

-Εργασία 3-

High Dynamic Range Imaging

1 Multi-frame HDR imaging

Με σκοπό να υπολογίσουμε τις (σταθερές) ποσότητες ακτινοβολίας E που φτάνουν σε κάθε αισθητήρα pixel (i, j) μιας κάμερας, όταν φωτογραφίζεται μια συγκεκριμένη σκηνή με διαφορετικούς χρόνους έκθεσης t , η σχέση που δίνει την τιμή της φωτεινότητας του κάθε pixel (ισχύει για κάθε κανάλι) είναι:

$$Z_{ij}^k = f(t^k E_{ij})$$

όπου f η συνάρτηση απόκρισης της κάμερας, t^k είναι ο χρόνος έκθεσης της k -οστής λήψης, E_{ij} η ποσότητα της ακτινοβολίας στο pixel (i, j) , και Z_{ij}^k η φωτεινότητα του προηγούμενου στην k -οστή λήψη.

Θεωρώντας αρχικά την συνάρτηση απόκρισης f γραμμική και παίρνοντας τον λογάριθμο της προηγούμενης σχέσης:

$$\ln(E_{ij}) = \ln(Z_{ij}^k) - \ln(t^k)$$

Λαμβάνοντας υπόψη όλες τις λήψεις η εκτίμηση για κάθε στοιχείο $\ln(E_{ij})$ μπορεί να προκύψει ως:

$$\ln(E_{ij}) = \frac{\sum_{k=1}^K w(Z_{ij}^k) (\ln(Z_{ij}^k) - \ln(t^k))}{\sum_{k=1}^K w(Z_{ij}^k)}$$

όπου W συνάρτηση βάρους, που δίνει μεγαλύτερη βαρύτητα στα pixels που βρίσκονται στο μέσο της καμπύλης απόκρισης, λιγότερο σε αυτά που είναι στα άκρα και μηδενικό βάρος σε pixels που ξεπερνούν κάποια κατώφλια Z_{min}, Z_{max} πέρα από τα οποία θεωρούμε ότι υπάρχει κορεσμός.

Δημιουργείται επομένως η συνάρτηση

```
function radiancemap = mergeLDRstack(imgStack, exposureTimes,
weightingFcn)
```

η οποία λαμβάνει ως είσοδο τις τιμές της φωτεινότητας ($imgStack - Z_{ij}^k$) για κάθε pixel και κάθε διαφορετική λήψη (για ένα κανάλι των αρχικών εικόνων κάθε φορά), τους χρόνους έκθεσης ($exposureTimes - t^k$) και την επιλογή της συνάρτησης βάρους που θα χρησιμοποιηθεί ($weightingFcn$) και επιστρέφει τον πίνακα με τις εκτιμήσεις των $\ln(E_{ij})$ για τα αντίστοιχα pixels.

Αρχικά εντοπίζονται τα pixels που δεν έχουν κορεστεί και για αυτά υπολογίζονται τα αντίστοιχα βάρη και τοποθετούνται στον πίνακα $WZ(i, j, k)$ ανάλογα με την συνάρτηση βάρους που επιλέγεται. Υπολογίζονται οι τιμές των λογαρίθμων των χρόνων εκθέσεων $T(i, j, k)$ (επαναλαμβάνονται για διαφορετικά pixel για κάθε έκθεση).

Υπολογίζονται οι λογάριθμοι της φωτεινότητας κάθε pixel για κάθε έκθεση $G(i, j, k)$.

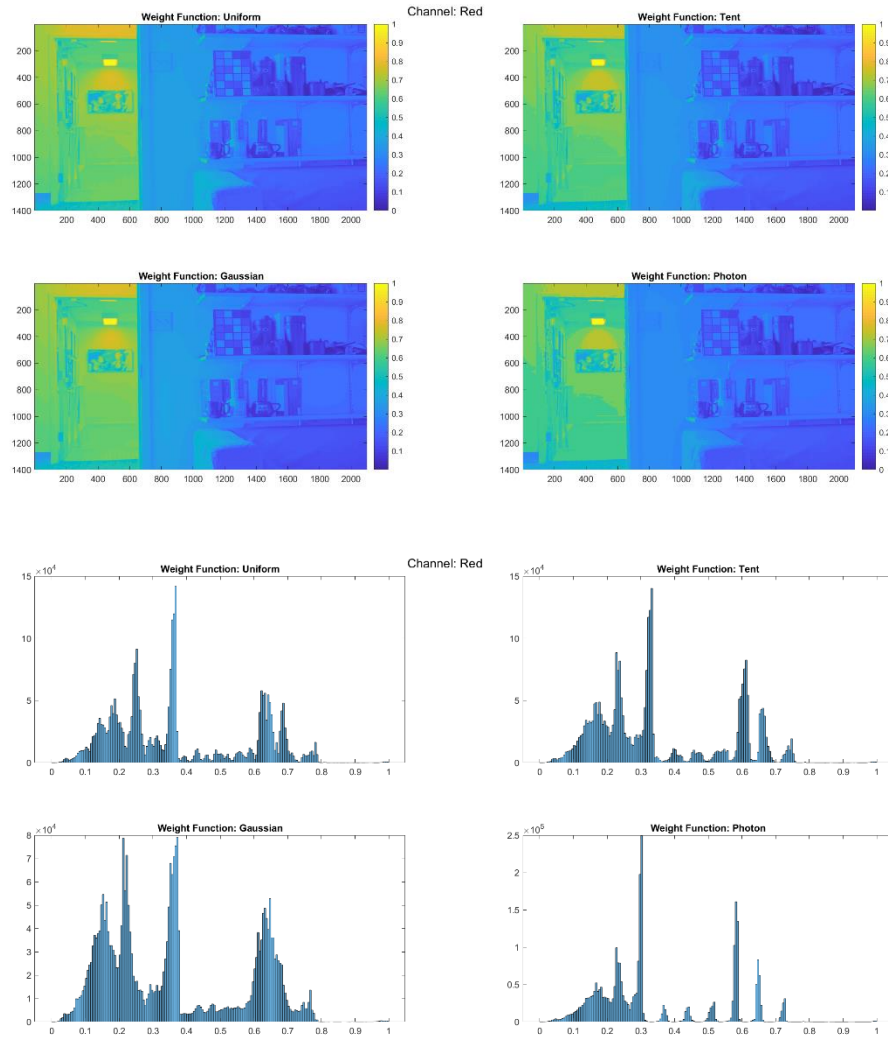
Δεδομένου πως επιλέγεται $Z_{min} > 0$ σε κάθε περίπτωση, σε κάθε pixel με φωτεινότητα $Z = 0 \Rightarrow G(Z) = -\infty, W(Z) = 0$ (κορεσμένο pixel) ανατίθεται η τιμή $G(Z) = 0$ για να μην προκύψουν υπολογιστικά σφάλματα (η συνεισφορά του όρου στο άθροισμα του αριθμητή θα είναι μηδενική λόγω του βάρους).

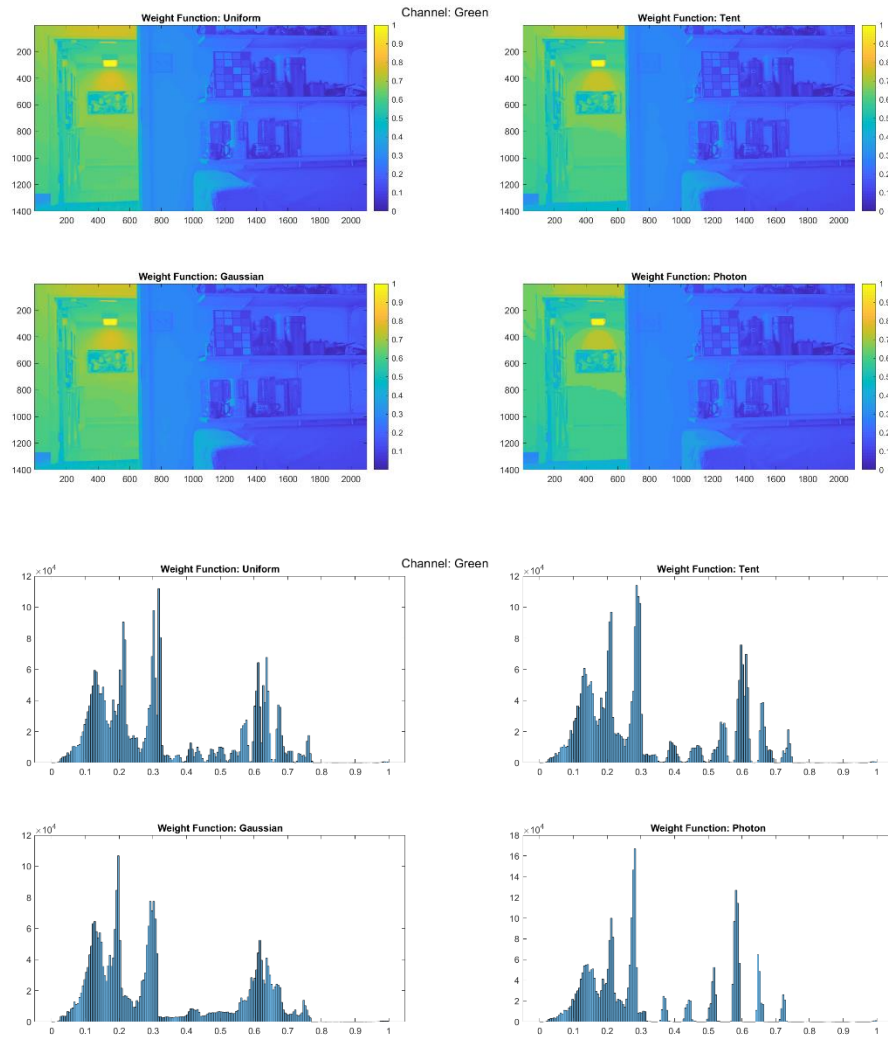
Υπολογίζεται ο πίνακας $\frac{\sum_k WZ(i, j, k) \odot (G(i, j, k) - T(i, j, k))}{\sum_k WZ(i, j, k)}$, ο οποίος ταυτίζεται με την ζητούμενη έξοδο, εκτός από τα στοιχεία (i, j) στα οποία έχουν ανατεθεί μηδενικά βάρη για όλες τις λήψεις (με \odot συμβολίζεται το *Hadamard Product*). Για κάθε ένα από αυτά ελέγχεται:

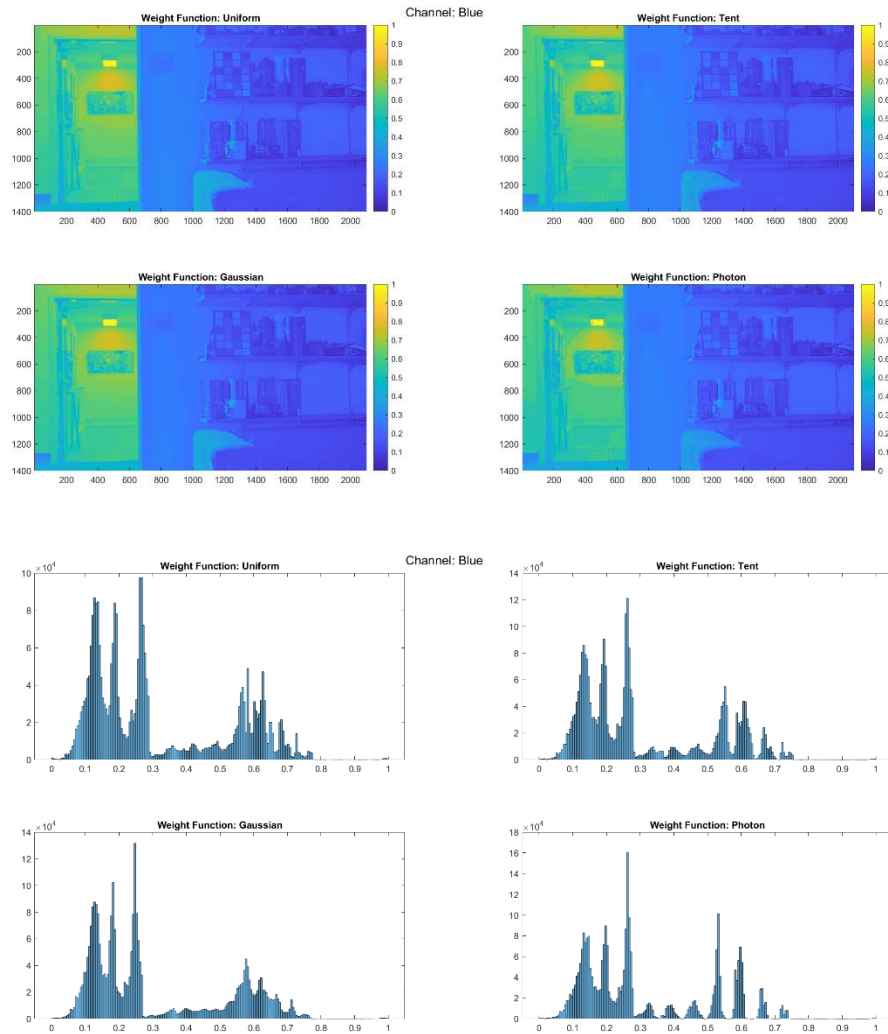
- αν είναι υποεκτεθειμένο σε όλες τις λήψεις, οπότε του ανατίθεται η ελάχιστη τιμή των στοιχείων της εξόδου
- αν είναι υπερεκτεθειμένο σε όλες τις λήψεις, οπότε του ανατίθεται η μέγιστη των στοιχείων της εξόδου

ενώ σε άλλη περίπτωση του ανατίθεται ο μέσος όρος του ελάχιστου και του μέγιστου στοιχείου της εξόδου.

Παρακάτω παρουσιάζονται τα αποτελέσματα της προηγούμενης διεργασίας για κάθε κανάλι και κάθε επιλογή συνάρτησης βάρους, όπου έχουν επιλεγεί $Z_{min} = 0.05$ (δίνει σημαντικά καλύτερα οπτικά αποτελέσματα από το προτεινόμενο 0.01) και $Z_{max} = 0.99$





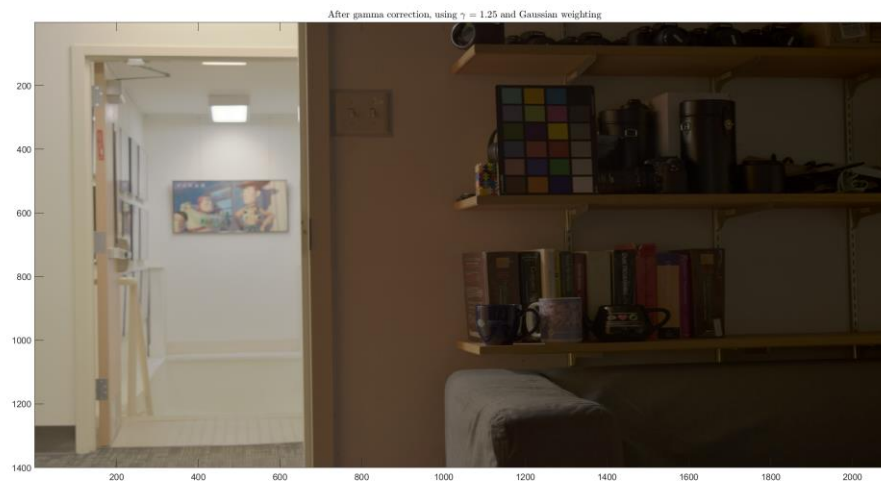


Για κάθε απεικόνιση εμφανίζονται διαφορετικά πλήθη από κυρίαρχες κορυφές στα ιστογράμματα. Για να αποθηκευτεί η τελική απεικόνιση με τρόπο ώστε να μην αλλοιώνεται σημαντικό μέρος πληροφορίας θα πρέπει κάθε τιμή να μπορεί να διαχωριστεί σε ποια από αυτές τις κυρίαρχες κορυφές ανήκει. Επομένως για παράδειγμα, στην περίπτωση της *Gaussian weighting function* εμφανίζονται 5 κυρίαρχες κορυφές, οπότε χρειάζονται τουλάχιστον 3 bit πληροφορίας.

Επίσης θα αρκούσε να χρησιμοποιηθούν τόσες εικόνες όσες το πλήθος των κυρίαρχων κορυφών (υπο την υπόθεση ότι κάθε εικόνα θα συνεισφέρει κυρίως σε διαφορετική κορυφή από τις άλλες) ώστε να εκτιμηθεί ικανοποιητικά η HDR εικόνα.

Χρησιμοποιώντας και πάλι την περίπτωση της *Gaussian weighting function* και χρησιμοποιώντας τις εικόνες [6, 9, 11, 13, 15] (5 σε αριθμό, όσες οι κυρίαρχες κορυφές και επιλεγμένες ώστε να αντιστοιχούν σε διαφορετικές εκθέσεις) η τελική HDR απεικόνιση που λαμβάνουμε χρησιμοποιώντας $Z_{min} = 0.05$, $Z_{max} = 0.99$ και μετά την χρήση Gamma Correction με $\gamma = 1.25$ (παρουσιάζεται στην

επόμενη ενότητα) φαίνεται παρακάτω:



2 Tone Mapping

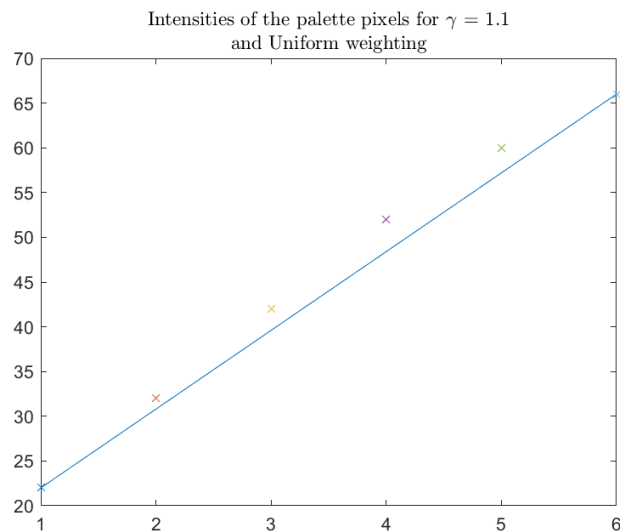
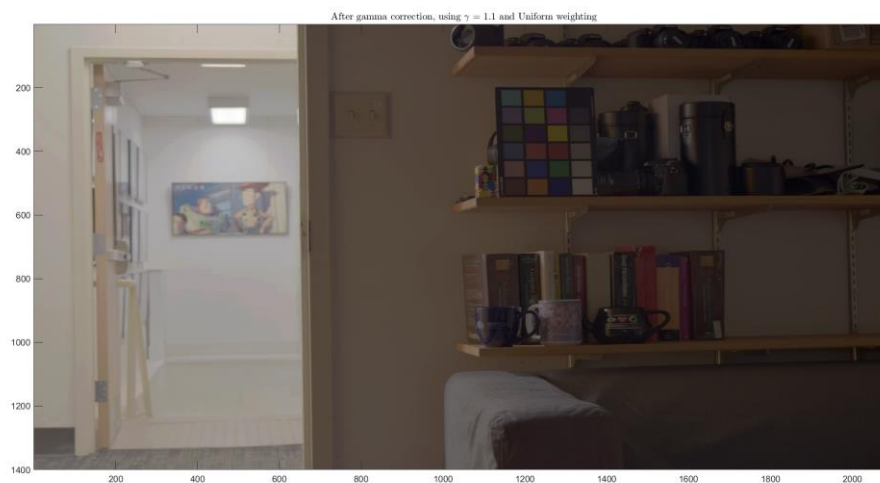
Για την καλύτερη απεικόνιση των προηγούμενων αποτελεσμάτων χρησιμοποιείται η μέθοδος απεικόνισης τόνου gamma correction, σύμφωνα με την οποία για μια εικόνα εισόδου $V_{in} \in [0,1]$, η εικόνα εισόδου προκύπτει ως $V_{out} = V_{in}^\gamma \in [0,1]$.

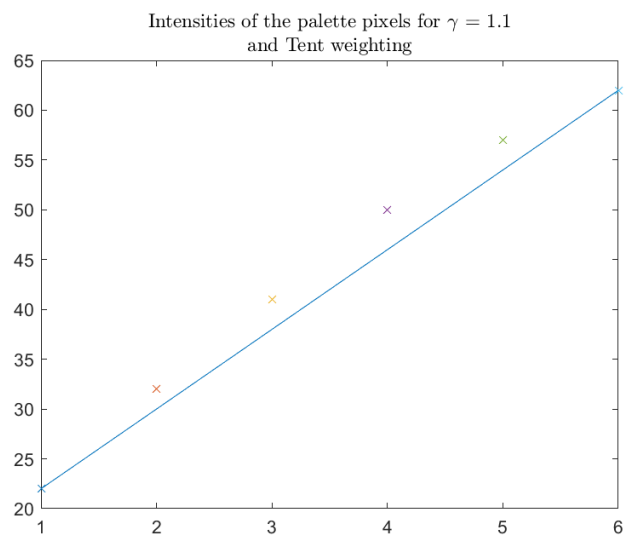
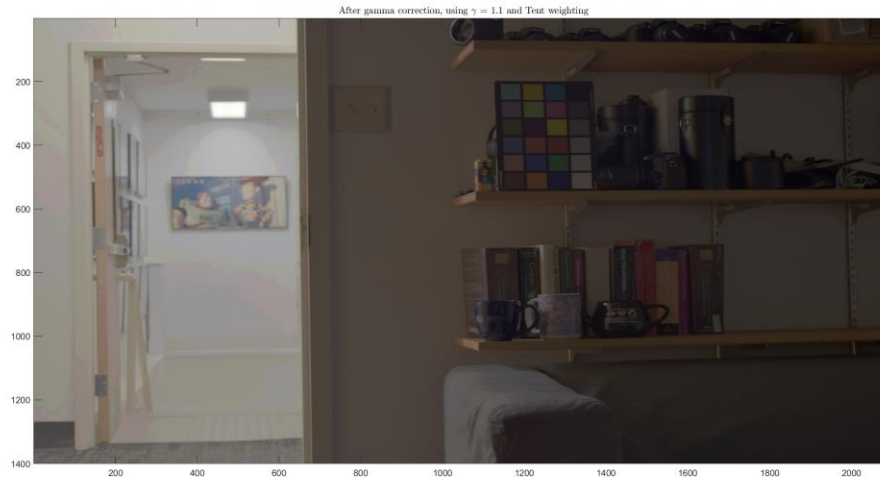
Δημιουργείται η συνάρτηση

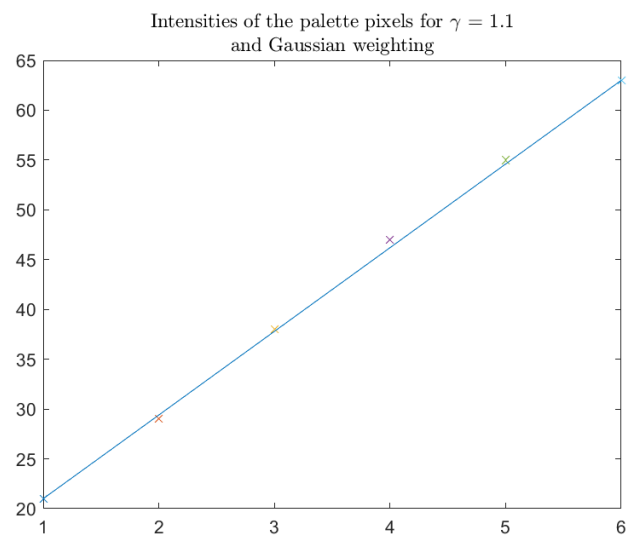
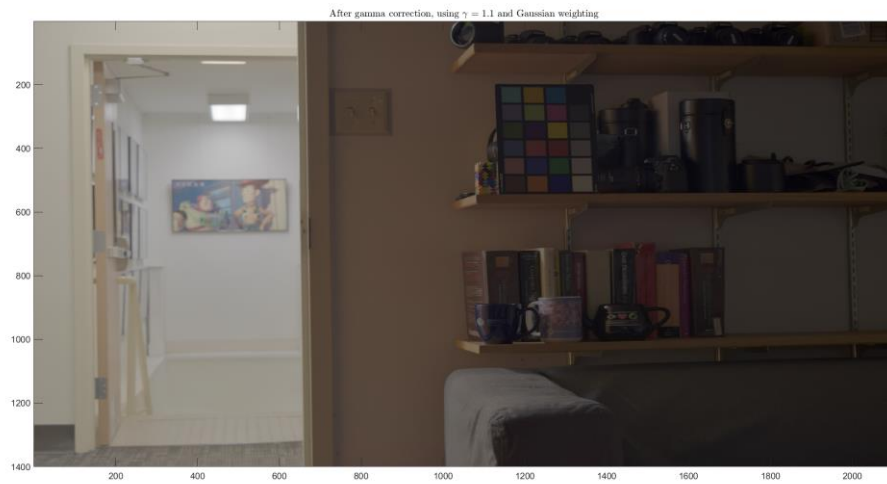
```
function tonedImage = toneMapping(radianceMap, gamma)
```

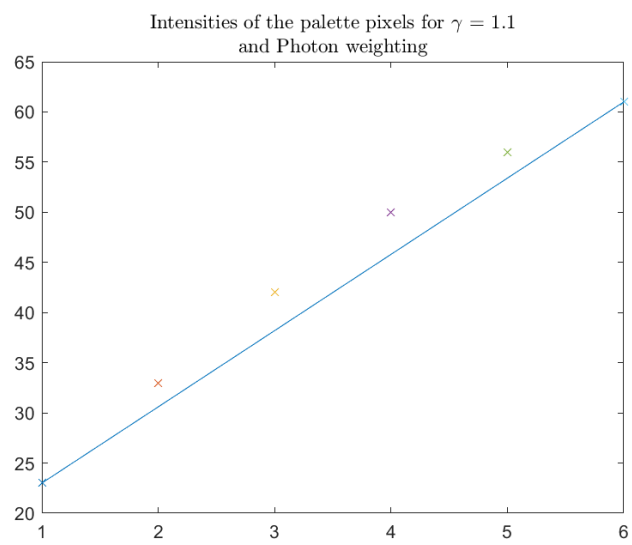
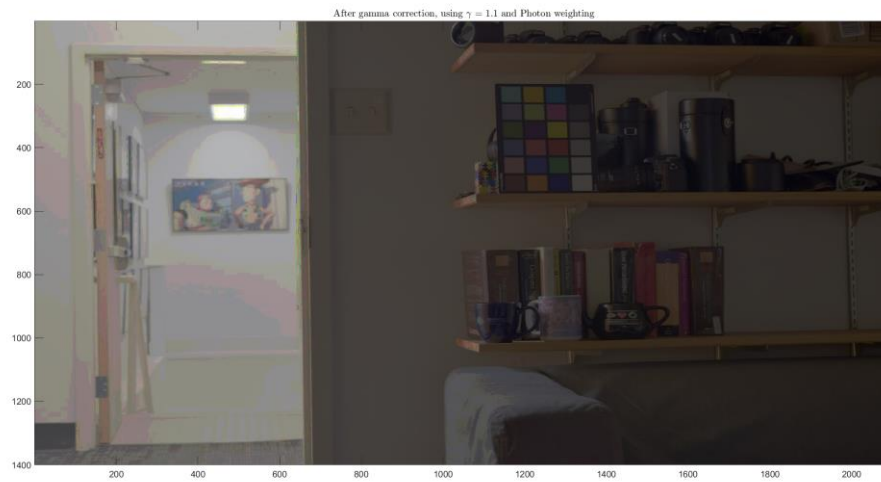
η οποία δέχεται μια εικόνα εισόδου στο διάστημα $[0,1]$, υπολογίζει τις τιμές εξόδου σύμφωνα με την προηγούμενη σχέση, κάνει rescale το αποτέλεσμα στο διάστημα $[0,255]$ και κβαντίζει την εικόνα εξόδου με 8bit για κάθε δείγμα.

Τα αποτελέσματα της μεθόδου με είσοδο τα αποτελέσματα της προηγούμενης ενότητας για κάθε συνάρτηση βάρους παρουσιάζονται παρακάτω όπου έχει επιλεγεί $\gamma = 1.1$, το οποίο φαίνεται να δίνει γενικά οπτικά καλά αποτελέσματα, ενώ παρουσιάζονται και τα γραφήματα για τις φωτεινότητες των pixel της παλέτας.

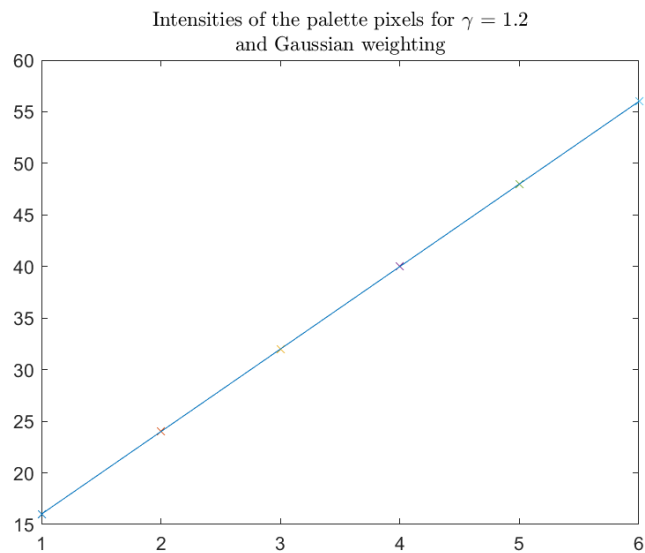
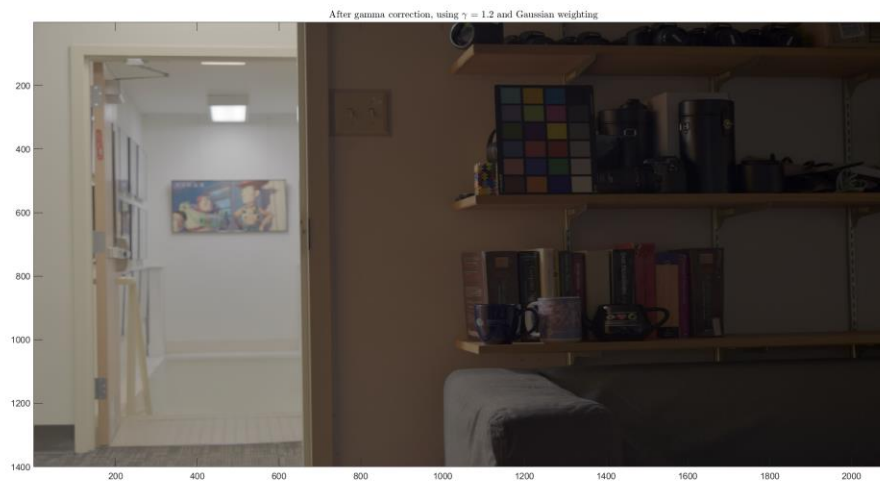








Αξιοσημείωτο είναι ότι για την περίπτωση εισόδου των αποτελεσμάτων της προηγούμενης ενότητας χρησιμοποιώντας *Gaussian weighting function* και $\gamma = 1.2$ φαίνεται η σχέση για τις φωτεινότητες των pixel της παλέτας να είναι πραγματικά γραμμική. Το σχετικό αποτέλεσμα παρουσιάζεται παρακάτω:



3 Radiometric calibration - Εκτίμηση καμπύλης απόκρισης κάμερας

Δεδομένου ότι η καμπύλη απόκρισης μιας κάμερας είναι γενικά μια μη γραμμική συνάρτηση (σε αντίθεση με τη θεώρηση στις προηγούμενες ενότητες), ότι οι τιμές της ακτινοβολίας σε κάθε pixel E_{ij} είναι σταθερές ως προς τις διαφορετικές λήψεις/εικόνες και ότι οι τιμές της φωτεινότητας των pixel (μιας εικόνας 8bit, όπως στην συγκεκριμένη περίπτωση) είναι οι ακέραιοι στο διάστημα $[0, 255]$, μπορούμε να εκτιμήσουμε τις 256 τιμές της συνάρτησης $g = \ln f^{-1}$ λύνοντας το πρόβλημα βελτιστοποίησης (για κάθε κανάλι) :

$$\min_{g, E_{ij}} \sum_{i,j} \sum_k \{w(Z_{ij}^k)[g(Z_{ij}^k) - \ln(E_{ij}) - \ln(t^k)]\}^2 + \lambda \sum_{z=0}^{255} w(z) \nabla^2 g(z)$$

όπου ελαχιστοποιείται μια κυρτή συνάρτηση κόστους και εξασφαλίζεται η βέλτιστη ικανοποίηση της σχέσης

$$\ln(E_{ij}) = \ln(Z_{ij}^k) - \ln(t^k)$$

σύμφωνα με την μέθοδο ελαχίστων τετραγώνων, ενώ ο δεύτερος όρος εξασφαλίζει την ομαλότητα της συνάρτησης g .

Η λαπλασιανή που εμφανίζεται στην συνάρτηση κόστους μπορεί να υλοποιηθεί κατά τα γνωστά ως $\nabla^2 g(z) = g(z+1) - 2g(z) + g(z-1)$.

Έχοντας εκτιμήσει τις τιμές της συνάρτησης g μπορούν να επαναληφθούν τα βήματα της πρώτης ενότητας, με την διαφορά ότι πλέον δεν γίνεται υπόθεση γραμμικότητας της f , αλλά χρησιμοποιούνται οι τιμές της g για να υπολογιστούν τελικά οι τιμές $\ln(E_{ij})$.

Δημιουργείται έτσι η συνάρτηση

```
function responseCurve = estimateResponseCurve(imgStack, exposureTimes, smoothingLamda, weightingFcn, resize_factor, Zmin, Zmax)
```

όπου λαμβάνονται ως είσοδοι οι τιμές της φωτεινότητας ($imgStack - Z_{ij}^k$) για κάθε pixel και κάθε διαφορετική λήψη (για ένα κανάλι των αρχικών εικόνων κάθε φορά), οι χρόνοι έκθεσης ($exposureTimes - t^k$), η επιλογή της παραμέτρου λ του προβλήματος βελτιστοποίησης ($smoothingLamda$), η επιλογή της συνάρτησης βάρους που θα χρησιμοποιηθεί ($weightingFcn$), ένας παράγοντας ($resize_factor$) σύμφωνα με τον οποίο υποδειγματοληπτούνται οι αρχικές εικόνες για υπολογιστικούς/πρακτικούς λόγους και τα κατώφλια Z_{min}, Z_{max} που χρησιμοποιούνται στις συναρτήσεις βάρους. Ως έξοδος προκύπτει το διάνυμα των τιμών της συνάρτησης g .

Υπολογίζεται το διάνυμα W με τα βάρη για τις τιμές του $Z \in [0, 255]$ και οι λογάριθμοι των χρόνων έκθεσης και το πρόβλημα βελτιστοποίησης λύνεται χρησιμοποιώντας την συνάρτηση

```
function [g, lE]=gsolve(Z, B, l, w, flag)
```

η οποία αποτελεί μια ελαφρώς τροποποιημένη εκδοχή αυτής που προτείνεται στο paper των *Debevec, Malik*.

Ως είσοδο η τελευταία συνάρτηση λαμβάνει τις τιμές της φωτεινότητας των υποδειγματοληπτημένων εικόνων για κάθε λήψη (Z), όπου οι χωρικές διαστάσεις έχουν συμπυκνωθεί σε μία γραμμική διάσταση και η δεύτερη διάσταση του πίνακα Z αφορά τις λήψεις με διαφορετικό χρόνο έκθεσης, τις τιμές των λογαρίθμων των χρόνων έκθεσης (B), την παράμετρο λ του προβλήματος βελτιστοποίησης (l), το διάνυμα που περιέχει τα βάρη (w) και μια σημαία που ενεργοποιείται όταν χρησιμοποιείται η συνάρτηση βάρους W_{photon} .

Δημιουργείται ο πίνακας A που περιέχει τις $M \cdot N \cdot K + 256 + 1$ παραμέτρους (μία για κάθε ένα από τα $M \cdot N$ pixels των K υποδειγματοληπτημένων εικόνων + 256 για κάθε εξίσωση εξομάλυνσης + 1 για να εξασφαλίζεται ότι το κέντρο της κατανομής της g θα αντιστοιχίζεται στη μονάδα) των $256 + M \cdot N$ αγνώστων (οι 256 τιμές της συνάρτησης g και οι $M \cdot N$ τιμές των $\ln(E_{ij})$) και το διάνυμα b με τις γνωστές παραμέτρους $\ln(t^k)$.

Λύνεται έτσι το (υπερπροσδιορισμένο) σύστημα εξισώσεων $Ax = b$ σύμφωνα με την μέθοδο των ελαχίστων τετραγώνων και επιστρέφονται οι εκτιμήσεις για τις τιμές της συνάρτησης g , μέσω των οποίων επαναπροσδιορίζονται οι τιμές $\ln(E_{ij})$ για τις αρχικές (πριν την υποδειγματοληψία) εικόνες, όπως έγινε και στην 1η ενότητα, δηλαδή:

$$\ln(E_{ij}) = \frac{\sum_{k=1}^K w(Z_{ij}^k)(g(Z_{ij}^k) - \ln(t^k))}{\sum_{k=1}^K w(Z_{ij}^k)}$$

όπου χρησιμοποιούνται παρόμοιες με την 1η ενότητα συναρτήσεις βάρους (οι σχέσεις έχουν μετασχηματιστεί κατάλληλα για να λαμβάνουν τιμές στο διάστημα $[0, 255]$) και χρησιμοποιούνται και πάλι κατώφλια στις συναρτήσεις βάρους για να αποφευχθούν φαινόμενα “καψίματος” στην τελική εικόνα (τα ίδια κατώφλια χρησιμοποιούνται και κατά την εκτίμηση των τιμών της συνάρτησης g)

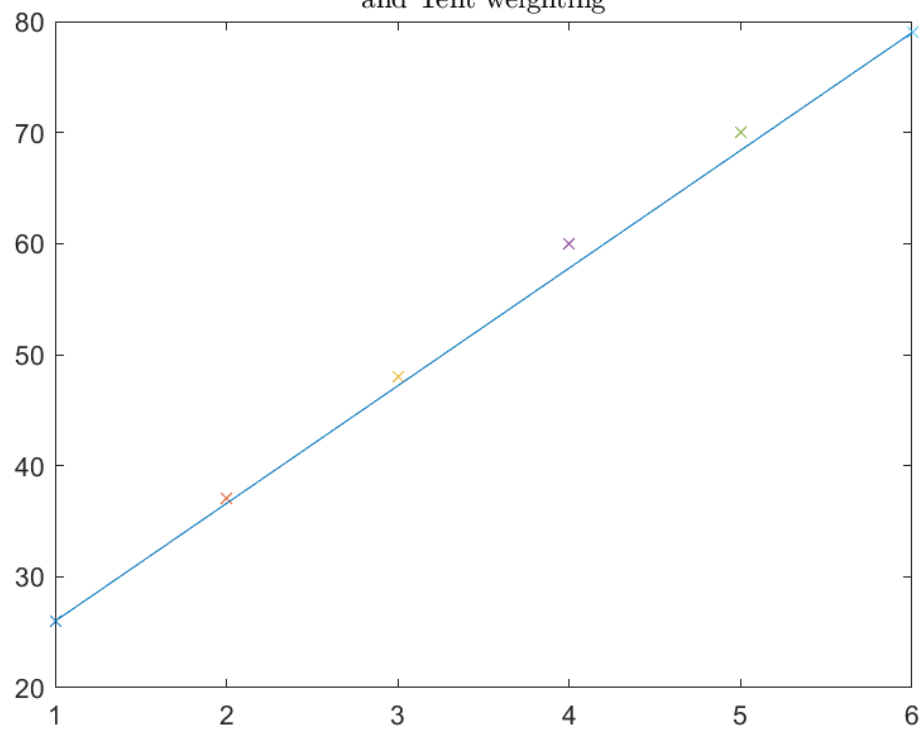
Παρακάτω παρουσιάζονται τα αποτελέσματα του αλγορίθμου για το πρώτο σετ εικόνων όταν χρησιμοποιείται παράγοντας υποδειγματοληψίας $\frac{1}{64}$, για επιλογή της παραμέτρου $\lambda = 300$ και $Z_{min} = 13 \cong 0.05 \cdot 255$, $Z_{max} = 252 \cong 0.99 \cdot 255$:



όπου δεν έχει χρησιμοποιηθεί ουσιαστικά Tone Mapping ($\gamma = 1$), αφού οι φωτεινότητες των pixel της

παλέτας προσεγγίζουν αρκετά καλά την γραμμική σχέση:

Intensities of the palette pixels for $\gamma = 1$
and Tent weighting



Παρακάτω παρουσιάζονται τα αποτελέσματα του αλγορίθμου για το δεύτερο σετ εικόνων όταν χρησιμοποιείται παράγοντας υποδειγματοληψίας $\frac{1}{32}$, για επιλογή της παραμέτρου $\lambda = 100$ και και $Z_{min} = 13 \cong 0.05 \cdot 255$, $Z_{max} = 252 \cong 0.99 \cdot 255$:

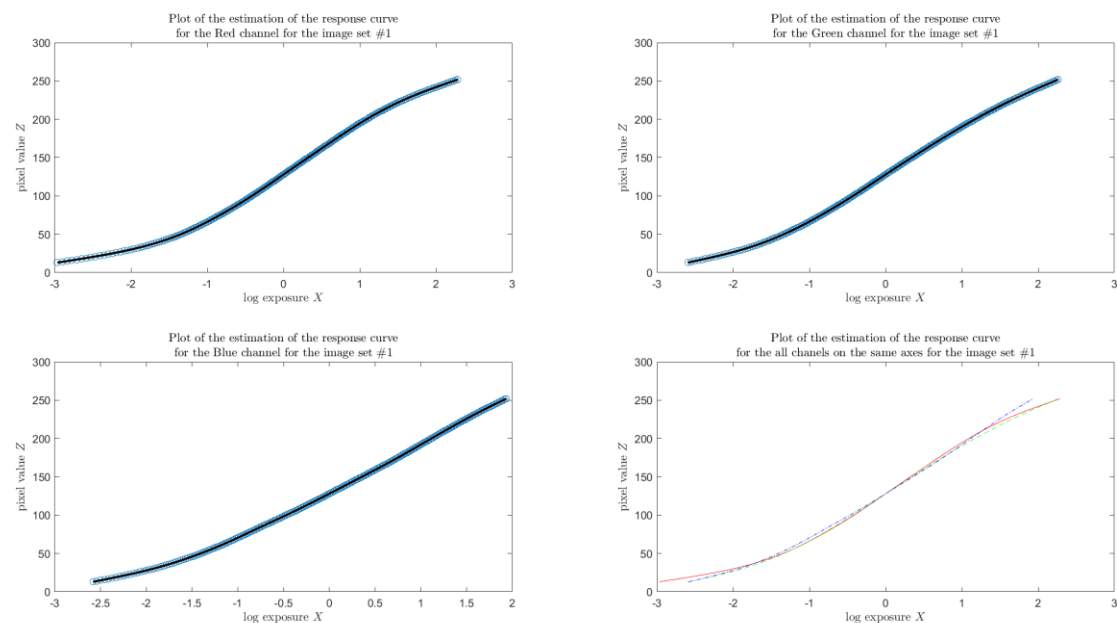


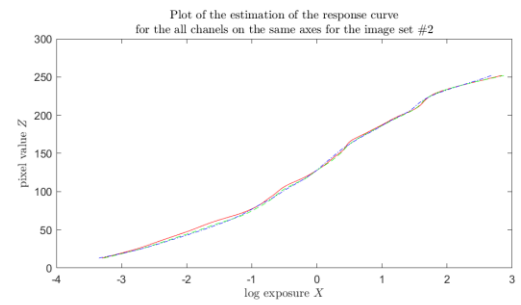
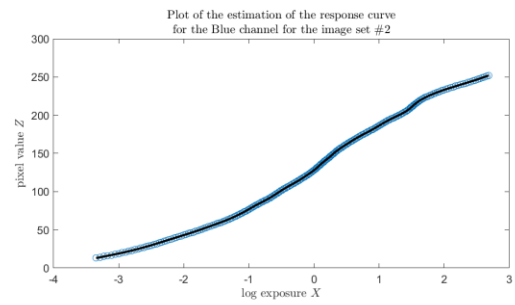
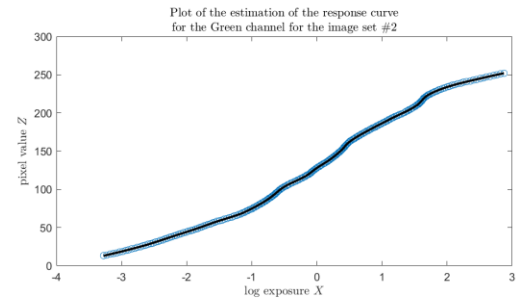
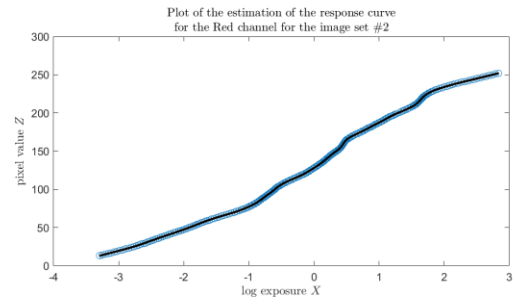
όπου έχει χρησιμοποιηθεί Tone Mapping με Gamma correction και $\gamma = 0.8$.

Είναι εμφανές ότι το λάθος orientation της μίας εικόνας του set επηρεάζει σημαντικά την τελική HDR εικόνα.

Και στις δύο παραπάνω περιπτώσεις έχει χρησιμοποιηθεί *Tent weighting function*, η οποία συμφωνεί με το paper των *Debevec, Malik* και διαπιστώθηκε ότι δίνει καλύτερα αποτελέσματα.

Παρακάτω παρουσιάζονται οι εκτιμήσεις των συναρτήσεων απόκρισης για κάθε σετ εικόνων και για κάθε κανάλι:





4 Βελτίωση HDR απεικόνισης με image alignment

Με σκοπό να διορθωθεί η τελευταία αστοχία που παρατηρείται για το δεύτερο σετ εικόνων χρησιμοποιούνται μέθοδοι παρόμοιες με αυτές που χρησιμοποιήθηκαν στην 1^η εργασία.

Συγκεκριμένα επιλέγεται η τελευταία (εκτεθειμένη για περισσότερο χρόνο) φωτογραφία ως αναφορά και βάσει αυτής γίνεται ο μετασχηματισμός της εικόνας που γνωρίζουμε ότι έχει περιστραφεί.

Εντοπίζονται τα ιδιάζοντα σημεία/ σημεία ενδιαφέροντος, σύμφωνα με τον αλγόριθμο *SURF* και για αυτά υπολογίζονται οι σχετικοί περιγραφείς.

Γίνεται ταίριασμα των διανυσμάτων των περιγραφέων και από τα ζεύγη των διανυσμάτων που προκύπτουν όμοια, υπολογίζεται ο μετασχηματισμός και τελικά η ζητούμενη εικόνα.

Το αποτέλεσμα που λαμβάνουμε χρησιμοποιώντας αυτή, αντί της εικόνας απ την οποία προήλθε, στο σετ των προηγούμενων εικόνων και επαναλαμβάνοντας τον αλγόριθμο της προηγούμενης ενότητας παρουσιάζεται παρακάτω:



όπου έχει επιλεγθεί και πάλι *Tent weighting function*, παράγοντας υποδειγματοληψίας $\frac{1}{32}$, $\lambda = 100$, $Z_{min} = 13 \cong 0.05 \cdot 255$, $Z_{max} = 252 \cong 0.99 \cdot 255$ και έχει γίνει Gamma Correction με $\gamma = 0.8$.

Το αποτέλεσμα αυτό προφανώς υπερέχει συγκριτικά με αυτό της προηγούμενης ενότητας.