

Ψηφιακή Επεξεργασία Εικόνας

-Εργασία 1: Από τον RGB αισθητήρα στη μνήμη

Α. Ντελόπουλος

Άνοιξη 2023

1 Εισαγωγικά

1.1 RAW format

Η ψηφιακή καταγραφή έγχρωμων εικόνων με τη χρήση RGB αισθητήρων υλοποιείται από πλέγμα αισθητήρων τοποθετημένων σε μια προτυποποιημένη διάταξη όπως για παράδειγμα αυτή του πρότυπου Bayer (βλ Εικόνα 1)

Καθένας από τους τρεις τύπους αισθητήρων έχει διαφορετικό Quantum Efficiency. Στο Σχήμα 2 εμφανίζεται το QE των τριών αισθητήρων ενός διαδεδομένου CMOS που χρησιμοποιείται σε φωτογραφικές μηχανές της Nikon. Συνεπώς κατά τη λήψη της φωτογραφίας σε κάθε θέση του πλέγματος καταγράφεται ένα μόνο από τα τρία κανάλια χρώματος R, G ή B. Ο παραγόμενος πίνακας καταγραφών αποτελεί την εικόνα σε RAW FORMAT. Η raw εικόνα έχει διαστάσεις $M_0 \times N_0$ ακριβώς ίσες με τις διαστάσεις του πλέγματος. Οι RGB εικόνες που συνήθως χρησιμοποιούμε αποτελούν προϊόν μετέπειτα υπολογισμών που περιγράφονται στην Ενότητα 1.3.

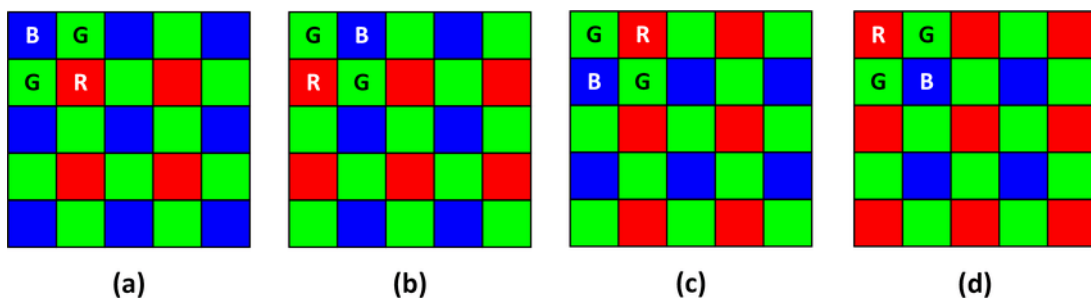
Μία πολύ καλή περιγραφή του RAW Format και της διαδικασίας που απαιτείται για την μετατροπή του σε RGB θα βρείτε στο [4]. Οι επόμενες ενότητες έχουν σε σημαντικό βαθμό στηριχθεί στις πληροφορίες που περιέχονται εκεί.

1.2 Εναλλακτικά format αποθήκευσης RAW εικόνων. Το Digital Negative της Adobe

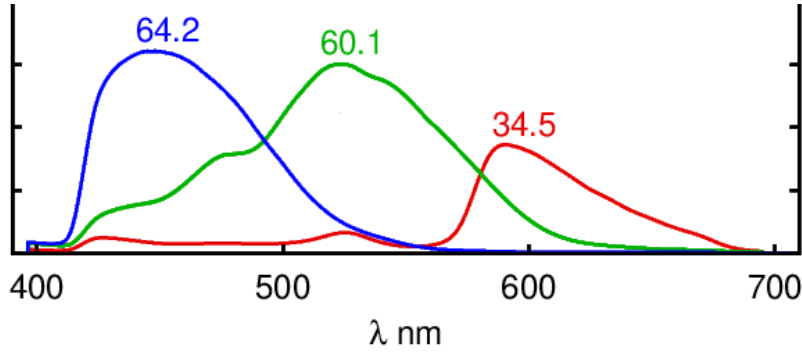
Ο ακριβής τρόπος με τον οποίο αποθηκεύεται το RAW FORMAT μαζί με τα συνοδευτικά metadata διαφοροποιείται μεταξύ των διαφόρων κατασκευαστών φωτογραφικών μηχανών. Η Nikon χρησιμοποιεί τα αρχεία .NEF η Canon τα .CR2 κ.λπ. Η Adobe έχει προτείνει ένα κοινό ανοιχτό format που το ονομάζει Digital Negative [1] και τα αντίστοιχα αρχεία έχουν κατάληξη .DNG. Επιπλέον έχει κατασκευάσει έναν ελεύθερα διαθέσιμο μετατροπέα από τα Format των κύριων κατασκευαστών σε .DNG.[2]. Στην παρούσα εργασία θα χρησιμοποιήσουμε εικόνες σε RAW FORMAT που έχει ήδη μετατραπεί σε .DNG.

1.3 Μετατροπή RAW εικόνας σε RGB

Κατά τη μετατροπή της RAW εικόνας σε RGB απαιτούνται τρία βασικά στάδια: (i) η ρύθμιση του white balance, (ii) η διαδικασία παρεμβολής τιμών R, G και B ώστε κάθε δείγμα της παραγόμενης εικόνας να έχει τιμή για κάθε ένα από τα τρία κανάλια R/G/B και (iii) η μετατροπή από το color space της κάμερας στο color space αναπαραγωγής (display) της εικόνας.



Σχήμα 1: Bayer patterns: (a) BGGR, (b) GBRG, (c) GRBG, and (d) RGGB.



Σχήμα 2: Το Quantum Efficiency των CMOS αισθητήρων που χρησιμοποιεί η Nikon στα μοντέλα DSLR D700, D800

1.3.1 White Balance

Λευκές περιοχές ενός αντικειμένου θα πρέπει να φαίνονται λευκές, δηλαδή θα πρέπει το καταγραφόμενο χρώμα a να είναι της μορφής

$$a = \lambda \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \lambda \in [0, 1]$$

Επειδή όμως το φάσμα του προσπίπτοντος φωτός δεν είναι πάντα ισοκατανομημένο στις περιοχές R, G, B μπορεί η παραπάνω συνθήκη να ανατραπεί οπότε χρειάζεται να πολλαπλασιαστούν με κατάλληλους διορθωτικούς συντελεστές τα 2 από τα τρία κανάλια. Συνήθως διατηρείται αμετάβλητο το G και πολλαπλασιάζονται με διορθωτικούς συντελεστές τα άλλα δύο (R, B). Οι συντελεστές που χρησιμοποιούνται είναι κοινοί για όλα τα pixels της εικόνας καθώς υιοθετείται η απλούστευση ότι το φάσμα του προσπίπτοντος φωτός είναι ίδιο σε όλα τα εικονιζόμενα σημεία. Οι συντελεστές white balance είναι αποθηκευμένοι στα metadata της εικόνας.

1.3.2 Παρεμβολή

Με τη χρήση παρεμβολής υπολογίζεται η τιμή του καναλιού $C \in \{R, G, B\}$ σε μία θέση x, y με βάση τις γνωστές τιμές του ίδιου καναλιού σε γειτονικές θέσεις. Ο αλγόριθμος παρεμβολής ποικίλλει από την απλή "αντιγραφή" της τιμής του πλησιέστερου γείτονα ως τη χρήση πολυωνυμικής παρεμβολής 2Δ. Οι θέσεις x, y επιλέγονται σε ένα πλέγμα σημείων με πυκνότητα που επιλέγεται από τον χρήστη και δεν είναι απαραίτητο να συμπίπτουν με θέσεις του πρωτογενούς πλέγματος του αισθητήρα.

1.3.3 Μετατροπή color space

Ο κατασκευαστής αποθηκεύει στα metadata τον 3×3 πίνακα $T_{XYZ \rightarrow Cam}$ που συνδέει το 3Δ colorspace της κάμερας με το πρότυπο colorspace XYZ:

$$C_{cam} = \begin{bmatrix} R_{cam} \\ G_{cam} \\ B_{cam} \end{bmatrix} = T_{XYZ \rightarrow Cam} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = C_{XYZ} \quad (1)$$

Από την άλλη μεριά το πρότυπο CIE [3] ορίζει τον αντίστοιχο πίνακα $T_{XYZ \rightarrow RGB}$ μετατροπής του XYZ στο τυποποιημένο σύστημα RGB που χρησιμοποιούν οι συσκευές προβολής (οθόνες, προβολείς)

$$C_{linear} \begin{bmatrix} R_{linear} \\ G_{linear} \\ B_{linear} \end{bmatrix} = T_{XYZ \rightarrow RGB} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = C_{XYZ}$$

Ο πίνακας έχει τις τιμές:

$$T_{XYZ \rightarrow RGB} = \begin{bmatrix} +3.2406 & -1.5372 & -0.4986 \\ -0.9689 & +1.8758 & +0.0415 \\ +0.0557 & -0.2040 & +1.0570 \end{bmatrix}$$

Το πρότυπο προβλέπει μια ακόμη μη γραμμική διόρθωση που περιγράφεται για κάθε κανάλι από την εξίσωση

$$C_{\text{sRGB}} = \begin{cases} 12.92C_{\text{linear}}, & C_{\text{linear}} \leq 0.0031308 \\ 1.055C_{\text{linear}}^{1/2.4} - 0.055, & C_{\text{linear}} > 0.0031308 \end{cases}, C_{\text{linear}} = R_{\text{linear}}, G_{\text{linear}}, B_{\text{linear}}$$

η απλούστερα -χωρίς σοβαρή απόκλιση

$$C_{\text{sRGB}} = C_{\text{linear}}^{1/2.2}$$

2 Παραδοτέα

Για τις ανάγκες της εργασίας θα πρέπει να κατασκευάσετε

1. Ένα σύνολο συναρτήσεων σε matlab που θα υλοποιούν τη μετατροπή μίας εικόνας από RAW DGN format σε RGB
2. Ένα πρόγραμμα επίδειξης με το όνομα `demo.m`
3. Ένα report που θα εξηγείτε τη λειτουργία των συναρτήσεων, θα επιδεικνύετε και θα σχολιάζετε τα αποτελέσματα και θα περιγράφετε πιθανές παραδοχές που κάνατε κατά την υλοποίηση

2.1 readrng.m

Να κατασκευάσετε συνάρτηση

```
1 function [rawim,XYZ2Cam,wbc coeffs ] = readrng(filename)
```

που δέχεται ως είσοδο το όνομα (ή path) `filename` ενός αρχείου `.DNG` και επιστρέφει τον πίνακα `rawim` με τις μετρήσεις του αισθητήρα, τον 3×3 πίνακα `XYZ2Cam` της εξ. 1 και το 1×3 διάνυσμα `wbc coeffs` με τους συντελεστές διόρθωσης του white balance.

Δεδομένου ότι τα `.DNG` αρχεία της Adobe έχουν τη δομή αρχείων `tiff` μπορείτε να χρησιμοποιήσετε τη βιβλιοθήκη `LibTIFF` μέσω της `matlab`.

Με τις ακόλουθες εντολές μπορείτε να διαβάσετε την RAW εικόνα

```
1 obj = Tiff(filename, 'r');
2 offsets = getTag(obj, 'SubIFD');
3 setSubDirectory(obj, offsets(1));
4 rawim = read(obj);
5 close(obj);
```

και με τις παρακάτω διαβάζετε τα χρήσιμα metadata

```
1 meta_info = imfinfo(filename);
2 % (x_origin,y_origin) is the upper left corner of the useful part of the
   sensor and consequently of the array rawim
3 y_origin = meta_info.SubIFDs{1}.ActiveArea(1)+1;
4 x_origin = meta_info.SubIFDs{1}.ActiveArea(2)+1;
5 % width and height of the image (the useful part of array rawim)
6 width = meta_info.SubIFDs{1}.DefaultCropSize(1);
7 height = meta_info.SubIFDs{1}.DefaultCropSize(2);
8
9 blacklevel = meta_info.SubIFDs{1}.BlackLevel(1); % sensor value corresponding
   to black
10 whitelevel = meta_info.SubIFDs{1}.WhiteLevel; % sensor value corresponding to
   white
11
12 wbc coeffs = (meta_info.AsShotNeutral).^ -1;
13 wbc coeffs = wbc coeffs/wbc coeffs(2); % green channel will be left unchanged
```

```

14
15 XYZ2Cam = meta_info.ColorMatrix2;
16 XYZ2Cam = reshape(XYZ2Cam,3,3)';

```

H `readrng.m` θα πρέπει να ενσωματώσει τις παραπάνω εντολές και τελικά να δώσει στην έξοδο μόνο το χρήσιμο μέρος του πίνακα `rawim` με τιμές που θα έχουν υποστεί σημειακό μετασχηματισμό έτσι ώστε το `blacklevel` να οδηγηθεί στο 0 και το `whitelevel` στο 1.

Σημείωση: Επειδή λόγω θορύβου μπορεί παρά τις προδιαγραφές των `blacklevel` και `whitelevel` η εικόνα να έχει τιμές και έξω από τα όρια, μετά την εφαρμογή του μετασχηματισμού να κάνετε αποκοπή όσων τιμών τύχει να βρεθούν έξω από το διάστημα $[0, 1]$ (`rawim = max(0,min(rawim,1));`)

2.2 dng2rgb.m

```

1 function [Csrgb, Clinear, Cxyz, Ccam] = dng2rgb(rawim, XYZ2Cam, wbc coeffs,
    bayertype, method, M, N)

```

που δέχεται σαν είσοδο τον $M_0 \times N_0$ πίνακα `rawim` με τις τιμές των δειγμάτων μιας εικόνας σε θέσεις που περιγράφονται από το φίλτρο Bayer (σχήμα 1).

Οι μεταβλητές εισόδου `XYZ2Cam` και `wbc coeffs` έχουν παραχθεί από τη συνάρτηση `readrng.m` ενώ η μεταβλητή `bayertype` λαμβάνει τιμές BGGR, GBGR, GRBG, RGGB για να δηλώσουν την υποκατηγορία του Bayer pattern όπως φαίνεται στο Σχήμα 1.

Η συνάρτηση επιστρέφει τις εικόνες `Csrgb`, `Clinear`, `Cxyz`, `Ccam` διάστασης $M \times N$ που αντιστοιχούν στις εικόνες C_{sRGB} , C_{linear} , C_{XYZ} , C_{cam} της Ενότητας 1.3.3

Για να το πετύχει αυτό, η συνάρτηση κατασκευάζει ένα νέο grid συντεταγμένων $M \times N$ (για κάθε χρώμα), έτσι ώστε καθένα από τα τέσσερα γωνιακά σημεία του να συμπίπτουν με τα αντίστοιχα γωνιακά σημεία του grid της εικόνας εισόδου. Θεωρείστε ότι οι φυσικές διαστάσεις της εικόνας δεν μεταβάλλονται. Στη συνέχεια, χρησιμοποιεί τη μέθοδο παρεμβολής `method` για να “χρωματίσει” κάθε σημείο του νέου grid.

Όταν η μεταβλητή εισόδου `method` έχει την τιμή 'nearest' κατά την παρεμβολή θα πρέπει να χρησιμοποιείται η μέθοδος του κοντινότερου γείτονα. Όταν η ίδια μεταβλητή δέχεται την τιμή 'linear' θα πρέπει να χρησιμοποιείται διγραμμική παρεμβολή (bilinear interpolation). Δεν μπορείτε να χρησιμοποιήσετε έτοιμες συναρτήσεις για παρεμβολή, όπως πχ η `interp1`. Ο χειρισμός κάθε καναλιού (χρώματος) θα είναι ανεξάρτητος.

Κατά τη μετατροπή από RAW σε RGB θα πρέπει να εφαρμόσετε τους μετασχηματισμούς που περιγράφηκαν στις Ενότητες 1.3.1 και 1.3.3.

2.3 demo.m

Με το script `demo.m` να επιδείξετε τη λειτουργία των άλλων δύο συναρτήσεων χρησιμοποιώντας την εικόνα `RawImage.dng` που σας δίνεται. Οι εικόνες που θα περιλάβετε στην αναφορά της εργασίας θα πρέπει να παραχθούν από το script αυτό. Για κάθε εικόνα εξόδου η `demo.m` πρέπει να παράγει και τα αντίστοιχα ιστογράμματα χωριστά για κάθε κανάλι της.

3 Για την υποβολή της εργασίας

Παραδώστε μία αναφορά με τις περιγραφές και τα συμπεράσματα που σας ζητούνται στην εκφώνηση. Η αναφορά θα πρέπει να επιδεικνύει την ορθή λειτουργία του κώδικά σας στην εικόνα που σας δίνεται και να παρουσιάζει και σχολιάζει τις εικόνες και τα ιστογράμματα που παράγονται από το πρόγραμμα επίδειξης.

Ο κώδικας θα πρέπει να είναι σχολιασμένος ώστε να είναι κατανοητό τι ακριβώς λειτουργία επιτελεί (σε θεωρητικό επίπεδο, όχι σε επίπεδο κλίσης συναρτήσεων). Επίσης, ο κώδικας θα πρέπει να εκτελείται και να υπολογίζει τα σωστά αποτελέσματα για οποιαδήποτε είσοδο πληροί τις υποθέσεις της εκφώνησης, και όχι μόνο για την εικόνα που σας δίνεται.

Απαραίτητες προϋποθέσεις για την βαθμολόγηση της εργασίας σας είναι ο κώδικας να εκτελείται χωρίς σφάλμα, καθώς και να τηρούνται τα ακόλουθα:

- Υποβάλετε ένα και μόνο αρχείο, τύπου zip.

- Το όνομα του αρχείου πρέπει να είναι AEM.zip, όπου AEM είναι τα τέσσερα ψηφία του Α.Ε.Μ. του φοιτητή της ομάδας.
- Το προς υποβολή αρχείο πρέπει να περιέχει τα αρχεία κώδικα Matlab και το αρχείο report.pdf το οποίο θα είναι η αναφορά της εργασίας.
- Η αναφορά πρέπει να είναι ένα αρχείο τύπου PDF, και να έχει όνομα report.pdf.
- Όλα τα αρχεία κώδικα πρέπει να είναι αρχεία κειμένου τύπου UTF-8, και να έχουν κατάληξη m.
- Το αρχείο τύπου zip που θα υποβάλετε δεν πρέπει να περιέχει κανέναν φάκελο.
- Μην υποβάλετε τις εικόνες που σας δίνονται για πειραματισμό.
- Μην υποβάλετε αρχεία που δεν χρειάζονται για την λειτουργία του κώδικά σας, ή φακέλους/αρχεία που δημιουργεί το λειτουργικό σας, πχ “Thumbs.db”, “.DS_Store”, “.directory”.
- Για την ονομασία των αρχείων που περιέχονται στο προς υποβολή αρχείο, χρησιμοποιείτε μόνο αγγλικούς χαρακτήρες, και όχι ελληνικούς ή άλλα σύμβολα, πχ “#”, “\$”, “%” κλπ.

Αναφορές

- [1] Adobe Systems Inc. *Digital Negative Specifications*. http://www.images.adobe.com/www.adobe.com/content/dam/Adobe/en/products/photoshop/pdfs/dng_spec_1.4.0.0.pdf, 2012.
- [2] Adobe Systems Inc. *Camera raw, DNG : Downloads*. <http://www.adobe.com/products/photoshop/extend.displayTab2.html>, downloads, 2013.
- [3] T. Smith and J. Guild, *The C.I.E. colorimetric standards and their use*, Transactions of the Optical Society, vol 33, no 3, p. 73, 1931.
- [4] Rob Sumner, *Processing RAW Images in MATLAB*, https://rcsumner.net/raw_guide/RAWguide.pdf, May 19, 2014