



ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ UNIVERSITY OF WEST ATTICA

Πολυτεχνική Σχολή

Τμήμα Μηχανικών Πληροφορικής και Υπολογιστών

Όραση Υπολογιστών

Εξαμηνιαία Εργασία

ΕΥΦΡΟΣΥΝΗ ΒΑΡΣΟΥ 21390021
ΑΓΓΕΛΟΣ ΚΩΝΣΤΑΝΤΙΝΟΣ ΜΕΝΤΖΕΛΟΣ 21390132

1. Θεωρητικό Μέρος

1. Θόρυβος

- Ποιά είναι η πιθανότητα μια τυχαία μεταβλητή X η οποία ακολουθεί την κανονική κατανομή της εξίσωσης (1) με μέση τιμή μ_0 μηδέν και μοναδιαία διασπορά σ_0^2 να πάρει την τιμή μηδέν;

$$f_X(x; \mu_0, \sigma_0^2) = \frac{1}{\sigma_0 \sqrt{2\pi}} \exp\left(-\frac{(x - \mu_0)^2}{2\sigma_0^2}\right) \quad (1)$$

Αντικαθιστούμε στην παραπάνω εξίσωση (όπως περιγράφεται και στην εκφώνηση) $\mu_0 = 0$ και $\sigma_0^2 = 1 \rightarrow \sigma_0 = 1$

$$f_X(x; 0, 1) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$$

Η παραπάνω σχέση περιγράφει την τυποποιημένη συνεχή κανονική κατανομή. Στις συνεχείς κατανομές, η πιθανότητα να πάρει η μεταβλητή X ακριβώς μία τιμή είναι πάντα μηδέν. Πιο συγκεκριμένα, η πιθανότητα ενός σημείου x αφορά το εμβαδόν κάτω από την καμπύλη της κατανομής. Όμως για ένα μοναδικό σημείο, αυτό το εμβαδόν είναι 0 όπως περιγράφεται λύνοντας και το παρακάτω ολοκλήρωμα:

$$P(X = 0) = \int_0^0 f(x) dx = 0$$

Συνεπώς, το X είναι μια συνεχή τυχαία μεταβλητή και η πιθανότητα να πάρει μια ακριβή τιμή είναι πάντα μηδέν. Οι πιθανότητες στις συνεχείς κατανομές μετριοούνται σε διαστήματα υπολογίζοντας το ολοκλήρωμα/εμβαδόν μεταξύ δύο σημείων:

$$P(a \leq X \leq b) = \int_a^b f(x) dx$$

- Με ποιόν μετασχηματισμό μετατρέπεται ο πολλαπλασιαστικός θόρυβος $\eta[i, j]$ της εξίσωσης (2) σε προσθετικό; Τι πλεονεκτήματα και ποιο μειονέκτημα έχει; Στην προηγούμενη εξίσωση γ_0 είναι μια παράμετρος του θορύβου, $s[i, j]$ η αρχική εικόνα, και $x[i, j]$ η ενθόρυβη.

$$x[i, j] = s[i, j] \eta[i, j]^{\gamma_0} \quad (2)$$

Ο πολλαπλασιαστικός θόρυβος $\eta[i, j]$ της εξίσωσης (2) μετατρέπεται σε προσθετικό χρησιμοποιώντας τον λογαριθμικό μετασχηματισμό "log-transform" ως εξής:

$$\log x[i, j] = \log s[i, j] + \gamma_0 \log \eta[i, j]$$

Έτσι, ο θόρυβος που ήταν πολλαπλασιαστικός ($\eta[i, j]^{\gamma_0}$) γίνεται προσθετικός ($\gamma_0 \log \eta[i, j]$), σε λογαριθμικό χώρο. Κάποια από τα πλεονεκτήματα είναι ότι ο προσθετικός θόρυβος είναι πιο εύκολος να αναλυθεί με κλασικές τεχνικές φιλτραρίσματος (π.χ. Gaussian). Επιπλέον, ο log transform συμπιέζει τις υψηλές τιμές φωτεινότητας και ενισχύει τις χαμηλές με αποτέλεσμα να βελτιώνει την αντίθεση της εικόνας. Το μόνο μειονέκτημα που έχει είναι ότι ο λογάριθμος δεν εφαρμόζεται σε μηδενικές ή αρνητικές τιμές. Αυτό σημαίνει ότι πρέπει να προστεθεί μια μικρή θετική σταθερά πριν τον μετασχηματισμό, που αυτή μπορεί να επηρεάσει την ακρίβεια της επεξεργασίας.

2. Παράγωγος συνάρτησης φωτεινότητας

Στον υπολογισμό της κατεύθυνσης της παραγώγου της συνάρτησης φωτεινότητας μιας εικόνας, ποιό είναι ένα πιθανό πρόβλημα στον αριθμητικό υπολογισμό της και πώς αυτό μπορεί να αντιμετωπιστεί;

Ένα πιθανό πρόβλημα στον αριθμητικό υπολογισμό της κατεύθυνσης της παραγώγου (gradient direction) είναι η διαίρεση με το μηδέν ή με πολύ μικρές τιμές όταν ο οριζόντιος gradient (G_x) είναι μηδέν ή κοντά στο μηδέν. Αυτό μπορεί να οδηγήσει σε αποτελέσματα με άπειρο ή αριθμητική αστάθεια. Συνεπώς, για να αντιμετωπιστεί το πρόβλημα αντί για απλή διαίρεση G_y/G_x , χρησιμοποιούμε τη συνάρτηση $\arctan2(G_y, G_x)$, η οποία χειρίζεται σωστά όλες τις περιπτώσεις (και όταν το G_x είναι μηδέν) και επιστρέφει τη σωστή γωνία κατεύθυνσης του gradient χωρίς αριθμητικά σφάλματα.

3. Εξομαλυσμένα ελάχιστα τετράγωνα

Στην περίπτωση όπου ένα σήμα (μονοδιάστατο χάριν απλότητας) ανακτάται μέσω εξομαλυσμένων ελαχίστων τετραγώνων όπως στην εξίσωση (3), τι συνέπειες θα είχε η επιλογή μιας αρνητικής τιμής για την υπερπαράμετρο ρ_0 ;

$$s_{LS} = \operatorname{argmin}[\|x - As\|_2^2 + \rho_0 \|s\|_2^2] \quad (3)$$

Στην παραπάνω εξίσωση υπολογίζεται το σήμα s_{LS} , το οποίο προσπαθεί να προσεγγίσει τα δεδομένα x μέσω του γραμμικού μοντέλου A_s αλλά επειδή προστίθεται ο όρος κανονικοποίησης $\rho_0 \|s\|_2^2$ προσθέτει ένα επιπλέον κόστος. Πιο συγκεκριμένα, η παράμετρος ρ_0 απλά ελέγχει τον βαθμό της κανονικοποίησης της εξίσωσης.

Όταν $\rho_0 > 0$, προστίθεται το κόστος του όρου κανονικοποίησης για μεγάλες τιμές των συντελεστών του s και ενισχύονται οι λύσεις που είναι πιο «εξομαλυσμένες». Αντίθετα, όταν $\rho_0 < 0$ τότε ο όρος κανονικοποίησης δεν αποτελεί πλέον κόστος αλλά κέρδος για μεγάλες τιμές του $\|s\|_2^2$.

Οι αρνητικές τιμές του ρ_0 έχει σαν συνέπεια η εξίσωση να γίνεται μη-κυρτή. Αυτό σημαίνει πως η λύση της εξίσωσης μπορεί να μην είναι μοναδική ή ακόμα και να μην βρεθεί ελάχιστο. Επιπλέον, οι αρνητικές τιμές κάνουν την εξίσωση να επιλέγει λύσεις που έχουν υψηλή ενέργεια. Με αποτέλεσμα, να ενισχύεται ο θόρυβος των x δεδομένων και προκαλώντας υπερεκπαίδευση και απώλεια γενίκευσης. Γενικά, η ρ_0 πρέπει να είναι μη αρνητική για να διασφαλιστεί η ευστάθεια, η γενίκευση και η ορθότητα της εξίσωσης.

4. Συμπίεση εικόνας

Ποιές είναι οι διαφορές όταν μια εικόνα συμπιέζεται μέσω παραγοντοποίησης ιδιοζουσών τιμών (SVD) και μέσω διδιάστατου διακριτού μετασχηματισμού συνημιτόνου (DCT2); Για να είναι ολοκληρωμένη η απάντησή σας, σκεφτείτε κατ' ελάχιστον την ερμηνεία των δύο μεθοδολογιών, τον χώρο τον οποίο καταλαμβάνουν οι αντίστοιχες συμπίεσμένες μορφές, και την πολυπλοκότητα υπολογισμού κάθε μετασχηματισμού. Επιπλέον διαφορές θα προσμετρηθούν θετικά.

Έχουμε πολλές επιμέρους διαφορές ανάμεσα στις μεθόδους συμπίεσης εικόνων των SVD (Singular Value Decomposition) και DCT2 (2D Discrete Cosine Transform). Όσον αφορά την ερμηνεία των δύο μεθοδολογιών, η SVD αναλύει την εικόνα ως άθροισμα βαθμίδων μέσω παραγοντοποίησης σε τρεις πίνακες (U, Σ, V^T) ενώ η DCT2 την μετασχηματίζει σε συχνότητες (DC/AC). Επιπλέον, η SVD επιτυγχάνει την συμπίεση κρατώντας μόνο τις μεγαλύτερες ιδιοτιμές (singular values) ενώ η DCT2 την επιτυγχάνει αγνοώντας ή κβαντίζοντας τις υψηλές συχνότητες (θόρυβος). Τέλος, γενικά η SVD είναι μια μέθοδος που συμπιέζει την εικόνα πλήρως από μια γενική μαθηματική προσέγγισή χωρίς να λαμβάνει υπόψιν κάποιο ειδικό χαρακτηριστικό της εκάστοτε

εικόνας. Αντίθετα, η DCT2 εκμεταλλεύεται τα χαρακτηριστικά της εικόνας λαμβάνοντας υπόψιν ότι έχουν ενέργεια που συγκεντρώνεται σε χαμηλές συχνότητες.

Όσον αφορά τις διαφορές των μεθόδων στο χώρο τον οποίο καταλαμβάνουν οι αντίστοιχες συμπιεσμένες μορφές, η SVD για να κάνει συμπίεση με k ιδιοτιμές αποθηκεύει k στήλες στο U και στο V αλλά επειδή στην συμπίεση κρατούνται μόνο οι μεγαλύτερες ιδιοτιμές δεν υπάρχει γενικά κάποια απώλεια πληροφορίας. Η DCT2 για να κάνει συμπίεση διατηρεί έναν πίνακα συντελεστών DCT οι οποίοι όπως είπαμε προηγουμένως μηδενίζονται ή κβαντίζονται. Επίσης, εδώ κρατούνται μόνο οι πιο σημαντικοί συντελεστές (συνήθως σε μορφή μπλοκ).

Όσον αφορά τις διαφορές των μεθόδων στην υπολογιστική πολυπλοκότητα που διαθέτουν, η SVD έχει μεγάλη πολυπλοκότητα ($O(MN^2)$ για κάποια $M \times N$ εικόνα) που αυξάνεται ειδικότερα πιο ραγδαία όσο πιο πολλά pixels (δηλαδή M και N) έχει μια εικόνα, κάτι που δεν την κάνει ιδανική για εφαρμογές που κάνουν συμπίεση σε πραγματικό χρόνο και σε μεγάλες προφανώς εικόνες. Αντίθετα, η DCT2 είναι πιο αποδοτική με λογαριθμογραμμική πολυπλοκότητα ($O(N^2 \log N)$) και αποτελεί γενικά βάση για πολλά πρότυπα συμπίεσης όπως π.χ. στις εικόνες .jpg.

2. Προγραμματιστικό μέρος

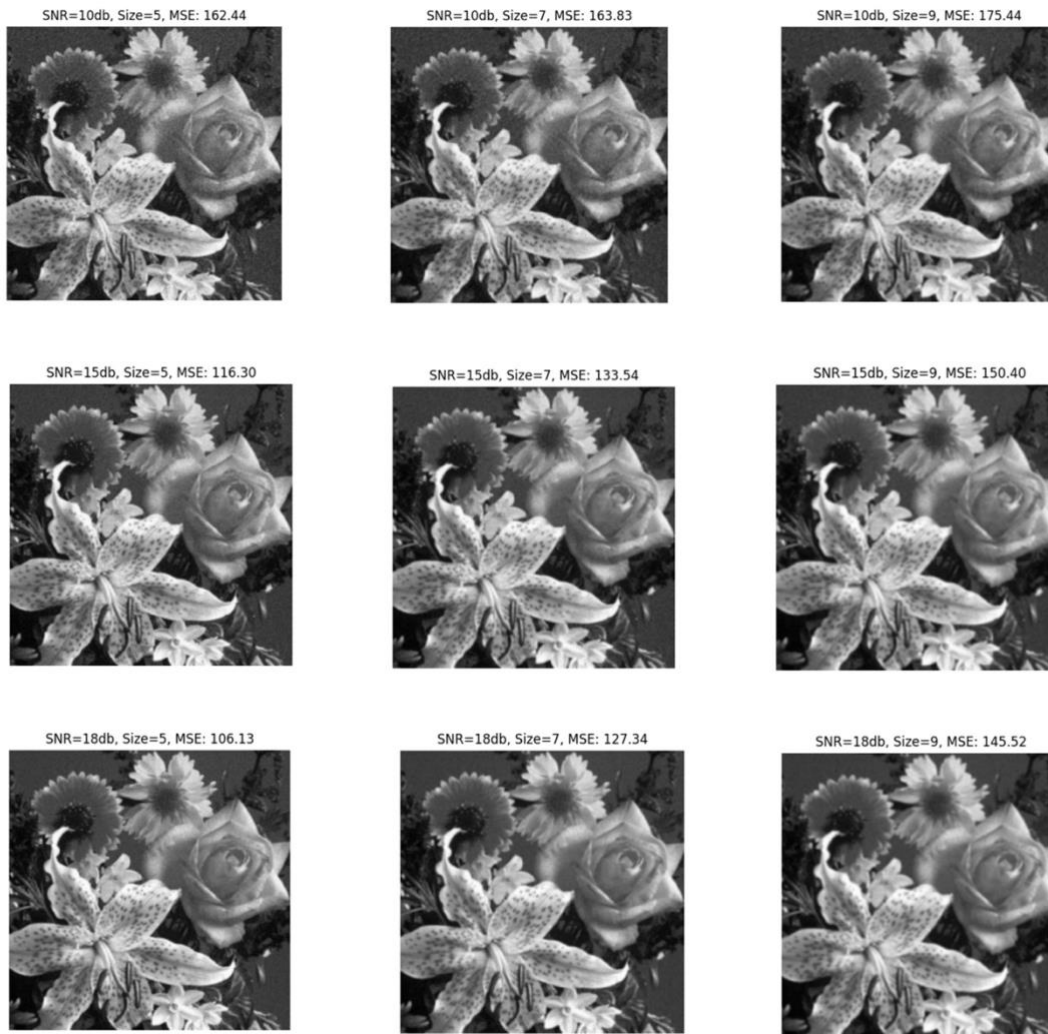
1. Μάσκα μέσου όρου

Εφαρμόστε την μάσκα του μεσαίου όρου στην εικόνα flowers.jpg για μήκος πέντε, επτά, και εννέα για θόρυβο AWGN με λόγο σήματος προς θόρυβο 10, 15, και 18 dB. Τι παρατηρείτε; Ποιά η απόκλιση από την αρχική εικόνα ως προς το μέσο τετραγωνικό σφάλμα;

Από τις παρακάτω εικόνες, παρατηρούμε ότι όσο αυξάνεται το μέγεθος της μάσκας μέσου όρου (kernel size 5, 7, 9), η κάθε εικόνα γίνεται πιο θολή και χάνει λεπτομέρειες. Αυτό συμβαίνει επειδή η μεγαλύτερη μάσκα εξομαλύνει περισσότερο τις τιμές των γειτονικών pixels, μειώνοντας τον θόρυβο αλλά και την οξύτητα της εικόνας.

Επίσης παρατηρούμε ότι, όσο μεγαλύτερο είναι το SNR (δηλαδή όσο μικρότερος είναι ο θόρυβος που προστίθεται στην εικόνα), τόσο καλύτερη είναι η ποιότητα του φίλτρου και ταυτόχρονα μικρότερη το MSE από την αρχική εικόνα. Για παράδειγμα, για $SNR=18dB$ και για kernel size 5×5 , το MSE είναι μικρότερο ($MSE = 106.21$) από ότι για $SNR=10dB$ και kernel size 9×9 ($MSE = 175.25$).

Συνεπώς, μικρότερο MSE παρατηρείται όταν έχουμε υψηλό SNR και μικρό kernel size. Αντίθετα, μεγαλύτερο MSE εμφανίζεται όταν έχουμε χαμηλό SNR και μεγάλο kernel size.



Εικόνα 1: flowers.jpg – AWGN θόρυβος

3. Όξυνση εικόνας

Εφαρμόστε τον τελεστή όξυνσης στις εικόνες bridge.tiff, im1.jpg, και im2.jpg. Τι παρατηρείτε; Ποιά είναι η διαφορά από τις αντίστοιχες αρχικές εικόνες ως προς το μέσο τετραγωνικό σφάλμα;

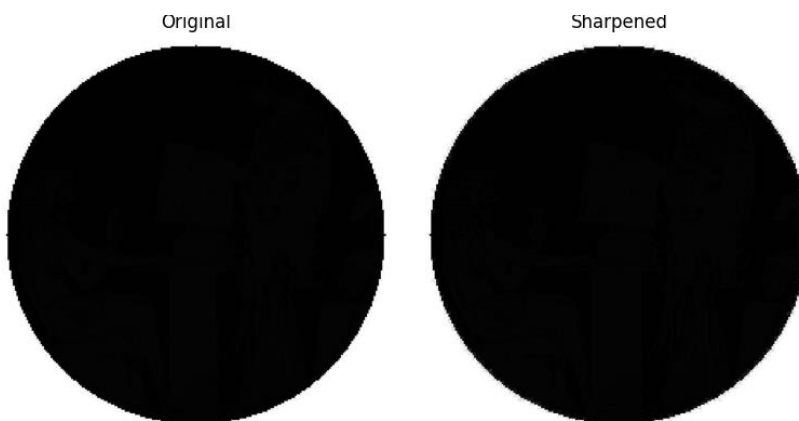
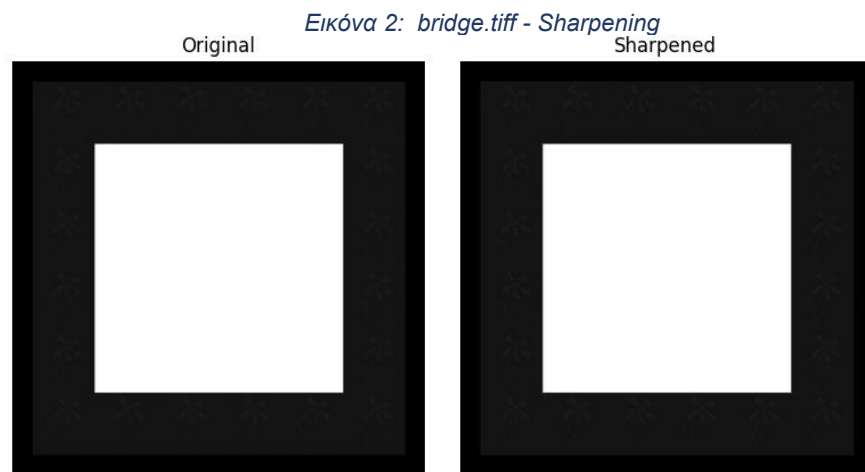
Από τις παρακάτω εικόνες παρατηρούμε ότι, μετά την εφαρμογή του τελεστή όξυνσης, διαπιστώνουμε ότι οι λεπτομέρειες και οι ακμές έχουν ενισχυθεί σε σχέση με τις αρχικές εικόνες.

Στην εικόνα bridge.tif, το mse (1558.73) είναι αρκετά υψηλό, το οποίο δείχνει ότι η όξυνση προκάλεσε σημαντικές αλλαγές στην εικόνα (αφου διαφέρει κατά πολύ από την αρχική). Αυτό πιθανόν οφείλεται στο ότι η εικόνα έχει πολλές λεπτομέρειες, οπότε ο τελεστής όξυνσης ενισχύει έντονα τις διαφορές, αυξάνοντας το σφάλμα σε σχέση με την αρχική.

Στην εικόνα m1.jpg, το MSE είναι πολύ μικρό (3.32), που σημαίνει ότι η όξυνση δεν άλλαξε σημαντικά την εικόνα. Αυτό συμβαίνει επειδή η εικόνα έχει ήδη υψηλή αντίθεση και λίγες λεπτομέρειες, οπότε ο τελεστής δεν έχει μεγάλο αποτέλεσμα.

Στην εικόνα m2.jpg, το MSE είναι ικανοποιητικό (17.77), που σημαίνει ότι η όξυνση είχε μέτρια επίδραση καθώς ενισχύσε τις ακμές χωρίς να αλλοιώσει υπερβολικά την αρχική εικόνα.

Συνεπώς, η όξυνση ενισχύει τις ακμές και τις λεπτομέρειες σε όλες τις εικόνες, αλλά το μέγεθος της διαφοράς MSE εξαρτάται από το αρχικό περιεχόμενο της εικόνας. Σε εικόνες με πολλές λεπτομέρειες, το MSE αυξάνεται σημαντικά, ενώ σε πιο απλές εικόνες με απλά σχήματα η διαφορά είναι μικρή.

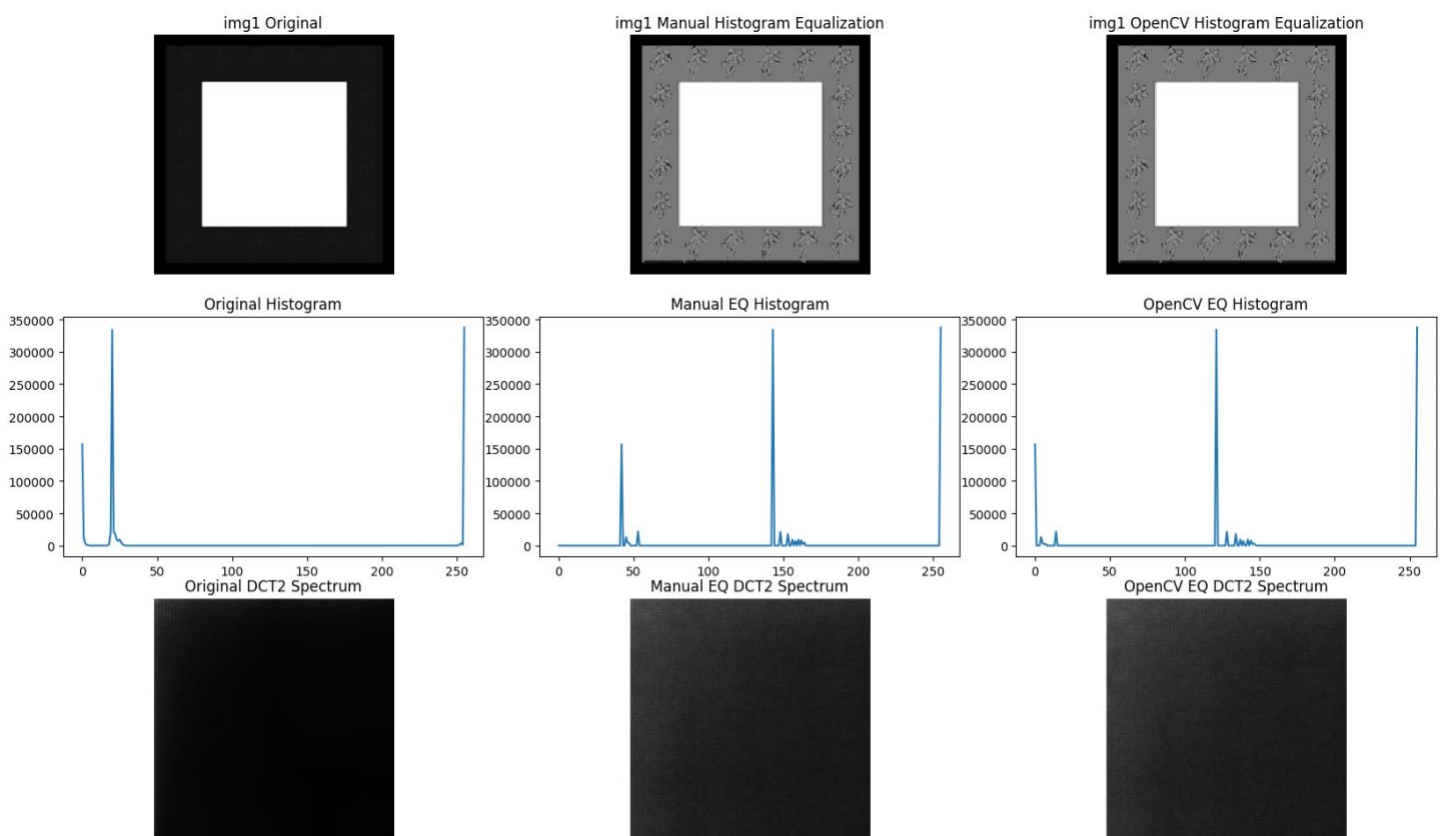


Εικόνα 4: im2.jpg - Sharpening

6. Ιστόγραμμα εικόνας

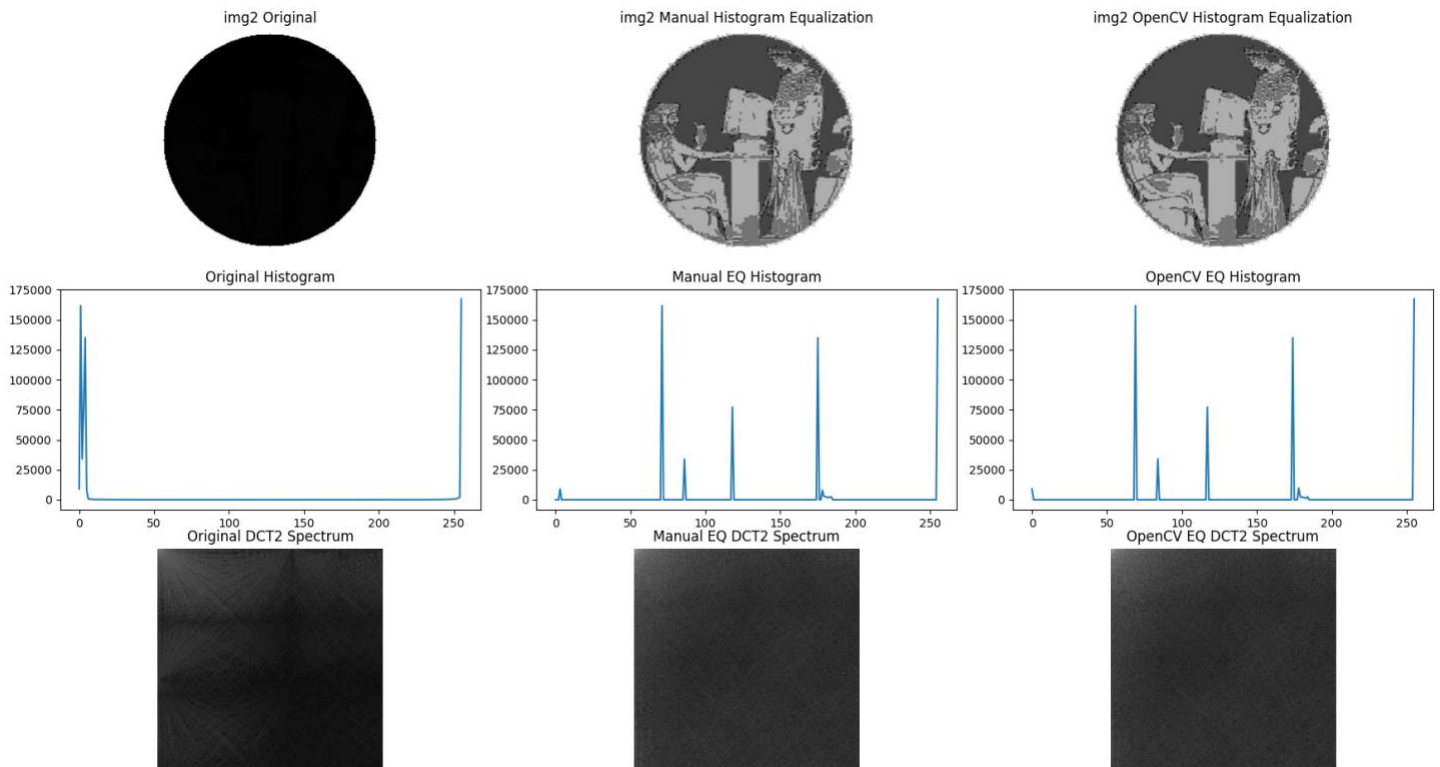
Να γίνει εξίσωση ιστογράμματος στις εικόνες `im1.jpg` και `im2.jpg` τόσο με δικό σας κώδικα όσο και με την μέθοδο `equalizeHist`. Συγκρίνετε τα αποτελέσματα μεταξύ τους. Να τυπώσετε το αρχικό ιστόγραμμα και το τελικό ιστόγραμμα καθώς και το αρχικό και τελικό φάσμα του DCT2 και στις δύο περιπτώσεις. Τι παρατηρείτε; Δώστε προσοχή στην δομή των εικόνων.

Για την πρώτη εικόνα παρατηρούμε ότι το αρχικό Ιστόγραμμα είναι πολύ συγκεντρωμένο στα άκρα (μαύρο και λευκό), λόγω των καθαρών περιοχών (μαύρο πλαίσιο και λευκό εσωτερικό). Μετά από την εξίσωση την δικιά μας και του OpenCV δημιουργούνται επιπλέον εντάσεις (αποχρώσεις του γκρι), κάτι που βοηθά να αναδειχθούν τα δέντράκια στο μαύρο φόντο. Στο φάσμα DCT2 πριν την εξισώσεις, βλέπουμε ότι η ενέργεια εντοπίζεται χαμηλά διότι είναι μια απλή εικόνα με ελάχιστες λεπτομέρειες. Μετά τις εξισώσεις του OpenCV και την δικιά μας, οι λεπτομέρειες του φόντου γίνονται πιο έντονες και το φάσμα περιέχει περισσότερες υψηλές συχνότητες. Οι δύο εξισώσεις δίνουν παρόμοια αποτελέσματα, με μικρές διαφορές στην κατανομή των τιμών όπως είναι και λογικό διότι με την βιβλιοθήκη της OpenCV έχει περισσότερη ακρίβεια.



Εικόνα 5: `im1.jpg` - Ιστόγραμμα & Φάσμα DCT2

Για την δεύτερη εικόνα παρατηρούμε ότι το αρχικό ιστόγραμμα είναι πολύ συγκεντρωμένο στα σκοτεινά επίπεδα, κάτι που δείχνει πως η εικόνα είναι ιδιαίτερα σκοτεινή και με χαμηλή ορατότητα στις φωτεινές περιοχές. Μετά την εξίσωση, τόσο με τη δικιά μας όσο και με του OpenCV, το ιστόγραμμα απλώνεται σε όλο το εύρος των φωτεινότητων. Αυτό έχει ως αποτέλεσμα να αναδεικνύονται οι λεπτομέρειες της εικόνας που δείχνει δύο ανθρώπους και έναν ηλεκτρονικό υπολογιστή. Στο φάσμα DCT2 πριν την εξίσωση, παρατηρείται ότι η ενέργεια βρίσκεται κυρίως σε χαμηλές συχνότητες, κάτι που δικαιολογείται από τη γενικά σκοτεινή εμφάνιση της εικόνας. Μετά τις εξισώσεις, τόσο της βιβλιοθήκης όσο και τη δική μας, εμφανίζονται περισσότερα υψηλά στοιχεία στο φάσμα, γεγονός που φανερώνει την ενίσχυση των λεπτομερειών και την καλύτερη απόδοση των υφών. Οι δύο μέθοδοι παράγουν παρόμοια αποτελέσματα, με μικρές διαφορές στην κατανομή των τιμών, κάτι αναμενόμενο, καθώς η υλοποίηση της OpenCV προσφέρει μεγαλύτερη ακρίβεια.

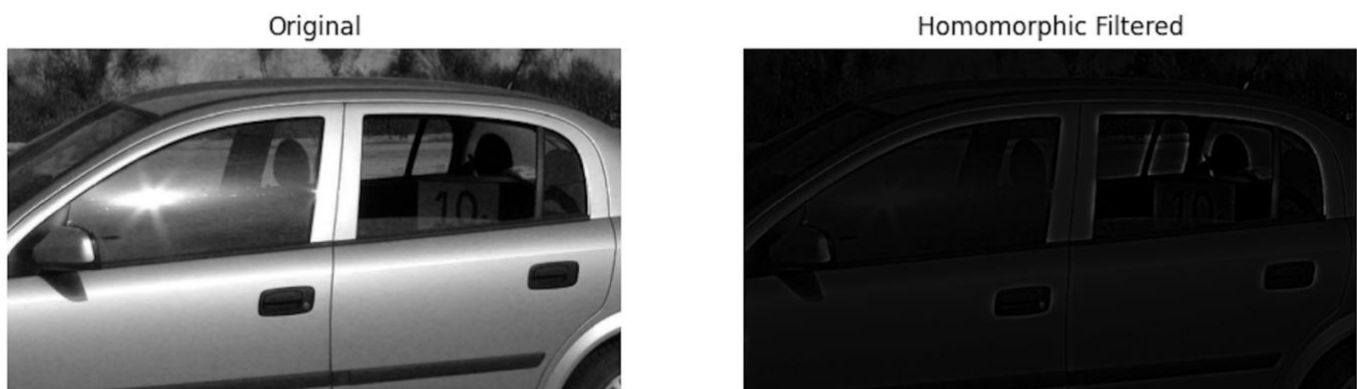


Εικόνα 6: *im2.jpg* - Ιστόγραμμα & Φάσμα DCT2

7. Ομοιομορφικό φιλτράρισμα

Να εφαρμοστεί η τεχνική του ομοιομορφικού φιλτραρίσματος στην εικόνα *car.jpg* με ένα χωρικό φίλτρο Butterworth πρώτης τάξεως με την προσθήκη μιας μικρής σταθεράς για την ενίσχυση των χαμηλών συχνοτήτων. Τι παρατηρείτε;

Παρατηρούμε ότι μετά την εφαρμογή του ομοιομορφικού φίλτρου, η εικόνα έχει γίνει συνολικά πιο σκοτεινή, αλλά έχουν ενισχυθεί σημαντικά οι λεπτομέρειες και τα περιγράμματα του αυτοκινήτου. Οι περιοχές που πριν ήταν υπερφωτισμένες από τον ήλιο όπως η αντανάκλαση στο παράθυρο του οδηγού έχουν περιοριστεί, ενώ περιοχές με μικρές διαφορές στη φωτεινότητα, όπως οι γραμμές της πόρτας, τα τζάμια και οι σκιές στο εσωτερικό του αυτοκινήτου, φαίνονται τώρα πιο καθαρά. Επίσης, παρατηρούμε ότι η επίδραση του φωτός και των σκιών έχει μειωθεί, κάτι που βοηθά να αναδειχθούν λεπτομέρειες που στην αρχική εικόνα ήταν δύσκολο να φανούν.



Εικόνα 7: *car.jpg* - Ομοιομορφικό φιλτράρισμα