

Όραση Υπολογιστών: Εργασία 1

Η παρούσα εργασία έχει ως στόχο την ανάλυση της δομής εγγράφων και χρησιμοποιεί δεδομένα εικόνων που περιέχουν σκαναρισμένα έγγραφα και βρίσκονται στον παρακάτω σύνδεσμο :

<https://vc.ee.duth.gr:6960/index.php/s/ICvWKEq6Y0VHTTr5>

Στην παραπάνω συλλογή υπάρχουν δύο είδη εικόνων. Οι πρωτότυπες με την ένδειξη 'original' καθώς και αυτές στις οποίες έχει προστεθεί θόρυβος τύπου "αλατιού και πιπεριού" (salt and pepper noise) με την ένδειξη 'noise'.

Ο κάθε φοιτητής θα πρέπει να αναπτύξει μια ολοκληρωμένη μεθοδολογία και να την υλοποιήσει σε κώδικα, ώστε στις οκτώ εικόνες να επιτύχει τα εξής:

1. **Ανίχνευση όλων των υπο-περιοχών** κειμένου του εγγράφου. Ως υπο-περιοχή ορίζεται ένα τμήμα του εγγράφου με κείμενο, το οποίο διακρίνεται ως προς τη θέση του από τα υπόλοιπα μέρη του εγγράφου (π.χ. αρίθμηση σελίδας, υποσέλιδο, παράγραφος, τίτλος κ.α.). Το πρόγραμμα θα πρέπει να **εμφανίζει και να αποθηκεύει την εικόνα εγγράφου** στην οποία να είναι εμφανείς οι διαφορετικές υπο-περιοχές της. Συγκεκριμένα, για κάθε υπο-περιοχή να σχεδιαστεί ένα **περιβάλλον κουτί** (bounding box) καθώς και ένας **μοναδικός αύξων αριθμός**.
2. Για κάθε μια υπο-περιοχή να μετρηθούν και να εμφανίζονται ως έξοδος του προγράμματος τα ακόλουθα μεγέθη:
 - a. Η **επιφάνεια της υπο-περιοχής που καταλαμβάνεται από κείμενο**, που ορίζεται ως ο αριθμός των εικονοστοιχείων που ανήκουν σε γράμμα (σκουρόχρωμα) και όχι στη σελίδα (ανοιχτόχρωμα)
 - b. Η **επιφάνεια του περιβάλλοντος κουτιού** (bounding box) της υπο-περιοχής.
 - c. Ο **αριθμός των λέξεων** που εμπεριέχονται στην υπο-περιοχή
 - d. Η **μέση τιμή διαβάθμισης του γκρι** των εικονοστοιχείων που περιέχονται στα περιβάλλοντα κουτιά (bounding boxes) των αντικειμένων, με τέτοιο τρόπο ώστε η ταχύτητα εκτέλεσης υπολογισμού να είναι ανεξάρτητη του μεγέθους της υπο-περιοχής.

Ο κώδικας που θα παραχθεί θα πρέπει να συνοδεύεται από αναφορά η οποία θα πρέπει να πληροί τα ακόλουθα:

1. Να περιγράφονται και να αναλύονται τα βήματα και οι επιμέρους μέθοδοι που απαρτίζουν την ολοκληρωμένη μεθοδολογία.
2. Να παρουσιάζονται και να σχολιάζονται τα αποτελέσματα των ενδιάμεσων βημάτων, τα οποία αιτιολογούν την επιλογή της συγκεκριμένης μεθόδου.
3. Να παρουσιάζονται τα τελικά αποτελέσματα.
4. Να δικαιολογούνται πιθανές αστοχίες της μεθοδολογίας με συγκεκριμένα χαρακτηριστικά των εικόνων που τις προκαλούν.
5. Να συγκρίνεται ποιοτικά η επίδοση της εφαρμογής στις περιπτώσεις ύπαρξης και μη θορύβου.

ΥΠΟΔΕΙΞΗ

Για την επίτευξη των παραπάνω ζητούμενων, προτείνεται :

1. Στην αρχική εικόνα να εφαρμόσετε γραμμικό ή μη γραμμικό φίλτρο απόρριψης θορύβου της επιλογής σας. Το βήμα αυτό θα πρέπει να υλοποιηθεί **ΧΩΡΙΣ** τη χρήση των αντίστοιχων συναρτήσεων της OpenCV.

2. Στην εικόνα του αποτελέσματος του παραπάνω βήματος να εφαρμόσετε κατάλληλο κατώφλι για τη μετατροπή της εικόνας διαβάθμισης του γκρι σε δυαδική εικόνα, χρησιμοποιώντας τη συνάρτηση 'cv2.threshold'
3. Για να γίνουν οι ζητούμενες μετρήσεις θα πρέπει να εφαρμόσετε κατάλληλη μεθοδολογία που να βρίσκει συνδεδεμένα αντικείμενα, που υλοποιείται στη συνάρτηση 'cv2.connectedComponents'.
4. Για την εύρεση του περιβάλλοντος κουτιού ενός συνόλου (λευκών) εικονοστοιχείων, μπορείτε να χρησιμοποιήσετε τη συνάρτηση 'cv2.boundingRect'.
5. Για την σχεδίαση παραλληλόγραμμων και κειμένου σε μια εικόνα, μπορείτε να χρησιμοποιήσετε τις συναρτήσεις 'cv2.rectangle' και 'cv2.putText', αντίστοιχα.

Στο παρακάτω σχήμα φαίνεται ένα υπόδειγμα της παραγόμενης εξόδου του προγράμματος προς ανάπτυξη:

