

Υπολογιστική Όραση (ΜΥΕ046)

1η σειρά ασκήσεων

Ακαδημαϊκό έτος: 2019-2020

Άσκηση 1

Το αρχείο *ask1.py* περιέχει τον κώδικα που εκτελεί κατωφλίωση – *thresholding* σε εικόνα εισόδου. Το πρόγραμμα δέχεται ως *command line arguments* την εικόνα εισόδου – “*trikoupi6.png*” στην οποία εφαρμόζει την κατωφλίωση, το αρχείο εξόδου – “*k.png*” στο οποίο αποθηκεύει την κατωφλιωμένη εικόνα και το κατώφλι - *threshold k*, το οποίο παίρνει τιμές από τα διαστήματα $[0,51]$, $[52,102]$, $[103,153]$, $[154,204]$, $[205,255]$ για την αντίστοιχη επανάληψη.

Αρχικά, ελέγχεται εάν η εικόνα εισόδου είναι έγχρωμη ή grayscale. Στην περίπτωση που η εικόνα είναι έγχρωμη καλείται η ρουτίνα *convert*, η οποία επιστρέφει την αντίστοιχη grayscale εικόνα. Η μετατροπή γίνεται θέτοντας στο *pixel* τον μέσο όρο των καναλιών Red, Green, Blue.

Η ρουτίνα *main* καλείται με όρισμα το *k* και την grayscale εικόνα και ελέγχει την τιμή έντασης κάθε *pixel*. Εάν η τιμή έντασης είναι τουλάχιστον ίση με το *k* που θέσαμε, το *pixel* παίρνει την τιμή 255 που αντιστοιχεί στο άσπρο χρώμα, διαφορετικά παίρνει την τιμή 0 που αντιστοιχεί στο μαύρο. Η επιλογή έγινε έτσι ώστε το άσπρο χρώμα να αντιστοιχεί στο φόντο και το μαύρο στο κείμενο. Η *main* επιστρέφει την εικόνα και αποθηκεύεται στο αρχείο εξόδου.

Οι μεταβλητές με ονομασία τύπου *image* αναφέρονται σε εικόνα, ενώ οι τύπου *img* αναφέρονται σε πίνακα που αναπαριστά εικόνα.

Το αρχείο *ask1.ipynb* εκτελείται σε *Jupyter notebook* και περιλαμβάνει παράδειγμα χρήσης του κώδικα στο χειρόγραφο που δίνεται για $k=28,83,115,187,233$.

Έπειτα από πειραματική αξιολόγηση του χειρόγραφου για διάφορες τιμές του *k*, παρατηρούμε ότι:

- ο όλες οι δυνατές τιμές του *k* για το ίδιο διάστημα, επιστρέφουν την ίδια κατωφλιωμένη εικόνα
- ο όσο αυξάνεται η τιμή του *k* μεταξύ των διαφορετικών διαστημάτων όλο και περισσότερα *pixel* παίρνουν μαύρο χρώμα.

Άσκηση 2

Το αρχείο *ask2.py* περιέχει τον κώδικα που εφαρμόζει αφινικό μετασχηματισμό σε εικόνα εισόδου. Το πρόγραμμα δέχεται ως *command line arguments* την εικόνα εισόδου – “*mri.png*” την οποία μετασχηματίζει, το αρχείο εξόδου – “*out2_x*.png*” στο οποίο αποθηκεύει τη μετασχηματισμένη εικόνα και τις 6 τιμές για το μετασχηματισμό.

Στον πίνακα Tt μεταφέρω την αρχή των αξόνων στο μέσο της εικόνας και στον T αποθηκεύω τις 6 τιμές του μετασχηματισμού. Στη *main* εκτελείται ο αφινικός μετασχηματισμός, πολλαπλασιάζοντας τον T με τον Tt και στη συνέχεια καλείται η *nnInterpolation*, η οποία εκτελεί παρεμβολή κοντινότερου γείτονα, χρησιμοποιώντας τον αντίστροφο του T , T_{inv} .

Οι μεταβλητές με ονομασία τύπου *image* αναφέρονται σε εικόνα, ενώ οι τύπου *img* αναφέρονται σε πίνακα που αναπαριστά εικόνα.

Το αρχείο *ask2.ipynb* εκτελείται σε *Jupyter notebook* και περιλαμβάνει παράδειγμα χρήσης του κώδικα στο χειρόγραφο που δίνεται για

- $(\alpha_1 \alpha_2 \alpha_3 \alpha_4 \alpha_5 \alpha_6) = (0 \ 2 \ 0 \ -2 \ 0 \ 0)$, για *scale 2x, rotate 90°*
- $(\alpha_1 \alpha_2 \alpha_3 \alpha_4 \alpha_5 \alpha_6) = (0 \ 2 \ 0 \ -2 \ 1 \ 0)$, για *scale 2x, shear 0.5 στον άξονα x, rotate 90°*
- $(\alpha_1 \alpha_2 \alpha_3 \alpha_4 \alpha_5 \alpha_6) = (1.42 \ 1.42 \ 0 \ -0.71 \ 2.13 \ 0)$, για *scale 2x, shear 0.5 στον άξονα x, rotate 45°*

Ακολουθώντας τις ίδιες εντολές αποθήκευσης εικόνων με την πρώτη άσκηση, εν τέλει στην έξοδο παίρνω μία μαύρη εικόνα. Στη γραμμή 13 θα δείτε σε σχόλιο *#astype(np.uint32)*. Το *notebook* έχει εκτελεστεί με αυτή την εντολή, γιατί διαφορετικά προκύπτει το παρακάτω μήνυμα λάθους

```
raise OSError("cannot write mode %s as PNG" % mode)
OSError: cannot write mode F as PNG
```

Επομένως, το *notebook* με τα παραδείγματα εκτέλεσης περιέχουν μαύρες εικόνες.

Υπάρχει πρόβλημα και στην αποθήκευση των μετασχηματισμένων εικόνων που προκύπτουν (ακόμα και 'χεροκίνητα') εκτελώντας την εντολή από το *command prompt*, πχ για τον πρώτο μετασχηματισμό που επιλέγω

```
Python3 ask2.py mri.png out2_1.png 0 2 0 -2 0 0
```

Γι' αυτό αφήνω την εντολή *show*, στην περίπτωση που εκτελέσετε το πρόγραμμα να προβάλλονται οι μετασχηματισμένες εικόνες.

Άσκηση 3

(α) Να αποδείξετε ότι η συνέλιξη 1Δ σήματος είναι γραμμικός μετασχηματισμός.

(β) Να αποδείξετε ότι ο Διακριτός μετασχηματισμός Fourier 1Δ σήματος είναι γραμμικός μετασχηματισμός.

Ο Διακριτός μετασχηματισμός Fourier 1Δ για ένα σήμα $x(n)$ ορίζεται ως

$$X(k) = \sum_{n=0}^{N-1} x(n) w_N^{kn}, \text{ όπου } w_N = e^{-j2\pi/N}$$

Το διάνυσμα που αντιστοιχεί στο σήμα $x(n)$ ορίζεται ως

$$\overline{x_N} = [x(0) \ x(1) \ \dots \ x(N-1)]^T$$

και το διάνυσμα του μετασχηματισμού ορίζεται ως

$$\overline{X_N} = [X(0) \ X(1) \ \dots \ X(N-1)]^T$$

Αναπαριστώ τις τιμές του w_N με έναν πίνακα, οποίος ορίζεται ως

$$\overline{\overline{w_N}} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & w_N & \dots & w_N^{N-1} \\ 1 & w_N^2 & \dots & w_N^{2(N-1)} \\ \vdots & & \ddots & \vdots \\ 1 & w_N^{N-1} & \dots & w_N^{(N-1)(N-1)} \end{bmatrix}$$

$$\text{Άρα, } \overline{\overline{w_N}}(k,n) = w_N^{kn} \text{ και } X(k) = \sum_{n=0}^{N-1} x(n)w_N^{kn} = \overline{\overline{w_N}}^{(k)} \overline{\overline{x_N}}$$

$$\text{όπου } \overline{\overline{w_N}}^{(k)} \begin{bmatrix} w_N^0 & w_N^k & \dots & w_N^{(N-1)k} \end{bmatrix}$$

$$\text{Επομένως, } \overline{\overline{X_N}} = \overline{\overline{w_N}} = \overline{\overline{x_N}}$$