

ΨΗΦΙΑΚΗ ΕΠΕΞΕΡΓΑΣΙΑ ΕΙΚΟΝΑΣ ΚΑΙ ΕΦΑΡΜΟΓΕΣ

Εργασία #1: Εισαγωγή, μετασχηματισμοί,
βελτιστοποίηση και τμηματοποίηση

it2022134
Εξάρχου Άθως

Πρέπει να έχετε εγκατεστημένες τις εξής βιβλιοθήκες (με pip install):

numpy, matplotlib, scikit-learn, pillow, scikit-image, opencv-python

Άσκηση #1, Κβάντιση:

Όσο αυξάνεται το K , δηλαδή όσο περισσότερα χρώματα χρησιμοποιούνται για την κβάντιση το MSE μειώνεται σημαντικά. Αυτό δείχνει ότι η κβαντισμένη εικόνα προσεγγίζει καλύτερα την αρχική.

- ❖ $K = 5$:
 - Πολύ απλοποιημένη εικόνα.
 - Μεγάλο σφάλμα (595.70), σημαντική απώλεια λεπτομερειών και χρωματικής ποικιλίας.
 - Η εικόνα μοιάζει αφύσικη.
- ❖ $K = 20$:
 - Το σφάλμα μειώνεται κατά ~75% ($MSE=158.03$), και η εικόνα διατηρεί βασικά χαρακτηριστικά, αλλά με περιορισμένη χρωματική ακρίβεια.
- ❖ $K = 200$:
 - Πιο ρεαλιστική απεικόνιση, αρκετά κοντά στην αρχική.
 - Το MSE είναι πλέον πολύ μικρό (25.84), και η οπτική ποιότητα είναι ικανοποιητική.
- ❖ $K = 1000$:
 - Αν όχι καθόλου, σχεδόν μη αντιληπτή διαφορά από την αρχική.
 - Πολύ μικρό MSE (8.90), αλλά το επιπλέον όφελος σε ποιότητα είναι ασυγκρίτως μικρότερο σε σχέση με την αύξηση του αριθμού χρωμάτων, ή τουλάχιστον δεν φαίνεται με το γυμνό μάτι.

Αρχική Εικόνα



$K=5$
 $MSE=595.70$



K=20
MSE=158.03



K=200
MSE=25.84



K=1000
MSE=8.90



Άσκηση #2, Μετασχηματισμός Fourier:

A. Υλοποίησα τον 2D μετασχηματισμό Fourier (DFT) σε εικόνα αποχρώσεων του γκρι και απεικόνισα το φάσμα πλάτους, το οποίο δείχνει την ένταση των συχνοτήτων, και το φάσμα φάσης, το οποίο καθορίζει τη χωρική κατανομή των στοιχείων της εικόνας. Η φάση δύσκολα είναι κατανοητή οπτικά, αλλά παίζει κρίσιμο ρόλο στην ανακατασκευή της εικόνας.

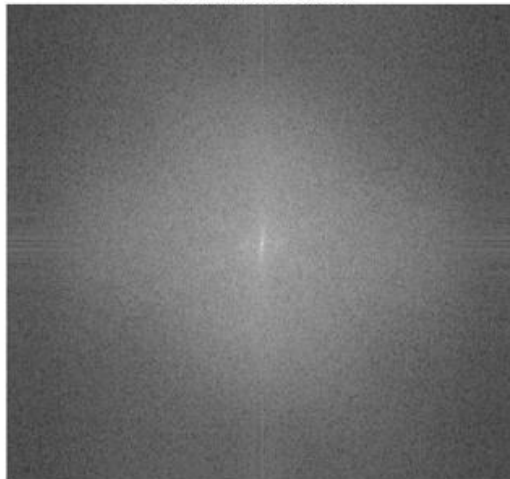
B. Το φάσμα φάσης τροποποίησα μέσω συμμετρίας ως προς τον οριζόντιο άξονα (np.flipud). Η μεταβολή αυτή δεν επηρέασε την απεικόνιση του ίδιου του φάσματος φάσης (οπτικά), όμως είχε σημαντικές επιπτώσεις κατά την ανακατασκευή της εικόνας.

Γ. Πραγματοποίησα τον αντίστροφο μετασχηματισμό Fourier με το τροποποιημένο φάσμα (ifftshift, ifft2). Ως αποτέλεσμα, η τροποποιημένη εικόνα έχει αντιστραφεί ως προς τον οριζόντιο άξονα και παρουσιάζει αλλοιώσεις στη γεωμετρία, καθώς η αλλαγή φάσης επηρεάζει τη διάταξη των συχνοτήτων στον χώρο και παρουσιάζει παραμόρφωση, επειδή η φάση είναι κρίσιμη για τη διατήρηση της πληροφορίας της εικόνας.

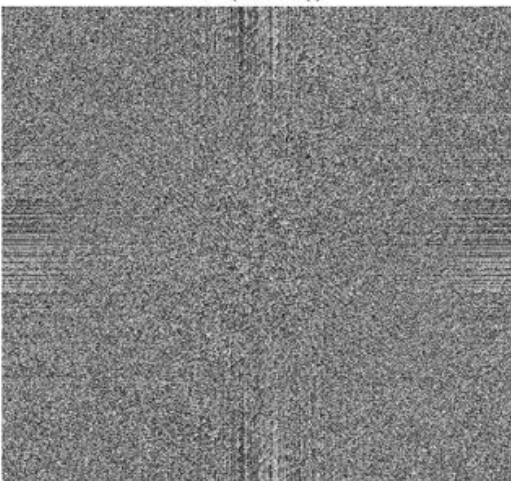
Αρχική εικόνα



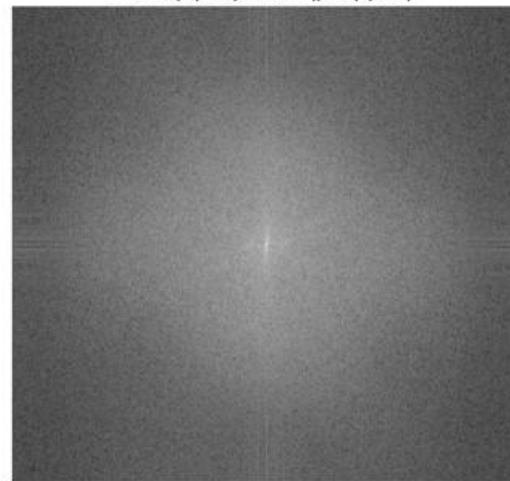
Φάσμα Πλάτους (log)



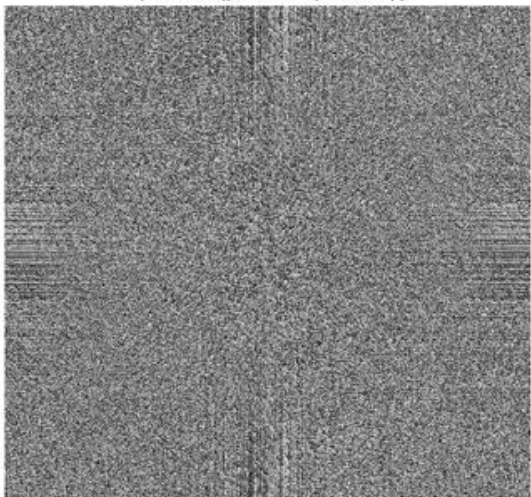
Φάσμα Φάσης



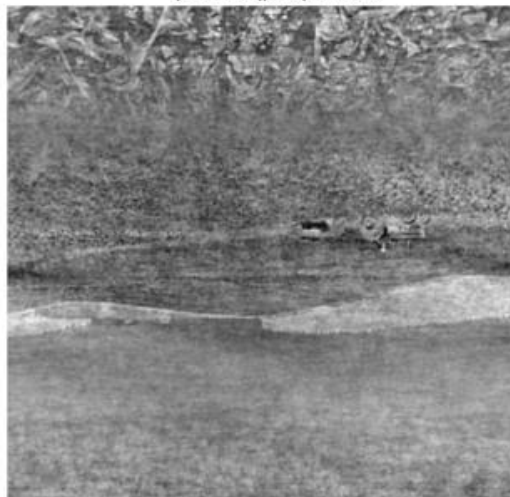
Πλάτος (με τροποποιημένη φάση)



Τροποποιημένο Φάσμα Φάσης



Τροποποιημένη Εικόνα



Άσκηση #3, Φιλτράρισμα:

Χρησιμοποιώ τα φίλτρα Sobel, Canny και LoG, τα οποία διδαχτήκαμε στην αναπληρωματική 'Διάλεξη 5' και αποδείχτηκαν αποτελεσματικά στη πράξη, για τα ερωτήματα Α, Β και Γ αντίστοιχα.

Page | 4

Α. Χρησιμοποιώ το φίλτρο Sobel για να εντοπίσω τις κύριες ακμές που αντιστοιχούν στα βασικά χαρακτηριστικά του προσώπου (π.χ. περίγραμμα προσώπου, μάτια, στόμα). Από το αποτέλεσμα φαίνεται ότι οι ακμές γύρω από τα κύρια χαρακτηριστικά του προσώπου τονίζονται, ενώ περιοχές με μικρότερη αντίθεση ή πιο ομαλές περιοχές δεν καταγράφονται με την ίδια ένταση. Θεώρησα ως βέλτιστο κατώφλι το $T = 65$, ώστε να εξαλειφθεί ο θόρυβος και να εντοπιστούν πιο καθαρά τα βασικά περιγράμματα του προσώπου. Στην εκτεταμένη μείωση του κατωφλίου, καταγράφεται πολύς θόρυβος και στην αύξηση χάνονται σημαντικές λεπτομέρειες.

Β. Μάθαμε πως το φίλτρο Canny βασίζεται σε Gaussian smoothing και εύρεση τοπικών μέγιστων. Έτσι, με την εφαρμογή των παραμέτρων $T1 = 5$, $T2 = 10$ και $\sigma = 2.0$ στην εικόνα 'fruits.jpg', το φίλτρο Canny απομονώνει τα πιο έντονα και σημαντικά περιγράμματα των φρούτων στην εικόνα. Ωστόσο, παρατήρησα και κάποια μικρότερα περιγράμματα ή θόρυβο, που προέρχονται από το χαμηλότερο κατώφλι $T1 (5)$. Αν χρησιμοποιούσα υψηλότερο $T1$, θα είχε απορριφθεί περισσότερος θόρυβος, αλλά ίσως και κάποιες λιγότερο έντονες ακμές. Αντιθέτως, το χαμηλό $T2 (10)$ μπορεί να επιτρέπει την ανίχνευση αρκετών ακμών, ενώ αν ήταν μεγαλύτερο θα οδηγούσε σε πιο σφιχτό περιορισμό και λιγότερες, πιο ισχυρές ακμές, αλλά έτσι χάνεται σημαντική πληροφορία και δεν θα ήταν δυνατό να διακρίνω ποιο φρούτο είναι το καθένα.

Γ. Το φίλτρο LoG εφαρμόζει Gaussian blur και μετά Laplacian. Επέλεξα $\sigma=1.5$ για να επιτύχω θόλωση που θα μειώσει τον θόρυβο χωρίς να καταστρέψει τις σημαντικές λεπτομέρειες. Αν το σ ήταν μικρότερο, η εικόνα θα είχε πιο λεπτομερές θόλωμα, ενώ αν ήταν μεγαλύτερο, θα αφαιρούσε ουσιαστικές λεπτομέρειες. Επέλεξα $T=0.7$ για να φιλτράρω τις πιο έντονες περιοχές από το LoG. Ένα χαμηλότερο κατώφλι (π.χ., 0.4) θα διατηρούσε περισσότερες λεπτομέρειες, αλλά θα εισήγαγε θόρυβο. Ένα υψηλότερο κατώφλι (π.χ., 0.9) θα διατηρούσε μόνο τις πιο έντονες περιοχές της εικόνας. Εφάρμοσα κανονικοποίηση στην εικόνα για να περιορίσω τις τιμές σε ένα εύρος από 0 έως 1 , έτσι ώστε να εφαρμοστεί το κατώφλι σε όλη την εικόνα με ομοιόμορφο τρόπο. Η εικόνα δείχνει καθαρά τις περιοχές που περιέχουν τις έντονες λεπτομέρειες του φύλλου, όπως τις ακμές, τους νευρώνες και τις αλλαγές στην υφή.

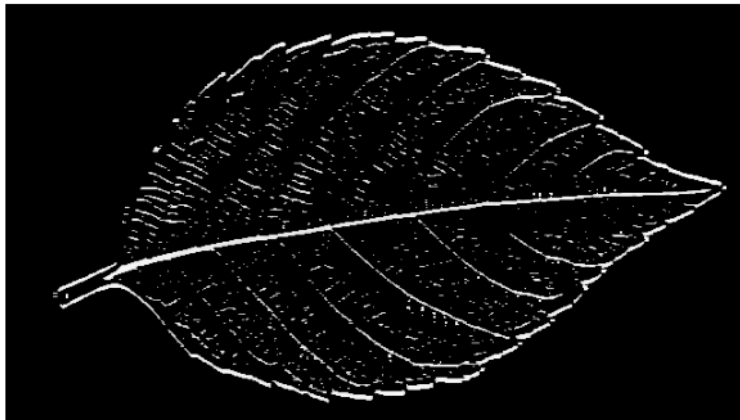
Α. Κύρια στοιχεία (Sobel)



Β. Περίγραμμα αντικειμένων (Canny)



Γ. Λεπτομέρειες αντικειμένων (LoG)



Άσκηση #4, Βελτιστοποίηση:

Gaussian θόρυβος

- Για την εικόνα 'lenna-n1' επέλεξα Gaussian φίλτρο για την αφαίρεση του Gaussian θορύβου με τυπική απόκλιση $\sigma = 2.95$. Αυτό είχε ως αποτέλεσμα η τιμή του SSIM για τον Gaussian θόρυβο να είναι $SSIM = 0.6944$, η οποία δεν είναι πολύ μεγάλη, αλλά το φίλτρο κατάφερε να μειώσει τον θόρυβο χωρίς σημαντική αλλοίωση των βασικών χαρακτηριστικών της εικόνας. Βέβαια, στο μάτι η εικόνα φαίνεται καλύτερα όταν $\sigma = 1$, αλλά η τιμή του SSIM μειώνεται δραστικά (0.4382) και προτίμησα το βέλτιστο SSIM (0.6944).

lenna-n1 (Gaussian noise)



Filtered (SSIM: 0.6944)



Salt & Pepper θόρυβος

- Στην εικόνα 'lenna-n2', ο θόρυβος Salt & Pepper είναι κρουστικός θόρυβος, οπότε διάλεξα το φίλτρο μεσαίας τιμής που διδάχτηκε στη 'Διάλεξη 3'. Ο θόρυβος παρουσιάζεται σαν αλατοπίπερο πάνω στην εικόνα, κάτι το οποίο το median filter είναι πολύ ικανό να αφαιρέσει, καθώς αντικαθιστά κάθε pixel με την "μέση" τιμή της περιοχής γύρω του. Ως αποτέλεσμα, η τιμή του SSIM είναι 0,8474, δείχνοντας ότι το φίλτρο ήταν εξαιρετικά αποτελεσματικό στην απομάκρυνση του θορύβου χωρίς σχεδόν καθόλου απώλεια ποιότητας σε σχέση με την αρχική εικόνα (lenna).

lenna-n2 (Salt & Pepper noise)



Filtered (SSIM: 0.8474)



Περιοδικός θόρυβος

- Για την εικόνα 'lenna-n3', η προσέγγιση για την αφαίρεση του περιοδικού θορύβου ήταν πολύ πιο περίπλοκη σε σχέση με τους δυο προηγούμενους θορύβους, καθώς απαιτούσε την ανάλυση του φάσματος συχνοτήτων της εικόνας μέσω FFT. Στην εικόνα με περιοδικό θόρυβο, οι παράλληλες γραμμές με κλίση $\sim 45^\circ$ αντιπροσωπεύονται σε συγκεκριμένες περιοχές του φάσματος συχνοτήτων.
- Μετέτρεψα την εικόνα στο χώρο των συχνοτήτων χρησιμοποιώντας FFT. Στη συνέχεια, εφάρμοσα το FFT Shift, ώστε το κέντρο του φάσματος να βρίσκεται στο κέντρο της εικόνας. Υπολόγισα το Magnitude Spectrum, το οποίο απεικονίζει την ισχύ κάθε συχνότητας στην εικόνα. Την ανίχνευση κορυφών στο magnitude spectrum πραγματοποιήσα χρησιμοποιώντας τη μέθοδο peak_local_max, η οποία ανιχνεύει περιοχές με υψηλή ένταση που ενδέχεται να αντιπροσωπεύουν τις περιοδικές συχνότητες του θορύβου. Με την παράμετρο threshold_rel=0,5, οι κορυφές που έχουν τουλάχιστον το 50% της

μέγιστης έντασης του φάσματος θεωρούνται σημαντικές και ενδέχεται να αντιπροσωπεύουν περιοδικό θόρυβο. Αν η τιμή του threshold είναι πολύ μικρή, θα μπορούσαν να ανιχνευθούν και ασταθείς ή αδιάφορες κορυφές, ενώ αν είναι πολύ μεγάλη, θα μπορούσαν να χαθούν σημαντικές περιοχές του θορύβου.

- Μετά από αρκετό trial and error, δημιούργησα μάσκα Notch Mask, η οποία εξαλείφει τις περιοχές στο φάσμα που σχετίζονται με τον περιοδικό θόρυβο. Η μάσκα εφαρμόζεται στο φάσμα, αφαιρώντας όσο γίνεται τις περιοδικές συχνότητες. Για να διασφαλιστεί η συμμετρία του θορύβου, εφαρμόσα τη μάσκα και στα συμμετρικά σημεία του φάσματος. Με την παράμετρο notch_radius=12, επηρεάζω άμεσα την ακρίβεια της αφαίρεσης του θορύβου: μικρότερο radius μπορεί να αφήσει λίγο θόρυβο, ενώ μεγαλύτερο μπορεί να καταστρέψει στοιχεία εικόνας. Στη διαδικασία ανίχνευσης των κορυφών, οι ανιχνευμένες κορυφές εντοπίζονται σε περιοχές του φάσματος που δείχνουν περιοδικό θόρυβο. Οι περιοχές αυτές είναι κοντά στην κεντρική συχνότητα του φάσματος (crow, ccol), και το threshold_rel καθορίζει πόσο ευαίσθητο είναι το φίλτρο στις περιοχές που εντοπίζονται ως θόρυβος.
- Μετά την εφαρμογή της μάσκας, υπολόγισα την Inverse FFT για να μετατραπεί το φάσμα πίσω στον χώρο εικόνας. Το αποτέλεσμα είναι μια εικόνα με ελάχιστο θόρυβο, στην οποία οι περιοδικές γραμμές του θορύβου έχουν αφαιρεθεί όσο το δυνατόν περισσότερο. Με το IFFT shift επαναφέρεται το φάσμα στην αρχική του μορφή πριν την εφαρμογή της μάσκας.
- Τέλος, εφαρμόζω ένα ελαφρύ φίλτρο Gaussian με μικρή τιμή τυπικής απόκλισης σ , για να απαλυνθούν οι όποιες ανωμαλίες στο σήμα μετά την εφαρμογή της μάσκας, κάτι που έχει αποτέλεσμα, αφού χωρίς αυτό, το SSIM θα ήταν μόνο 0,6719. Με την παράμετρο (σ) sigma_blur=1, καθορίζω πόσο ισχυρό είναι το θόλωμα που εφαρμόζεται στην αποθορυβοποιημένη εικόνα. Η εικόνα παραμένει αρκετά καθαρή χωρίς να έχει αφαιρεθεί σημαντική πληροφορία.

Η τελική τιμή του SSIM είναι 0,7212, το οποίο θεωρώ είναι πολύ ικανοποιητικό, καθώς οι γραμμές έχουν μειωθεί με ελάχιστη απώλεια πληροφορίας και άλλα φίλτρα που δοκίμασα ήταν ασυγκρίτως χαμηλότερης ποιότητας και τιμής SSIM.

Για την απαλοιφή του περιοδικού θορύβου, φάνηκαν χρήσιμες οι εξής ιστοσελίδες:

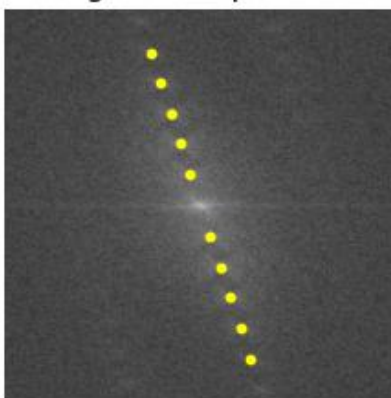
https://docs.opencv.org/4.x/d2/d0b/tutorial_periodic_noise_removing_filter.html

https://github.com/imdeep2905/Notch-Filter-for-Image-Processing/blob/master/filters/notch_filters.py

lenna-n3 (Periodic noise)



Magnitude Spectrum



Filtered (SSIM: 0.7212)



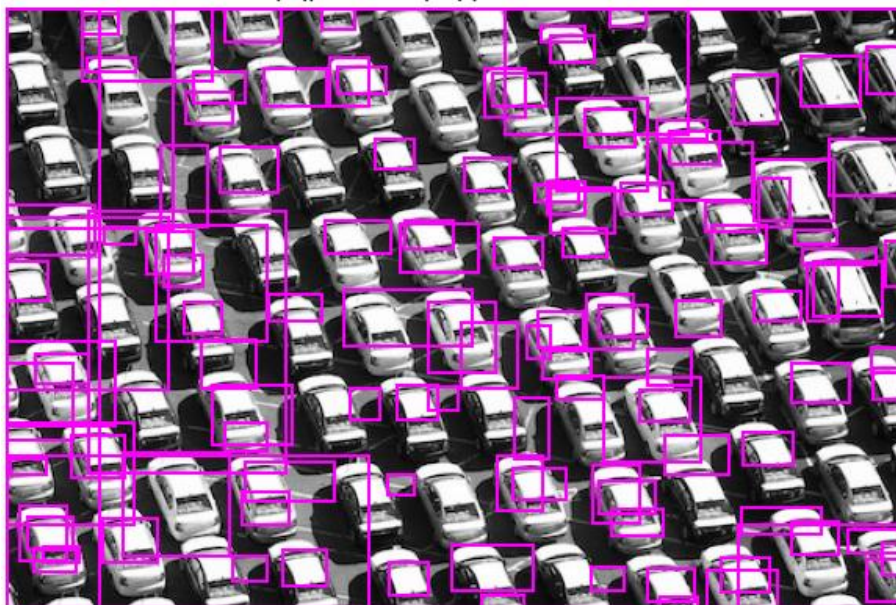
Άσκηση #5, Τμηματοποίηση:

Για την τμηματοποίηση της εικόνας 'parking-lot', χρησιμοποίησα τον αλγόριθμο Watershed που είδαμε στο lab03_3, σε συνδυασμό με φίλτρο Sobel για τον υπολογισμό του gradient. Με χρήση του Watershed, η εικόνα χωρίζεται σε διάφορες περιοχές, οι οποίες αντιστοιχούν σε πιθανά αντικείμενα, με βάση τις αλλαγές έντασης (edges). Χρησιμοποιώ τον αλγόριθμο Sobel για την εκτίμηση των ορίων των αντικειμένων στην εικόνα. Το φίλτρο Sobel υπολογίζει το gradient της εικόνας, εντοπίζοντας τα σημεία όπου η ένταση αλλάζει απότομα (όρια αντικειμένων). Με χαμηλή ένταση, υποδεικνύονται τα όρια των αντικειμένων (markers), όπου πάνω στα οποία συν του gradient εφαρμόζεται ο Watershed. Από την τμηματοποίηση αυτή, ορίζονται τα Bounding Boxes, δηλαδή τα περιθώρια γύρω από κάθε ανιχνευμένο αντικείμενο.

Page | 8

Το αποτέλεσμα της τμηματοποίησης είναι ικανοποιητικό, εφόσον ο αλγόριθμος Watershed διαχώρισε αποτελεσματικά τα αντικείμενα της εικόνας. Τα bounding boxes είναι τοποθετημένα γύρω από τα αυτοκίνητα και συγκεκριμένων έντονων χαρακτηριστικών τους, επισημαίνοντας την περιοχή που καταλαμβάνει κάθε αντικείμενο. Ο αλγόριθμος Watershed λειτούργησε καλά στην απομόνωση των περιοχών με απότομες αλλαγές στην ένταση, τα οποία αντιστοιχούν συνήθως στα όρια των αντικειμένων. Περιοχές μικρού μεγέθους, όπως πολύ μικρά αντικείμενα αγνοούνται, καθώς έχει τεθεί όριο μεγέθους για την εμφάνιση των bounding boxes (ελάχιστο εμβαδόν 100 pixel). Βέβαια, κάποια αντικείμενα ομαδοποιούνται και ένας μικρός αριθμός αυτοκινήτων μπορεί να μην έχει bounding box γύρω του, ο αλγόριθμος Watershed λοιπόν δεν είναι τέλειος, παρόλα αυτά δουλεύει ικανοποιητικά καλά.

Τμηματοποίηση με Watershed

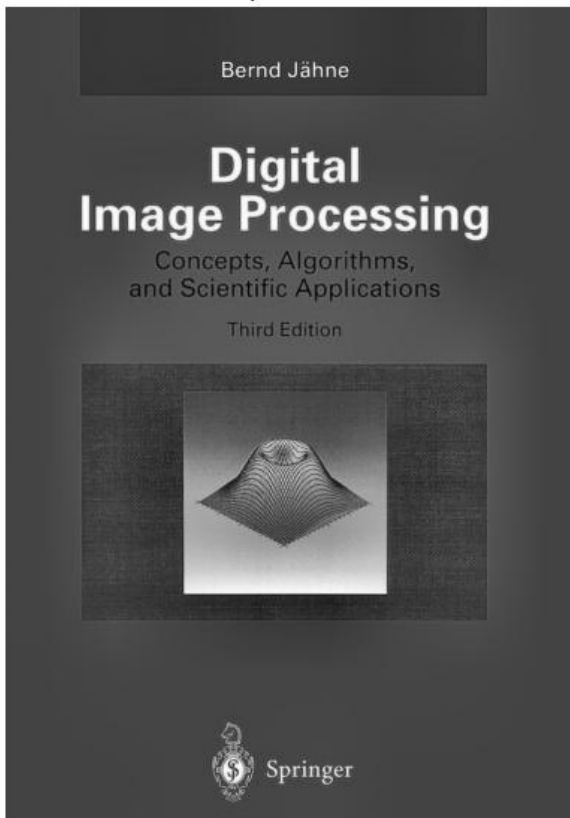


Άσκηση #6, Τμηματοποίηση 2:

Αφού μετέτρεψα την εικόνα 'book-cover' σε αποχρώσεις του γκρι, έκανα ισοστάθμιση του ιστογράμματος (`exposure.equalize_adapthist`) για καλύτερη αντίθεση και εφάρμοσα τοπικό κατώφλι για προσαρμογή κάθε περιοχής. Μετά ανέστρεψα την δυαδική εικόνα, φέρνοντας έτσι τους χαρακτήρες στο προσκήνιο και όλα τα υπόλοιπα στο παρασκήνιο. Στο προσκήνιο περιλαμβάνονται τα σημεία της εικόνας που αντιστοιχούν στους εντοπισμένους χαρακτήρες, ενώ το παρασκήνιο περιέχει τα υπόλοιπα στοιχεία. Επιπλέον, με την `'remove_small_objects'`, αφαίρεσα μικρά ανιχνευμένα αντικείμενα-θόρυβο και βελτίωσα τη μάσκα. Ως αποτέλεσμα, τα γράμματα έχουν εντοπιστεί με ικανοποιητική ακρίβεια και τα σχήματα κρατήσει τα βασικά τους χαρακτηριστικά.

Page | 9

Γκρι εικόνα



Δυαδική μάσκα χαρακτήρων

