

Μάθημα: Υπολογιστική όραση

## Εργασία #1

Ημερομηνία παράδοσης: 12/01/2024

**Άσκηση #1, Κβάντιση:** Θεωρήστε την εικόνα αποχρώσεων του γκρι (grayscale image) “cameraman.bmp”.

Ζητούμενα:

- A. Σχεδιάστε και υλοποιείτε (να μη γίνει χρήση έτοιμων συναρτήσεων κβάντισης) έναν ομοιόμορφο βαθμωτό κβαντιστή μέσου πατήματος με 7, 11, 15 και 19 στάθμες.
- B. Τυπώστε σε ένα ενιαίο σχήμα τη συνάρτηση μετασχηματισμού του κάθε κβαντιστή
- Γ. Εφαρμόστε τον κάθε κβαντιστή στην εξεταζόμενη εικόνα και τυπώστε τα αποτελέσματα σε ένα ενιαίο σχήμα. Τι παρατηρείτε?
- Δ. Υπολογίστε για κάθε περίπτωση το μέσο τετραγωνικό σφάλμα κβάντισης.

**Άσκηση #2, Μετασχηματισμός Fourier:** Θεωρήστε την εικόνα αποχρώσεων του γκρι (grayscale image) “cameraman.bmp”.

Ζητούμενα:

- A. Υπολογίστε τον 2Δ μετασχηματισμό Fourier (DFT) της εικόνας.
- B. Υπολογίστε και τυπώστε το φάσμα πλάτους και φάσης του μετασχηματισμού.
- Γ. Θεωρήστε διαδοχικά μόνο το 20%, 40%, 60%, 80% των συντελεστών (χαμηλότερες συχνότητες) του DFT ανά άξονα και ανακατασκευάστε (με χρήση του αντίστροφου μετασχηματισμού Fourier) την αρχική εικόνα. Τυπώστε σε ένα ενιαίο σχήμα τις ανακατασκευασμένες εικόνες. Τι παρατηρείτε?
- Δ. Υπολογίστε το μέσο τετραγωνικό σφάλμα ανακατασκευής σε κάθε περίπτωση.

**Άσκηση #3, Μετασχηματισμός συνημίτονου:** Θεωρήστε την εικόνα αποχρώσεων του γκρι (grayscale image) “cameraman.bmp”.

Ζητούμενα:

- A. Υλοποιείτε τον ευθύ και τον αντίστροφο 2Δ μετασχηματισμό συνημίτονου (DCT-IDCT) της εικόνας (να μη γίνει χρήση έτοιμων συναρτήσεων υπολογισμού του DCT-IDCT).
- B. Υπολογίστε και τυπώστε το φάσμα του μετασχηματισμού.
- Γ. Θεωρήστε διαδοχικά μόνο το 20%, 40%, 60%, 80% των συντελεστών (χαμηλότερες συχνότητες) του DCT ανά άξονα και ανακατασκευάστε (με χρήση του αντίστροφου μετασχηματισμού συνημίτονου) την αρχική εικόνα. Τυπώστε σε ένα ενιαίο σχήμα τις ανακατασκευασμένες εικόνες. Τι παρατηρείτε?
- Δ. Υπολογίστε το μέσο τετραγωνικό σφάλμα ανακατασκευής σε κάθε περίπτωση.

**Άσκηση #4, Φιλτράρισμα στο πεδίο του χώρου:** Θεωρήστε την εικόνα αποχρώσεων του γκρι (grayscale image) “lenna.bmp”. Προσθέστε στην εικόνα κρουστικό θόρυβο (“salt & pepper”) με πυκνότητα 0.05 (δηλαδή ποσοστό κάλυψης της επιφάνειας της εικόνας ίσο με 5%).



Ζητούμενα:

- A. Τυπώστε την εικόνα πριν και μετά την προσθήκη του θορύβου.
- B. Θεωρήστε διαδοχικά φίλτρα μέσης τιμής διαστάσεων  $3 \times 3$ ,  $5 \times 5$  και  $7 \times 7$  στο πεδίο του χώρου. Εφαρμόστε ξεχωριστά το κάθε φίλτρο στην εικόνα με θόρυβο (χρήση τελεστή συνέλιξης).
- Γ. Τυπώστε σε ένα ενιαίο σχήμα την προκύπτουσα φιλτραρισμένη εικόνα για κάθε περίπτωση χρησιμοποιούμενου φίλτρου και συγκρίνετε με την αρχική εικόνα. Τι παρατηρείτε?

**Άσκηση #5, Φιλτράρισμα στο πεδίο της συχνότητας:** Θεωρήστε την εικόνα αποχρώσεων του γκρι (grayscale image) “lenna.bmp”. Προσθέστε στην εικόνα Γκαουσιανό θόρυβο με μέση τιμή μηδέν και διακύμανση 0.01.

Ζητούμενα:

- A. Τυπώστε την εικόνα πριν και μετά την προσθήκη του θορύβου.
- B. Θεωρήστε διαδοχικά χαμηλοπερατά φίλτρα Butterworth 3<sup>ης</sup>, 5<sup>ης</sup> και 7<sup>ης</sup> τάξης στο πεδίο της συχνότητας. Εφαρμόστε ξεχωριστά το κάθε φίλτρο στην εικόνα με θόρυβο (χρήση ευθύ και αντίστροφου μετασχηματισμού Fourier).
- Γ. Τυπώστε σε ένα ενιαίο σχήμα την προκύπτουσα φιλτραρισμένη εικόνα για κάθε περίπτωση χρησιμοποιούμενου φίλτρου και συγκρίνετε με την αρχική εικόνα. Τι παρατηρείτε? Υπενθυμίζεται ότι το μιγαδικό μέρος της απόκρισης που προκύπτει από την εφαρμογή του αντίστροφου μετασχηματισμού Fourier μπορεί να αγνοηθεί.

**Άσκηση #6, Τμηματοποίηση εικόνας:** Θεωρήστε την έγχρωμη εικόνα “butterfly.jpg”. Θεωρήστε τη μέθοδο τμηματοποίησης ‘Mean shift’ με Γκαουσιανού τύπου πυρήνα με την ακόλουθη απλουστευμένη παραμετροποίηση:

- Μετασχηματισμός της έγχρωμης εικόνας στον χρωματικό χώρο CIELAB (που παρουσιάζει ικανοποιητικά γραμμική αντίληψη της αλλαγής των χρωμάτων). Δηλαδή κάθε εικονοστοιχείο  $c_i(x,y)$  αναπαρίσταται από ένα 3Δ διάνυσμα χρωματικών συνιστωσών  $f_i(x,y)$ .
- Αρχικοποίηση αλγορίθμου: επιλογή ως αρχικών κέντρων  $m_j(x,y)$  των κλάσεων-αντικειμένων σημείων της εικόνας πάνω σε ένα ομοιόμορφο τετραγωνικό πλέγμα διαστάσεων  $30 \times 30$  που να καλύπτει ολόκληρη την επιφάνεια της εικόνας (κατά πλήρη αντιστοιχία με τη διαδικασία της τετραγωνικής δειγματοληψίας). Για κάθε κέντρο κλάσης  $m_j(x,y)$  υπολογίζεται και ένα αντίστοιχο μέσο διάνυσμα mean-shift  $f_j(x,y)$  στον 3Δ χρωματικό χώρο.
- Καθορισμός μεγέθους παραθύρου (2Δ επιφάνεια γύρω από ένα κεντρικό εικονοστοιχείο  $m_j(x,y)$  για το οποίο υπολογίζεται ένα μέσο διάνυσμα mean-shift  $f_j(x,y)$ ): κύκλος με ακτίνα  $r=20$  εικονοστοιχεία (pixels).
- Μέγιστη διαφορά στον χρωματικό χώρο ώστε ένα εικονοστοιχείο  $c_i(x,y)$  να συμμετέχει στον υπολογισμό του μέσου διανύσματος mean-shift  $f_j(x,y)$  για την εξεταζόμενη κλάση  $m_j(x,y)$ :  $d = \|f_j(x,y) - f_i(x,y)\| < 40$ .
- Κατά την επαναληπτική εκτέλεση του αλγορίθμου mean-shift, εάν 2 (ή περισσότερα) κεντρικά στοιχεία  $m_{j1}(x_1,y_1)$  και  $m_{j2}(x_2,y_2)$  συμπίσουν (δηλαδή  $(x_1,y_1) = (x_2,y_2)$ ) σε κάποιο βήμα εκτέλεσης, να συγχωνευτούν σε ένα μοναδικό κεντρικό σημείο (δηλαδή για κάθε συγχώνευση να μειωθεί το συνολικό πλήθος των κλάσεων κατά 1).
- Ο αλγόριθμος mean-shift να εκτελεστεί μέχρι να συγκλίνει (δηλαδή να μην παρατηρηθεί κάποια αλλαγή στα μέσα διανύσματα mean-shift  $f_j(x,y)$  σε δύο διαδοχικά στάδια εκτέλεσης) ή πραγματοποιηθούν  $T=50$  βήματα του αλγορίθμου.

Χρήσιμο υλικό για τον αλγόριθμο mean-shift:

- <https://courses.csail.mit.edu/6.869/handouts/PAMIMeanshift.pdf>
- [https://sbme-tutorials.github.io/2018/cv/notes/8\\_week8.html](https://sbme-tutorials.github.io/2018/cv/notes/8_week8.html)
- <https://www.ipb.uni-bonn.de/html/teaching/photo12-2021/2021-pho1-12-segmentation.pptx.pdf>
- [https://www.inf.tu-dresden.de/content/institutes/ki/is/VORTRAG/Vortrag\\_Huong\\_Nguyen.pdf](https://www.inf.tu-dresden.de/content/institutes/ki/is/VORTRAG/Vortrag_Huong_Nguyen.pdf)

Ζητούμενα:

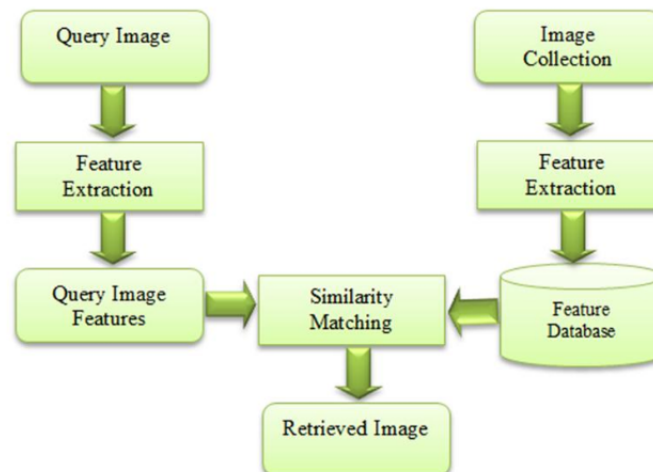
- Α. Υλοποιείτε την παραπάνω έκδοση του αλγορίθμου ‘Mean shift’.
- Β. Εφαρμόστε τον αλγόριθμο τμηματοποίησης στη δοθείσα έγχρωμη εικόνα και τυπώστε το αποτέλεσμα (μάσκα τμηματοποίησης) σε μορφή εικόνας αποχρώσεων του γκρι (κάθε αντικείμενο να αντιστοιχεί σε διαφορετική απόχρωση). Αναφέρατε το συνολικό πλήθος των τελικών αντικειμένων που έχουν ανιχνευτεί. Τι παρατηρείτε?
- Γ. Δοκιμάστε διαφορετικές τιμές για τις παραμέτρους  $r$ ,  $d$  και  $T$ , και υπολογίστε (3-4) ενδεικτικά αποτελέσματα τμηματοποίησης. Τι παρατηρείτε?

**Άσκηση #7, Ανίχνευση ακμών:** Θεωρήστε τις εικόνες αποχρώσεων του γκρι (grayscale image) “butterfly\_g.jpg”, “cameraman.bmp” και “lenna.bmp”.

Ζητούμενα:

- Ε. Υπολογίστε για την εικόνα “butterfly\_g.jpg” τις ακμές με χρήση των ακόλουθων μεθόδων που στηρίζονται στην χρήση πρώτων παραγώγων: Sobel, Roberts, Prewitt και Kirsch. Για τις μεθόδους των Sobel, Roberts και Prewitt να χρησιμοποιηθεί η μέθοδος καθολικής κατωφλίωσης του Otsu, ενώ για τη μέθοδο του Kirsch η βέλτιστη τιμή του κατωφλίου να προσδιοριστεί πειραματικά. Να αξιολογηθούν συγκριτικά ως προς την απόδοσή τους οι μέθοδοι.
- ΣΤ. Υπολογίστε για την εικόνα “butterfly\_g.jpg” τις ακμές με χρήση παραγώγων δευτέρου βαθμού (φίλτρο LoG). Υπολογίστε πειραματικά τις βέλτιστες τιμές των παραμέτρων της διακύμανσης και του κατωφλίου στην υλοποίηση του τελεστή LoG. Παραθέστε ενδεικτικά αποτελέσματα και αντίστοιχο σχολιασμό τους.
- Ζ. Υπολογίστε για την εικόνα “butterfly\_g.jpg” τις ακμές με χρήση της μεθόδου του “Canny” (θεωρώντας την ‘default’ παραμετροποίηση της μεθόδου, όπως παρέχεται από το πρόγραμμα Octave). Σχολιάστε το αποτέλεσμα ανίχνευσης.
- Η. Υπολογίστε για τις εικόνες “cameraman.bmp” και “lenna.bmp” τις ακμές με χρήση των μεθόδων που εφαρμόστηκαν στα βήματα (Α)-(Γ) και αξιολογήστε συγκριτικά την απόδοσή τους (παραθέστε μόνο τη βέλτιστη μάσκα ακμών για κάθε μέθοδο).

**Άσκηση #8, Εξαγωγή χαρακτηριστικών για ανάκτηση εικόνων:** Θεωρήστε τη γενική αρχιτεκτονική ενός συστήματος ανάκτησης εικόνων με βάση το περιεχόμενο (Content-Based Image Retrieval, CBIR) που φαίνεται στο παρακάτω σχήμα. Θεωρήστε και κατεβάστε το δημόσιο σύνολο εικόνων ‘[Flower Image Dataset](#)’, το οποίο περιέχει φωτογραφίες λουλουδιών που ανήκουν στις ακόλουθες δέκα κατηγορίες: α) Bougainvillea, β) Tulips, γ) Orchids, δ) Peonies, ε) Hydrangeas, ζ) Lilies, η) Gardenias, θ) Garden Roses, ι) Daisies και κ) Hibiscus.



### Ζητούμενα:

- Δ. Για κάθε εικόνα του συνόλου δεδομένων (αφού τη μετατρέψετε σε εικόνα αποχρώσεων του γκρι), εξάγετε τα ακόλουθα χαρακτηριστικά: A1) Κανονικοποιημένο ιστόγραμμα φωτεινότητας 256 κελιών (bins) και A2) Κανονικοποιημένο ιστόγραμμα τιμών χαρακτηριστικού υψής LBP 256 κελιών.
- Ε. Υλοποιήστε τις ακόλουθες δύο μετρικές (αν)ομοιότητας μεταξύ δύο διανυσμάτων χαρακτηριστικών  $\mathbf{f}_1(\mathbf{i})$  και  $\mathbf{f}_2(\mathbf{i})$  (μικρότερες τιμές της μετρικής υποδηλώνουν πιο 'όμοια' διανύσματα χαρακτηριστικών): B1)  $L_1 = \sum_i |f_1(i) - f_2(i)|$  και B2)  $L_2 = \sqrt{\sum_i |f_1(i) - f_2(i)|^2}$
- ΣΤ. Επιλέξτε τυχαία 5 εικόνες από διαφορετικές κατηγορίες. Για κάθε μια από αυτές τις εικόνες, θεωρήστε την επιλεγμένη εικόνα ως την εικόνα αναζήτησης (query image) και υπολογίστε/τυπώστε ιεραρχικά τις 10 πιο όμοιες εικόνες στο σύνολο δεδομένων (top-10 retrieval results). Στα top-10 αποτελέσματα δε θα πρέπει να υπάρχει η query image. Υπολογίστε τα top-10 αποτελέσματα για όλους τους δυνατούς συνδυασμούς χαρακτηριστικών (A1-A2) και μετρικών (B1-B2). Σχολιάστε και αξιολογήστε συγκριτικά το σύνολο των αποτελεσμάτων (π.χ. ποιος τύπος χαρακτηριστικού και ποια μετρική οδηγούν στα καλύτερα αποτελέσματα (είτε σε επίπεδο κατηγορίας λουλουδιών είτε συνολικά)?)