι η HOG, με kmeans ίσον με 20 και με την τεχνική Cosine, οπότε αυτή η τεχνική θα υλοποιηθεί για τις φωτογραφίες της τάξης.

Πρέπει να σημειωθεί ότι ήταν αδύνατη η δυναμική δημιουργία των relevants, καθώς δεν είχαμε ίσες εικόνες για όλα τα πρόσωπα σε αυτήν την περίπτωση. Στην υπόλοιπη άσκηση είχαμε 10 relevant εικόνες για κάθε πρόσωπο. Οπότε αν δοκιμάσετε να τρέξετε τα final αρχεία με τα πρόσωπα των φοιτητών, δεν θα τρέξει καθώς είναι αδύνατη η δυναμική εκτίμηση των relevants (ίσως με machine learning να γινόταν). Η υλοποίηση έχει γίνει στο αρχείο **custom\_class\_final\_k20\_cosine.m**.

Ακόμα ένα πρόβλημα στο ότι δεν έχουμε ίσα relevant, είναι ότι πρέπει να έχουμε ίσες διαστάσεις για πράξεις πινάκων, άρα αναγκάζομαστε να γεμίσουμε με μηδενικά τα κενά.

**To Mean Average Precision (MAP) βγαίνει 0.20**