

Μάθημα: Ψηφιακή επεξεργασία εικόνας και εφαρμογές

**Εργασία #2: Ανίχνευση ακμών, συμπίεση και εξαγωγή χαρακτηριστικών**

Ημερομηνία παράδοσης: 21/06/2024

**Άσκηση #1, Ανίχνευση ακμών:** Θεωρήστε τις εικόνες αποχρώσεων του γκρι (grayscale image) “bridge.bmp”, “girlface.bmp” και “lighthouse.bmp”.

Ζητούμενα:

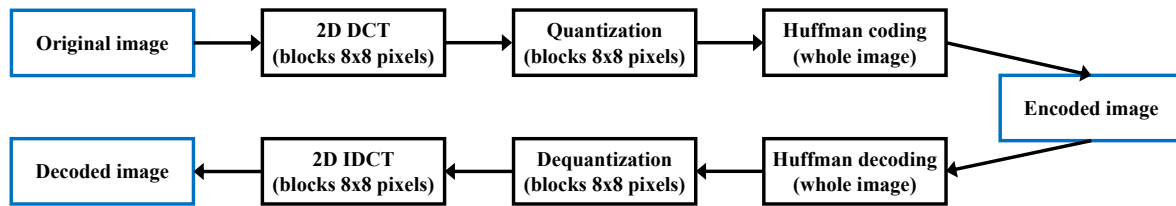
- A. Υπολογίστε για την εικόνα “lighthouse.bmp” τις ακμές με χρήση των ακόλουθων μεθόδων που στηρίζονται στην χρήση πρώτων παραγώγων: Sobel, Roberts, Prewitt και Kirsch. Για τις μεθόδους των Sobel, Roberts και Prewitt να χρησιμοποιηθεί η μέθοδος καθολικής κατωφλίωσης του Otsu, ενώ για τη μέθοδο του Kirsch η βέλτιστη τιμή του κατωφλίου να προσδιοριστεί πειραματικά. Να αξιολογηθούν συγκριτικά ως προς την απόδοσή τους οι μέθοδοι.
- B. Υπολογίστε για την εικόνα “lighthouse.bmp” 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.
- Γ. Υπολογίστε για την εικόνα “lighthouse.bmp” τις ακμές με χρήση της μεθόδου του “Canny” (θεωρώντας την ‘default’ παραμετροποίηση της μεθόδου, όπως παρέχεται από το πρόγραμμα Octave). Σχολιάστε το αποτέλεσμα ανίχνευσης.
- Δ. Υπολογίστε για τις εικόνες “bridge.bmp” και “girlface.bmp” τις ακμές με χρήση των μεθόδων που εφαρμόστηκαν στα βήματα (Α)-(Γ) και αξιολογήστε συγκριτικά την απόδοσή τους και για τις 3 εικόνες της εκφώνησης (παραθέστε μόνο τη βέλτιστη μάσκα ακμών για κάθε μέθοδο).

**Άσκηση #2, Κωδικοποίηση Huffman:** Θεωρήστε τις εικόνες αποχρώσεων του γκρι (grayscale image) “bridge.bmp”, “girlface.bmp” και “lighthouse.bmp”.

Ζητούμενα:

- A. Υλοποιήστε και εφαρμόστε τη μέθοδο κωδικοποίησης (χωρίς απώλειες) Huffman για τις εικόνες “bridge.bmp”, “girlface.bmp” και “lighthouse.bmp”. Να μη γίνει χρήση έτοιμων συναρτήσεων-υλοποιήσεων της μεθόδου Huffman.
- B. Για κάθε εικόνα παραθέστε: 1) Τις κωδικές λέξεις που προκύπτουν (αντιστοιχία τιμών φωτεινότητας με δυαδικές κωδικές λέξεις), 2) Το μέσο μήκος κωδικής λέξης, 3) Την τιμή της εντροπίας, και 4) Το λόγο συμπίεσης που επιτυγχάνεται.
- Γ. Σχολιάστε-αξιολογήστε συγκριτικά την απόδοση του αλγορίθμου (όπως υπολογίστηκε στο βήμα (B)) για τις διάφορες εικόνες.

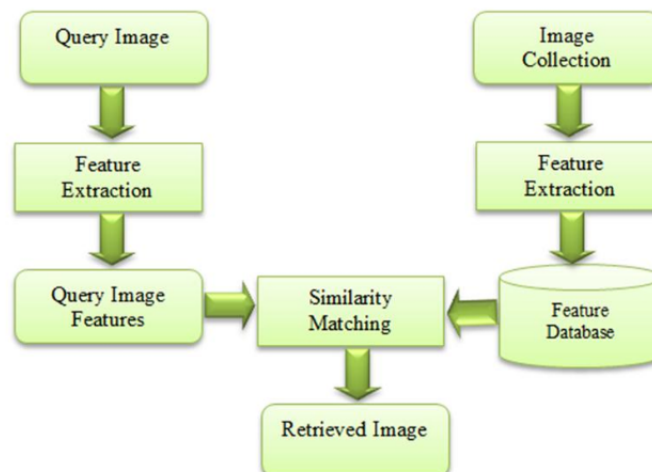
**Άσκηση #3, Συμπίεση εικόνων σύμφωνα με το πρότυπο JPEG:** Θεωρήστε τις εικόνες αποχρώσεων του γκρι (grayscale images) “bridge.bmp”, “girlface.bmp” και “lighthouse.bmp”. Επίσης, θεωρήστε την ακόλουθη παραλλαγή του πρότυπου JPEG για συμπίεση εικόνων (στην παρούσα εργασία θα ληφθούν υπόψη μόνο εικόνες αποχρώσεων του γκρι).



**Ζητούμενα:**

- A. Υλοποιήστε τον κωδικοποιητή και τον αποκωδικοποιητή εικόνων κατά το πρότυπο JPEG (που απεικονίζεται στο παραπάνω σχήμα). Θεωρήστε δύο εναλλακτικές επιλογές για τον πίνακα κβάντισης (συγκεκριμένα τους πίνακες Q10 και Q50).
- B. Κωδικοποιήστε τις εικόνες “bridge.bmp”, “girlface.bmp” και “lighthouse.bmp” (θεωρώντας και τις δύο επιλογές για τον πίνακα κβάντισης). Υπολογίστε το μέσο μήκος κωδικής λέξης (αλγόριθμος Huffman) και το λόγο συμπίεσης που επιτυγχάνεται για κάθε εικόνα σε κάθε περίπτωση.
- Γ. Για τις εικόνες “bridge.bmp”, “girlface.bmp” και “lighthouse.bmp” τυπώστε τις αποκωδικοποιημένες εκδόσεις τους (θεωρώντας διαδοχικά και τις δύο επιλογές για τον πίνακα κβάντισης) και προσδιορίστε την ποσότητα του θορύβου (μετρική Peak signal-to-noise ratio (PSNR)) που εισάγεται σε κάθε περίπτωση.

**Άσκηση #4, 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 ‘Villains - Image Classification’<sup>1</sup>, το οποίο περιέχει φωτογραφίες από χαρακτήρες κόμικ που ανήκουν στις ακόλουθες πέντε κατηγορίες: α) Darth Vader, β) Green Goblin, γ) Joker, δ) Thanos, και ε) Venom.



**Ζητούμενα:**

- A. 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 υφής LBP 256 κελιών.
- B. 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

<sup>1</sup> <https://www.kaggle.com/datasets/rogeriovaz/villains-image-classification?resource=download>



‘όμοια’ διανύσματα χαρακτηριστικών): B1)  $L_1 = \sum_i |f_1(i) - f_2(i)|$  και B2)  $L_2 = \sqrt{\sum_i |f_1(i) - f_2(i)|^2}$

- Γ. 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 5 πιο όμοιες εικόνες στο σύνολο δεδομένων (top-5 retrieval results). Στα top-5 αποτελέσματα δε θα πρέπει να υπάρχει η query image. Υπολογίστε τα top-5 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 χαρακτήρων είτε συνολικά)?)