Blue text on a black background

Description automatically generated

***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 µ0 μηδέν και μοναδιαία διασπορά σ02 να πάρει την τιμή μηδέν;**

|  |  |
| --- | --- |
|  | (1) |

Αντικαθιστούμε στην παραπάνω εξίσωση (όπως περιγράφεται και στην εκφώνηση) μ0 = 0 και σ02= 1 → σ0= 1

H παραπάνω σχέση περιγράφει την τυποποιημένη συνεχής κανονική κατανομή. Στις συνεχείς κατανομές, η πιθανότητα να πάρει η μεταβλητή Χ 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:

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

* **Με ποιόν μετασχηματισμό μετατρέπεται ο πολλαπλασιαστικός θόρυβος η[i,j] της εξίσωσης (2) σε προσθετικό; Τι πλεονεκτήματα και ποιό μειονέκτημα έχει; Στην προηγούμενη εξίσωση γ0 είναι μια παράμετρος του θορύβου, s[i,j] η αρχική εικόνα, και x[i,j] η ενθόρυβη.**

|  |  |
| --- | --- |
|  | (2) |
| Ο πολλαπλασιαστικός θόρυβος η[i,j] της εξίσωσης (2) μετατρέπεται σε προσθετικό χρησιμοποιώντας τον λογαριθμικό μετασχηματισμό (“log-transform”) ως εξής:  Έτσι, ο θόρυβος που ήταν πολλαπλασιαστικός (​) γίνεται προσθετικός (), 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 (Gx) είναι μηδέν ή κοντά στο μηδέν. Αυτό μπορεί να οδηγήσει σε απειρίες ή αριθμητική αστάθεια.

Για να αντιμετωπιστεί το πρόβλημα αντί για απλή διαίρεση Gy/Gx, χρησιμοποιούμε τη συνάρτηση arctan2(Gy, Gx), η οποία χειρίζεται σωστά όλες τις περιπτώσεις (και όταν το Gx είναι μηδέν) και επιστρέφει τη σωστή γωνία κατεύθυνσης του gradient χωρίς αριθμητικά σφάλματα.

## 3. Εξομαλυμένα ελάχιστα τετράγωνα

**Στην περίπτωση όπου ένα σήμα (μονοδιάστατο χάριν απλότητας) ανακτάται μέσω εξομαλυμένων ελαχίστων τετραγώνων όπως στην εξίσωση (3), τι συνέπειες θα είχε η επιλογή μιας αρνητικής τιμής για την υπερπαράμετρο ρ0;**

|  |  |
| --- | --- |
|  | (3) |
| Στην παραπάνω εξίσωση υπολογίζεται το σήμα sLS, το οποίο προσπαθεί να προσεγγίσει τα δεδομένα x μέσω του γραμμικού μοντέλου Αs αλλά επειδή προστίθεται ο όρος κανονικοποιήσης προσθέτει ένα επιπλέον κόστος. Πιο συγκεκριμένα, η παράμετρος ρ0 απλά ελέγχει τον βαθμό της κανονικοποίησης της εξίσωσης.  Όταν ρ0>0, προστίθεται το κόστος του όρου κανονικοποιήσης για μεγάλες τιμές των συντελεστών του s και ενισχύονται οι λύσεις που είναι πιο «εξομαλυμένες». Αντίθετα, όταν ρ0 < 0 τότε ο όρος κανονικοποιήσης δεν αποτελεί πλέον κόστος αλλά κέρδος για μεγάλες τιμές του .  Οι αρνητικές τιμές του ρ0 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 ρ0​ 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 παραγοντοποιήσεως ιδιαζουσών 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 συμπιεσμένες μορφές, και την πολυπλοκότητα υπολογισμού κάθε μετασχηματισμού. Επιπλέον διαφορές θα προσμετρηθούν θετικά.**

Οσόν αναφορά για την ερμηνεία των μεθοδολογιών έχουμε ότι:

* SVD (Singular Value Decomposition):
  + Η εικόνα αναλύεται ως άθροισμα βαθμίδων μέσω παραγοντοποίησης σε τρεις πίνακες (U, Σ, Vᵗ).
  + Η συμπίεση επιτυγχάνεται κρατώντας μόνο τις μεγαλύτερες ιδιοτιμές (singular values), δηλαδή τις σημαντικότερες συνιστώσες πληροφορίας.
  + Είναι μια καθαρά μαθηματική προσέγγιση, χωρίς να εκμεταλλεύεται ειδικά τα χαρακτηριστικά των εικόνων.
* DCT2 (2D Discrete Cosine Transform):
  + Η εικόνα μετασχηματίζεται σε συχνότητες (DC/AC components).
  + Η συμπίεση επιτυγχάνεται αγνοώντας ή κβαντίζοντας τις υψηλές συχνότητες (που συνήθως αντιστοιχούν σε λεπτομέρειες ή θόρυβο).
  + Εκμεταλλεύεται το γεγονός ότι οι περισσότερες φυσικές εικόνες έχουν ενέργεια συγκεντρωμένη στις χαμηλές συχνότητες.

Σχετικά με τον χώρο που καταλαμβάνουν οι συμπιεσμένες μορφές έχουμε ότι:

* SVD:
  + Για συμπίεση με k ιδιοτιμές, αποθηκεύονται k στήλες του U, k στήλες του V και k ιδιοτιμές (συνολικά k·(M+N+1) για εικόνα MxN).
  + Δεν υπάρχει απώλεια πληροφορίας αν κρατηθούν όλες οι ιδιοτιμές, αλλά για συμπίεση κρατάμε μόνο τις μεγαλύτερες.
* DCT2:
  + Η συμπιεσμένη μορφή είναι ένας πίνακας συντελεστών DCT, όπου οι περισσότεροι (συνήθως οι υψηλές συχνότητες) μπορούν να μηδενιστούν ή να κβαντιστούν.
  + Η αποθήκευση γίνεται συνήθως σε μορφή μπλοκ (π.χ. JPEG 8x8), και η ανακατασκευή απαιτεί μόνο τους σημαντικούς συντελεστές.

Στην υπολογιστική πολυπλοκότητα έχουμε ότι:

* SVD:
  + Πολύ υψηλή πολυπλοκότητα (O(MN²) για MxN εικόνα), ειδικά για μεγάλες εικόνες.
  + Δεν είναι πρακτική για πραγματικού χρόνου εφαρμογές ή μεγάλες εικόνες.
* DCT2:
  + Πολύ πιο αποδοτική (O(N² log N)), ειδικά με γρήγορους αλγορίθμους (Fast DCT).
  + Είναι η βάση για πρότυπα συμπίεσης όπως το JPEG.

Τέλος οι άλλες διαφορές είναι :

* SVD:
  + Παρέχει βέλτιστη προσέγγιση (σε νόρμα Frobenius) για δεδομένο αριθμό βαθμίδων.
  + Δεν εκμεταλλεύεται τοπικά χαρακτηριστικά της εικόνας.
  + Μπορεί να χρησιμοποιηθεί και για ανίχνευση δομής ή μείωση διαστάσεων.
* DCT2:
  + Εκμεταλλεύεται τη χωρική συσχέτιση των pixel.
  + Εύκολη υλοποίηση σε hardware και software.
  + Ενδέχεται να προκαλέσει τεχνουργήματα (artifacts) σε υψηλή συμπίεση (π.χ. blockiness στο JPEG).

2. Προγραμματιστικό μέρος

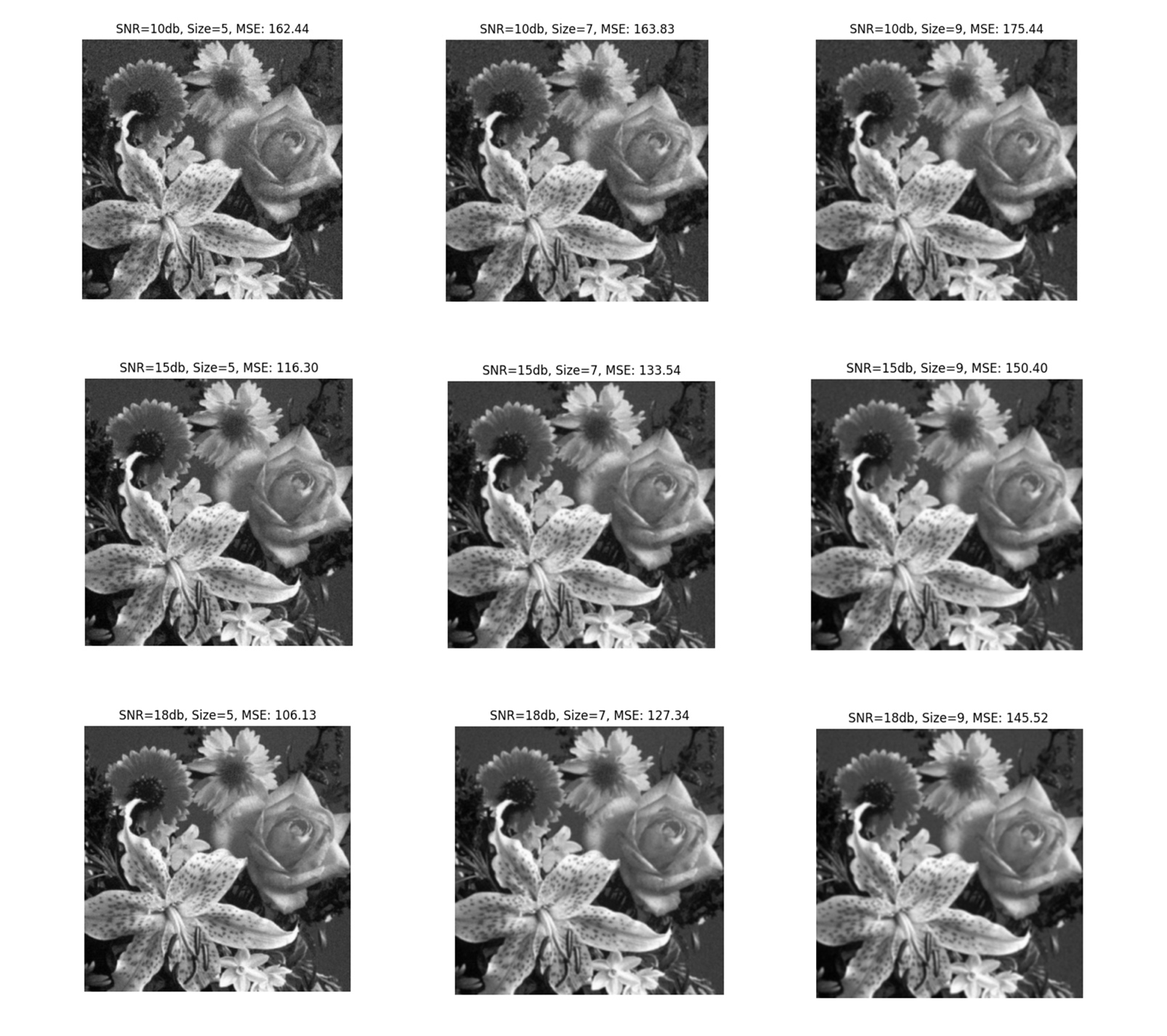
## 1. Μάσκα μέσου όρου

**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 5x5, το MSE είναι μικρότερο (MSE = 106.21) από ότι για SNR=10dB και kernel size 9x9 (MSE = 175.25).

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

****

*Εικόνα 1: flowers.jpg – AWGN θόρυβος*

3. ΄Οξυνση 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;**

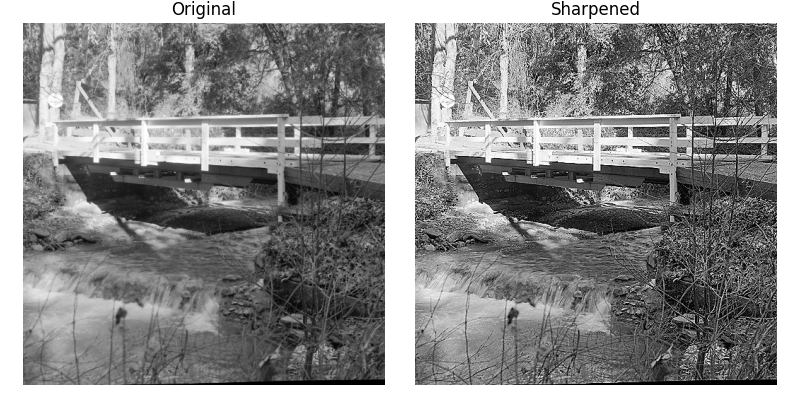
Από τις παρακάτω εικόνες παρατηρούμε ότι, μετά την εφαρμογή του τελεστή όξυνσης, διαπιστώνουμε ότι οι λεπτομέρειες και οι ακμές έχουν ενισχυθεί σε σχέση με τις αρχικές εικόνες.

Στην εικόνα bridge.tif, το mse (1558.73) είναι αρκέτα 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.

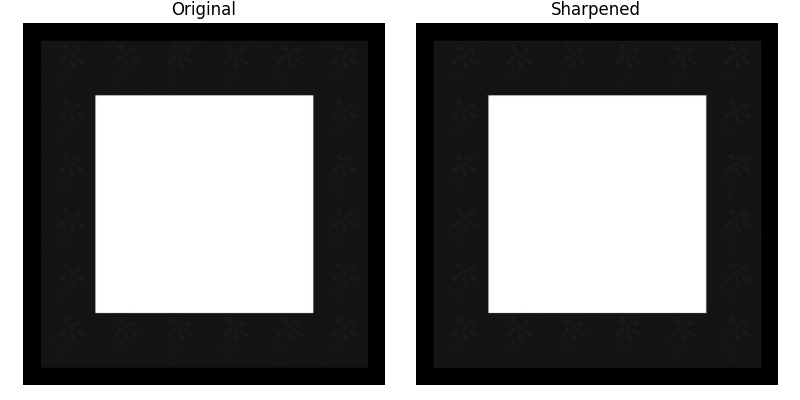
Στην εικόνα m1.jpg, το MSE είναι πολύ μικρό (3.32), που σημαίνει ότι η όξυνση δεν άλλαξε σημαντικά την εικόνα. Αυτό συμβαίνει επειδή η εικόνα έχει ήδη υψηλή αντίθεση και λίγες λεπτομέρειες, οπότε ο τελεστής δεν έχει μεγάλο αποτέλεσμα.

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.

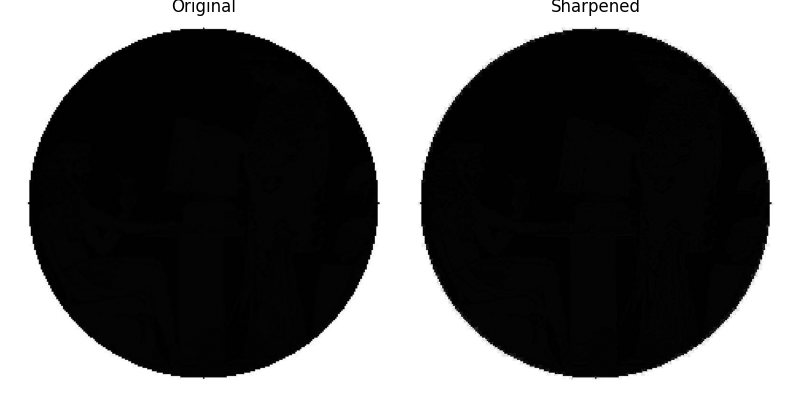
Συνεπώς, η όξυνση ενισχύει τις ακμές και τις λεπτομέρειες σε όλες τις εικόνες, αλλά το μέγεθος της διαφοράς MSE εξαρτάται από το αρχικό περιεχόμενο της εικόνας. Σε εικόνες με πολλές λεπτομέρειες, το MSE αυξάνεται σημαντικά, ενώ σε πιο απλές εικόνε με απλά σχήματα ς η διαφορά είναι μικρή.



*Εικόνα 2: bridge.tiff - Sharpening*



*Εικόνα 3: im1.jpg - Sharpening*

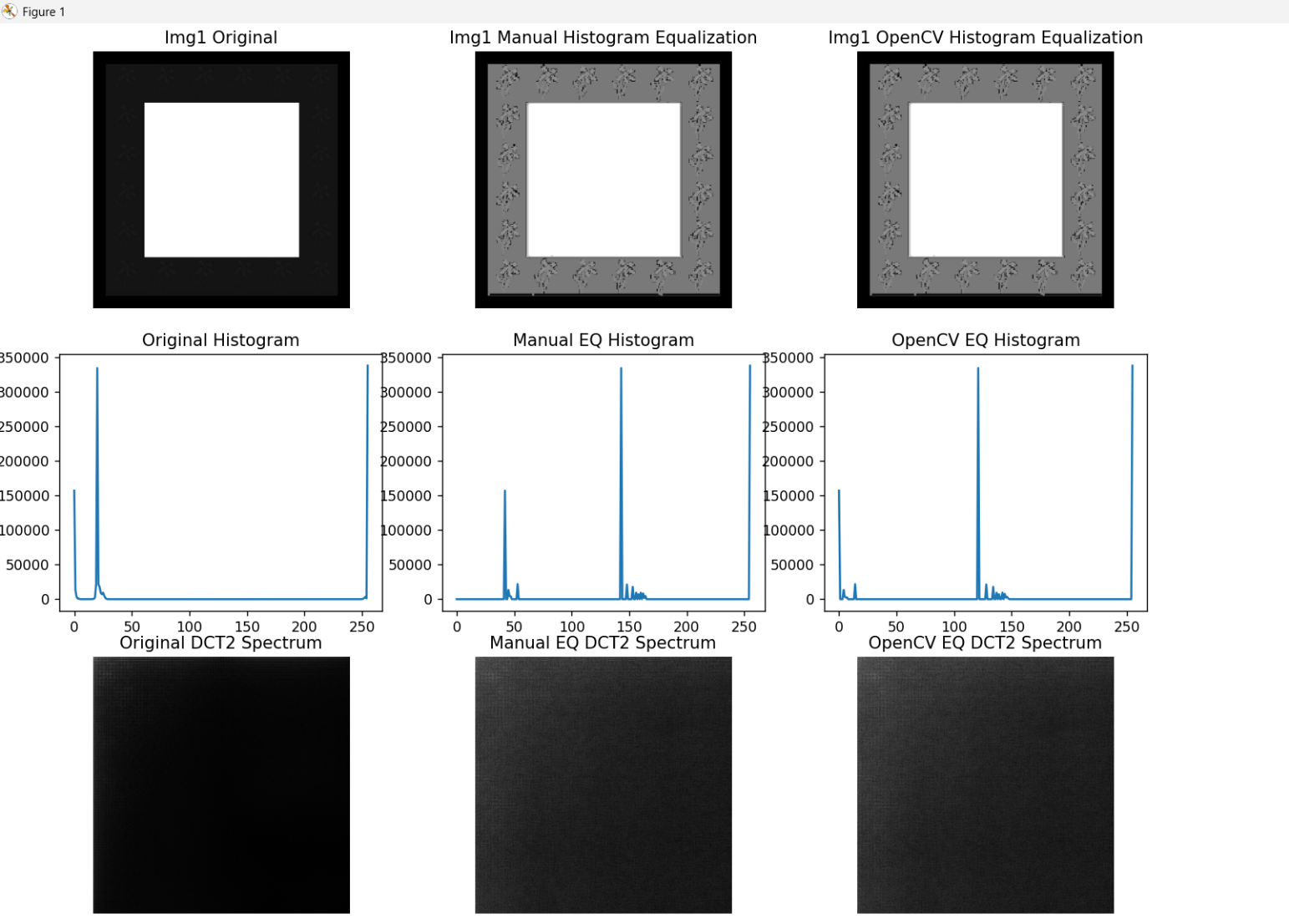


*Εικόνα 4: im2.jpg - Sharpening*

6. Ιστόγραμμα εικόνας

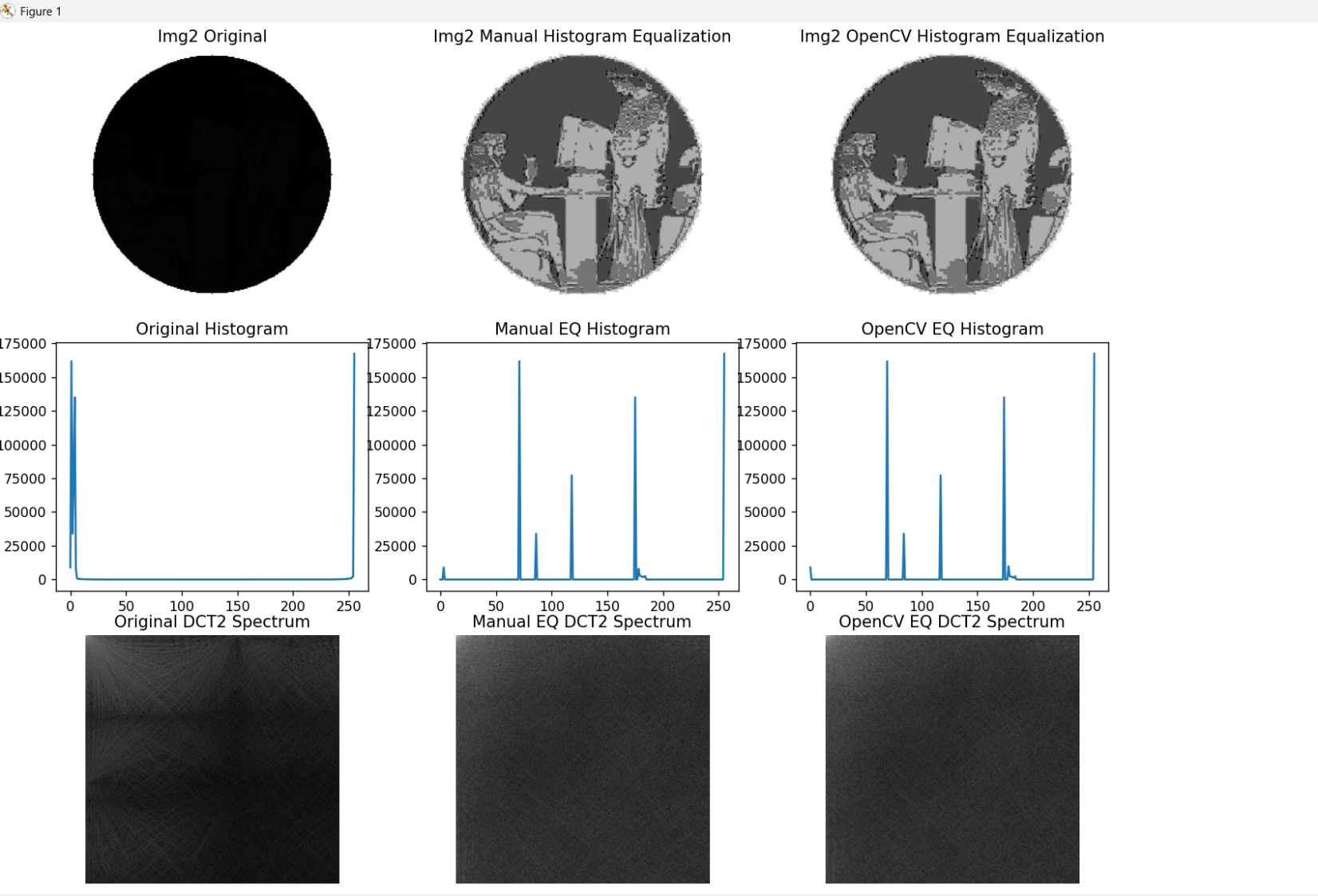
**Να γίνει εξίσωση ιστογράμματος στις εικόνες im1.jpg και im2.jpg τόσο με δικό σας κώδικα όσο και με την μέθοδο equalizeHist. Συγκρίνετε τα αποτελέσματα μεταξύ τους. Να τυπώσετε το αρχικό ιστόγραμμα και το τελικό ιστόγραμμα καθώς και το αρχικό και τελικό φάσμα του DCT2 και στις δύο περιπτώσεις. Τι παρατηρείτε;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

Για την πρώτη εικόνα παρατηρώ 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 μου και του OpenCV δημιουργούνται επιπλέον εντάσεις (αποχρώσεις του γκρί), κάτι που βοηθά να αναδειχθούν τα δεντράκια 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 τις του OpenCV και την δικιά μου 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 έχει περισσότερη ακρίβεια.

****

Για την δεύτερη εικόνα παρατηρώ 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 μου όσο και με του OpenCV, το ιστόγραμμα απλώνεται σε όλο το εύρος των φωτεινοτήτων. 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 μου, 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.

****

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;**

Παρατηρώ 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, παρατηρώ ότι η επίδραση του φωτός και των σκιών έχει μειωθεί, κάτι που βοηθά να αναδειχθούν λεπτομέρειες που στην αρχική εικόνα ήταν δύσκολο να φανούν. Για παράδειγμα, το νούμερο "10" στο εσωτερικό ήταν πιο θολό, ενώ τώρα διακρίνεται καθαρότερα.