

ΨΗΦΙΑΚΗ
ΕΠΕΞΕΡΓΑΣΙΑ
ΕΙΚΟΝΩΝ

Γιώργος Τζιρίτας, Καθηγητής

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
Τμήμα Επιστήμης Υπολογιστών

Φθινόπωρο 2016

Περιεχόμενα

1	Εισαγωγή	1
2	Σχηματισμός και αντίληψη εικόνων	3
2.1	Σχηματισμός εικόνων	3
2.2	Οπτική αντίληψη	4
3	2-Δ συνεχή σήματα	8
4	2-Δ δειγματοληψία	11
5	Μετασχηματισμός Fourier 2-Δ ακολουθιών	16
6	Κβαντισμός	22
6.1	Κριτήριο βασισμένο στην οπτική αντίληψη μιας αντίθεσης	22
6.2	Στατιστικό κριτήριο	23
7	2-Δ διακριτά συστήματα	26
7.1	Γραμμικά αμετάβλητα σε μετατόπιση συστήματα	26
7.2	2-Δ μετασχηματισμός Z	31
7.3	Ευστάθεια	33
8	2-Δ ορθομοναδιαίοι μετασχηματισμοί	37
8.1	Διακριτός μετασχηματισμός Fourier	39
8.2	Διακριτός μετασχηματισμός συνημιτόνου	39
8.3	Διακριτός μετασχηματισμός ημιτόνου	40
8.4	Διακριτός μετασχηματισμός Walsh-Hadamard	41
9	Βελτίωση εικόνων	43
9.1	Τονισμός της φωτεινής αντίθεσης	43
9.2	Μείωση θορύβου	45
9.3	Τονισμός ακμών	49
9.4	Μεγέθυνση και παρεμβολή	49
10	Τμηματοποίηση εικόνων	54
10.1	Ανίχνευση ακμών	54
10.1.1	Μέθοδος της κλίσης	55
10.1.2	Μέθοδος δευτέρων παραγώγων	59
10.2	Ανάλυση ιστογράμματος	60
10.3	Ανάλυση ενός μοντέλου Markov ή Gibbs	62

10.4 Τμηματοποίηση σε τοπολογικά συνεκτικές περιοχές	66
11 Μαθηματική μορφολογία	69
11.1 Μορφολογική επεξεργασία δυαδικών εικόνων	69
11.2 Μορφολογική επεξεργασία εικόνων φωτεινότητας	72
12 Αποκατάσταση εικόνων	74
12.1 Αντίστροφο φίλτρο	76
12.2 Ψευδο-αντίστροφο φίλτρο	77
12.3 Γενικευμένη αντιστροφή	77
12.4 Εξομάλυνση	78
12.5 Παραρτήματα	78
12.5.1 Κατάβαση στην κατεύθυνση της κλίσης	78
12.5.2 Αλγόριθμος των συζυγών κλίσεων	79
13 Επεξεργασία έγχρωμων εικόνων	80
13.1 Συστήματα χρωμάτων	82
13.1.1 HSI	82
13.1.2 C.I.E. Lab	82
13.1.3 YCbCr	83
13.2 Κβαντισμός χρώματος	83
13.3 Μείωση θορύβου	84
14 Ανάλυση σε ζώνες συχνότητων	85
15 Συμπίεση εικόνων	89
15.1 Διανυσματικός κβαντισμός	90
15.2 Συμπίεση με πρόβλεψη	91
15.3 Συμπίεση με χρήση 2-Δ μετασχηματισμού	93
15.3.1 Σταθερός ρυθμός πληροφορίας	93
15.3.2 Μεταβλητός και προσαρμοζόμενος ρυθμός πληροφορίας	94
15.4 Παράρτημα: Κωδικοποίηση Huffman	96

Κεφάλαιο 1

Εισαγωγή

Με την έλευση της ψηφιακής φωτογραφίας η ψηφιακή εικόνα είναι πλέον κτήμα του ευρέως κοινού. Η **Ψηφιακή Επεξεργασία Εικόνων**, με μια ιστορία περίπου πενήντα ετών, είναι ήδη μια ώριμη επιστημονική περιοχή. Ως αντικείμενο έχει εικόνες οποιασδήποτε προέλευσης, πρωτογενώς προσιτές ή μη στην ανθρώπινη όραση.

Αρχικά, περί τα μέσα της δεκαετίας του 1960, η ζήτηση για επεξεργασία εικόνων προέρχονταν από την εξερεύνηση του διαστήματος, και στη συνέχεια από τις αρχές της δεκαετίας του 1970 από την ιατρική απεικόνιση. Τα ζητήματα που ετίθεντο αφορούσαν κύρια στη βελτίωση και αποκατάσταση των εικόνων, διορθώνοντας αστοχίες και αδυναμίες των οπτικών συστημάτων κατά τη διαδικασία λήψης. Έκτοτε οι εφαρμογές βαίνουν αυξανόμενες και πολλαπλές, ενώ τα προβλήματα γίνονται πιο σύνθετα. Παράλληλα όμως αυξάνονται και οι δυνατότητες, αφού τόσο η υπολογιστική ισχύς μεγαθύνεται αδιάκοπα, όσο και οι τηλεπικοινωνίες επιτρέπουν τη μετάδοση με ταχύτατους ρυθμούς.

Στις ιατρικές εφαρμογές οι εικόνες προέρχονται από διάφορες ραδιολογικές πηγές ή μέσω υπερήχων, ή ακόμα προκύπτουν με χρήση μεθόδων υπολογιστικής τομογραφίας. Οι εφαρμογές αφορούν στον εντοπισμό πάνω στις εικόνες οργάνων ή ιστών και σε εξαγωγή πληροφορίας με μετρήσεις που αποσκοπούν στη βοήθεια στη διάγνωση.

Οι εφαρμογές παρακολούθησης του γήινου περιβάλλοντος μέσω δορυφόρων περιλαμβάνουν την χαρτογράφηση, την παρακολούθηση της εξέλιξης της χλωρίδας ή των συνεπειών της ανθρώπινης δραστηριότητας στο περιβάλλον, ή ακόμα τη μετεωρολογία. Οι εικόνες του διαστήματος συμβάλουν πάντοτε σημαντικά στην εξερεύνησή του.

Στις βιομηχανικές εφαρμογές περιλαμβάνονται ο αυτόματος έλεγχος ποιότητας προϊόντων και οι αντιληπτικές δεξιότητες ρομποτικών συστημάτων. Ο εντοπισμός και η αναγνώριση αντικειμένων επιτρέπει τη συναρμολόγησή τους από ένα ρομποτικό βραχίονα, ενώ ένα αυτόνομο κινούμενο ρομπότ έχει ανάγκη τη χαρτογράφηση του χώρου για να κινηθεί απρόσκοπτα.

Στις πολυμεσικές εφαρμογές έχουν σημασία η εξαγωγή αντικειμένων και χαρακτηριστικών τους για το χειρισμό των εικόνων σε γραφικά περιβάλλοντα, αλλά και για την περιγραφή του περιεχομένου των εικόνων. Σε διαδικτυακές και τηλεπικοινωνιακές εφαρμογές πρωταρχική σημασία έχει η συμπίεση των εικόνων για αποτελεσματική μετάδοση και αποθήκευση.

Η επεξεργασία των εικόνων, όπως προκύπτει από τη συνοπτική περιγραφή ανωτέρω, στοχεύει στη βελτίωση ή στην αποκατάσταση των εικόνων, στον προσδιορισμό τμημάτων που αντιστοιχούν σε αντικείμενα, σε μετρήσεις φωτομετρικών ή γεωμετρικών χαρακτηριστικών και στη συμπίεση. Συγγενείς επιστημονικές περιοχές είναι η γραφική και η υπολογιστική όραση. Η γραφική στοχεύει στη σύνθεση εικόνων φωτορεαλιστικών ή μη. Στην περίπτωση των φωτορεαλιστικών απεικονίσεων, η επεξεργασία των εικόνων θα μπορούσε να παράσχει πληροφορία για πιο πειστικές γραφικές αναπαραστάσεις. Η υπολογιστική όραση χρησιμοποιεί τις μεθοδολογίες

επεξεργασίας των εικόνων για την κατανόηση του περιεχομένου και την εξαγωγή πληροφορίας σε υψηλότερο επίπεδο με σκοπό την αντίληψη των αντικειμένων, της θέσης τους και της κίνησής τους.

Ακολουθεί βιβλιογραφία σχετική με το αντικείμενο του μαθήματος.

Βιβλιογραφία

- A. Bovik, ed., *Handbook of image and video processing*, 2nd ed., Academic Press, 2005.
- R. Bracewell, *Two-dimensional imaging*, Prentice Hall, 1995.
- K. Castleman, *Digital image processing*, 2nd ed., Prentice Hall, 1996.
- R. B. Fisher, et al., *Dictionary of computer vision and image processing*, J. Wiley and sons, 2005.
- R. Gonzalez and R. Woods, *Digital image processing*, 3rd ed., Addison-Wesley, 2008.
- R. Gonzalez, R. Woods and S. Eddins, *Digital image processing using MATLAB*, Prentice Hall, 2nd ed., 2009.
- S. G. Hoggar, *Mathematics of digital images*, Cambridge University Press, 2006.
- A. K. Jain, *Fundamentals of digital image processing*, Prentice Hall, 1989.
- J. Lim, *Two-dimensional signal and image processing*, Prentice Hall, 1990.
- D. Lyon, *Image processing in Java*, Prentice Hall, 1999.
- M. Petrou and P. Bosdogianni, *Image processing: The fundamentals*, J. Wiley and sons, 1999.
- M. Petrou and P. Garcia Sevilla, *Image processing, dealing with textures*, J. Wiley and sons, 2006.
- W. K. Pratt, *Digital image processing*, 4th ed., J. Wiley and sons, 2007.
- A. Rosenfeld and A. C. Kak, *Digital picture processing*, 2nd ed., Academic Press, 1982.
- J. Russ, *The image processing handbook*, CRC Press, 1995.
- J. Woods, *Multidimensional signal, image, and video processing and coding*, Academic Press, 2006.

Περιοδικά Τα ακόλουθα περιοδικά δημοσιεύουν άρθρα σχετικά με την επεξεργασία εικόνων.

- *IEEE Transactions on Image Processing*
- *IEEE Transactions on Pattern Analysis and Machine Intelligence*
- *IEEE Transactions on Medical Imaging*
- *Computer Vision and Image Understanding*
- *Signal Processing: Image Communication*
- *Journal of Visual Communication and Image Representation*
- *Image and Vision Computing*

Κεφάλαιο 2

Σχηματισμός και αντίληψη εικόνων

2.1 Σχηματισμός εικόνων

Πολλά φαινόμενα δεν είναι άμεσα παρατηρήσιμα ή είναι πρακτικά αδύνατο να μετρηθούν. Μπορούν όμως να παρατηρηθούν μέσα από μια **εικόνα**, που είναι 'η αναπαράσταση ενός όντος ή ενός αντικειμένου'. Ο σχηματισμός μιας εικόνας προϋποθέτει μια πηγή ακτινοβολίας, ένα δέκτη και ένα αντικείμενο.

Η κύρια πηγή ακτινοβολίας είναι η ηλεκτρομαγνητική. Το όλο φάσμα ακτινοβολίας εκτείνεται από τις ακτίνες Γάμμα μέχρι τα βραχέα ραδιοκύματα, με μήκη κύματος που καλύπτουν το φάσμα από 10^{-15} του μέτρου έως μερικά μέτρα, περνώντας από τις ακτίνες X, τις υπεριώδεις ακτίνες, την ορατή μέσω του ανθρώπινου οφθαλμού ακτινοβολία, τις υπέρυθρες ακτίνες και τα μικροκύματα. Ένα αορτικό αγγειογράφημα με ακτίνες X (μήκος κύματος περί τα 10^{-9} του μέτρου) δίδεται στο Σχήμα 2.1. Ο σχηματισμός εικόνων είναι δυνατός και με ακουστικά κύματα ή υπερήχους (Σχήμα 2.1b).

Οι εικόνες σχηματίζονται είτε με ανάκλαση της ακτινοβολίας σε επιφάνεια εξωτερική ή εσωτερική του αντικειμένου, είτε μετά από απορρόφηση μέρους της ακτινοβολίας από το σώμα που διαπερνάται και συλλογή της εναπομένουσας από το δέκτη. Είναι ακόμα δυνατό το απεικονιζόμενο αντικείμενο να αποτελεί πηγή ακτινοβολίας, φυσική ή τεχνητή.

Ο δέκτης συνίσταται κατά κανόνα από μία δισδιάστατη (2-Δ) επιφάνεια και μετρά την ένταση της ακτινοβολίας που προσπίπτει σ' αυτή την επιφάνεια. Η ακτινοβολία χαρακτηρίζεται από το μήκος κύματος και την ένταση, που μπορεί να μεταβάλλεται στο χρόνο. Η συνάρτηση της φωτεινής έντασης που προσπίπτει στο δέκτη, $\Phi(x, y, t, \lambda)$, εξαρτάται από το σημείο (x, y) , από τη χρονική στιγμή, t , και το μήκος κύματος, λ .

Στην πραγματικότητα ο δέκτης χαρακτηρίζεται από την ευαισθησία του, $E(\lambda)$, τη δεκτικότητα του δηλαδή σε ορισμένα μήκη κύματος. Υποθέτουμε ότι η ευαισθησία του δέκτη είναι ομοιογενής στο χώρο και στο χρόνο, μ' αποτέλεσμα την ακόλουθη έκφραση της φωτεινής έντασης

$$L(x, y, t) = \int \Phi(x, y, t, \lambda)E(\lambda)d\lambda. \quad (2.1)$$

Αν το παραπάνω ολοκλήρωμα καλύπτει ολόκληρο το φάσμα της ακτινοβολίας, ο δέκτης είναι μονοχρωματικός. Είναι επίσης δυνατό σε κάθε σημείο, και κάθε χρονική στιγμή, να είναι διαθέσιμοι περισσότεροι του ενός δέκτες, όπου ο καθένας καλύπτει ένα τμήμα του φάσματος. Στη συνέχεια θα περιορισθούμε σε στατικές εικόνες, δοσμένες δηλαδή για μια ορισμένη χρονική στιγμή.



(a)



(b)

Σχήμα 2.1: Αριστερά: Αορτικό αγγειογράφημα. Δεξιά: Υπερηχογράφημα εμβρύου.

2.2 Οπτική αντίληψη

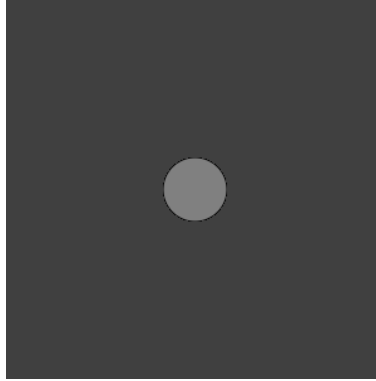
Στο ανθρώπινο μάτι η εικόνα σχηματίζεται στον αμφιβληστροειδή χιτώνα που αποτελείται από δύο ειδών υποδοχείς φωτός: τα ραβδία και τα κωνία. Τα ραβδία ενεργοποιούνται σε χαμηλό φωτισμό (σκοτοπική όραση) και είναι πολυπληθή, περί τα 100 εκατομμύρια. Τα ραβδία είναι κατανομημένα στο μέγιστο μέρος του χιτώνα και δε διακρίνουν χρωματικές διαφορές, δηλαδή πρόκειται για μονοχρωματικούς φωτοϋποδοχείς. Τα κωνία είναι λιγότερο ευαίσθητα στο φως, και κατά συνέπεια ενεργοποιούνται μόνο εφόσον η φωτεινή ένταση ξεπερνά κάποιο όριο (φωτοπική όραση). Το πλήθος των κωνίων είναι περί τα 6,5 εκατομμύρια και είναι ευαίσθητα στο χρώμα, διακρίνοντας τα τρία βασικά χρώματα: κόκκινο, πράσινο και μπλε.

Ακολουθώντας παρουσιάζουμε σε συντομία μερικές ιδιότητες της ανθρώπινης όρασης που μπορεί να είναι χρήσιμες για την επεξεργασία των εικόνων.

Κατ' αρχήν υπάρχουν όρια για να γίνει αντιληπτό ένα φωτεινό σήμα, για την ένταση, το μέγεθος στο χώρο (της τάξης του 1' στερεάς γωνίας), και τη διάρκεια (της τάξης των 40 msec). Υπάρχουν επίσης όρια για την αντίληψη μιας διαφοράς στη φωτεινή ένταση. Μία διαφορά στη φωτεινότητα, ΔL , γίνεται αντιληπτή σε σχέση με την περιβάλλουσα φωτεινότητα, L (Σχήμα 2.2), αν

$$\frac{\Delta L}{L} \geq C \approx 0,02. \quad (2.2)$$

Η σταθερά C ονομάζεται σταθερά του Weber, και ισχύει για μια ευρεία ζώνη τιμών της φωτεινότητας L . Αν η διαφορά ΔL παρατηρείται ανάμεσα σε δύο φωτεινούς στόχους που περιβάλλονται από φωτεινότητα L_0 , το όριο αντίληψης είναι μεγαλύτερο. Αυτή η εξάρτηση από το περιβάλλον δείχνει ότι το ανθρώπινο σύστημα όρασης επιτελεί μία επεξεργασία του σήματος της εικόνας στο χώρο. Το ίδιο συμπέρασμα εξάγεται από το φαινόμενο της ταυτόχρονης διαφοράς (Σχήμα 2.3).



Σχήμα 2.2: Πείραμα Weber

Η ίδια φωτεινή ένταση φαίνεται πιο φωτεινή αν περιβάλλεται από μια σκοτεινή περιοχή, απ' ό,τι φαίνεται αν περιβάλλεται από μια φωτεινή περιοχή.



Σχήμα 2.3: Φαινόμενο της ταυτόχρονης διαφοράς

Παρόμοιο συμπέρασμα εξάγεται επίσης από το φαινόμενο της ζώνης Mach (Σχήμα 2.4). Ενώ οι στήλες έχουν ομοιόμορφη φωτεινότητα, φαίνονται πιο φωτεινές από την πλευρά της πιο σκοτεινής γειτονικής στήλης, και πιο σκοτεινές από την πλευρά της πιο φωτεινής γειτονικής στήλης. Δηλαδή η φωτεινή αντίθεση τονίζεται χάρη στην επεξεργασία από το ανθρώπινο σύστημα όρασης. Άλλα πειράματα έδειξαν ότι το ανθρώπινο σύστημα όρασης συμπεριφέρεται στις χωρικές συχνότητες σαν ένα ζωνοπερατό σύστημα, με μέγιστη απόκριση στις μέσες χωρικές συχνότητες (της τάξης των 10 κύκλων ανά μοίρα), και πρακτικά αποκόπτει τις χωρικές συχνότητες πάνω από 30 κύκλους ανά μοίρα και τις χρονικές συχνότητες πάνω από 30 κύκλους ανά δευτερόλεπτο (Σχήμα 2.6). Στο Σχήμα 2.5 δίδονται διεγέρσεις με αρμονικά σήματα. Τα δύο σήματα στην κάθε γραμμή είναι στην ίδια οριζόντια χωρική συχνότητα και σταθερά στην κατακόρυφη κατεύθυνση. Διαφέρουν ως προς τη φωτεινή αντίθεση που είναι 10 φορές μεγαλύτερη στην αριστερή στήλη. Η φωτεινή αντίθεση ορίζεται ως

$$C = \frac{L_{\max} - L_{\min}}{L_{\max} + L_{\min}}.$$



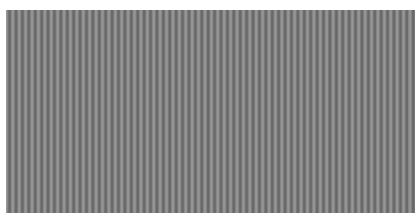
Σχήμα 2.4: Ζώνη Mach



(a)



(b)

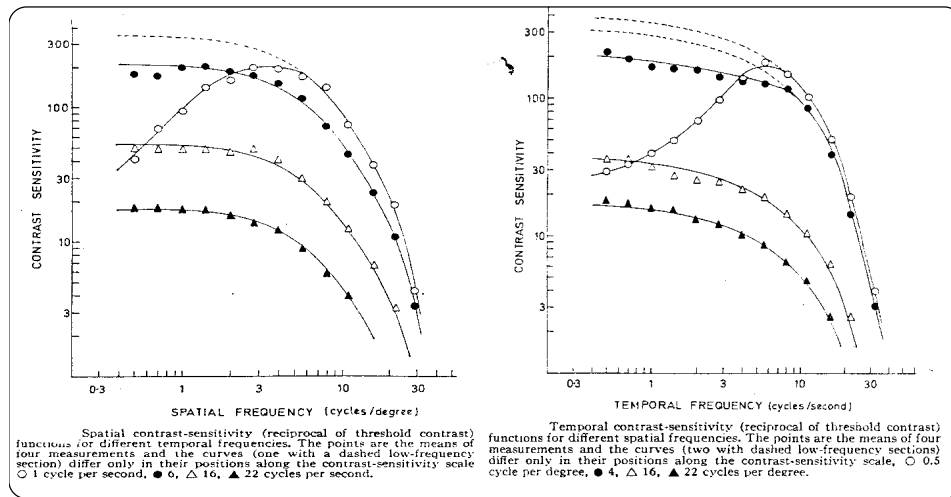


(c)



(d)

Σχήμα 2.5: Αρμονική διέγερση.



Σχήμα 2.6: Απόκριση του ανθρώπινου συστήματος όρασης στις χωρικές και στις χρονικές συχνότητες (Robson,1966).

Κεφάλαιο 3

2-Δ συνεχή σήματα

Μία συνεχής εικόνα μπορεί να παρασταθεί από ένα 2-Δ συνεχές σήμα, δηλαδή από μία συνάρτηση, $f(x, y)$. Ιδιαίτερης σημασίας για την επεξεργασία των σημάτων είναι ο μετασχηματισμός Fourier, που δίδει την παράσταση των σημάτων στο πεδίο των συχνοτήτων. Ο μετασχηματισμός Fourier ενός 2-Δ συνεχούς σήματος δίδεται ως εξής

$$F(u, v) = \mathcal{F}[f(x, y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-j2\pi(ux+vy)} dx dy. \quad (3.1)$$

Ικανή συνθήκη ύπαρξης του μετασχηματισμού είναι η ύπαρξη του ολοκληρώματος κατ' απόλυτο τιμή,

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |f(x, y)| dx dy < \infty.$$

Ο αντίστροφος μετασχηματισμός Fourier δίδεται ως εξής

$$f(x, y) = \mathcal{F}^{-1}[F(u, v)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v) e^{j2\pi(ux+vy)} du dv. \quad (3.2)$$

Αν (x, y) είναι οι συντεταγμένες στο 2-Δ χώρο, τότε (u, v) είναι συχνότητες στο χώρο. Η μονάδα μέτρησης της συχνότητας είναι αντίστοιχη της μονάδας μέτρησης των αποστάσεων στο επίπεδο των (x, y) . Πολύ συχνά όμως, εφόσον πρόκειται για σήματα εικόνων, οι συντεταγμένες (x, y) κανονικοποιούνται ως προς τη γωνία όρασης, με αποτέλεσμα οι συχνότητες (u, v) να μετρώνται σε κύκλους, δηλαδή περιοδικές χωρικές μεταβολές, ανά μοίρα της γωνίας όρασης.

Για τη μελέτη τόσο της δειγματοληψίας, όσο και των ιδιοτήτων των γραμμικών συστημάτων επεξεργασίας 2-Δ συνεχών σημάτων, είναι πολύ χρήσιμη η εισαγωγή της 2-Δ κατανομής Dirac

$$\delta(x, y) = 0, |x| + |y| \neq 0 \quad (3.3)$$

$$\lim_{\epsilon \rightarrow 0} \int_{-\epsilon}^{\epsilon} \int_{-\epsilon}^{\epsilon} \delta(x, y) dx dy = 1 \quad (3.4)$$

Ο μετασχηματισμός Fourier της κατανομής Dirac είναι ίσος με τη μονάδα για όλες τις συχνότητες

$$\mathcal{F}[\delta(x, y)] = 1 \quad (3.5)$$

Η κατανομή Dirac μπορεί να εκφράσει τη λήψη ενός δείγματος από ένα σήμα

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x', y') \delta(x - x', y - y') dx' dy' = f(x, y) \quad (3.6)$$

Δίδονται στη συνέχεια μερικές βασικές ιδιότητες του μετασχηματισμού Fourier (Εξίσωση 3.1)

1. Μετατόπιση

$$\mathcal{F}[f(x - x_0, y - y_0)] = e^{-j2\pi(x_0u + y_0v)} F(u, v) \quad (3.7)$$

2. Περιστροφή

$$\mathcal{F}[f(y, -x)] = F(v, -u) \quad (3.8)$$

3. Διαχωριστικότητα

$$\text{Αν } f(x, y) = f_1(x)f_2(y), \text{ τότε } F(u, v) = F_1(u)F_2(v) \quad (3.9)$$

όπου $F_1(u)$ (αντίστοιχα $F_2(v)$) είναι ο μετασχηματισμός Fourier της $f_1(x)$ (αντίστοιχα $f_2(y)$).

4. Αλλαγή κλίμακας

$$\mathcal{F}[f(ax, by)] = \frac{1}{|ab|} F\left(\frac{u}{a}, \frac{v}{b}\right) \quad (3.10)$$

5. Παραγωγή

$$\mathcal{F}[f_x(x, y)] = j2\pi u F(u, v) \quad \text{και} \quad \mathcal{F}[f_y(x, y)] = j2\pi v F(u, v) \quad (3.11)$$

6. Συνέλιξη

Η 2-Δ συνέλιξη της $h(., .)$ με την $f(., .)$ ορίζεται ως ακολούθως

$$g(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x', y') f(x - x', y - y') dx' dy' \quad (3.12)$$

Ο μετασχηματισμός Fourier $G(u, v)$ της $g(x, y)$ έχει ως εξής

$$G(u, v) = H(u, v)F(u, v) \quad (3.13)$$

όπου $H(u, v)$ είναι ο μετασχηματισμός Fourier της $h(x, y)$.

7. Εσωτερικό γινόμενο

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) g^*(x, y) dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v) G^*(u, v) du dv \quad (3.14)$$

Δίδονται ακόμη στη συνέχεια οι μετασχηματισμοί Fourier τριών χαρακτηριστικών συναρτήσεων.

1. Μιγαδική εκθετική συνάρτηση

$$\mathcal{F}[e^{j2\pi(u_0x + v_0y)}] = \delta(u - u_0, v - v_0) \quad (3.15)$$

2. Ορθογώνιο παραλληλεπίπεδο

Η συνάρτηση ορθογώνιου παραλληλεπίπεδου ορίζεται ως εξής

$$\Pi_2(x, y) = \begin{cases} 1 & \text{αν } |x| \leq 0,5 \text{ και } |y| \leq 0,5 \\ 0 & \text{αν } |x| > 0,5 \text{ είτε } |y| > 0,5 \end{cases} \quad (3.16)$$

Τότε έχουμε

$$\mathcal{F}[\Pi_2(x, y)] = \frac{\sin \pi u}{\pi u} \frac{\sin \pi v}{\pi v} \quad (3.17)$$

3. Συνάρτηση Gauss

$$\mathcal{F}[e^{-\pi(x^2+y^2)}] = e^{-\pi(u^2+v^2)} \quad (3.18)$$

Ασκήσεις

1. Βρείτε τους μετασχηματισμούς Fourier των σημάτων

- a) $\sin 2\pi x\eta_1 \cos 2\pi y\eta_2$
- b) $\cos 2\pi(x\eta_1 + y\eta_2)$

2. Εάν $F(u, v)$ είναι ο μετασχηματισμός Fourier του 2-Δ σήματος $f(x, y)$, να ευρεθούν οι μετασχηματισμοί Fourier των δύο πρώτων μερικών παραγώγων του $f(x, y)$. Ο Λαπλασιανός τελεστής ορίζεται ως το άθροισμα των δύο δεύτερων μερικών παραγώγων ως προς τις δύο μεταβλητές. Ποιός είναι ο μετασχηματισμός Fourier της Λαπλασιανής του $f(x, y)$;

3. Ορίζεται το ζεύγος μέσω των τιμών του 2-Δ σήματος $f(x, y)$ ως ακολούθως

$$(x_0, y_0) = \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f(x, y) dx dy, \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f(x, y) dx dy \right)$$

Το ισοδύναμο χωρικό εύρος 2-Δ σήματος ορίζεται ως εξής

$$E_x = \sqrt{\frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} ((x - x_0)^2 + (y - y_0)^2) f^2(x, y) dx dy}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f^2(x, y) dx dy}}$$

Εάν $F(u, v)$ είναι ο μετασχηματισμός Fourier του σήματος $f(x, y)$, ορίζεται το ισοδύναμο εύρος συχνοτήτων του σήματος

$$E_v = \sqrt{\frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} ((u - u_0)^2 + (v - v_0)^2) F^2(u, v) du dv}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F^2(u, v) du dv}}$$

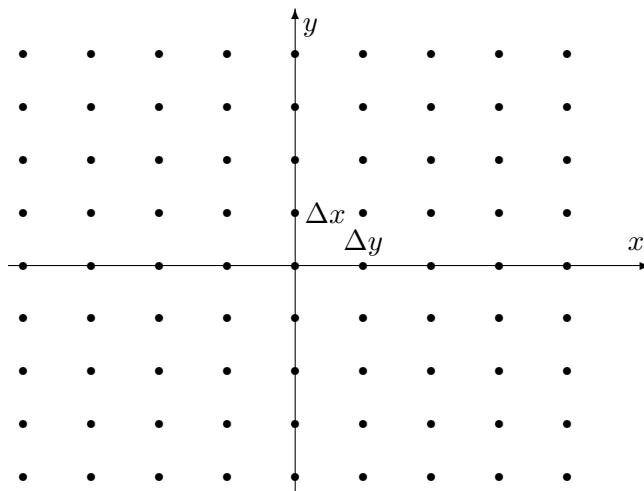
όπου τό ζεύγος (u_0, v_0) ορίζεται όπως παραπάνω για το $F(u, v)$. Να ευρεθεί το ισοδύναμο χωρικό εύρος και το ισοδύναμο εύρος συχνοτήτων του σήματος

$$f(x, y) = \exp(-\pi(x^2 + y^2))$$

Κεφάλαιο 4

2-Δ δειγματοληψία

Έστω ένα 2-Δ ορθογώνιο πλέγμα σημείων περιοδικά διαταγμένων στο επίπεδο των πραγματικών αριθμών (Σχήμα 4.1). Το πλέγμα αυτό δίδει τα σημεία όπου θα ληφθούν δείγματα από το 2-Δ συνεχές σήμα. Ας είναι Δx (αντίστοιχα Δy) η περίοδος της δειγματοληψίας ως προς x (αντίστοιχα y). Με βάση την ιδιότητα της κατανομής Dirac, που δίδεται στην Εξίσωση (3.6),



Σχήμα 4.1: 2-Δ ορθογώνιο πλέγμα σημείων

για εξαγωγή δειγμάτων από ένα 2-Δ σήμα, ορίζεται η ‘συνάρτηση’ δειγματοληψίας

$$s(x, y) = \Delta x \Delta y \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \delta(x - m\Delta x, y - n\Delta y). \quad (4.1)$$

Το αποτέλεσμα της δειγματοληψίας του σήματος $f(x, y)$ είναι το γινόμενο αυτού του σήματος με τη ‘συνάρτηση’ δειγματοληψίας

$$f_s(x, y) = s(x, y)f(x, y), \quad (4.2)$$

που δίδει

$$f_s(x, y) = \Delta x \Delta y \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} f(m\Delta x, n\Delta y) \delta(x - m\Delta x, y - n\Delta y). \quad (4.3)$$

Η δειγματοληψία συνεπάγεται περιοδικοποίηση στο πεδίο των συχνοτήτων. Πράγματι ο μετασχηματισμός Fourier του σήματος $f_s(x, y)$ προκύπτει από τη συνέλιξη του $F(u, v)$ με το $S(u, v)$, το μετασχηματισμό Fourier της $s(x, y)$. Επειδή η $s(x, y)$ είναι περιοδική συνάρτηση, μπορεί να παρασταθεί με τη βοήθεια μιας σειράς Fourier, που είναι η ακόλουθη

$$s(x, y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} e^{j2\pi(m\frac{x}{\Delta x} + n\frac{y}{\Delta y})}. \quad (4.4)$$

Πράγματι οι συντελεστές της σειράς Fourier είναι ίσοι με τη μονάδα

$$\begin{aligned} \frac{1}{\Delta x \Delta y} \int_{-\Delta x/2}^{\Delta x/2} \int_{-\Delta y/2}^{\Delta y/2} s(x, y) e^{-j2\pi(m\frac{x}{\Delta x} + n\frac{y}{\Delta y})} dx dy = \\ \int_{-\Delta x/2}^{\Delta x/2} \int_{-\Delta y/2}^{\Delta y/2} \delta(x, y) e^{-j2\pi(m\frac{x}{\Delta x} + n\frac{y}{\Delta y})} dx dy = 1. \end{aligned}$$

Στηριζόμενοι τώρα στην Εξίσωση (3.15), βρίσκουμε

$$S(u, v) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \delta(u - \frac{m}{\Delta x}, v - \frac{n}{\Delta y}). \quad (4.5)$$

Επομένως το αποτέλεσμα της συνέλιξης, με βάση και την Εξίσωση (3.6), είναι

$$F_s(u, v) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} F(u - \frac{m}{\Delta x}, v - \frac{n}{\Delta y}). \quad (4.6)$$

Η συνάρτηση αυτή είναι προφανώς περιοδική, με περίοδο τη συχνότητα δειγματοληψίας $\nu_x = \frac{1}{\Delta x}$ (αντίστοιχα $\nu_y = \frac{1}{\Delta y}$) για τη συχνότητα u (αντίστοιχα v).

Στο ζήτημα αν είναι δυνατό να ανακατασκευασθεί το συνεχές σήμα από τα περιοδικά του δείγματα όπως ορίστηκαν προηγούμενα απαντά το θεώρημα της δειγματοληψίας.

Θεώρημα της 2- Δ δειγματοληψίας

Αν το σήμα $f(x, y)$ είναι πεπερασμένης ζώνης συχνοτήτων, δηλαδή, αν

$$F(u, v) = 0, |u| > u_M, |v| > v_M, \quad (4.7)$$

τότε αρκεί οι συχνότητες δειγματοληψίας να είναι τουλάχιστον ίσες με τις συχνότητες Nyquist,

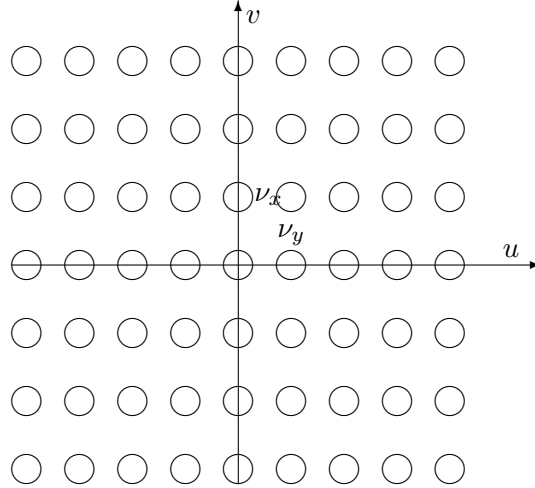
$$\nu_x \geq 2u_M, \nu_y \geq 2v_M, \quad (4.8)$$

για να μπορεί να αποκατασταθεί τέλεια το συνεχές σήμα από το διακριτό που δίδει η δειγματοληψία.

Στο Σχήμα 4.2 φαίνεται η περιοδικοποίηση λόγω δειγματοληψίας στο επίπεδο των συχνοτήτων, σε συνθήκες ισχύος του θεωρήματος της δειγματοληψίας. Ο κύκλος που περιβάλλει την αρχή των αξόνων αντιστοιχεί στο χωρίο όπου εκτείνεται με μη μηδενικές τιμές ο μετασχηματισμός Fourier του σήματος.

Η ανακατασκευή του συνεχούς σήματος επιτυγχάνεται χρησιμοποιώντας ένα φίλτρο που δίδει τέλεια την κύρια περίοδο του $F_s(u, v)$

$$H(u, v) = \begin{cases} 1, & |u| \leq \frac{\nu_x}{2} \quad \text{και} \quad |v| \leq \frac{\nu_y}{2} \\ 0, & |u| > \frac{\nu_x}{2} \quad \text{είτε} \quad |v| > \frac{\nu_y}{2} \end{cases} \quad (4.9)$$



Σχήμα 4.2: Η περιοδικοποίηση στο επίπεδο Fourier

Προφανώς έχουμε

$$F(u, v) = H(u, v)F_s(u, v). \quad (4.10)$$

Ο αντίστροφος μετασχηματισμός Fourier του $H(u, v)$ είναι (βλέπε Εξίσωση (3.16))

$$h(x, y) = \frac{1}{\Delta x \Delta y} \operatorname{sinc} \frac{\pi x}{\Delta x} \operatorname{sinc} \frac{\pi y}{\Delta y}, \quad (4.11)$$

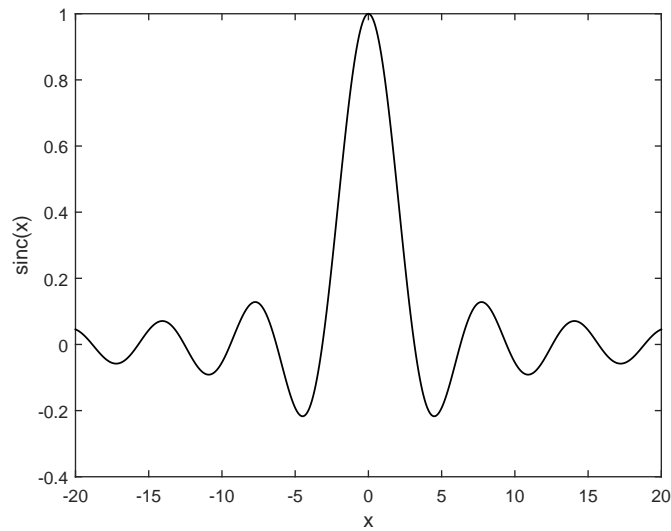
όπου $\operatorname{sinc} x = \frac{\sin x}{x}$ (Σχήμα 4.3). Οπότε η παρεμβολή για οποιαδήποτε τιμή των x και y δίδεται από την ακόλουθη συνέλιξη

$$f(x, y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} f(m\Delta x, n\Delta y) \operatorname{sinc} \frac{\pi(x - m\Delta x)}{\Delta x} \operatorname{sinc} \frac{\pi(y - n\Delta y)}{\Delta y}. \quad (4.12)$$

Αν οι συχνότητες δειγματοληψίας είναι κάτω από τις συχνότητες Nyquist, τότε προκύπτουν επικαλύψεις στο φάσμα των συχνοτήτων, με αποτέλεσμα να είναι αδύνατο να ανακτηθεί τέλεια το συνεχές σήμα. Οι συχνότητες που λόγω επικάλυψης αλλάζουν θέση ονομάζονται ψευδώνυμες, και εισάγουν παραμόρφωση στο συνεχές σήμα. Συνεπώς οι Εξισώσεις (4.10) και (4.12) δεν ισχύουν. Για τον περιορισμό αυτής της παραμόρφωσης συνιστάται η χρήση του φίλτρου $H(u, v)$ πριν τη δειγματοληψία.

Παραμόρφωση μπορεί επίσης να εισαχθεί, έστω κι αν ικανοποιούνται οι συνθήκες του θεωρήματος της δειγματοληψίας, αν το φίλτρο της παρεμβολής δεν είναι ιδανικό. Αν δηλαδή δεν χρησιμοποιηθεί το φίλτρο της Εξίσωσης (4.11), που δεν είναι πρακτικά υλοποιήσιμο, αλλά κάποιο άλλο που να το προσεγγίζει, έχοντας για παράδειγμα πεπερασμένη απόκριση στο χώρο. Τέτοια φίλτρα παρεμβολής είναι τα ακόλουθα:

- $h_0(x, y) = \frac{1}{\Delta x \Delta y} \Pi_2\left(\frac{x}{\Delta x}, \frac{y}{\Delta y}\right)$
- $h_1(x, y) = h_0(x, y) * h_0(x, y) = \begin{cases} \frac{1}{\Delta x \Delta y} \left(1 - \frac{|x|}{\Delta x}\right) \left(1 - \frac{|y|}{\Delta y}\right), & |x| < \Delta x \text{ και } |y| < \Delta y \\ 0 & \text{αλλού} \end{cases}$
- $h_2(x, y) = h_1(x, y) * h_0(x, y)$



Σχήμα 4.3: Συνάρτηση sinc

Στο Σχήμα 4.4 δείχνεται η διεργασία της δειγματοληψίας και ο ρόλος του βαθυπερατού φίλτρου. Εδώ η αρχική εικόνα είναι επίσης ορισμένη πάνω σε ένα διακριτό πλέγμα, και υφίσταται υποδειγματοληψία με ρυθμό $1/4$, δηλαδή λαμβάνοντας ένα στα τέσσερα δείγματα τόσο στην οριζόντια, όσο και στην κατακόρυφη κατεύθυνση. Δείχνεται το αποτέλεσμα όταν προηγείται της δειγματοληψίας η χρήση βαθυπερατού φίλτρου, που φυσικά δεν είναι το ιδανικό. Αν δεν χρησιμοποιηθεί το φίλτρο, παρατηρείται η εμφάνιση ψευδώνυμων συχνοτήτων, που βέβαια δεν διορθώνονται, αν το βαθυπερατό φίλτρο ακολουθήσει τη δειγματοληψία.

Άσκηση

Το 2-Δ σήμα $f(x, y) = \cos 4\pi x \cos 6\pi y$ υφίσταται δειγματοληψία με $\Delta x = \Delta y = 0,5$, και με $\Delta x = \Delta y = 0,2$. Το φίλτρο ανακατασκευής είναι ένα ιδεατό βαθυπερατό φίλτρο με εύρη ζώνης $(\frac{1}{2\Delta x}, \frac{1}{2\Delta y})$. Ποία είναι η ανακατασκευαζόμενη εικόνα σε κάθε περίπτωση;



(a)



(b)



(c)



(d)

Σχήμα 4.4: Ανω αριστερά : εικόνα “Barbara”. Ανω δεξιά : χρήση βαθυπερατού φίλτρου και μετά δειγματοληψία. Κάτω αριστερά : δειγματοληψία χωρίς να προηγηθεί φίλτρο. Κάτω δεξιά : δειγματοληψία και μετά χρήση βαθυπερατού φίλτρου.

Κεφάλαιο 5

Μετασχηματισμός Fourier 2-Δ ακολουθιών

Ας είναι $x(m, n)$ μία 2-Δ ακολουθία ορισμένη στο επίπεδο των ακεραίων αριθμών \mathbb{Z}^2 . Ο μετασχηματισμός Fourier της ακολουθίας $x(m, n)$ είναι μια περιοδική συνάρτηση των συχνοτήτων (u, v) με περίοδο 1 και για τις δύο συχνοτήτες,

$$X(u, v) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} x(m, n) e^{-j2\pi(um+vn)}. \quad (5.1)$$

Η ύπαρξη του μετασχηματισμού προϋποθέτει ότι οι τιμές της 2-Δ ακολουθίας αθροίζονται κατ' απόλυτη τιμή

$$\sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} |x(m, n)| < \infty. \quad (5.2)$$

Ο αντίστροφος μετασχηματισμός Fourier δίδει την αρχική ακολουθία

$$x(m, n) = \int_{-1/2}^{1/2} \int_{-1/2}^{1/2} X(u, v) e^{j2\pi(um+vn)} du dv. \quad (5.3)$$

Εάν η 2-Δ ακολουθία είναι πραγματική, ο μετασχηματισμός Fourier έχει την ιδιότητα της συζυγούς συμμετρίας,

$$X(-u, -v) = X^*(u, v). \quad (5.4)$$

Στο Σχήμα 5.1 δίδεται ο μετασχηματισμός Fourier ενός περιοδικού σήματος, χωριστά το μέτρο και η φάση. Οι χωρικές συχνοτήτες, λόγω της περιοδικότητας είναι διακριτές, ενώ επιβεβαιώνεται η συζυγής συμμετρία, αφού το μέτρο είναι άρτια συνάρτηση των συχνοτήτων και η φάση περιττή.

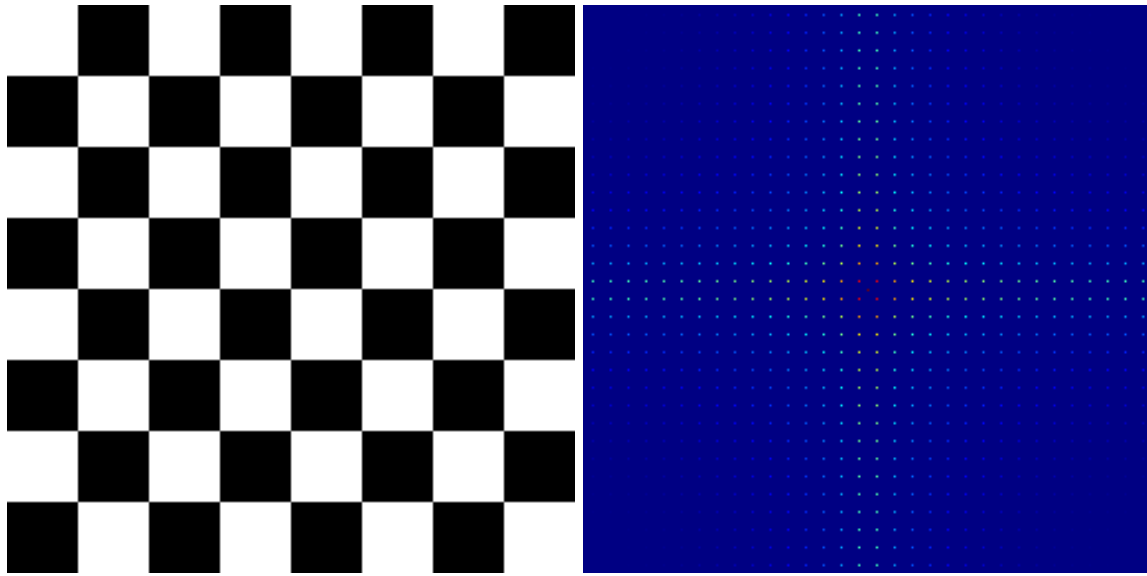
Δίδονται στη συνέχεια μερικές από τις ιδιότητες του μετασχηματισμού Fourier, παρόμοιες με τις ιδιότητες του μετασχηματισμού Fourier 2-Δ συνεχών σημάτων.

1. Γραμμικότητα

Κάθε γραμμικός συνδυασμός σημάτων μετασχηματίζεται στον αντίστοιχο γραμμικό συνδυασμό των μετασχηματισμών Fourier των σημάτων.

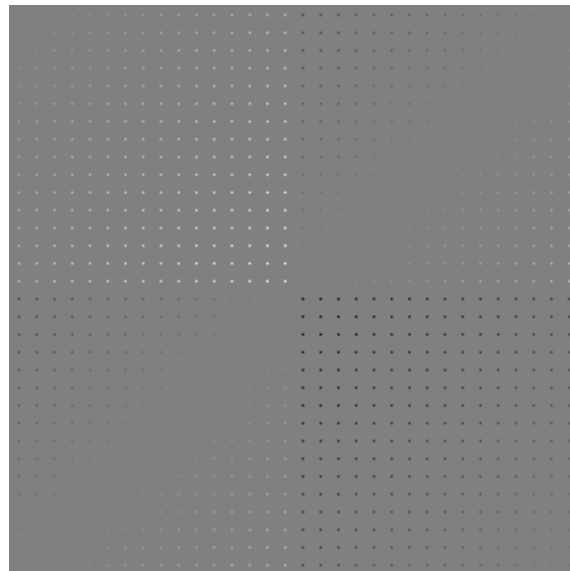
2. Μετατόπιση

$$\sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} x(m - m_0, n - n_0) e^{-j2\pi(um+vn)} = e^{-j2\pi(um_0+vn_0)} X(u, v) \quad (5.5)$$



(a)

(b)



(c)

Σχήμα 5.1: Μετασχηματισμός Fourier ενός 2-Δ περιοδικού σήματος. Το μπλε χρώμα στο μέτρο αντιστοιχεί σε μηδενική τιμή. Η φάση κυμαίνεται στο διάστημα $(-\pi, \pi]$ και απεικονίζεται σε αποχρώσεις του γκρι.

3. Διαχωρισσιμότητα

$$\text{Αν } x(m, n) = x_1(m)x_2(n), \text{ τότε } X(u, v) = X_1(u)X_2(v), \quad (5.6)$$

όπου $X_1(u)$ (αντίστοιχα $X_2(v)$) είναι ο μετασχηματισμός Fourier της $x_1(m)$ (αντίστοιχα $x_2(n)$).

4. Σταθερή τιμή

$$X(0, 0) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} x(m, n). \quad (5.7)$$

5. Διαμόρφωση

Ο πολλαπλασιασμός με μία εκθετική αρμονική συνάρτηση,

$$y(m, n) = x(m, n)e^{j2\pi(u_0m+v_0n)}, \quad (5.8)$$

συνεπάγεται μετατόπιση στο πεδίο των χωρικών συχνοτήτων,

$$Y(u, v) = X(u - u_0, v - v_0). \quad (5.9)$$

6. Συνέλιξη

Η 2-Δ συνέλιξη της $h(., .)$ με την $x(., .)$ ορίζεται ως ακολούθως

$$y(m, n) = \sum_{m'=-\infty}^{\infty} \sum_{n'=-\infty}^{\infty} h(m', n')x(m - m', n - n'). \quad (5.10)$$

Ο μετασχηματισμός Fourier $Y(u, v)$ της $y(m, n)$ έχει ως εξής

$$Y(u, v) = H(u, v)X(u, v). \quad (5.11)$$

όπου $H(u, v)$ είναι ο μετασχηματισμός Fourier της $h(m, n)$.

7. Πολλαπλασιασμός

Ο μετασχηματισμός Fourier του γινομένου δύο 2-Δ ακολουθιών, $x(m, n)$ και $y(m, n)$, δίδεται ως εξής

$$\sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} x(m, n)y(m, n)e^{-j2\pi(um+vn)} = \int_{-1/2}^{1/2} \int_{-1/2}^{1/2} X(u', v')Y(u - u', v - v')du'dv', \quad (5.12)$$

όπου $X(u, v)$ και $Y(u, v)$ είναι όπως παραπάνω.

8. Εσωτερικό γινόμενο

$$\sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} x(m, n)y^*(m, n) = \int_{-1/2}^{1/2} \int_{-1/2}^{1/2} X(u, v)Y^*(u, v)dudv. \quad (5.13)$$

Κατά συνέπεια η ενέργεια ενός σήματος διατηρείται μετά το μετασχηματισμό,

$$\sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} |x(m, n)|^2 = \int_{-1/2}^{1/2} \int_{-1/2}^{1/2} |X(u, v)|^2dudv. \quad (5.14)$$

Ακολουθούν παραδείγματα αναλυτικών εκφράσεων για το μετασχηματισμό Fourier.

Παράδειγμα 5.1. Ας είναι

$$x(m, n) = \begin{cases} 1 & m^2 + n^2 = 0 \\ 1/2 & m^2 + n^2 = 1 \\ 1/4 & m^2 + n^2 = 2 \\ 0 & m^2 + n^2 > 2 \end{cases} \quad (5.15)$$

2-Δ διακριτό σήμα πεπερασμένης έκτασης. Διαπιστώνουμε ότι το σήμα είναι διαχωρίσιμο με

$$x_1(m) = x_2(m) = \begin{cases} 1 & |m| = 0 \\ 1/2 & |m| = 1 \\ 0 & |m| > 1 \end{cases}$$

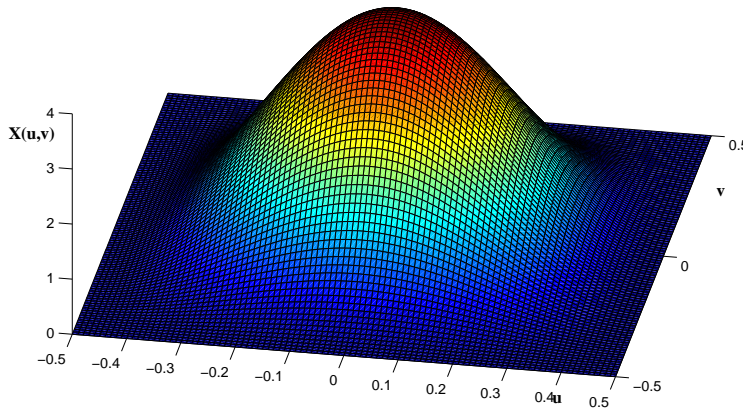
Άμεσα από τον ορισμό του μετασχηματισμού προκύπτει ότι

$$X_1(u) = e^{j2\pi u} + 1 + e^{-j2\pi u} = 1 + 2 \cos 2\pi u = 2 \cos^2 \pi u. \quad (5.16)$$

Άρα για το 2-Δ σήμα, λόγω διαχωρισιμότητας, θα ισχύει

$$X(u, v) = 4 \cos^2 \pi u \cos^2 \pi v. \quad (5.17)$$

Στο Σχήμα 5.2 δίδεται ο μετασχηματισμός Fourier.



Σχήμα 5.2: Μετασχηματισμός Fourier της πεπερασμένης έκτασης ακολουθίας του παραδείγματος 5.1.

Παράδειγμα 5.2. Ας είναι

$$x(m, n) = \begin{cases} 1 & |m| \leq N \text{ και } |n| \leq N \\ 0 & |m| > N \text{ είτε } |n| > N \end{cases} \quad (5.18)$$

2-Δ διακριτό σήμα πεπερασμένης έκτασης. Διαπιστώνουμε ότι το σήμα είναι διαχωρίσιμο με

$$x_1(m) = x_2(m) = \begin{cases} 1 & |m| \leq N \\ 0 & |m| > N \end{cases}$$

Άμεσα από τον ορισμό του μετασχηματισμού προκύπτει ότι

$$X_1(u) = \sum_{m=-N}^N e^{-j2\pi mu} = e^{-j2\pi Nu} \sum_{m=0}^{2N} e^{j2\pi mu} = e^{-j2\pi Nu} \frac{e^{j2\pi(2N+1)u} - 1}{e^{j2\pi u} - 1}.$$

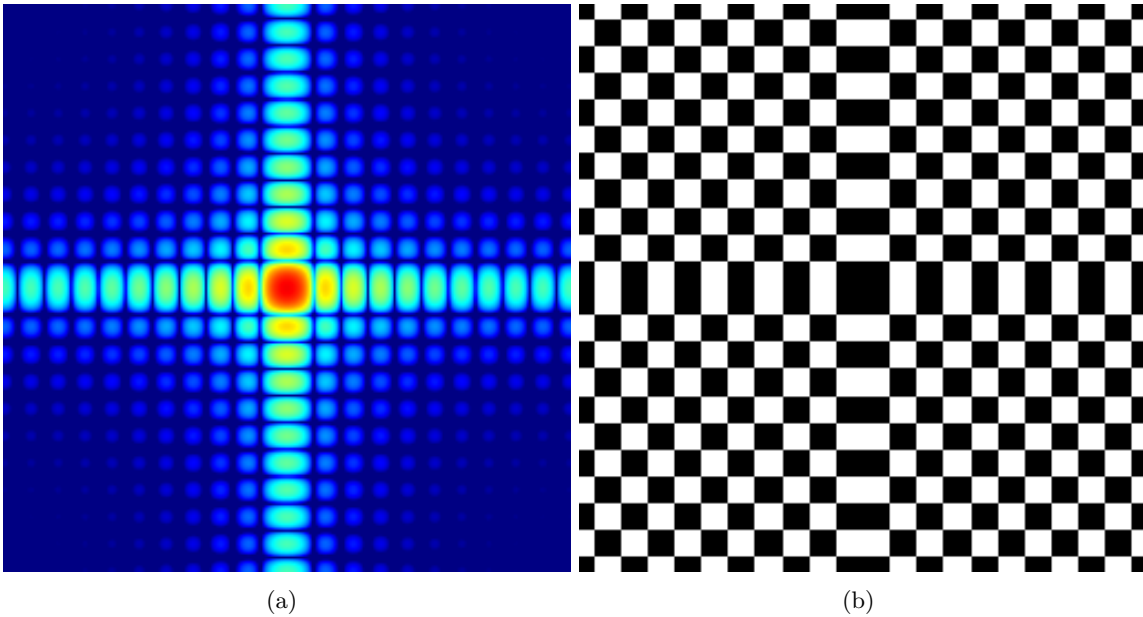
Οπότε

$$X_1(u) = \frac{e^{j\pi(2N+1)u} - e^{-j\pi(2N+1)u}}{e^{j\pi u} - e^{-j\pi u}} = \frac{\sin \pi(2N+1)u}{\sin \pi u}. \quad (5.19)$$

Άρα για το 2-Δ σήμα, λόγω διαχωρισιμότητας, θα ισχύει

$$X(u, v) = \frac{\sin \pi(2N+1)u \sin \pi(2N+1)v}{\sin \pi u \sin \pi v}. \quad (5.20)$$

Στο Σχήμα 5.3 δίδεται αυτός ο μετασχηματισμός Fourier, χωριστά το μέτρο και η φάση. Η φάση λαμβάνει δύο τιμές, 0 ή π .



Σχήμα 5.3: Μετασχηματισμός Fourier μιας οπτικής διέργησης τετράγωνου σχήματος. Το μέτρο απεικονίζεται με χρώματα σε κλίμακα που πηγαίνει από το μπλε στο κόκκινο για τις μεγαλύτερες τιμές.

Παράδειγμα 5.3. Ας είναι

$$x(m, n) = \alpha^{|m|} \beta^{|n|}, \quad |\alpha| < 1, \quad |\beta| < 1 \quad (5.21)$$

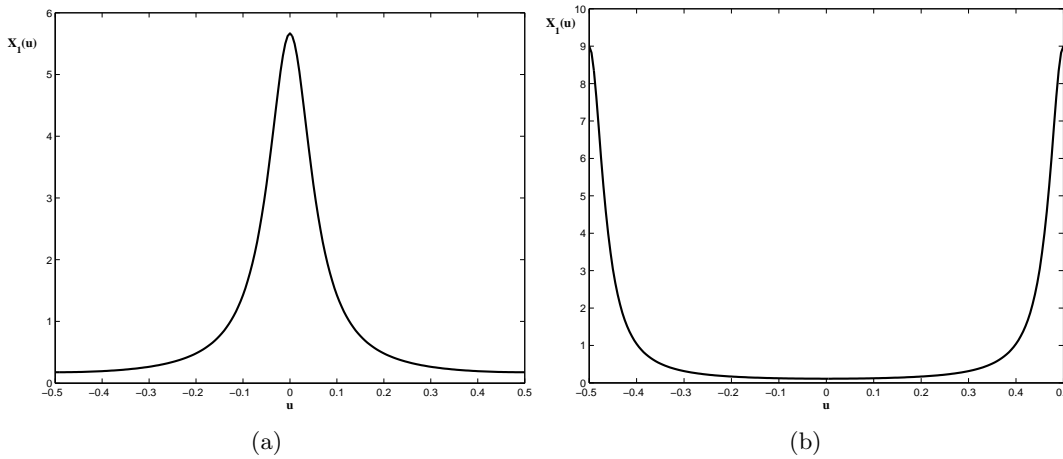
2-Δ διακριτό σήμα. Αφού το σήμα είναι διαχωρίσιμο, αρκεί να ευρεθούν οι δύο μονοδιάστατοι μετασχηματισμοί Fourier. Θα έχουμε

$$X_1(u) = \sum_{m=-\infty}^{\infty} \alpha^{|m|} e^{-j2\pi um} = \sum_{m=-\infty}^0 \alpha^{|m|} e^{-j2\pi um} + \sum_{m=0}^{\infty} \alpha^{|m|} e^{-j2\pi um} - 1$$

$$\begin{aligned}
&= \sum_{m=-\infty}^0 (\alpha e^{j2\pi u})^{-m} + \sum_{m=0}^{\infty} (\alpha e^{-j2\pi u})^m - 1 = \sum_{m=0}^{\infty} (\alpha e^{j2\pi u})^m + \sum_{m=0}^{\infty} (\alpha e^{-j2\pi u})^m - 1 \\
&= \frac{1}{1 - \alpha e^{j2\pi u}} + \frac{1}{1 - \alpha e^{-j2\pi u}} - 1 = \frac{1 - \alpha^2}{(1 - \alpha e^{j2\pi u})(1 - \alpha e^{-j2\pi u})} \\
&= \frac{1 - \alpha^2}{1 - 2\alpha \cos(2\pi u) + \alpha^2} \tag{5.22}
\end{aligned}$$

Οι τιμές του μετασχηματισμού δίδονται γραφικά στο Σχήμα 5.4 για δύο διαφορετικές τιμές του α , μία θετική και μία αρνητική. Τελικά ο μετασχηματισμός Fourier της $x(m, n)$ είναι

$$X(u, v) = \frac{(1 - \alpha^2)(1 - \beta^2)}{(1 - 2\alpha \cos(2\pi u) + \alpha^2)(1 - 2\beta \cos(2\pi v) + \beta^2)} \tag{5.23}$$



Σχήμα 5.4: Μετασχηματισμός Fourier εκθετικά φθίνουσας ακολουθίας για (a) $\alpha = 0.7$, (b) $\alpha = -0.8$.

Άσκηση

Δίδονται τα ακόλουθα σήματα

$$\bullet f_1(m, n) = \begin{cases} 1 & |m| + |n| = 0 \\ 1/2 & |m| + |n| = 1 \\ 1/4 & |m| + |n| = 2 \\ 0 & |m| + |n| > 2 \end{cases}$$

$$\bullet f_2(m, n) = \begin{cases} 1/8 & m = -1 \text{ και } |n| = 1 \\ 1/4 & m = -1 \text{ και } n = 0 \\ -1/4 & m = 1 \text{ και } n = 0 \\ -1/8 & m = 1 \text{ και } |n| = 1 \\ 0 & |m| \neq 1 \text{ είτε } |n| > 1 \end{cases}$$

Για καθένα από τα σήματα αυτά ζητείται εάν είναι διαχωρίσιμο, και να ευρεθεί ο μετασχηματισμός Fourier.

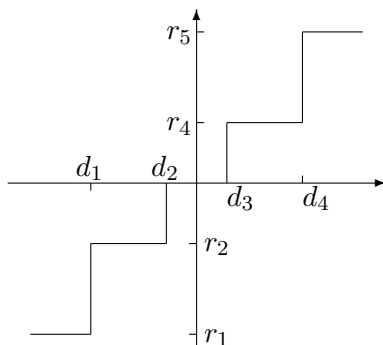
Κεφάλαιο 6

Κβαντισμός

Ένας κβαντιστής (Σχήμα 6.1) αντιστοιχεί σε μία συνεχή μεταβλητή X μία διακριτή τιμή από ένα πεπερασμένο σύνολο $\{r_1, r_2, \dots, r_N\}$ αντιπροσωπευτικών τιμών. Για τον προσδιορισμό του κβαντιστή απαιτείται αφενός το σύνολο των αντιπροσωπευτικών τιμών και αφετέρου το σύνολο των σημείων απόφασης $\{d_1, d_2, \dots, d_{N-1}\}$ που χωρίζουν τον άξονα των πραγματικών αριθμών στα ακόλουθα διαστήματα

$$\mathbb{R} = (-\infty, d_1] \cup (d_1, d_2] \cup \dots \cup (d_{N-1}, \infty). \quad (6.1)$$

Για το σχεδιασμό ενός κβαντιστή παρουσιάζονται στη συνέχεια δύο προσεγγίσεις του θέματος.



Σχήμα 6.1: Ένας συμμετρικός κβαντιστής

Η πρώτη στηρίζεται στα συμπεράσματα των πειραμάτων του Weber, ενώ η δεύτερη βασίζεται σε καθαρά στατιστικά κριτήρια.

6.1 Κριτήριο βασισμένο στην οπτική αντίληψη μιας αντίθεσης

Ο νόμος του Weber για την αντίληψη της φωτεινής αντίθεσης (Εξίσωση (2.2)» μπορεί να χρησιμοποιηθεί για τον καθορισμό ενός κβαντιστή για δοσμένο διάστημα τιμών. Ας είναι L_0 (αντίστοιχα L_T) η ελάχιστη (αντίστοιχα μέγιστη) τιμή της φωτεινής έντασης που ζητείται η προσέγγισή της από ένα πεπερασμένο σύνολο διακριτών αντιπροσωπευτικών τιμών. Αν υποθέσουμε ότι οι παραπάνω ακραίες τιμές ανήκουν στο διάστημα που ισχύει η σταθερά του Weber C , για να μην είναι ορατή η παραμόρφωση που εισάγει ο κβαντισμός, ο απαιτούμενος αριθμός αντιπροσωπευτικών τιμών είναι τουλάχιστον

$$N = \frac{\log \frac{L_T}{L_0}}{\log(1 + C)} + 1. \quad (6.2)$$

Μία συνηθισμένη τιμή για το λόγο L_T/L_0 είναι 100. Αν επομένως ληφθεί $C = 0,02$, προκύπτει ότι απαιτούνται 234 επίπεδα κβαντισμού, που σημαίνει ότι 8 bits είναι αρκετά κατά κανόνα για να παρασταθούν οι τιμές μιας εικόνας. Αυτή η μέθοδος κβαντισμού ισοδυναμεί με ομοιόμορφο κβαντισμό του λογάριθμου της φωτεινής έντασης.

6.2 Στατιστικό κριτήριο

Ας είναι X μία τυχαία μεταβλητή, της οποίας η συνάρτηση πυκνότητας πιθανότητας είναι $p(x)$. Ζητείται ο προσδιορισμός τόσο των επιπέδων κβαντισμού, όσο και των σημείων απόφασης, για δοσμένο αριθμό N διαστημάτων και επιπέδων κβαντισμού, με τρόπο ώστε το μέσο τετραγωνικό σφάλμα να ελαχιστοποιείται

$$D = E\{(X - Q(X))^2\} = \sum_{i=1}^N \int_{d_{i-1}}^{d_i} (x - r_i)^2 p(x) dx, \quad (6.3)$$

όπου $Q(\cdot)$ είναι ο τελεστής του κβαντισμού, και $d_0 = -\infty$, $d_N = \infty$.

Οι Lloyd και Max έδωσαν τις αναγκαίες συνθήκες για την ελαχιστοποίηση του D , που συνίστανται σε δύο συστήματα εξισώσεων. Το πρώτο σύστημα δίδει την καλύτερη αντιπροσώπευση, με δοσμένα τα διαστήματα

$$r_i = E\{X|X \in (d_{i-1}, d_i)\} = \frac{\int_{d_{i-1}}^{d_i} xp(x) dx}{\int_{d_{i-1}}^{d_i} p(x) dx}; i = 1, \dots, N. \quad (6.4)$$

Το δεύτερο σύστημα δίδει την καλύτερη τμηματοποίηση του άξονα των πραγματικών αριθμών με δοσμένα τα επίπεδα αντιπροσώπευσης

$$d_i = \frac{r_i + r_{i+1}}{2}; i = 1, \dots, N - 1. \quad (6.5)$$

Αναλυτικές εκφράσεις για τη λύση του παραπάνω συστήματος εξισώσεων λαμβάνονται μόνο για ειδικές περιπτώσεις κατανομών και αριθμού επιπέδων κβαντισμού.

Παράδειγμα 6.1. Ας υποθεθεί ότι η κατανομή είναι ομοιόμορφη στο διάστημα $[L_0, L_T]$. Τότε από την εξίσωση (6.4) προκύπτει ότι

$$r_i = \frac{d_{i-1} + d_i}{2}.$$

Οπότε από το σύνολο των εξισώσεων, που σ' αυτή την περίπτωση είναι όλες γραμμικές, έχουμε την ακόλουθη λύση για τα όρια των διαστημάτων

$$d_0 = L_0, d_i - d_{i-1} = \frac{L_T - L_0}{N} (i = 1, \dots, N).$$

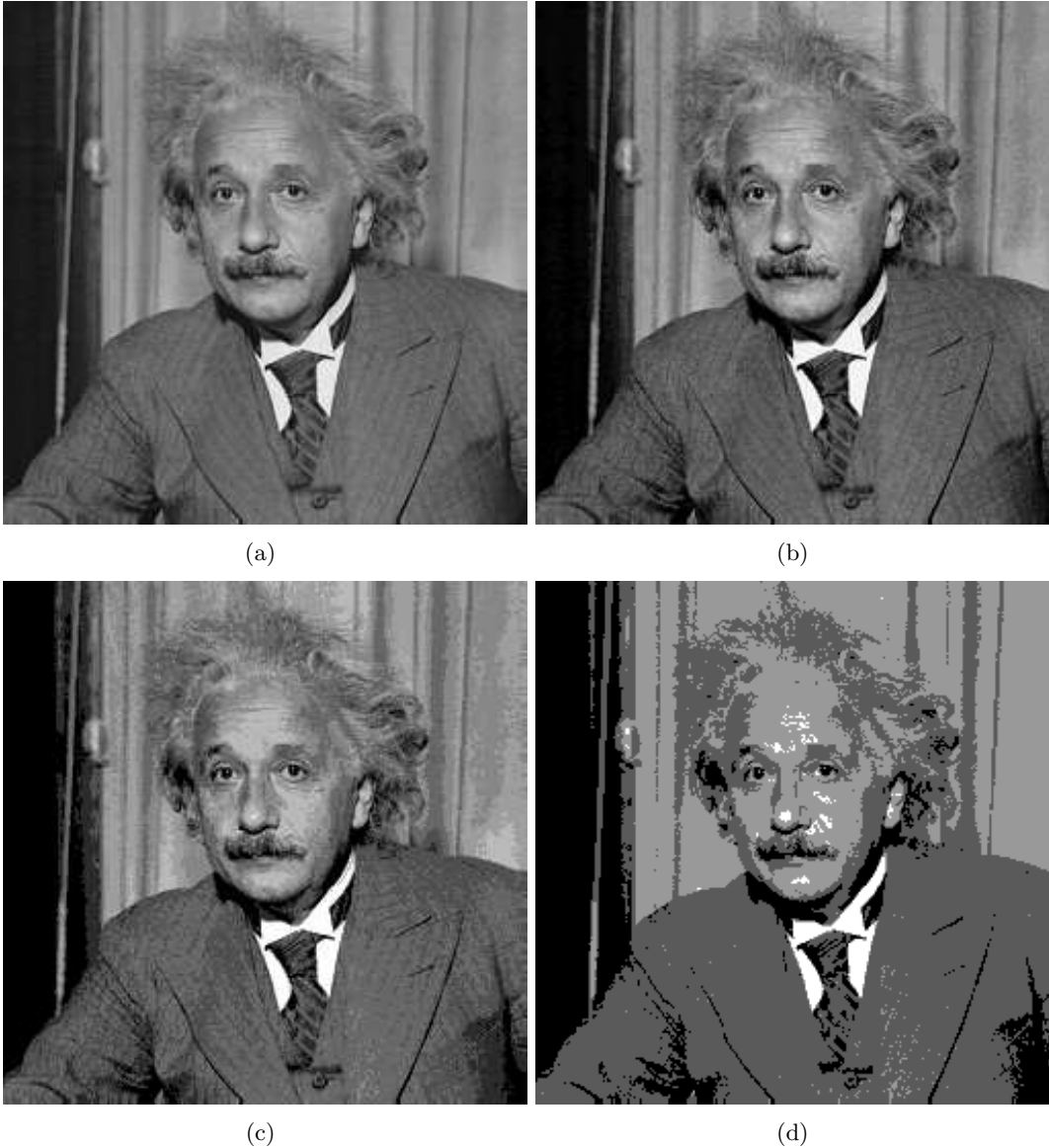
Τα επίπεδα κβαντισμού θα είναι

$$r_i = L_0 + \left(i - \frac{1}{2}\right) \frac{L_T - L_0}{N}, i = 1, \dots, N.$$

Άρα τα διαστήματα και τα επίπεδα κβαντισμού κατανέμονται ομοιόμορφα στην περίπτωση της ομοιόμορφης κατανομής. Το ελάχιστο μέσο τετραγωνικό σφάλμα είναι

$$D_{\min} = \frac{(L_T - L_0)^2}{12N^2}.$$

Στη γενική περίπτωση το παραπάνω σύστημα εξισώσεων (6.4) και (6.5) μπορεί να επιλυθεί μόνο με αριθμητικές επαναληπτικές μεθόδους. Μετά από κάποιες αρχικές τιμές για τα επίπεδα χβαντισμού, χρησιμοποιούνται διαδοχικά, και επαναληπτικά μέχρι τη σύγκλιση, οι εξισώσεις (6.5) και (6.4). Πρέπει να σημειωθεί ότι η επαναληπτική επίλυση του συστήματος των εξισώσεων δεν εξασφαλίζει πάντοτε την ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος. Μπορεί η σύγκλιση να οδηγήσει σε ένα τοπικό ελάχιστο, όταν οι παραπάνω συνθήκες είναι μόνο αναγκαίες, αλλά όχι και ικανές για την ελαχιστοποίηση του D . Στο Σχήμα 6.2 δίδονται αποτελέσματα χβαντισμού με χρήση της παραπάνω μεθόδου, όπου η κατανομή πιθανότητας για τις τιμές της έντασης προκύπτει εμπειρικά.



Σχήμα 6.2: Κβαντισμός σε 256, 16, 8 και 4 επίπεδα προσαρμοσμένα στην κατανομή των αποχρώσεων της εικόνας.

Ο Lloyd-Max κβαντιστής έχει τις ακόλουθες ιδιότητες.

- Η μέση κβαντισμένη τιμή είναι ίση με τη μέση τιμή της μεταβλητής

$$E\{Q(X)\} = E\{X\}.$$

- Η κβαντισμένη τιμή και το σφάλμα κβαντισμού είναι ασυσχέτιστες μεταβλητές

$$E\{(X - Q(X))Q(X)\} = 0.$$

- Η διασπορά του σφάλματος κβαντισμού είναι ίση με τη διαφορά της διασποράς της μεταβλητής μείον τη διασπορά των κβαντισμένων τιμών

$$E\{(X - Q(X))^2\} = \int_{-\infty}^{\infty} x^2 p(x) dx - \sum_{i=1}^N P_i r_i^2,$$

$$\text{όπου } P_i = \int_{d_{i-1}}^{d_i} p(x) dx.$$

- Ασυμπτωτικά, δηλαδή για μεγάλο N , η διασπορά του σφάλματος κβαντισμού μειώνεται όπως το $1/N^2$,

$$D \approx \frac{\left(\int_{-\infty}^{\infty} (p(x))^{1/3} dx\right)^3}{12N^2}.$$

Αν η τυχαία μεταβλητή X κατανέμεται ομοιόμορφα, τα διαστήματα κβαντισμού είναι ισομήκη, και τα επίπεδα κβαντισμού κατανέμονται ομοιόμορφα, όπως αποδείχθηκε στο Παράδειγμα 6.1. Ο ομοιόμορφος κβαντιστής προσδιορίζεται με τη χρήση μίας μόνο παραμέτρου, το βήμα του κβαντισμού. Γι' αυτό το λόγο θα μπορούσε να έχει ενδιαφέρον και στην περίπτωση μεταβλητών που δεν κατανέμονται ομοιόμορφα. Αν μία μεταβλητή κατανέμεται στο σύνολο των πραγματικών αριθμών, υπάρχει ένα βέλτιστο βήμα κβαντισμού που ελαχιστοποιεί το μέσο τετραγωνικό σφάλμα. Παρουσιάζεται στη συνέχεια το πρόβλημα βελτιστοποίησης στην περίπτωση μίας συμμετρικής μεταβλητής, όταν $p(x) = p(-x)$, με συνέπεια ο κβαντιστής να είναι συμμετρικός, κι όταν επιπλέον ο αριθμός των επιπέδων κβαντισμού είναι άρτιος.

Ας είναι Δ το βήμα του κβαντιστή και $2N$ ο αριθμός των επιπέδων κβαντισμού. Θα έχουμε

$$\left. \begin{aligned} r_i &= (2i - \text{sgn}(i)) \frac{\Delta}{2}, & i = \pm 1, \dots, \pm N \\ d_i &= i\Delta, & i = 0, \pm 1, \dots, \pm(N-1) \end{aligned} \right\} \quad (6.6)$$

Το μέσο τετραγωνικό σφάλμα είναι

$$D = 2 \sum_{i=1}^{N-1} \int_{d_{i-1}}^{d_i} \left(x - (2i-1) \frac{\Delta}{2}\right)^2 p(x) dx + 2 \int_{d_{N-1}}^{\infty} \left(x - (2N-1) \frac{\Delta}{2}\right)^2 p(x) dx. \quad (6.7)$$

Η λύση ως προς Δ ευρίσκεται με αριθμητικές μεθόδους.

Άσκηση

Δίδεται τυχαία μεταβλητή X με συνάρτηση πυκνότητας πιθανότητας

$$p(x) = \frac{1}{\sigma\sqrt{2}} \exp\left(-\frac{\sqrt{2}|x|}{\sigma}\right).$$

1. Να ευρεθεί ο βέλτιστος κατά Lloyd-Max κβαντιστής για $N = 2, 3$ και 4 επίπεδα κβαντισμού. Ποιό είναι το μέσο τετραγωνικό σφάλμα σε κάθε μία από τις παραπάνω περιπτώσεις;
2. Τα ίδια ερωτήματα τίθενται για το βέλτιστο ομοιόμορφο κβαντιστή.

Κεφάλαιο 7

2-Δ διακριτά συστήματα

Ένα 2-Δ διακριτό σύστημα (ή φίλτρο) δέχεται στην είσοδό του ένα 2-Δ διακριτό σήμα $x(m, n)$ και δίδει σαν απόκριση ένα μοναδικό 2-Δ διακριτό σήμα $y(m, n)$

$$y(m, n) = \mathcal{H}[x(m, n)]. \quad (7.1)$$

Ένα σύστημα ονομάζεται γραμμικό, αν, και μόνο αν, η απόκριση του συστήματος σε κάθε γραμμικό συνδυασμό εισόδων $x_1(m, n)$ και $x_2(m, n)$ είναι ο ίδιος γραμμικός συνδυασμός των αντίστοιχων εξόδων $y_1(m, n)$ και $y_2(m, n)$

$$\mathcal{H}[a_1x_1(m, n) + a_2x_2(m, n)] = a_1y_1(m, n) + a_2y_2(m, n), \forall a_1, a_2 \in \mathbb{R}. \quad (7.2)$$

Ορίζεται η ακολουθία $\delta(m, n)$ του Kronecker κατά τον ακόλουθο τρόπο

$$\delta(m, n) = \begin{cases} 1 & (m, n) = (0, 0) \\ 0 & (m, n) \neq (0, 0) \end{cases} \quad (7.3)$$

Οποιοδήποτε σήμα επομένως θα μπορεί να θεωρηθεί ως γραμμικός συνδυασμός των δειγμάτων του,

$$x(m, n) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \delta(m-k, n-l)x(k, l).$$

Άρα η έξοδος ενός γραμμικού συστήματος γράφεται ως εξής

$$y(m, n) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} h(m-k, n-l; m, n)x(k, l). \quad (7.4)$$

7.1 Γραμμικά αμετάβλητα σε μετατόπιση συστήματα

Στην περίπτωση που η απόκριση του συστήματος είναι αμετάβλητη σε μια μετατόπιση της εισόδου, δηλαδή αν μία μετατόπιση της εισόδου συνεπάγεται την ίδια ακριβώς μετατόπιση της εξόδου, χωρίς καμία άλλη αλλαγή, τότε η σχέση εισόδου-εξόδου γίνεται

$$y(m, n) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} h(m-k, n-l)x(k, l). \quad (7.5)$$

Μπορεί κανείς εύκολα να διαπιστώσει ότι $h(m, n)$ είναι η απόκριση του συστήματος στην ακολουθία $\delta(m, n)$.

Γι' αυτό το λόγο η 2-Δ ακολουθία $h(m, n)$ ονομάζεται κρουστική απόκριση του συστήματος (ή φίλτρου). Η έξοδος επομένως ενός γραμμικού αμετάβλητου σε μετατόπιση συστήματος προκύπτει από τη συνέλιξη της εισόδου με την κρουστική απόκριση του συστήματος.

Η συνέλιξη έχει την ιδιότητα της συμμετρικότητας. Άρα ισχύει επίσης

$$y(m, n) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} h(k, l)x(m-k, n-l).$$

Η συνέλιξη έχει και την προσεταιριστική ιδιότητα, που σημαίνει ότι δύο συστήματα στη σειρά ισοδυναμούν με ένα σύστημα με κρουστική απόκριση τη συνέλιξη των δύο αντίστοιχων κρουστικών αποκρίσεων. Η συνέλιξη έχει επίσης την επιμεριστική ιδιότητα, που σημαίνει πως αν δύο συστήματα χρησιμοποιηθούν παράλληλα, με την ίδια είσοδο, και υπερτεθούν οι δύο έξοδοι, είναι ταυτόσημο μ' ένα σύστημα του οποίου η κρουστική απόκριση είναι το άθροισμα των δύο κρουστικών αποκρίσεων.

Αν η κρουστική απόκριση ενός συστήματος είναι διαχωρίσιμη

$$h(m, n) = h_1(m)h_2(n), \quad (7.6)$$

η απόκριση του συστήματος μπορεί να υπολογισθεί χωριστά στους δύο δείκτες (m, n) , με αποτέλεσμα οικονομία και απλοποίηση των υπολογισμών.

Παράδειγμα 7.1. Το σύστημα με κρουστική απόκριση

$$h(m, n) = \begin{cases} 1/6 & m = -1 \text{ και } |n| \leq 1 \\ -1/6 & m = 1 \text{ και } |n| \leq 1 \\ 0 & |m| \neq 1 \text{ είτε } |n| > 1 \end{cases}$$

είναι διαχωρίσιμο, με

$$h_1(m) = \begin{cases} 1/2 & m = -1 \\ -1/2 & m = 1 \\ 0 & |m| \neq 1 \end{cases} \text{ και } h_2(n) = \begin{cases} 1/3 & |n| \leq 1 \\ 0 & |n| > 1 \end{cases}$$

Το πραγματικό μέρος του μετασχηματισμού Fourier της $h(m, n)$ είναι μηδέν, ενώ το φανταστικό μέρος δίδεται στο Σχήμα 7.1.

Ένα γραμμικό σύστημα ονομάζεται αμετάβλητο στην περιστροφή, αν η κρουστική του απόκριση είναι συνάρτηση μόνο μιας μεταβλητής, της απόστασης d από την αρχή των αξόνων

$$h(m, n) = h_0(d). \quad (7.7)$$

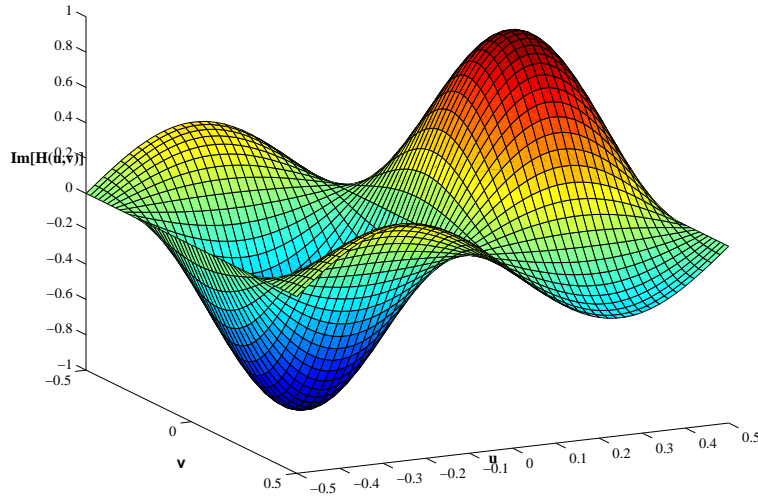
Η ακριβής έκφραση της $h(m, n)$ εξαρτάται από τον ορισμό της απόστασης. Δυνατές εκφράσεις είναι οι ακόλουθες: $d = \sqrt{m^2 + n^2}$ ή $d = |m| + |n|$.

Παράδειγμα 7.2. Το σύστημα με κρουστική απόκριση

$$h(m, n) = \begin{cases} 1 & m^2 + n^2 = 0 \\ 1/2 & m^2 + n^2 = 1 \\ 1/4 & m^2 + n^2 = 2 \\ 0 & m^2 + n^2 > 2 \end{cases}$$

είναι αμετάβλητο στην περιστροφή. Είναι επίσης διαχωρίσιμο με

$$h_1(m) = \begin{cases} 1 & m = 0 \\ 1/2 & |m| = 1 \\ 0 & |m| > 1 \end{cases}$$



Σχήμα 7.1: Μετασχηματισμός Fourier της κρουστικής απόκρισης του παραδείγματος 7.1.

και $h_2(n) = h_1(n)$. Ο μετασχηματισμός Fourier της κρουστικής απόκρισης του συστήματος αυτού δίδεται στο Σχήμα 5.2.

Λέμε ότι ένα φίλτρο είναι πεπερασμένης κρουστικής απόκρισης, αν η $h(m, n)$ παίρνει μη μηδενικές τιμές μόνο σ' ένα πεπερασμένο σύνολο σημείων (m, n) , αν δηλαδή

$$h(m, n) = 0, |m| > M, |n| > N. \quad (7.8)$$

Και τα δύο συστήματα που δόθηκαν στα δύο παραπάνω παραδείγματα έχουν πεπερασμένη κρουστική απόκριση ($M = N = 1$). Στην αντίθετη περίπτωση το φίλτρο ονομάζεται άπειρης κρουστικής απόκρισης.

Μεταξύ των φίλτρων με άπειρης έκτασης κρουστική απόκριση ευρέως χρησιμοποιείται αυτό που φέρει το όνομα του Gauss λόγω της έκφρασης του ορισμού του

$$h(m, n) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{m^2 + n^2}{2\sigma^2}\right). \quad (7.9)$$

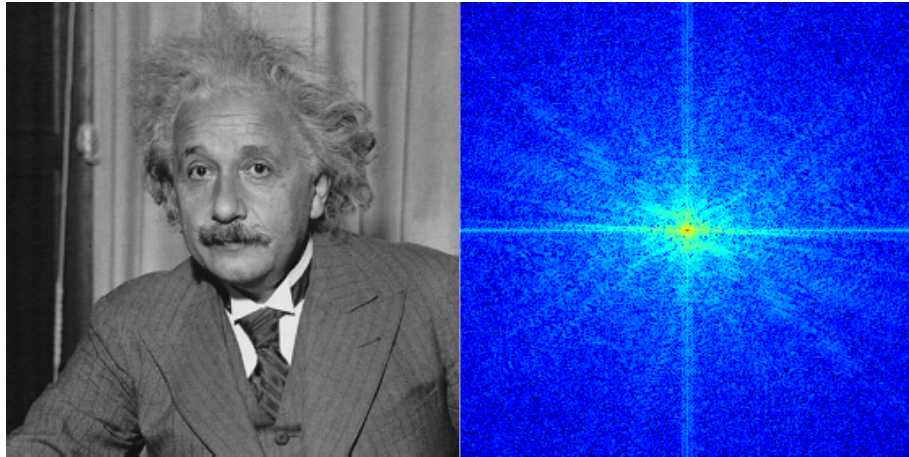
Το σύστημα αυτό είναι αναλλοίωτο στην περιστροφή. Η απόκρισή του μπορεί να υπολογισθεί προσεγγίζοντας την κρουστική του απόκριση σε πεπερασμένη έκταση, ή με προσέγγιση με αναδρομικές σχέσεις, όπως αυτές που παρουσιάζονται παρακάτω. Υπάρχει επίσης η δυνατότητα υλοποίησης στο πεδίο των χωρικών συχνοτήτων. Θα μπορούσε σε αυτή την περίπτωση να χρησιμοποιηθεί ο μετασχηματισμός Fourier όπως δίδεται κατωτέρω

$$H(u, v) = \exp\left(-\frac{2\sigma^2(u^2 + v^2)}{\pi}\right). \quad (7.10)$$

Στο Σχήμα 7.2 δίδεται το αποτέλεσμα του φίλτρου με $\sigma = 1$ και $\sigma = 2$, καθώς και οι αντίστοιχοι μετασχηματισμοί Fourier.

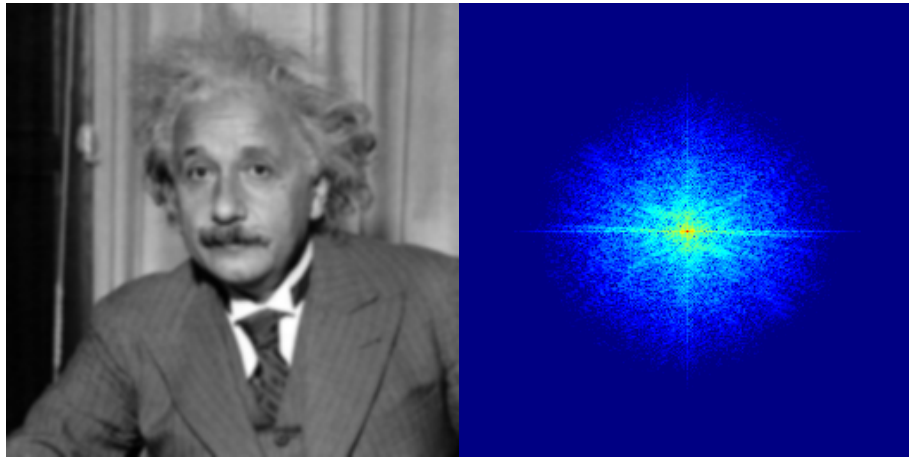
Παράδειγμα 7.3. Το σύστημα με κρουστική απόκριση

$$h(m, n) = \alpha^{|m|} \beta^{|n|}, (m, n) \in \mathbb{Z}^2$$



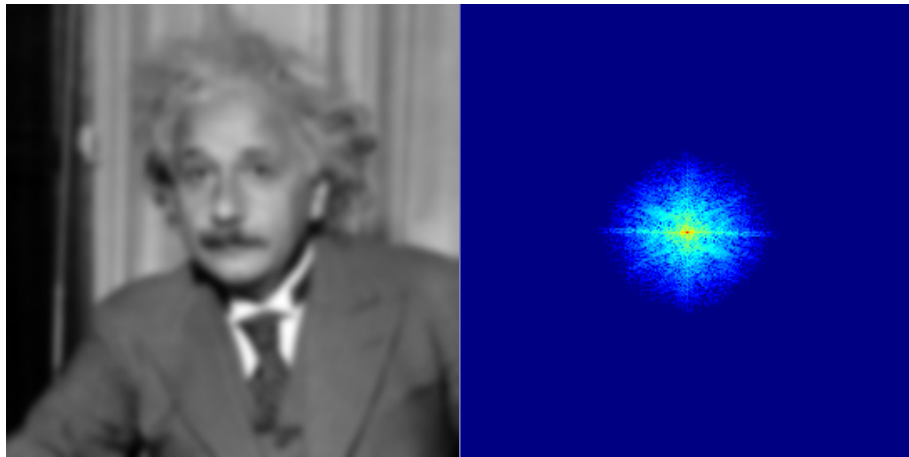
(a)

(b)



(c)

(d)



(e)

(f)

Σχήμα 7.2: Φίλτρο Gauss. Στην αριστερή στήλη δίδεται η αρχική εικόνα και το αποτέλεσμα του φίλτρου για $\sigma = 1$ και $\sigma = 2$. Στη δεξιά στήλη δίδεται ο αντίστοιχος μετασχηματισμός Fourier σε αναπαράσταση με χρωματική κλίμακα όπου το μπλε αντιστοιχεί σε χαμηλές τιμές.

είναι διαχωρίσιμο και εκτείνεται σε όλο το επίπεδο των ακεραίων. Αν $\alpha = \beta$, το σύστημα είναι αμετάβλητο στην περιστροφή, με απόσταση $d = |m| + |n|$.

Τα φίλτρα πεπερασμένης χροστικής απόκρισης δεν παρουσιάζουν καμία δυσκολία υλοποίησης. Αντίθετα τα φίλτρα άπειρης χροστικής απόκρισης μπορούν να υλοποιηθούν μόνο αν μπορούν να περιγραφούν με τη βοήθεια μιας αναδρομικής σχέσης, δηλαδή μιας 2- Δ εξίσωσης διαφοράς, όπως η ακόλουθη

$$y(m, n) = \sum_{(k,l) \in D_a} a(k, l)y(m - k, n - l) + \sum_{(k,l) \in D_b} b(k, l)x(m - k, n - l), \quad (7.11)$$

όπου τα σύνολα D_a και D_b είναι πεπερασμένα. Η παραπάνω σχέση ορίζει μοναδικά ένα γραμμικό σύστημα μόνο εφόσον ορισθούν οι οριακές συνθήκες. Το γραμμικό σύστημα που ορίζεται μ' αυτόν τον τρόπο είναι αμετάβλητο σε μετατόπιση υπό τις προϋποθέσεις που ακολουθούν. Κατ' αρχήν απαιτείται η εξασφάλιση της δυνατότητας αναδρομικών υπολογισμών, θέμα που αναπτύσσεται με λεπτομέρεια στην επόμενη παράγραφο. Στη συνέχεια πρέπει να προσδιορισθεί η οριακή περιοχή, η οποία προκύπτει ως το συμπλήρωμα της περιοχής της απόκρισης. Η περιοχή της απόκρισης προσδιορίζεται μέσω της περιοχής της χροστικής απόκρισης και της περιοχής της εισόδου $x(m, n)$, όπως αυτή υπεισέρχεται στο δεύτερο μέρος της Εξίσωσης (7.11). Η απόκριση στην οριακή περιοχή πρέπει να παίρνει μηδενικές τιμές, για να είναι το σύστημα που ορίζει η 2- Δ εξίσωση διαφοράς αμετάβλητο σε μετατόπιση, και να χαρακτηρίζεται από τη συνέλιξη χροστικής απόκρισης και εισόδου.

Για να είναι η σχέση (7.11) αναδρομική απαιτείται κατ' αρχήν να ορισθεί μία σχέση διάταξης στο επίπεδο των ακεραίων αριθμών, με δοσμένο ότι δεν υπάρχει φυσική διάταξη για ζεύγη ακεραίων αριθμών. Ας υποθέσουμε ότι υιοθετείται η λεξικογραφική διάταξη, με τη σειρά που εμφανίζονται οι δύο δείκτες που προσδιορίζουν ένα σημείο. Η σχέση (7.11) είναι αναδρομική, αν, για μιά δοσμένη διάταξη, το σύνολο D_a είναι τέτοιο ώστε το σημείο (m, n) να έπεται όλων των σημείων $(m - k, n - l)$, όπου $(k, l) \in D_a$. Για τη λεξικογραφική διάταξη, και για L_1, L_2 ακέραιους αριθμούς, έχουμε

$$D_a = \{(k, l) : k = 0, 0 < l \leq L_2 \text{ είτε } 0 < k \leq K, L_1 \leq l \leq L_2\}. \quad (7.12)$$

Όποια κι αν είναι η θεωρούμενη διάταξη, για να είναι η σχέση (7.11) αναδρομική, απαιτείται όλα τα σημεία του D_a να εγκλείονται σ' ένα τομέα με κορυφή το σημείο $(0, 0)$ και άνοιγμα αυστηρά μικρότερο από π , κι επιπλέον $(0, 0) \notin D_a$.

Ας θεωρήσουμε τώρα το καθαρά αναδρομικό φίλτρο που δίδεται από την ακόλουθη 2- Δ εξίσωση διαφοράς

$$y(m, n) = \sum_{(k,l) \in D_a} a(k, l)y(m - k, n - l) + x(m, n). \quad (7.13)$$

Η χροστική απόκριση αυτού του φίλτρου είναι λύση της εξίσωσης

$$h(m, n) = \sum_{(k,l) \in D_a} a(k, l)h(m - k, n - l) + \delta(m, n). \quad (7.14)$$

Μία τέτοια χροστική απόκριση μπορεί να πάρει μη μηδενικές τιμές σε όλα τα σημεία του ελάχιστου τομέα που με κορυφή το σημείο $(0, 0)$ εγκλείει όλα τα σημεία του D_a . Είναι τότε δυνατό να διακρίνουμε δύο περιπτώσεις. Αν $L_1 = 0$, ο μέγιστος τέτοιος τομέας έχει άνοιγμα $\pi/2$, και τότε αναφερόμαστε σε ένα φίλτρο ενός τέταρτου του επιπέδου. Στην αντίθετη περίπτωση αναφερόμαστε σ' ένα φίλτρο ενός μη συμμετρικού ημιεπίπεδου.

Παράδειγμα 7.4. Η σχέση εισόδου/εξόδου

$$y(m, n) = \alpha y(m - 1, n) + \beta y(m, n - 1) + x(m, n)$$

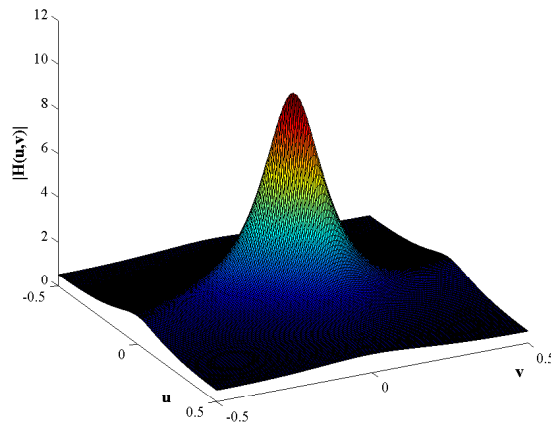
ορίζει ένα φίλτρο ενός τετάρτου του επιπέδου, ενώ η σχέση εισόδου/εξόδου

$$y(m, n) = \alpha y(m - 1, n + 1) + \beta y(m, n - 1) + x(m, n)$$

ορίζει ένα φίλτρο ενός μη συμμετρικού ημιεπιπέδου. Ο μετασχηματισμός Fourier της κρουστικής απόκρισης του φίλτρου που ορίζεται στο τέταρτο του επιπέδου είναι

$$H(u, v) = \frac{1}{1 - \alpha \exp(-2\pi ju) - \beta \exp(-2\pi jv)}$$

Για $\alpha = 0,55$ και $\beta = 0,35$ στο Σχήμα 7.3 δίδεται η γραφική παράσταση του μέτρου του μετασχηματισμού Fourier.



Σχήμα 7.3: Μετασχηματισμός Fourier της κρουστικής απόκρισης του παραδείγματος 7.4.

7.2 2-Δ μετασχηματισμός Z

Πριν προχωρήσουμε στο θέμα της ευστάθειας ενός φίλτρου με άπειρης έκτασης κρουστική απόκριση, και για τις ανάγκες της μελέτης της ευστάθειας μεταξύ άλλων, δίνουμε στη συνέχεια τον ορισμό και μερικές ιδιότητες του 2-Δ μετασχηματισμού Z. Ο 2-Δ μετασχηματισμός Z αποτελεί γενίκευση του μετασχηματισμού Fourier 2-Δ διακριτών ακολουθιών. Ο μετασχηματισμός Z του 2-Δ διακριτού σήματος $x(m, n)$ ορίζεται ως εξής

$$X(z_1, z_2) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} x(m, n) z_1^{-m} z_2^{-n}, \quad (7.15)$$

όπου z_1 και z_2 είναι μιγαδικές μεταβλητές. Το σύνολο των τιμών των z_1 και z_2 για τις οποίες το άθροισμα (7.15) υπάρχει, ονομάζεται περιοχή σύγκλισης. Πρόκειται για το σύνολο των τιμών

των z_1 και z_2 για τα οποία υπάρχει το ακόλουθο άθροισμα

$$\sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} |x(m, n)| |z_1|^{-m} |z_2|^{-n} < \infty. \quad (7.16)$$

Επομένως η περιοχή σύγκλισης ορίζεται στο επίπεδο $(|z_1|, |z_2|)$. Για 2-Δ σήματα πεπερασμένης έκτασης η περιοχή σύγκλισης είναι ολόκληρο το επίπεδο $(|z_1|, |z_2|)$, εκτός ίσως των σημείων όπου $|z_1| = 0, |z_1| = \infty, |z_2| = 0, |z_2| = \infty$. Για 2-Δ σήματα περιορισμένα στο πρώτο τέταρτο του επιπέδου, αν το $(|z'_1|, |z'_2|)$ ανήκει στην περιοχή σύγκλισης, τότε κάθε σημείο τέτοιο ώστε $|z_1| \geq |z'_1|$ και $|z_2| \geq |z'_2|$ ανήκει επίσης. Ο αντίστροφος μετασχηματισμός Z δίδει την αρχική ακολουθία

$$x(m, n) = \frac{1}{(j2\pi)^2} \oint \oint X(z_1, z_2) z_1^{m-1} z_2^{n-1} dz_1 dz_2, \quad (7.17)$$

όπου η ολοκλήρωση γίνεται με φορά αντίστροφη εκείνης των δεικτών ενός ρολογιού πάνω σε κλειστές καμπύλες που ανήκουν στην περιοχή σύγκλισης και περικλείουν την αρχή $(0, 0)$. Για την αντιστροφή ιδιαίτερα χρήσιμος είναι ο τύπος του Cauchy,

$$f(z_0) = \frac{1}{j2\pi} \oint \frac{f(z) dz}{z - z_0},$$

όπου η κλειστή καμπύλη και η ολοκλήρωση είναι ως ανωτέρω κι επιπλέον η συνάρτηση $f(z)$ είναι αναλυτική, δηλαδή παραγωγίζεται σε όλα τα σημεία του κλειστού συνόλου που ορίζει η καμπύλη του ολοκληρώματος. Ο ανωτέρω τύπος επεκτείνεται σε παραγώγους οποιασδήποτε τάξης,

$$f^{(n)}(z_0) = \frac{n!}{j2\pi} \oint \frac{f(z) dz}{(z - z_0)^{n+1}}, \quad (7.18)$$

όπου n μη αρνητικός ακέραιος.

Παράδειγμα 7.5. Ας θεωρήσουμε το μετασχηματισμό

$$H(z_1, z_2) = \frac{1}{1 - \alpha z_1^{-1} - \beta z_2^{-1}}. \quad (7.19)$$

Η Εξίσωση (7.17) σε συνδυασμό με το γενικό τύπο του Cauchy της Εξίσωσης (7.18) δίδει την χρουστική απόκριση του συστήματος

$$h(m, n) = \frac{(m+n)!}{m!n!} \alpha^m \beta^n, \quad m \geq 0, n \geq 0. \quad (7.20)$$

Ας έλθουμε τώρα σε μερικές ιδιότητες του μετασχηματισμού Z . Κατ' αρχή ο μετασχηματισμός είναι γραμμικός.

1. Μετατόπιση

$$\sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} x(m - m_0, n - n_0) z_1^{-m} z_2^{-n} = z_1^{-m_0} z_2^{-n_0} X(z_1, z_2) \quad (7.21)$$

2. Διαχωρισιμότητα

$$\text{Αν } x(m, n) = x_1(m)x_2(n), \text{ τότε } X(z_1, z_2) = X_1(z_1)X_2(z_2) \quad (7.22)$$

όπου $X_1(z_1)$ (αντίστοιχα $X_2(z_2)$) είναι ο μετασχηματισμός Z της $x_1(m)$ (αντίστοιχα $x_2(n)$).

3. Συμμετρική ακολουθία

Αν η ακολουθία $y(m, n)$ είναι συμμετρική της $x(m, n)$, δηλαδή αν $y(m, n) = x(-m, -n)$, τότε

$$Y(z_1, z_2) = X(z_1^{-1}, z_2^{-1}) \quad (7.23)$$

4. Συνέλιξη

Ο μετασχηματισμός Z της $y(m, n)$, που προκύπτει ως συνέλιξη της $h(., .)$ με τη $x(., .)$, που δίδεται στην Εξίσωση (7.5), έχει ως εξής

$$Y(z_1, z_2) = H(z_1, z_2)X(z_1, z_2), \quad (7.24)$$

όπου $H(z_1, z_2)$ είναι ο μετασχηματισμός Z της $h(m, n)$, και ονομάζεται συνάρτηση μεταφοράς του συστήματος.

7.3 Ευστάθεια

Ένα σύστημα ονομάζεται ευσταθές, αν για μια φραγμένη είσοδο ($|x(m, n)| < \infty, \forall(m, n)$), η έξοδος είναι φραγμένη ($|y(m, n)| < \infty, \forall(m, n)$). Αυτό ισοδυναμεί με την ύπαρξη του ακόλουθου αθροίσματος

$$\sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} |h(m, n)| < \infty. \quad (7.25)$$

Πράγματι, θα έχουμε

$$|y(m, n)| < \sum_{m'=-\infty}^{\infty} \sum_{n'=-\infty}^{\infty} |h(m', n')x(m-m', n-n')| < \max |x(m, n)| \sum_{m'=-\infty}^{\infty} \sum_{n'=-\infty}^{\infty} |h(m', n')| < \infty.$$

Ισχύει και το αντίστροφο, δηλαδή η συνθήκη της Εξίσωσης (7.25) είναι αναγκαία για την ευστάθεια. Ας θεωρήσουμε ως σήμα εισόδου το $x(m, n) = \text{sign}(h(-m, -n))$. Τότε θα έχουμε

$$|y(0, 0)| = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \text{sign}(h(m, n))h(m, n) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} |h(m, n)| < \infty.$$

Με βάση τον ορισμό της περιοχής σύγκλισης της Εξίσωσης (7.16), η ευστάθεια ενός συστήματος εξαρτάται από το εάν η περιοχή σύγκλισης για το μετασχηματισμό Z της κρουστικής απόκρισης του συστήματος περιλαμβάνει τους μοναδιαίους κύκλους $|z_1| = 1$ και $|z_2| = 1$. Είναι φανερό από τον ορισμό της ευστάθειας, ότι όλα τα φίλτρα με πεπερασμένη κρουστική απόκριση είναι ευσταθή. Ας θεωρήσουμε τώρα τα φίλτρα με άπειρη κρουστική απόκριση, που όμως περιορίζεται στο τέταρτο του επιπέδου, για τα οποία δίδεται μία αναδρομική σχέση ως ακολούθως

$$y(m, n) = \sum_{\substack{k=0 \\ |k|+|l| \neq 0}}^K \sum_{l=0}^L a(k, l)y(m-k, n-l) + x(m, n). \quad (7.26)$$

Εφαρμόζοντας την ιδιότητα της μετατόπισης για το μετασχηματισμό Z που δίδεται στην Εξίσωση (7.21), προκύπτει ότι η συνάρτηση μεταφοράς αυτού του συστήματος είναι ίση με το αντίστροφο ενός πολυωνύμου των δύο μιγαδικών μεταβλητών z_1^{-1} και z_2^{-1} ,

$$H(z_1, z_2) = \frac{1}{A(z_1, z_2)} = \frac{1}{1 - \sum_{\substack{k=0 \\ |k|+|l| \neq 0}}^K \sum_{l=0}^L a(k, l)z_1^{-k}z_2^{-l}}. \quad (7.27)$$

Τότε η συνθήκη ευστάθειας μπορεί να εκφραστεί με βάση τη θέση των ριζών του πολυωνύμου $A(z_1, z_2)$. Ένα φίλτρο με συνάρτηση μεταφοράς που δίδεται από την Εξίσωση (7.27) είναι ευσταθές, εάν, και μόνο εάν,

$$A(z_1, z_2) \neq 0 \quad \text{για} \quad |z_1| \geq 1, |z_2| \geq 1 \quad (7.28)$$

εάν, και μόνο εάν,

$$\begin{aligned} A(z_1, z_2) \neq 0 & \quad \text{για} \quad |z_1| \geq 1, |z_2| = 1 \\ A(z_1, z_2) \neq 0 & \quad \text{για} \quad |z_1| = 1, |z_2| \geq 1 \end{aligned} \quad (7.29)$$

εάν, και μόνο εάν,

$$\begin{aligned} A(z_1, z_2) \neq 0 & \quad \text{για} \quad |z_1| \geq 1, |z_2| = 1 \\ A(z_1, z_2) \neq 0 & \quad \text{για} \quad z_1 = 1, |z_2| \geq 1 \end{aligned} \quad (7.30)$$

εάν, και μόνο εάν,

$$\begin{aligned} A(z_1, z_2) \neq 0 & \quad \text{για} \quad |z_1| = |z_2| = 1 \\ A(z_1, 1) \neq 0 & \quad \text{για} \quad |z_1| \geq 1 \\ A(1, z_2) \neq 0 & \quad \text{για} \quad |z_2| \geq 1 \end{aligned} \quad (7.31)$$

Παράδειγμα 7.6. Ας θεωρήσουμε το πολυώνυμο

$$A(z_1, z_2) = 1 - \alpha z_1^{-1} - \beta z_2^{-1} \quad (7.32)$$

Οι ικανές και αναγκαίες συνθήκες ευστάθειας (7.29) γράφονται για την περίπτωση του παραδείγματος ως εξής

$$\begin{aligned} |\alpha| &< |1 - \beta e^{-j\omega_2}| \quad \forall \omega_2 \\ |\beta| &< |1 - \alpha e^{-j\omega_1}| \quad \forall \omega_1 \end{aligned} \quad (7.33)$$

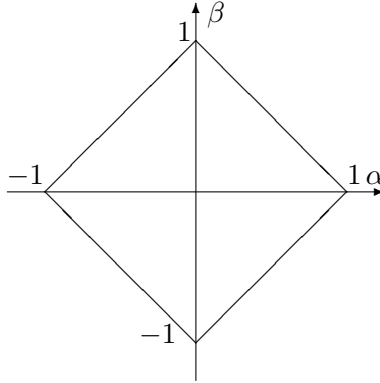
Από τις δύο αυτές ανισότητες προκύπτει ότι, αν το σύστημα είναι ευσταθές, θα πρέπει,

$$|\alpha \cos \omega_1 + \beta \cos \omega_2| < 1, \quad \forall \omega_1, \forall \omega_2 \quad (7.34)$$

Κατά συνέπεια, αν το σύστημα είναι ευσταθές θα έχουμε

$$|\alpha| + |\beta| < 1. \quad (7.35)$$

Αν αυτή η σχέση ισχύει μπορεί κανείς να διαπιστώσει ότι θα ισχύουν και οι σχέσεις (7.33), κι επομένως η σχέση (7.35) είναι ικανή και αναγκαία, για την ευστάθεια του συστήματος του παραδείγματος. Η συνθήκη αυτή δίδεται γραφικά στο Σχήμα 7.4. Το σύστημα είναι ευσταθές για τιμές των α και β στο εσωτερικό του τετραγώνου.



Σχήμα 7.4: Συνθήκη για ευστάθεια δισδιάστατης πρώτης τάξης αναδρομής.

Ασκήσεις

1. Για καθένα από τα παρακάτω συστήματα προσδιορίστε εάν το σύστημα είναι ή όχι γραμμικό, είναι ή όχι αμετάβλητο κατά την μετατόπιση. Εάν το σύστημα είναι γραμμικό και αμετάβλητο κατά την μετατόπιση, βρείτε την κρουστική απόκρισή του.

(a) $y(m, n) = x(m, n) + 1$

(b) $y(m, n) = x(2m, 2n)$

(c) $y(m, n) = \begin{cases} x(m/2, n/2) & m = 2k, n = 2l \\ 0 & \text{αλλιώς} \end{cases}$

(d) $y(m, n) = (m^2 + n^2)x(m, n)$

(e) $y(m, n) = \sum_{m'=-\infty}^m \sum_{n'=-\infty}^n x(m', n')$

(f) $y(m, n) = \exp(2x(m, n))$

(g) $y(m, n) = \sum_{m'=-1}^1 \sum_{n'=-1}^1 x(m', n')$

(h) $y(m, n) = \sum_{m'=0}^{M-1} \sum_{n'=0}^{N-1} x(m', n') \exp(-j \frac{2\pi m m'}{M}) \exp(-j \frac{2\pi n n'}{N})$

2. Δίδεται το ακόλουθο 2-Δ διακριτό σήμα

$$u(m, n) = \begin{cases} 1 & m \geq 0 \text{ και } n \geq 0 \\ 0 & \text{αλλιώς} \end{cases}$$

και το ακόλουθο 2-Δ φίλτρο

$$h(m, n) = \begin{cases} -4 & m^2 + n^2 = 0 \\ 1 & m^2 + n^2 = 1 \\ 0 & m^2 + n^2 > 1 \end{cases}$$

Να ευρεθεί η απόκριση του φίλτρου $h(m, n)$ στο σήμα $u(m, n)$. Να ευρεθεί επίσης ο μετασχηματισμός Z του $h(m, n)$.

3. Δίδεται το ακόλουθο δισδιάστατο φίλτρο με κρουστική απόκριση

$$h(m, n) = \begin{cases} \alpha^{-m+|n|} & m < 0, \forall n \\ 0 & m = 0, \forall n \\ -\alpha^{m+|n|} & m > 0, \forall n \end{cases}$$

Ποιά είναι απόκριση του φίλτρου στα ακόλουθα σήματα:

(a) $x(m, n) = 1, \forall m, n$

(b) $x(m, n) = \begin{cases} 1 & m \geq 0, n \geq 0 \\ 0 & \text{αλλού} \end{cases}$

Ποιά είναι η συνθήκη για το α ώστε να υπάρχουν αυτές οι αποκρίσεις; Πώς μπορεί να υλοποιηθεί αυτό το φίλτρο; Να δοθούν ο μετασχηματισμός Fourier και ο μετασχηματισμός Z του φίλτρου.

4. Εάν $X(z_1, z_2)$ είναι ο μετασχηματισμός Z της $x(m, n)$, και εάν $x(m, n) = x(-m, -n)$, να δειχθεί ότι $X(z_1, z_2) = X(\frac{1}{z_1}, \frac{1}{z_2})$.
5. Δίδεται η κρουστική απόκριση ενός 2-Δ συστήματος

$$h(m, n) = \begin{cases} \alpha^m & m \geq n \geq 0 \\ 0 & \text{αλλού} \end{cases}$$

Να ευρεθεί ο μετασχηματισμός Z της $h(m, n)$ και η περιοχή σύγκλισης. Ποιά είναι η συνθήκη ευστάθειας του συστήματος;

6. Δίδεται

$$x(m, n) = y(a_{11}m + a_{12}n, a_{21}m + a_{22}n)$$

όπου για σημεία του σήματος $y(m, n)$ που δεν έχουν αντίστοιχο $x(m, n)$, λαμβάνεται η τιμή 0. Να αποδειχθεί ότι

$$Y(z_1, z_2) = X(z_1^{a_{11}} z_2^{a_{21}}, z_1^{a_{12}} z_2^{a_{22}}).$$

7. Ας υποτεθεί ότι τα διανύσματα (N_{11}, N_{12}) και (N_{21}, N_{22}) προσδιορίζουν τα όρια του τομέα όπου ορίζεται η κρουστική απόκριση ενός 2-Δ συστήματος, με άνοιγμα αυστηρά μικρότερο του π . Δίδεται ότι: $D = N_{11}N_{22} - N_{12}N_{21} > 0$.
- (a) Να ευρεθεί ένας τρόπος γραμμικής απεικόνισης των σημείων του παραπάνω τομέα στο πρώτο τέταρτο του επιπέδου.
- (b) Αιτιολογείστε γιατί η συνθήκη ευστάθειας είναι ισοδύναμη για τις δύο περιπτώσεις πεδίου ορισμού.
- (c) Δίδεται ο μετασχηματισμός Z ενός 2-Δ συστήματος

$$H(z_1, z_2) = \frac{1}{1 - az_1^{-1}z_2 - bz_2^{-1}}.$$

Ποιά είναι η συνθήκη ευστάθειας του συστήματος;

- (d) Ποιά είναι η συνθήκη ευστάθειας για ένα σύστημα οριζόμενο στο δεύτερο τέταρτο του επιπέδου ($m \geq 0, n \leq 0$), αν

$$H(z_1, z_2) = \frac{1}{A(z_1, z_2)}$$

Κεφάλαιο 8

2-Δ ορθομοναδιαίοι μετασχηματισμοί

Αντικείμενο αυτού του κεφαλαίου είναι οι διακριτοί μετασχηματισμοί, που χρησιμεύουν στην εξαγωγή ορισμένων χαρακτηριστικών από τις εικόνες, όπως επίσης στην παράσταση στο πεδίο των συχνοτήτων, και γι' αυτό ακριβώς το λόγο χρησιμοποιούνται στη συμπίεση των εικόνων. Οι 2-Δ μετασχηματισμοί γίνονται σε ορθογώνια ή τετράγωνα μπλοκ. Για λόγους απλότητας θα περιορισθούμε εδώ στην περίπτωση των τετράγωνων μπλοκ.

Ας θεωρήσουμε ότι ένα τετράγωνο μπλοκ παριστάνεται από ένα τετραγωνικό πίνακα X , διαστάσεων $N \times N$. Ας είναι E_N ο διανυσματικός χώρος των τετραγωνικών πινάκων διαστάσεων $N \times N$, με εσωτερικό γινόμενο ορισμένο κατά τον ακόλουθο τρόπο

$$\langle A, B \rangle = \text{tr}[B^\dagger A] = \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} a(m, n) b^*(m, n), \quad (8.1)$$

όπου $A \in E_N, B \in E_N$, και B^\dagger είναι ο ανάστροφος και συζυγής πίνακας του B . Ένας γραμμικός μετασχηματισμός είναι μία γραμμική απεικόνιση του E_N στον E_N , που ορίζεται ως εξής

$$E_N \ni X \xrightarrow{\mathcal{T}} Y \in E_N$$
$$y(m, n) = \sum_{i=0}^{N-1} \sum_{k=0}^{N-1} t^*(m, n; i, k) x(i, k). \quad (8.2)$$

Σύμφωνα με τον ορισμό του εσωτερικού γινομένου (8.1) μπορούμε να γράψουμε

$$y(m, n) = \langle X, \Phi(m, n) \rangle, \quad (8.3)$$

όπου το τυπικό στοιχείο του πίνακα $\Phi(m, n)$ είναι $\{t(m, n; i, k)\}$. Τα στοιχεία του πίνακα Y ονομάζονται συντελεστές του μετασχηματισμού. Οι πίνακες $\Phi(m, n)$ συνιστούν τη βάση του μετασχηματισμού.

Υποθέτουμε ότι ο μετασχηματισμός \mathcal{T} είναι αντιστρέψιμος, δηλαδή ότι υπάρχει μετασχηματισμός, \mathcal{T}^{-1} , τέτοιος ώστε

$$\mathcal{T}\mathcal{T}^{-1} = \mathcal{T}^{-1}\mathcal{T} = \mathcal{I}, \quad (8.4)$$

όπου \mathcal{I} είναι ο μοναδιαίος μετασχηματισμός του διανυσματικού χώρου E_N . Ο ανάστροφος και συζυγής μετασχηματισμός, \mathcal{T}^\dagger , έχει τυπικό στοιχείο $\{t^*(i, k; m, n)\}$. Στη συνέχεια θα

περιορισθούμε στους ορθομοναδιαίους μετασχηματισμούς, οι οποίοι είναι αντιστρέψιμοι, και ο αντίστροφος είναι ίσος με τον ανάστροφο και συζυγή

$$\mathcal{T}^{-1} = \mathcal{T}^\dagger. \quad (8.5)$$

Σύμφωνα με τους παραπάνω ορισμούς προκύπτει ότι η βάση ενός ορθομοναδιαίου μετασχηματισμού είναι ορθοκανονική

$$\langle \Phi(m, n), \Phi(m', n') \rangle = \delta(m - m', n - n'). \quad (8.6)$$

Μπορεί επίσης να αποδειχθεί ότι η βάση είναι πλήρης. Επομένως ένας ορθομοναδιαίος μετασχηματισμός συνιστά μία ανάλυση σε μία πλήρη και ορθοκανονική βάση.

Ιδιαίτερο ενδιαφέρον έχουν οι διαχωρίσιμοι μετασχηματισμοί, για τους οποίους ισχύει

$$t(m, n; i, k) = t_1(m, i)t_2(n, k), \forall i, k, m, n \quad (8.7)$$

Σε αυτή την περίπτωση δύο τετραγωνικοί πίνακες, T_1 και T_2 , αρκούν για τον ορισμό του μετασχηματισμού, που δίδεται με τη βοήθεια δύο γινομένων πινάκων

$$Y = T_1^* X T_2^\dagger. \quad (8.8)$$

Όσον αφορά το μετασχηματισμό \mathcal{T} , είναι ίσος με το γινόμενο Kronecker των δύο συζυγών πινάκων

$$\mathcal{T} = T_1^* \otimes T_2^*. \quad (8.9)$$

Το γινόμενο Kronecker δύο τετραγωνικών πινάκων A και B ορίζεται ως εξής

$$A \otimes B = \begin{bmatrix} a(0,0)B & \cdots & a(0, N-1)B \\ \vdots & & \vdots \\ a(N-1,0)B & \cdots & a(N-1, N-1)B \end{bmatrix} \quad (8.10)$$

Ο αντίστροφος μετασχηματισμός είναι επίσης διαχωρίσιμος, και εφόσον είναι ορθομοναδιαίος δίδεται ως ακολούθως

$$\mathcal{T}^{-1} = T_1^t \otimes T_2^t, \quad (8.11)$$

οπότε θα έχουμε

$$X = T_1^t Y T_2. \quad (8.12)$$

Η πλέον συνηθισμένη περίπτωση είναι όταν $T_1 = T_2 = T$, οπότε είναι

$$Y = T^* X T^\dagger \quad \text{και} \quad X = T^t Y T. \quad (8.13)$$

Ο αριθμός των πολλαπλασιασμών και προσθέσεων που απαιτούνται στην περίπτωση ενός διαχωρίσιμου μετασχηματισμού είναι της τάξης $2N^3$, σε σχέση με N^4 που είναι στη γενική περίπτωση.

Κάθε ορθομοναδιαίος μετασχηματισμός έχει την ιδιότητα της διατήρησης της ενέργειας

$$\sum_{m=0}^{N-1} \sum_{n=0}^{N-1} y^2(m, n) = \|Y\|^2 = \langle \mathcal{T}X, \mathcal{T}X \rangle = \langle \mathcal{T}^\dagger \mathcal{T}X, X \rangle = \langle X, X \rangle = \|X\|^2 = \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} x^2(m, n). \quad (8.14)$$

Στη συνέχεια δίνουμε τους πλέον χρησιμοποιούμενους ορθομοναδιαίους μετασχηματισμούς.

8.1 Διακριτός μετασχηματισμός Fourier

Το τυπικό στοιχείο του διακριτού μετασχηματισμού Fourier (DFT, Discrete Fourier Transform) είναι το εξής

$$t(m, n; i, k) = \frac{1}{N} \exp\left(\frac{2\pi j}{N}(im + kn)\right) \quad (8.15)$$

Από τον ορισμό είναι προφανές ότι ο DFT είναι διαχωρίσιμος, και επίσης είναι εύκολο να αποδειχθεί ότι είναι ορθομοναδιαίος. Ο DFT είναι ευρύτατα χρησιμοποιούμενος μετασχηματισμός, για τον οποίο υπάρχουν γρήγοροι αλγόριθμοι υπολογισμού. Χάρη στη διαχωρισιμότητα, ο 2-Δ DFT μπορεί να υπολογισθεί με χρήση του 1-Δ FFT (Fast Fourier Transform), οπότε ο αριθμός των πολλαπλασιασμών ανέρχεται σε $N^2 \log_2 N$. Εφαρμόζοντας την ίδια μέθοδο με τον 1-Δ FFT κατ' ευθείαν στο 2-Δ DFT, μπορεί να μειωθεί ακόμη κατά 25% ο αριθμός των πολλαπλασιασμών. Γι' αυτό θέτουμε: $W_N = e^{-\frac{j2\pi}{N}}$ και γράφουμε:

$$\begin{aligned} S_{00}(m, n) &= \sum_{i=0}^{N/2-1} \sum_{k=0}^{N/2-1} x(2i, 2k) W_N^{2im+2kn} \\ S_{01}(m, n) &= \sum_{i=0}^{N/2-1} \sum_{k=0}^{N/2-1} x(2i, 2k+1) W_N^{2im+2kn} \\ S_{10}(m, n) &= \sum_{i=0}^{N/2-1} \sum_{k=0}^{N/2-1} x(2i+1, 2k) W_N^{2im+2kn} \\ S_{11}(m, n) &= \sum_{i=0}^{N/2-1} \sum_{k=0}^{N/2-1} x(2i+1, 2k+1) W_N^{2im+2kn} \end{aligned} \quad (8.16)$$

Μπορούμε τότε να γράψουμε

$$\begin{aligned} y(m, n) &= S_{00}(m, n) + W_N^n S_{01}(m, n) + W_N^m S_{10}(m, n) + W_N^{m+n} S_{11}(m, n) \\ y\left(m + \frac{N}{2}, n\right) &= S_{00}(m, n) + W_N^n S_{01}(m, n) - W_N^m S_{10}(m, n) - W_N^{m+n} S_{11}(m, n) \\ y\left(m, n + \frac{N}{2}\right) &= S_{00}(m, n) - W_N^n S_{01}(m, n) + W_N^m S_{10}(m, n) - W_N^{m+n} S_{11}(m, n) \\ y\left(m + \frac{N}{2}, n + \frac{N}{2}\right) &= S_{00}(m, n) - W_N^n S_{01}(m, n) - W_N^m S_{10}(m, n) + W_N^{m+n} S_{11}(m, n) \end{aligned} \quad (8.17)$$

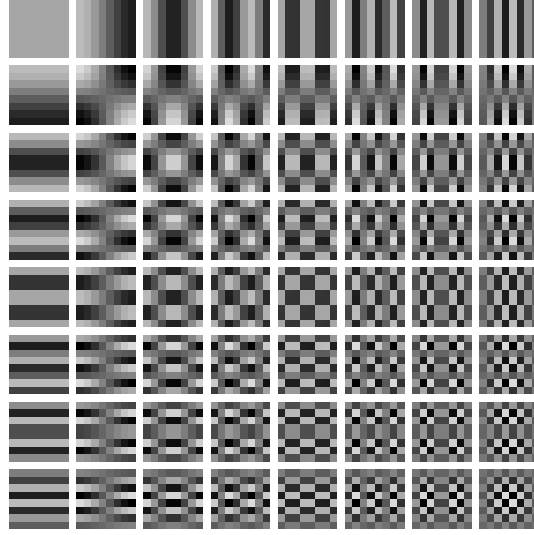
Το πλήθος υποδιαρέσεων αυτού του τύπου είναι $\log_2 N$, και για κάθε τέσσερις συντελεστές του μετασχηματισμού απαιτούνται τρεις πολλαπλασιασμοί. Επομένως ο αριθμός των πολλαπλασιασμών θα είναι $\frac{3N^2}{4} \log_2 N$.

8.2 Διακριτός μετασχηματισμός συνημιτόνου

Το τυπικό στοιχείο του διακριτού μετασχηματισμού συνημιτόνου (DCT, Discrete Cosine Transform) είναι το εξής

$$t(m, n; i, k) = \frac{c(m)c(n)}{N} \cos \frac{(2i+1)m\pi}{2N} \cos \frac{(2k+1)n\pi}{2N}, \quad (8.18)$$

όπου $c(m) = \begin{cases} 1 & m=0 \\ \sqrt{2} & m \neq 0 \end{cases}$. Ο DCT είναι ένας διαχωρίσιμος ορθομοναδιαίος μετασχηματισμός, που χρησιμοποιείται ευρύτατα για τη συμπίεση των εικόνων. Ο μετασχηματισμός συνημιτόνου είναι ιδιαίτερα αποτελεσματικός στη συμπίεση, γιατί προσεγγίζει τις επιδόσεις του βέλτιστου γραμμικού μετασχηματισμού, εφόσον τα δεδομένα είναι ισχυρά συσχετισμένα.



Σχήμα 8.1: Η βάση του 2-Δ διακριτού μετασχηματισμού συνημιτόνου

Ο διακριτός μετασχηματισμός συνημιτόνου συνδέεται άμεσα με το διακριτό μετασχηματισμό Fourier. Ας είναι

$$x_1(m, n) = \begin{cases} x(m, n) & 0 \leq m < N, 0 \leq n < N \\ x(m, 2N - 1 - n) & 0 \leq m < N, N \leq n < 2N \\ x(2N - 1 - m, n) & N \leq m < 2N, 0 \leq n < N \\ x(2N - 1 - m, 2N - 1 - n) & N \leq m < 2N, N \leq n < 2N \end{cases} \quad (8.19)$$

και $y_1(m, n)$ ο διακριτός μετασχηματισμός Fourier του $x_1(m, n)$. Τότε αποδεικνύεται ότι, αν είναι $y(m, n)$ ο διακριτός μετασχηματισμός συνημιτόνου του $x(m, n)$, ισχύει

$$y(m, n) = \frac{1}{4} c(m) c(n) W_{2N}^{\frac{m+n}{2}} y_1(m, n), 0 \leq m < N, 0 \leq n < N \quad (8.20)$$

Οπότε για τον υπολογισμό του 2-Δ DCT μπορεί να χρησιμοποιηθεί ο 2-Δ FFT με διπλάσιο αριθμό σημείων ανά διάσταση. Ωστόσο υπάρχουν γρήγοροι αλγόριθμοι υπολογισμού που δεν απαιτούν τη χρήση του FFT με μικρότερο τελικά κόστος.

8.3 Διακριτός μετασχηματισμός ημιτόνου

Το τυπικό στοιχείο του διακριτού μετασχηματισμού ημιτόνου (DST, Discrete Sine Transform) είναι το εξής

$$t(m, n; i, k) = \frac{2}{N+1} \sin \frac{\pi(i+1)(m+1)}{N+1} \sin \frac{\pi(k+1)(n+1)}{N+1} \quad (8.21)$$

Είναι επίσης διαχωρίσιμος και ορθομοναδιαίος, και υπάρχουν γρήγοροι αλγόριθμοι για τον υπολογισμό του.

8.4 Διακριτός μετασχηματισμός Walsh-Hadamard

Το τυπικό στοιχείο του διακριτού μετασχηματισμού Walsh-Hadamard (WHT) ορίζεται με τις δυαδικές τιμές ± 1 , αν αγνοήσουμε τον παράγοντα κανονικοποίησης $1/N$. Ο υπολογισμός επομένως του WHT δεν απαιτεί πολλαπλασιασμούς. Το τυπικό στοιχείο είναι το εξής

$$t(m, n; i, k) = \frac{1}{N}(-1)^{p(m, n; i, k)} \quad (8.22)$$

όπου $N = 2^l$ και $p(m, n; i, k) = \sum_{r=0}^{l-1} (i_r m_r^{(-)} + k_r n_r^{(-)})$. i_r (αντίστοιχα k_r) είναι οι δυαδικές παραστάσεις των i (αντίστοιχα k). Τα $m_r^{(-)}$ και $n_r^{(-)}$ περιλαμβάνουν επίσης μία αντιστροφή των bits, κατά τον ακόλουθο τρόπο

$$\begin{aligned} m_0^{(-)} &= m_{l-1} \\ m_1^{(-)} &= m_{l-1} + m_{l-2} \\ &\vdots \\ m_{l-1}^{(-)} &= m_1 + m_0 \end{aligned} \quad (8.23)$$

Ο WHT είναι διαχωρίσιμος και ορθομοναδιαίος, και υπάρχουν γρήγοροι αλγόριθμοι για τον υπολογισμό του.

Άσκηση

Ας είναι $r(m, n)$ η συνάρτηση συμμεταβλητότητας ενός 2-Δ σήματος $x(m, n)$. Ο βέλτιστος γραμμικός ορθομοναδιαίος $N \times N$ μετασχηματισμός δίδεται από τη λύση των παρακάτω εξισώσεων:

$$\sum_{j=0}^{N-1} \sum_{l=0}^{N-1} r(i-j, k-l)t(m, n; j, l) = \lambda(m, n)t(m, n; i, k), \quad \text{για } 0 \leq i, k, m, n < N$$

Αν η συνάρτηση συμμεταβλητότητας είναι διαχωρίσιμη

$$r(m, n) = r_1(m)r_2(n),$$

τότε και ο μετασχηματισμός είναι διαχωρίσιμος

$$t(m, n; i, k) = t_1(m, i)t_2(n, k)$$

και προκύπτει ως εξής:

$$\sum_{j=0}^{N-1} r_1(i-j)t_1(m, j) = \lambda_1(m)t_1(m, i), \quad \text{για } 0 \leq m, i < N$$

$$\sum_{l=0}^{N-1} r_2(k-l)t_2(n, l) = \lambda_2(n)t_2(n, k), \quad \text{για } 0 \leq n, k < N$$

Δίδεται η ακόλουθη συνάρτηση συμμεταβλητότητας

$$r(m, n) = \sigma^2 \rho^{|m|+|n|}, \quad \text{με } 0 < \rho < 1$$

1. Να ευρεθεί ο 2×2 μετασχηματισμός που ικανοποιεί τις παραπάνω σχέσεις.
2. Να συγκριθεί το αποτέλεσμα με τους μετασχηματισμούς Fourier και συννημιτόνου.

Κεφάλαιο 9

Βελτίωση εικόνων

Η έννοια της ποιότητας των εικόνων όντας πολύ συχνά υποκειμενική, η βελτίωση ορίζεται με κριτήριο το ζητούμενο στόχο, που μπορεί να σχετίζεται απλά με την εμφάνιση των εικόνων. Έτσι η βελτίωση αναφέρεται, είτε στον τονισμό κάποιων χαρακτηριστικών της εικόνας, όπως η συνολική φωτεινή αντίθεση, ή τα όρια των τμημάτων της εικόνας, είτε στον περιορισμό του θορύβου, που ενδεχόμενα επηρεάζει την εμφάνιση μιας εικόνας ή παρενοχλεί την εξαγωγή χρήσιμης πληροφορίας από την εικόνα.

9.1 Τονισμός της φωτεινής αντίθεσης

Αυτή η επεξεργασία αφορά εικόνες με χαμηλή φωτεινή αντίθεση. Πρόκειται για ένα τελεστή σημείου, που τροποποιεί την τιμή της φωτεινής έντασης σημειακά. Η αρχική τιμή $x(m, n)$ τροποποιείται σε

$$y(m, n) = g(x(m, n)). \quad (9.1)$$

Ο τελεστής αυτός υλοποιείται σε κβαντισμένα δεδομένα με τη βοήθεια ενός πίνακα αντιστοιχίας. Συνέπεια της τροποποίησης των τιμών της φωτεινής έντασης είναι η τροποποίηση του ιστογράμματος.

Ας υποθέσουμε ότι είναι επιθυμητή η επέκταση του ιστογράμματος για τις τιμές της φωτεινής έντασης που περιλαμβάνονται στο διάστημα $[x_{min}, x_{max}]$, στο μέγιστο δυνατό διάστημα τιμών $[0, L]$. Τότε η γραμμική αντιστοιχία των τιμών γίνεται με τον ακόλουθο τρόπο

$$y = \begin{cases} 0 & x < x_{min} \\ \left[L \frac{x - x_{min}}{x_{max} - x_{min}} + 0,5 \right] & x_{min} \leq x \leq x_{max} \\ L & x > x_{max} \end{cases} \quad (9.2)$$

Στα Σχήματα 9.1 και 9.2(b) δίδονται αποτελέσματα του τονισμού της φωτεινής αντίθεσης με την παραπάνω σχέση. Θα μπορούσε επίσης να ορισθεί μια μη-γραμμική αντιστοίχιση, όπως

$$y = \left\lfloor L \left(\frac{x - x_{min}}{x_{max} - x_{min}} \right)^\gamma + 0,5 \right\rfloor, \quad (9.3)$$

όπου γ σταθερά που λαμβάνεται μικρότερη του 1 για να δώσει πιο φωτεινές τιμές και μεγαλύτερη του 1 για να δώσει πιο σκοτεινές τιμές στην εικόνα, ανάλογα με τις ανάγκες για τονισμό της φωτεινής αντίθεσης. Στο Σχήμα 9.3 δίδεται μία εικόνα μαγνητικού συντονισμού και το αποτέλεσμα του τονισμού της φωτεινής αντίθεσης με $\gamma = 0,5$.

Είναι επίσης δυνατό η τροποποίηση του ιστογράμματος να προσαρμοσθεί στα δεδομένα της εικόνας, καθορίζοντας την επιθυμητή μορφή του τροποποιημένου ιστογράμματος. Αν επιδιώκεται



(a)

(b)

Σχήμα 9.1: Τονισμός της φωτεινής αντίθεσης.



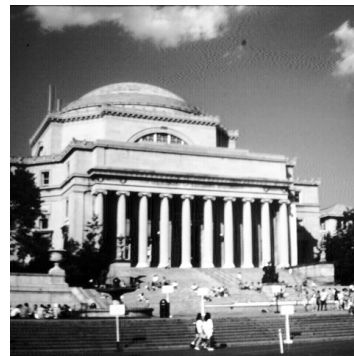
(a)



(b)

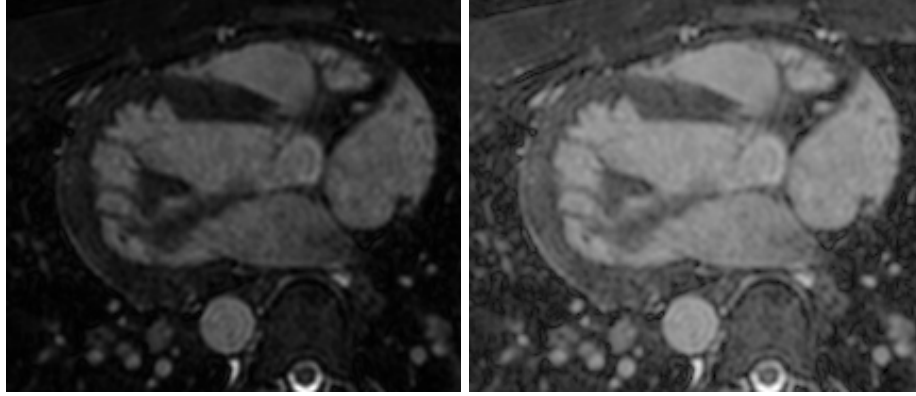


(c)



(d)

Σχήμα 9.2: Τονισμός της φωτεινής αντίθεσης (b) γραμμικά , (c) μη-γραμμικά ($\gamma = 0,7$) και (d) με εξισορρόπηση ιστογράμματος.



(a)

(b)

Σχήμα 9.3: Τονισμός της φωτεινής αντίθεσης ($\gamma = 0,5$).

στο τελικό ιστόγραμμα οι τιμές να είναι ομοιόμορφα κατανεμημένες, τότε αναφερόμαστε στην ισοστάθμιση του ιστογράμματος, που επιτυγχάνεται με τον ακόλουθο τρόπο

$$y = \left\lfloor L \frac{P(x) - P(x_{min})}{1 - P(x_{min})} + 0,5 \right\rfloor, \quad (9.4)$$

όπου $P(x) = \sum_{u \leq x} p(u)$, $p(\cdot)$ όντας το ιστόγραμμα της αρχικής εικόνας. Ωστόσο εικόνες με ισοσταθμισμένο ιστόγραμμα συχνότατα δεν φαίνονται ιδιαίτερα φυσικές, όπως για παράδειγμα στο Σχήμα 9.2(d). Γι' αυτό είναι προτιμότερη η τροποποίηση του ιστογράμματος ώστε να ακολουθεί μία κατανομή που να χρησιμοποιεί το διαθέσιμο εύρος τιμών και να είναι πυκνότερη περί τη μέση τιμή, όπως η κατανομή Gauss.

9.2 Μείωση θορύβου

Η μείωση του θορύβου μπορεί να επιτευχθεί χρησιμοποιώντας ένα γραμμικό φίλτρο με υψηλή απόκριση στις χαμηλές συχνότητες και αμετάβλητο τόσο σε μια μετάτωση, όσο και σε μια περιστροφή. Ας είναι $h(m, n)$ η κρουστική απόκριση του συστήματος, κι ας υποθέσουμε ότι τοπικά το σήμα είναι σταθερό, με τιμή μ , κι ότι η διασπορά του θορύβου είναι σ^2 . Για να μην υπάρξει μεταβολή της μέσης τιμής του σήματος απαιτείται όπως

$$\sum_{m=-M}^M \sum_{n=-N}^N h(m, n) = 1, \quad (9.5)$$

εφόσον το φίλτρο είναι πεπερασμένης κρουστικής απόκρισης και εκτείνεται συμμετρικά ως προς την αρχή. Αν υποθέσουμε ότι ο θόρυβος είναι λευκός, δηλαδή χωρικά ασυσχέτιστος, η διασπορά του θορύβου στην έξοδο του φίλτρου θα είναι

$$\sigma^2 \sum_{m=-M}^M \sum_{n=-N}^N h^2(m, n).$$

Αυτή η διασπορά είναι ελάχιστη, αν

$$h(m, n) = \frac{1}{(2M+1)(2N+1)}.$$

Πρόκειται επομένως για ένα φίλτρο μέσης τιμής με απόκριση

$$y(m, n) = \frac{1}{(2M+1)(2N+1)} \sum_{k=-M}^M \sum_{l=-N}^N x(m-k, n-l). \quad (9.6)$$

Η μέση τιμή ως γνωστόν ελαχιστοποιεί την τετραγωνική απόκλιση

$$\sum_{k=-M}^M \sum_{l=-N}^N (x(m-k, n-l) - y(m, n))^2.$$

Ο βαθμός μείωσης του θορύβου που επιτυγχάνεται είναι

$$\beta = \frac{1}{\sum_{m=-M}^M \sum_{n=-N}^N h^2(m, n)} = (2M+1)(2N+1). \quad (9.7)$$

Ωστόσο στις περιοχές όπου η υπόθεση σταθερού σήματος δεν ισχύει, το σήμα υφίσταται παραμόρφωση που είναι ιδιαίτερα αισθητή στις ακμές, δηλαδή στα όρια μεταξύ διαφόρων τμημάτων της εικόνας. Η δυσμενής συνέπεια της παραμόρφωσης μετριάζεται, αν οι συντελεστές του φίλτρου δεν είναι όλοι ίσοι. Αυτό ισοδυναμεί με την ελαχιστοποίηση μιας σταθμισμένης τετραγωνικής απόκλισης,

$$\sum_{k=-M}^M \sum_{l=-N}^N h(k, l)(x(m-k, n-l) - y(m, n))^2.$$

Παράδειγμα ενός τέτοιου φίλτρου, με πεπερασμένη κρουστική απόκριση, είναι το ακόλουθο

$$h(m, n) = \begin{cases} \frac{1}{4} & m^2 + n^2 = 0 \\ \frac{1}{8} & m^2 + n^2 = 1 \\ \frac{1}{16} & m^2 + n^2 = 2 \\ 0 & \text{άλλού} \end{cases} \quad (9.8)$$

Άλλο παράδειγμα, με άπειρη όμως κρουστική απόκριση, είναι το φίλτρο Gauss, που οφείλει το όνομά του στο γεγονός ότι η κρουστική απόκριση προέρχεται από την πυκνότητα πιθανότητας Gauss,

$$h(m, n) = \frac{1}{2\pi\alpha^2} \exp\left(-\frac{m^2 + n^2}{2\alpha^2}\right). \quad (9.9)$$

Στο Σχήμα 9.4 δίδεται το αποτέλεσμα της μείωσης του θορύβου με χρήση του φίλτρου Gauss, όπου $\alpha = \sqrt{2}$ προσαρμοσμένο στη διασπορά του θορύβου.

Επίσης θα μπορούσε να χρησιμοποιηθεί ένα σύστημα με κρουστική απόκριση

$$h(m, n) = \left(\frac{1-\alpha}{1+\alpha}\right)^2 \alpha^{|m|+|n|}, 0 < \alpha < 1. \quad (9.10)$$

Ο συντελεστής μείωσης θορύβου για το τελευταίο φίλτρο δίδεται στο Σχήμα 9.5. Βέβαια η παραμόρφωση των ακμών δεν αίρεται πλήρως. Καλύτερη εξισορρόπηση ανάμεσα στα αντιμαχόμενα αποτελέσματα της μείωσης του θορύβου και της διατήρησης αναλλοίωτου του σήματος επιτυγχάνεται με την προσαρμογή του φίλτρου στα δεδομένα της εικόνας.

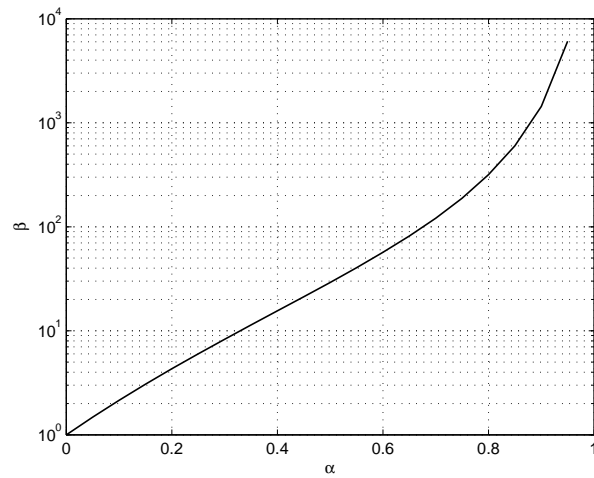
Η χρήση μη γραμμικών φίλτρων μπορεί επίσης να μειώσει το θόρυβο, με λιγότερες συνέπειες στο σήμα. Τέτοιο είναι το φίλτρο μεσαίας τιμής. Ας είναι \mathcal{D} ένα σύνολο σημείων με περιττό



(a)

(b)

Σχήμα 9.4: Μείωση του θορύβου με φίλτρο Gauss ($\alpha = \sqrt{2}$).



Σχήμα 9.5: Συντελεστής μείωσης θορύβου.

πληθικό αριθμό, $K = (2M + 1)(2N + 1)$, που είναι συμμετρικό ως προς το σημείο $(0,0)$, και περιλαμβάνει το σημείο $(0,0)$. Το φίλτρο μεσαίας τιμής δίνει την εξής απόκριση

$$y(m, n) = \text{Μεσαία}\{x(m - k, n - l) : (k, l) \in \mathcal{D}\}. \quad (9.11)$$

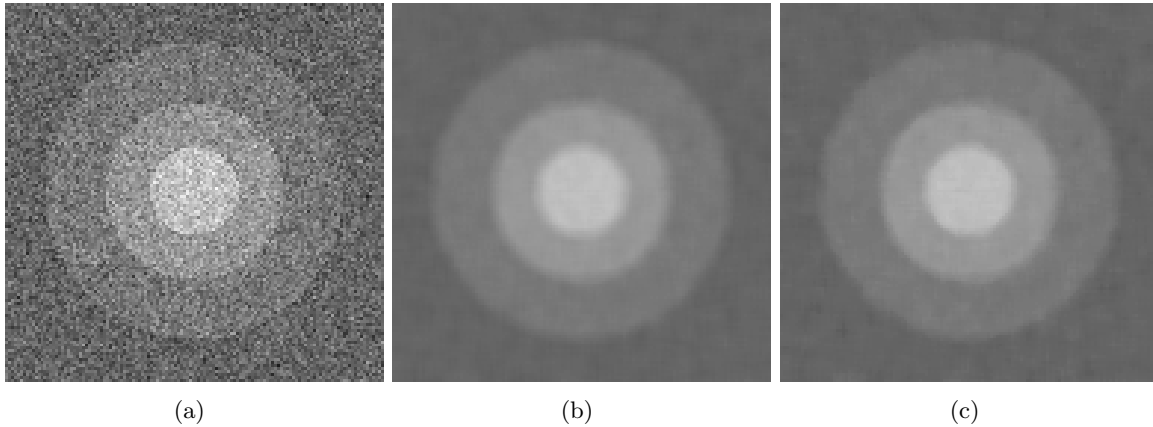
Η εύρεση της μεσαίας τιμής προκύπτει από τη διάταξη των τιμών του συνόλου $\{x(m - k, n - l) : (k, l) \in \mathcal{D}\}$. Η μεσαία τιμή έχει δείκτη στο διαταγμένο σύνολο $(K + 1)/2$, και ελαχιστοποιεί την κατ' απόλυτη τιμή απόκλιση

$$\sum_{k=-M}^M \sum_{l=-N}^N |y(m, n) - x(m - k, n - l)|.$$

Το φίλτρο μεσαίας τιμής έχει τις ακόλουθες ιδιότητες:

- $\text{Μεσαία}[\alpha x(m, n)] = \alpha y(m, n)$
- $\text{Μεσαία}[\alpha + x(m, n)] = \alpha + y(m, n)$
- $\text{Μεσαία}[x_1(m, n) + x_2(m, n)]$ όχι υποχρεωτικά ίση με $\text{Μεσαία}[x_1(m, n)] + \text{Μεσαία}[x_2(m, n)]$

Η τελευταία ιδιότητα δείχνει ότι το φίλτρο μεσαίας τιμής δεν είναι γραμμικό. Η μη γραμμικότητα επιτρέπει τη διατήρηση των ακμών της εικόνας με ταυτόχρονη μείωση του θορύβου. Στο



Σχήμα 9.6: Μείωση του θορύβου με φίλτρο μέσης τιμής και με φίλτρο μεσαίας τιμής.

Σχήμα 9.6 δίδεται μία συνθετική εικόνα και το αποτέλεσμα της μείωσης του θορύβου μ' ένα φίλτρο μέσης τιμής και μ' ένα φίλτρο μεσαίας τιμής και τα δύο διαστάσεων 11×11 .

Αν κάνουμε τις ίδιες υποθέσεις για το σήμα και το θόρυβο, όπως αυτές που αναφέρθηκαν στην περίπτωση των γραμμικών φίλτρων, κι αν επιπλέον υποθέσουμε ότι η συνάρτηση πυκνότητας πιθανότητας του θορύβου είναι συμμετρική, τότε η εμπειρική μεσαία τιμή είναι μια αμερόληπτη εκτιμήτρια της θεωρητικής μεσαίας τιμής, που επίσης ταυτίζεται με τη μέση τιμή. Αποδεικνύεται ότι η μείωση της διασποράς του θορύβου εξαρτάται από τη συνάρτηση πυκνότητας πιθανότητας του θορύβου. Αποδεικνύεται ότι ο συντελεστής μείωσης του θορύβου, β , θα είναι:

1. Αν ο θόρυβος κατανέμεται ομοιόμορφα, $\beta = \frac{K+2}{3}$
2. Αν ο θόρυβος ακολουθεί την κανονική κατανομή, $\beta \approx \frac{2(K-1)+\pi}{\pi}$

3. Αν ο θόρυβος ακολουθεί την κατανομή Laplace¹, $\beta \approx 2K - 1$

Ο μεγαλύτερος συντελεστής μείωσης θορύβου στην περίπτωση της κατανομής Laplace εξηγείται λόγω του γεγονότος ότι η μεσαία τιμή αποτελεί την εκτίμηση μέγιστης πιθανοφάνειας για τη θεωρητική μέση, και ταυτόχρονα μεσαία, τιμή της μεταβλητής. Το φίλτρο μεσαίας τιμής έχει επίσης πολύ καλές επιδόσεις στην περίπτωση του χροστικού θορύβου, που αλλοιώνει τις πραγματικές τιμές αθροιστικά και με ιδιαίτερα μεγάλες τιμές.

Ενδιαφέρον παρουσιάζει επίσης ο συνδυασμός των φίλτρων μέσης (ή σταθμισμένης μέσης) τιμής με το φίλτρο μεσαίας τιμής, ώστε να αξιοποιηθούν κατά το δυνατό οι επιθυμητές ιδιότητες και των δύο, χωρίς τα μειονεκτήματα. Τέτοιο είναι το τροποποιημένο φίλτρο απόρριψης, όπου η απόρριψη ορισμένων τιμών γίνεται με χρήση του φίλτρου μεσαίας τιμής, και η μέση τιμή, ενδεχόμενα σταθμισμένη, δίδει την απόκριση περιοριζόμενη μόνο στις εναπομένουσες τιμές,

$$y(m, n) = \frac{\sum_{k=-M}^M \sum_{l=-N}^N h(k, l) \phi(x(m-k, n-l) - \text{Μεσαία}[x(m, n)]) x(m-k, n-l)}{\sum_{k=-M}^M \sum_{l=-N}^N h(k, l) \phi(x(m-k, n-l) - \text{Μεσαία}[x(m, n)])} \quad (9.12)$$

όπου

$$\phi(x) = \begin{cases} 1 & |x| \leq \delta \\ 0 & |x| > \delta \end{cases}$$

Η τιμή του δ μπορεί να προσδιορισθεί με βάση τη διασπορά των τιμών, ή με βάση ένα ποσοστό απόρριψης.

9.3 Τονισμός ακμών

Ο τονισμός των ακμών μπορεί να γίνει είτε με κάποιο υψιπερατό φίλτρο, αν ο θόρυβος είναι χαμηλός, είτε με κάποιο ζωνοπερατό φίλτρο, αν ταυτόχρονα απαιτείται μείωση του θορύβου. Παράδειγμα υψιπερατού φίλτρου είναι η διακριτή υλοποίηση του λαπλασιανού (Laplacian) τελεστή

$$h(m, n) = \begin{cases} -4 & |m| + |n| = 0 \\ 1 & |m| + |n| = 1 \\ 0 & \text{αλλού} \end{cases} \quad (9.13)$$

Το φίλτρο με χροστική απόκριση $\delta(m, n) - \lambda h(m, n)$, όπου λ μία θετική σταθερά, δίδει στην έξοδο του μία εικόνα με τονισμένες ακμές. Ένα ζωνοπερατό φίλτρο μπορεί να υλοποιηθεί σαν μια διαφορά δύο βαθυπερατών φίλτρων. Συχνά ως βαθυπερατό φίλτρο χρησιμοποιείται το φίλτρο Gauss (Εξίσωση 9.9). Αρκεί να ληφθούν για τη διαφορά δύο διαφορετικές τιμές της παραμέτρου α . Όσο μεγαλύτερο είναι το α , τόσο πιο βαθυπερατό είναι το φίλτρο, δηλαδή τόσο πιο περιορισμένο είναι το εύρος των συχνοτήτων στην έξοδο του φίλτρου.

9.4 Μεγέθυνση και παρεμβολή

Πολλές φορές είναι επιθυμητή η μεγέθυνση ενός τμήματος μιας εικόνας. Η μεγέθυνση μπορεί να παρασταθεί με μία παρεμβολή μηδενικών στις θέσεις που λείπουν οι τιμές της εικόνας ακολουθούμενη από ένα βαθυπερατό φίλτρο. Αν το φίλτρο είναι διαχωρίσιμο, με 1- Δ χροστική απόκριση

¹Η συνάρτηση πυκνότητας πιθανότητας είναι

$$p(x) = \frac{1}{\sigma\sqrt{2}} \exp\left(-\frac{\sqrt{2}|x-\mu|}{\sigma}\right)$$

$h(\cdot)$, και ζητείται μεγέθυνση επί δύο για τις διαστάσεις της εικόνας, τότε το αποτέλεσμα της μεγέθυνσης είναι

$$y(m, n) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} h(m-2k)h(n-2l)x(k, l). \quad (9.14)$$

Η απόκριση στις συχνότητες είναι

$$Y(u, v) = H(u)H(v)X(2u, 2v). \quad (9.15)$$

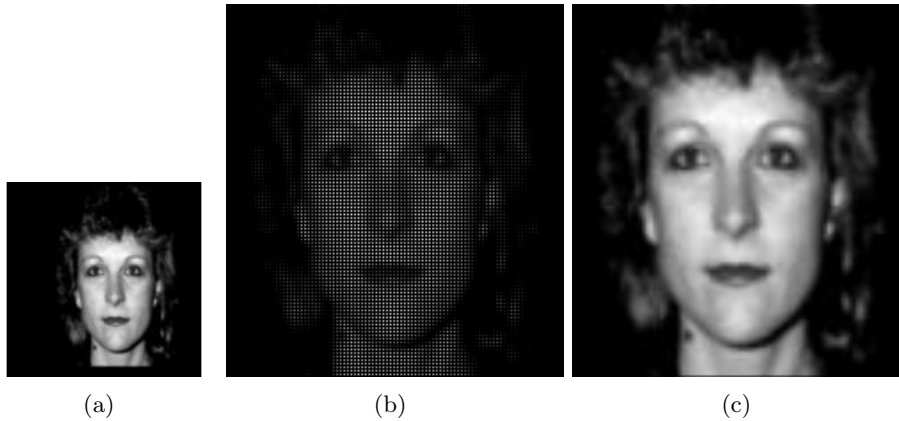
Αν $h(m) = 1$ για $m = 0$ ή $m = 1$, και $h(m) = 0$, αλλού, τότε η μεγέθυνση γίνεται με απλή αντιγραφή των τιμών της εικόνας. Διαφορετικά γίνεται παρεμβολή των τιμών της εικόνας στα σημεία όπου δεν διατίθενται τιμές. Κατάλληλο τέτοιο φίλτρο είναι αυτό της γραμμικής παρεμβολής

$$h(m, n) = \begin{cases} 1 & m^2 + n^2 = 0 \\ 0,5 & m^2 + n^2 = 1 \\ 0,25 & m^2 + n^2 = 2 \\ 0 & \text{αλλού} \end{cases} \quad (9.16)$$

Η σχέση εισόδου/εξόδου γίνεται

$$y(m, n) = \begin{cases} x(k, l) & m = 2k, n = 2l \\ \frac{1}{2}(x(k, l) + x(k+1, l)) & m = 2k+1, n = 2l \\ \frac{1}{2}(x(k, l) + x(k, l+1)) & m = 2k, n = 2l+1 \\ \frac{1}{4}(x(k, l) + x(k+1, l) + x(k, l+1) + x(k+1, l+1)) & m = 2k+1, n = 2l+1 \end{cases}$$

Στο Σχήμα 9.7 δίδεται το αποτέλεσμα της χρήσης του φίλτρου της εξίσωσης (9.16) για το διπλασιασμό του μεγέθους μιας εικόνας.



Σχήμα 9.7: Διπλασιασμός του μεγέθους επιλεγμένης περιοχής μιας εικόνας (c) μετά από παρεμβολή μηδενικών (b).

Παράρτημα: Διασπορά της μεσαίας τιμής

Η συνάρτηση πυκνότητας πιθανότητας της εκτιμήτριας της μεσαίας τιμής, M_k , είναι:

$$f_m(x) = \frac{(2k+1)!}{(k!)^2} F(x)^k (1-F(x))^k f(x),$$

όπου $f(x)$ είναι συνάρτηση πυκνότητας πιθανότητας της αρχικής μεταβλητής, $F(x)$ είναι η αθροιστική της πιθανότητα και $K = 2k + 1$ είναι ο αριθμός των δεδομένων.

Θεωρούμε για την αρχική μεταβλητή ομοιόμορφη κατανομή στο διάστημα $[-\frac{1}{2}, \frac{1}{2}]$. Ευρίσκουμε επομένως:

$$f_m(x) = \frac{(2k+1)!}{(k!)^2} \left(\frac{1}{4} - x^2\right)^k, |x| \leq \frac{1}{2}.$$

Λόγω συμμετρίας προκύπτει εύκολα ότι η προσδοκητή τιμή της εκτιμήτριας της μεσαίας τιμής είναι μηδέν.

Δεδομένου ότι η ολική πιθανότητα είναι 1 θα έχουμε:

$$\int_{-1/2}^{1/2} \frac{(2k+1)!}{(k!)^2} \left(\frac{1}{4} - x^2\right)^k dx = 1.$$

Οπότε η διασπορά της M_k θα είναι:

$$\begin{aligned} \text{var}\{M_k\} &= \int_{-1/2}^{1/2} \frac{(2k+1)!}{(k!)^2} x^2 \left(\frac{1}{4} - x^2\right)^k dx \\ &= \frac{(2k+1)!}{(k!)^2} \left(\frac{1}{4} \int_{-1/2}^{1/2} \left(\frac{1}{4} - x^2\right)^k dx - \int_{-1/2}^{1/2} \left(\frac{1}{4} - x^2\right)^{k+1} dx \right) \\ &= \frac{1}{4} - \frac{(2k+1)!}{(k!)^2} \frac{((k+1)!)^2}{(2k+3)!} = \frac{1}{4} - \frac{k+1}{2(2k+3)} = \frac{1}{4(2k+3)}. \end{aligned}$$

Αφού η διασπορά της αρχικής μεταβλητής είναι $\frac{1}{12}$, ο συντελεστής μείωσης του θορύβου με χρήση της μεσαίας τιμής θα είναι:

$$\beta = \frac{2k+3}{3} = \frac{K+2}{3}.$$

Ασκήσεις

1. Δίδεται μία εικόνα 8×8 σημείων ως ακολούθως.

0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1
1	1	1	2	2	2	3	4
1	1	1	2	2	2	3	4
1	1	1	2	2	3	3	5
1	1	1	2	2	3	3	5
1	1	2	2	2	3	4	6
1	1	2	2	2	3	4	7

Ζητείται η τροποποίηση των τιμών της εικόνας ώστε να εξισορροπηθεί το ιστόγραμμα στο διάστημα μεταξύ 0 και 7.

2. Θεωρείστε το γραμμικό φίλτρο του οποίου η απόκριση στο διδιάστατο $(2-\Delta)$ σήμα $x(m, n)$ είναι $y(m, n)$, ώστε

$$y(m, n) = \frac{1}{16} \sum_{k=-1}^1 \sum_{l=-1}^1 g(|k| + |l|) x(m - k, n - l),$$

όπου $g(0) = 4$, $g(1) = 2$ και $g(2) = 1$.

- (a) Είναι το παραπάνω φίλτρο αμετάβλητο σε μετατόπιση; Δώστε την κρουστική απόκριση του φίλτρου, και το μετασχηματισμό Fourier της κρουστικής απόκρισης.
 (b) Είναι το φίλτρο αυτό διαχωρίσιμο; Σε αριθμητική ακεραίων αριθμών, πως μπορεί να υλοποιηθεί το φίλτρο χωρίς πολλαπλασιασμούς; Πόσες προσθέσεις απαιτούνται ανά σημείο;

Δίδεται ότι η διασπορά του θορύβου στην έξοδο ενός φίλτρου $h(m, n)$ είναι

$$\sigma_h^2 = \sigma^2 \sum_{(m, n) \in Z^2} h^2(m, n),$$

όπου σ^2 είναι η διασπορά του θορύβου στην είσοδο του φίλτρου.

- (c) Ποιά είναι η διασπορά του θορύβου στην έξοδο του παραπάνω φίλτρου;
 (d) Ποιό ανεπιθύμητο αποτέλεσμα μπορεί να έχει η χρήση αυτού του φίλτρου για μείωση του θορύβου;

Για την τοπική προσαρμογή στις διαφορετικές περιοχές της εικόνας, θεωρείστε τη συνάρτηση $\phi(\psi) = \begin{cases} 0 & |\psi| > \delta \\ 1 & |\psi| \leq \delta \end{cases}$ και το φίλτρο με την ακόλουθη απόκριση

$$y(m, n) = \frac{\sum_{k=-1}^1 \sum_{l=-1}^1 \phi(x(m - k, n - l) - x(m, n)) g(|k| + |l|) x(m - k, n - l)}{\sum_{k=-1}^1 \sum_{l=-1}^1 \phi(x(m - k, n - l) - x(m, n)) g(|k| + |l|)}.$$

- (e) Είναι αυτό το φίλτρο γραμμικό;

3. Δίδονται τα ακόλουθα δισδιάστατα διακριτά σήματα $((m, n) \in \mathbb{Z}^2)$:

$$u(m, n) = \begin{cases} 1 & m \geq 0 \text{ και } n \geq 0 \\ 0 & \text{αλλιώς} \end{cases}$$

$$s_0(m, n) = \begin{cases} 1 & n \geq 0 \\ 0 & \text{αλλιώς} \end{cases}$$

$$s_1(m, n) = \begin{cases} 1 & m + n \geq 0 \\ 0 & \text{αλλιώς} \end{cases}$$

- (a) Να ευρεθεί η απόκριση στα παραπάνω σήματα του φίλτρου μεσαίας τιμής για τα ακόλουθα σύνολα σημείων: $\mathcal{D}_1 = \{(k, l) : k^2 + l^2 \leq 1\}$ και $\mathcal{D}_2 = \{(k, l) : k^2 + l^2 \leq 2\}$, και επίσης η απόκριση σ'ένα διαχωρίσιμο φίλτρο μεσαίας τιμής τριών σημείων για κάθε συντεταγμένη. Για ποιά φίλτρα τα παραπάνω σήματα παραμένουν αναλλοίωτα; Με παρουσία ανεξάρτητου αθροιστικού ομοιόμορφα κατανεμημένου θορύβου, ποιός είναι ο συντελεστής μείωσης του θορύβου για τις τρεις περιπτώσεις φίλτρου μεσαίας τιμής;
- (b) Να ευρεθεί η απόκριση στα παραπάνω σήματα των ακόλουθων δύο φίλτρων υλοποίησης του λαπλασιανού τελεστή:

$$h_1(m, n) = \begin{cases} -4 & m^2 + n^2 = 0 \\ 1 & m^2 + n^2 = 1 \\ 0 & m^2 + n^2 > 1 \end{cases}$$

$$h_2(m, n) = \begin{cases} -4 & m^2 + n^2 = 0 \\ 0,5 & m^2 + n^2 = 1 \text{ είτε } m^2 + n^2 = 2 \\ 0 & m^2 + n^2 > 2 \end{cases}$$

Συγκρίνετε τη συμπεριφορά των δύο φίλτρων σε γωνίες και την ευαισθησία τους στην κατεύθυνση μιας ευθείας. Στην περίπτωση ύπαρξης ανεξάρτητου αθροιστικού θορύβου, τα παραπάνω φίλτρα τον μειώνουν ή τον ενισχύουν, και με ποιό συντελεστή;

Κεφάλαιο 10

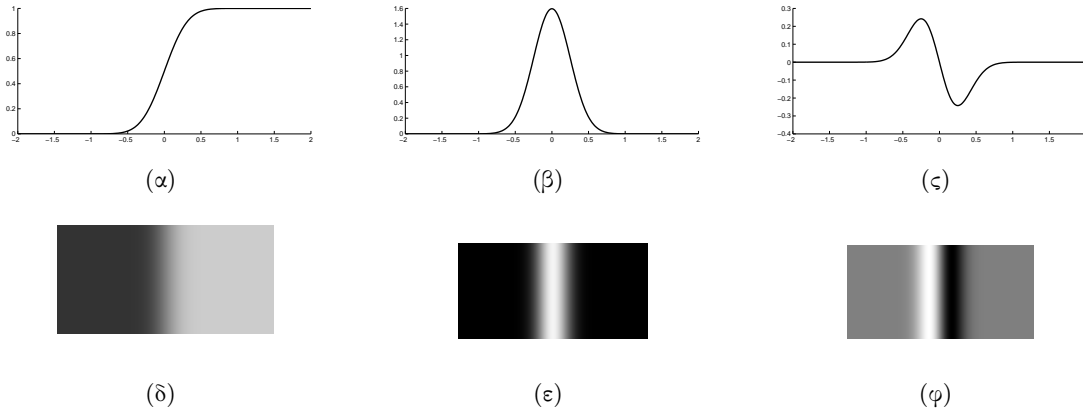
Τμηματοποίηση εικόνων

Η τμηματοποίηση των εικόνων είναι το πρώτο στάδιο στην ανάλυση των εικόνων. Η αναγνώριση αντικειμένων, η αναγνώριση γραπτών χαρακτήρων, η χαρτογράφηση, η ανάλυση μιας τρισδιάστατης σκηνής με τη στερεοσκοπία ή την ανάλυση της κίνησης βασίζονται κατά κανόνα στο αποτέλεσμα της τμηματοποίησης των εικόνων. Επομένως η ανάλυση των εικόνων θα είναι τόσο πιο ακριβής, όσο πιο επιτυχής είναι ο προσδιορισμός των τμημάτων της εικόνας, που αντιστοιχούν σε επιφάνειες ή τμήματα επιφανειών τρισδιάστατων αντικειμένων, ή ακόμα σε περιοχές με διαφορετική σύσταση στο εσωτερικό σωμάτων.

Δύο ισοδύναμοι και συμπληρωματικοί τρόποι προσδιορίζουν τα τμήματα μιας εικόνας: τα σύνορα και τα ομοιογενή χαρακτηριστικά του κάθε τμήματος. Σε αυτό το κεφάλαιο θα αναφερθούμε αποκλειστικά στην τμηματοποίηση με βάση τη φωτεινή ένταση. Επομένως τα σύνορα αντιστοιχούν σε ασυνέχειες της φωτεινής έντασης, ενώ τα τμήματα παρουσιάζουν ομοιογένεια της φωτεινής έντασης. Παρουσιάζονται στη συνέχεια τέσσερις προσεγγίσεις του προβλήματος της τμηματοποίησης των εικόνων. Στην ανίχνευση των ακμών αναζητούνται με τη χρήση αριθμητικής παραγωγίσις τα σημεία ασυνέχειας της φωτεινής έντασης. Στην ανάλυση του ιστογράμματος γίνεται κατάταξη των σημείων της εικόνας σε διαφορετικές κλάσεις ανάλογα με το διάστημα τιμών της φωτεινής έντασης. Η χρήση ενός μοντέλου Markov επιτρέπει να ληφθεί υπόψη η συνεκτικότητα των περιοχών μέσα από ένα κριτήριο στο σύνολο της εικόνας. Τέλος, η τμηματοποίηση μπορεί να γίνει χρησιμοποιώντας τοπικά ένα κριτήριο ομοιογένειας για την κατάτμηση μιας περιοχής ή τη συγχώνευση γειτονικών περιοχών.

10.1 Ανίχνευση ακμών

Ως σημείο ακμής ορίζεται εκείνο το σημείο όπου η μεταβολή της φωτεινής έντασης είναι τοπικά μέγιστη. Ένα σύνολο συνδεδεμένων σημείων ακμής αποτελεί ένα περίγραμμα. Σε κάθε σημείο ενός περιγράμματος μπορεί να ορισθεί ο προσανατολισμός του σε σχέση με το σύστημα συντεταγμένων του επιπέδου της εικόνας. Η φωτεινή ένταση παρουσιάζει τη μέγιστη μεταβολή στην κάθετη στο περίγραμμα κατεύθυνση. Με αυτή την έννοια η ανίχνευση ακμής ισοδυναμεί με τον εντοπισμό μιας μονοδιάστατης μεταβολής. Στο Σχήμα 10.1 δίδεται μια οριζόντια βαθμιαία μεταβολή και δείχνεται πώς η πρώτη ή η δεύτερη παράγωγος μπορούν να χρησιμοποιηθούν για την ανίχνευση και τον εντοπισμό της μεταβολής. Σε διδιάστατες εικόνες η ανίχνευση των σημείων ακμής μπορεί να γίνει με τη βοήθεια είτε του διανύσματος των δύο πρώτων μερικών παραγώγων της φωτεινής έντασης, που ονομάζεται κλίση, είτε του αθροίσματος των δύο δεύτερων μερικών παραγώγων της φωτεινής έντασης, που καλείται λαπλασιανός (Laplacian) τελεστής. Ο τελεστής αυτός είναι ισοτροπικός, δηλαδή η απόκρισή του είναι ανεξάρτητη από τον προσανατολισμό του



Σχήμα 10.1: Επάνω δίδεται η συνάρτηση μιας βαθμιαίας μεταβολής, η πρώτη και η δεύτερη παράγωγός της. Κάτω δίδεται μια εικόνα με οριζόντια βαθμιαία μεταβολή, η πρώτη οριζόντια παραγωγή και η δεύτερη οριζόντια παραγωγή.

περιγράμματος. Η ιστροπικότητα είναι σημαντική ιδιότητα για την ακρίβεια του εντοπισμού των περιγραμμάτων.

10.1.1 Μέθοδος της κλίσης

Το διάνυσμα της κλίσης της φωτεινής έντασης, $I(x, y)$, έχει σα συνιστώσες τις δύο πρώτες μερικές παραγώγους

$$\nabla I = \begin{bmatrix} I_x \\ I_y \end{bmatrix}$$

Για την ανίχνευση των ακμών χρησιμοποιείται το μέτρο του διανύσματος της κλίσης

$$\|\nabla I\| = \sqrt{I_x^2 + I_y^2}.$$

Αποδεικνύεται ότι το μέτρο του διανύσματος της κλίσης για συνεχή σήματα ικανοποιεί την ιδιότητα της ιστροπικότητας. Η κατεύθυνση του διανύσματος της κλίσης δίδεται από τη γωνία

$$\theta = \arctan \frac{I_y}{I_x}.$$

Η κλίση της φωτεινής έντασης σε μια δοσμένη κατεύθυνση, ϕ , είναι

$$I_\phi = I_x \cos \phi + I_y \sin \phi.$$

Η κλίση της φωτεινής έντασης είναι μέγιστη στην κατεύθυνση του διανύσματος της κλίσης, και μηδενική στην κάθετη σ' αυτήν κατεύθυνση. Σημεία ακμής ανιχνεύονται ακριβώς στην κατεύθυνση που η κλίση είναι μέγιστη, οπότε είναι ίση με το μέτρο του διανύσματος της κλίσης, $\|\nabla I\|$. Τελικά τα σημεία ακμής ανιχνεύονται ως τα τοπικά μέγιστα του μέτρου του διανύσματος της κλίσης, με τα τοπικά μέγιστα να ευρίσκονται στην κατεύθυνση αυτού του ίδιου διανύσματος.

Για την εφαρμογή της μεθόδου της κλίσης απαιτείται επομένως ο αριθμητικός υπολογισμός των πρώτων μερικών παραγώγων της φωτεινής έντασης. Εφόσον πρόκειται για μερική παραγωγή, το χρησιμοποιούμενο φίλτρο είναι μονοδιάστατο. Ωστόσο λόγω της ύπαρξης θορύβου, είναι αναγκαία επιπλέον η χρήση ενός βαθυπερατού φίλτρου στην άλλη κατεύθυνση, ως προς αυτή της

παραγωγίσιμης. Έτσι στην περίπτωση της οριζόντιας παραγωγίσιμης, πραγματοποιείται ολοκλήρωση στην κατακόρυφη κατεύθυνση, και αντίστροφα. Τα δύο φίλτρα χρησιμοποιούνται ξεχωριστά, ώστε το 2-Δ φίλτρο που προκύπτει να είναι διαχωρίσιμο. Ας είναι $g(\cdot)$ η χροστική απόκριση του φίλτρου παραγωγίσιμης, και $h(\cdot)$ η χροστική απόκριση του φίλτρου ολοκλήρωσης. Η ακολουθία $g(\cdot)$ είναι πάντοτε αντισυμμετρική

$$g(n) = -g(-n),$$

ενώ συνιστάται η ακολουθία $h(\cdot)$ να είναι συμμετρική

$$h(n) = h(-n).$$

Πολύ συχνά μεταξύ των δύο αυτών ακολουθιών υπάρχει ακριβώς η σχέση παραγωγίου και ολοκληρώματος

$$h(n) = \sum_{k=-\infty}^{n-1} g(k) + \frac{1}{2}g(n). \quad (10.1)$$

Παρουσιάζονται στη συνέχεια μερικά φίλτρα παραγωγίσιμης και ολοκλήρωσης με πεπερασμένη χροστική απόκριση. Το απλούστερο αντισυμμετρικό φίλτρο παραγωγίσιμης είναι το ακόλουθο

$$g(n) = \begin{cases} 1/2 & n = -1 \\ -1/2 & n = 1 \\ 0 & |n| \neq 1 \end{cases} \quad (10.2)$$

Το αντίστοιχο φίλτρο ολοκλήρωσης, με βάση την Εξίσωση (10.1), είναι

$$h(m) = \begin{cases} 1/4 & |m| = 1 \\ 1/2 & m = 0 \\ 0 & |m| > 1 \end{cases} \quad (10.3)$$

Το 2-Δ φίλτρο που προκύπτει είναι γνωστό ως το φίλτρο Sobel. Αποτελέσματα με χρήση αυτού του φίλτρου δίδονται στο Σχήμα 10.2.

Αν είναι

$$h(m) = \begin{cases} 1/3 & |m| \leq 1 \\ 0 & |m| > 1 \end{cases} \quad (10.4)$$

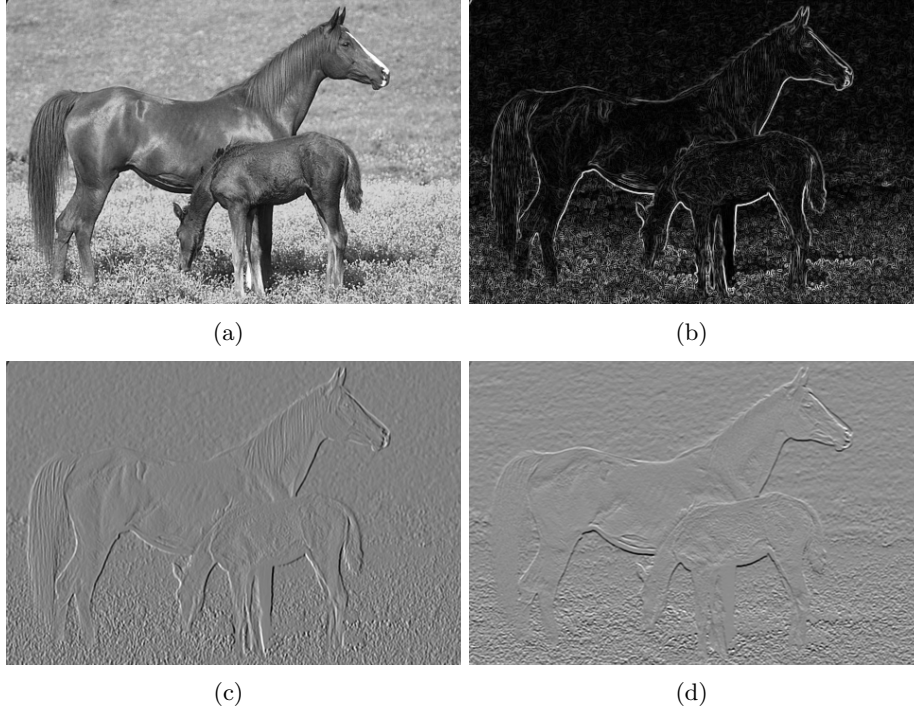
πρόκειται για το φίλτρο Prewitt. Αν λόγω θορύβου απαιτείται ένα φίλτρο ολοκλήρωσης με πιο τονισμένο το βαθυπερατό χαρακτήρα, αρκεί να πάρουμε τη συνέλιξη του (10.3) με τον εαυτό του, που δίνει το εξής αποτέλεσμα

$$h(m) = \begin{cases} 1/16 & |m| = 2 \\ 1/4 & |m| = 1 \\ 6/16 & m = 0 \\ 0 & |m| > 2 \end{cases}$$

Το φίλτρο παραγωγίσιμης που ικανοποιεί την Εξίσωση (10.1), είναι

$$g(n) = \begin{cases} 1/8 & n = -2 \\ 1/4 & n = -1 \\ -1/4 & n = 1 \\ -1/8 & n = 2 \\ 0 & n = 0 \text{ είτε } |n| > 2 \end{cases}$$

Το πρόβλημα της σχεδίασης του φίλτρου παραγωγίσιμης μπορεί να τεθεί σαν ένα πρόβλημα βελτιστοποίησης ορισμένων κριτηρίων που να εκφράζουν την καλή ανίχνευση ακμών με παρουσία



Σχήμα 10.2: Χρήση του φίλτρου Sobel για την ανίχνευση ακμών. Επάνω δίδεται η εικόνα και το μέτρο του διανύσματος της κλίσης. Κάτω αριστερά δίδεται η οριζόντια παραγωγή και κάτω δεξιά η κατακόρυφη.

θορύβου. Το πρόβλημα τέθηκε αρχικά από τον Canny και λύθηκε για ένα συνεχές μονοδιάστατο σήμα, όπου το μοντέλο της ασυνέχειας δίδεται από μια βηματική συνάρτηση, στην οποία προστίθεται λευκός θόρυβος με μηδενική μέση τιμή και διασπορά σ^2 . Ας υποθέσουμε ότι το σήμα έχει ως εξής

$$f(x) = \begin{cases} A & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (10.5)$$

Ζητείται το φίλτρο παραγωγής που ικανοποιεί τρία κριτήρια: καλής ανίχνευσης, μικρής τοπικής απόκλισης και μικρής παρενόχλησης από πολλαπλά μέγιστα. Η καλή ανίχνευση μετριέται με τη σηματοθορυβική σχέση στην έξοδο του φίλτρου στο σημείο της ασυνέχειας

$$\mathcal{R} = \frac{|\int_{-\infty}^{\infty} g(-x)f(x)dx|}{\sigma\sqrt{\int_{-\infty}^{\infty} g^2(x)dx}}. \quad (10.6)$$

Η σηματοθορυβική σχέση πρέπει να είναι η μέγιστη δυνατή. Η ασυνέχεια πρέπει επίσης να ανιχνεύεται όσο γίνεται πλησιέστερα στο σημείο της πραγματικής ασυνέχειας του σήματος. Έχει προταθεί το ακόλουθο κριτήριο καλού εντοπισμού

$$\mathcal{L} = \frac{|\int_{-\infty}^{\infty} f'(-x)g'(x)dx|}{\sigma\sqrt{\int_{-\infty}^{\infty} g'^2(x)dx}}. \quad (10.7)$$

Τέλος, αποδεικνύεται ότι η μέση απόσταση μεταξύ δύο μεγίστων στην έξοδο του φίλτρου είναι

$$\mathcal{D} = 2\pi \sqrt{\frac{\int_{-\infty}^{\infty} g'^2(x) dx}{\int_{-\infty}^{\infty} g''^2(x) dx}}. \quad (10.8)$$

Με δοσμένη την απαίτηση ικανοποίησης πολλαπλών κριτηρίων, έχει προταθεί ο ορισμός του βέλτιστου φίλτρου παραγωγίσις ως εκείνου που μεγιστοποιεί το γινόμενο \mathcal{RL} με μία συνθήκη για το \mathcal{D} . Κάτω από αυτούς τους όρους είναι δυνατή η λύση του προβλήματος βελτιστοποίησης για την περίπτωση του σήματος (10.5) και για πεπερασμένη ή άπειρη κρουστική απόκριση. Για την περίπτωση μιας πεπερασμένης, και οπωσδήποτε αντισυμμετρικής, κρουστικής απόκρισης, εκτεινόμενης στο διάστημα $[-M, M]$, ζητείται επομένως, για διάφορες τιμές του \mathcal{D}/M , η μεγιστοποίηση του

$$\frac{|g'(0)| \left| \int_{-M}^0 g(x) dx \right|}{\sqrt{\int_{-M}^M g^2(x) dx \int_{-M}^M g'^2(x) dx}}$$

Για χαμηλές τιμές του $\mathcal{D}/M (\approx 0, 1)$, η λύση που προκύπτει μπορεί να προσεγγισθεί από

$$g(x) = \begin{cases} 1/M^2 & -M \leq x < 0 \\ -1/M^2 & 0 < x \leq M \\ 0 & x = 0 \text{ είτε } |x| > M \end{cases} \quad (10.9)$$

Το αντίστοιχο φίλτρο ολοκλήρωσης είναι

$$h(x) = \begin{cases} \frac{M-|x|}{M^2} & |x| \leq M \\ 0 & |x| > M \end{cases} \quad (10.10)$$

Για υψηλές τιμές του $\mathcal{D}/M (\approx 1, 5)$, η λύση που προκύπτει μπορεί να προσεγγισθεί από την παράγωγο μιας συνάρτησης Gauss

$$g(x) = -\frac{x}{\sigma^3 \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}. \quad (10.11)$$

Προφανώς το φίλτρο αυτό έχει άπειρη κρουστική απόκριση. Αν το φίλτρο ζητηθεί κατευθείαν να έχει άπειρη κρουστική απόκριση, η λύση του προβλήματος βελτιστοποίησης δίδει

$$g(x) = -\frac{1}{4} \alpha^3 x e^{-\alpha|x|}. \quad (10.12)$$

Υψηλές τιμές του α αντιστοιχούν σε χαμηλές τιμές του \mathcal{D} , και αντίστροφα. Το αντίστοιχο φίλτρο ολοκλήρωσης είναι

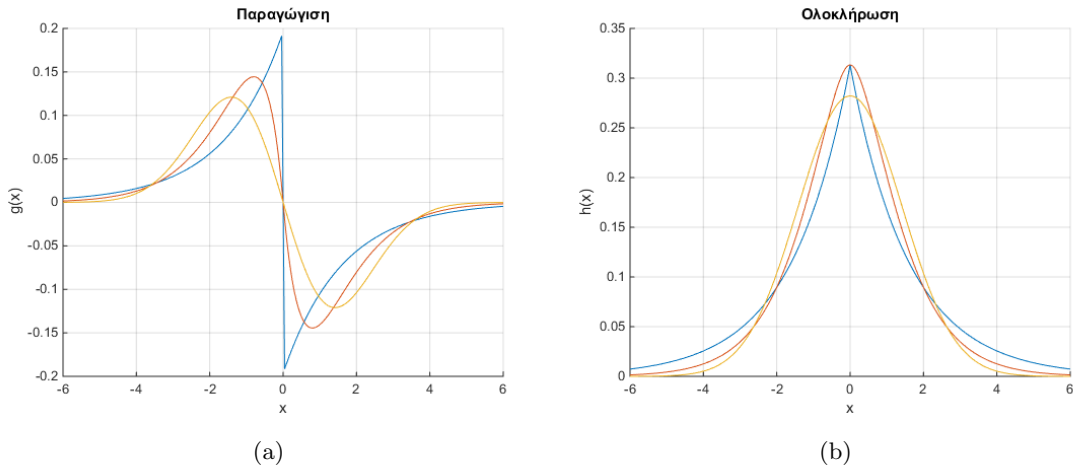
$$h(x) = \frac{1}{4} \alpha (1 + \alpha|x|) e^{-\alpha|x|}. \quad (10.13)$$

Άλλο χρησιμοποιούμενο ζεύγος φίλτρων με άπειρη κρουστική απόκριση είναι το ακόλουθο

$$g(x) = -\frac{\alpha^2}{2} \text{sign}(x) e^{-\alpha|x|} \quad \text{και} \quad h(x) = \frac{\alpha}{2} e^{-\alpha|x|} \quad (10.14)$$

με παρόμοιο ρόλο για το α . Όλα τα παραπάνω φίλτρα χρησιμοποιούνται φυσικά σε διακριτή μορφή, ανεξάρτητα αν είναι πεπερασμένες ή άπειρης κρουστικής απόκρισης. Για τα φίλτρα άπειρης κρουστικής απόκρισης υπάρχουν υλοποιήσεις μέσω αναδρομικών σχέσεων.

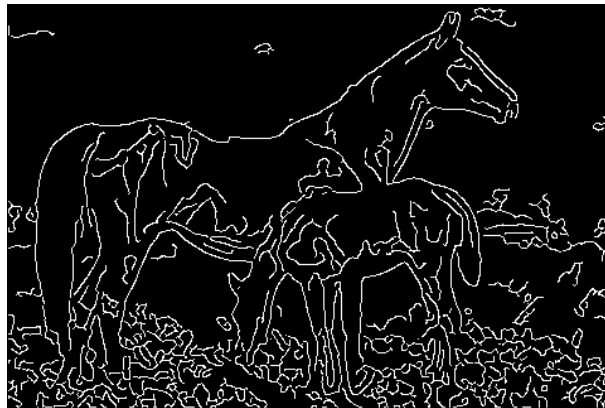
Στο Σχήμα 10.3(a) δίδονται τα παραπάνω φίλτρα παραγωγίσις, με τέτοιες τιμές για τις παραμέτρους, ώστε η σηματοθορυβική σχέση στην έξοδο του φίλτρου να είναι ίδια. Συγκεκριμένα θα



Σχήμα 10.3: Φίλτρα παραγωγίσις και ολοκλήρωσης.

είναι για την παράγωγο της συνάρτησης Gauss, $\sigma = \sqrt{2}$, για το φίλτρο της Εξίσωσης (10.12), που δίδεται με διακοπτόμενη γραμμή, $\alpha = \sqrt{\pi/2}$, και για το φίλτρο της Εξίσωσης (10.14), $\alpha = \frac{1}{2}\sqrt{\frac{\pi}{2}}$. Τα αντίστοιχα φίλτρα ολοκλήρωσης δίδονται στο Σχήμα 10.3(b).

Στο Σχήμα 10.4 δίδεται το αποτέλεσμα χρήσης της παραγωγού της συνάρτησης Gauss (φίλτρο Canny) με $\sigma = 1$ και με ανίχνευση που χρησιμοποιεί δύο κατώφλια και την τεχνική της υστέρησης. Με αυτή την τεχνική κατά την εφαρμογή του δεύτερου και μικρότερου κατωφλιού λαμβάνεται επίσης υπόψη η συνδεσιμότητα των σημείων ακμής.



Σχήμα 10.4: Αποτέλεσμα ανίχνευσης ακμών με τη μέθοδο του Canny.

10.1.2 Μέθοδος δευτέρων παραγώγων

Η διέλευση από το μηδέν των τιμών του αθροίσματος των δύο δευτέρων παραγώγων δίνει επίσης σημεία ακμών της φωτεινής έντασης. Ο τελεστής που δίνει αυτό το άθροισμα, ονομάζεται λαπλασιανός (Laplacian), και ορίζεται ως ακολούθως

$$\Delta I = I_{xx} + I_{yy}.$$

Αποδεικνύεται ότι ο λαπλασιανός τελεστής είναι ιστροπικός, αρκεί να είναι: $I_{xy} = I_{yx}$.

Απλά φίλτρα πεπερασμένης κρουστικής απόκρισης που χρησιμοποιούνται για την υλοποίηση του λαπλασιανού τελεστή είναι τα ακόλουθα:

$$g(m, n) = \begin{cases} -4 & m^2 + n^2 = 0 \\ 1 & m^2 + n^2 = 1 \\ 0 & m^2 + n^2 > 1 \end{cases} \quad (10.15)$$

και

$$g(m, n) = \begin{cases} -4 & m^2 + n^2 = 0 \\ 0,5 & m^2 + n^2 = 1 \text{ είτε } m^2 + n^2 = 2 \\ 0 & m^2 + n^2 > 2 \end{cases} \quad (10.16)$$

Ωστόσο αυτά τα φίλτρα χαρακτηρίζονται από ενίσχυση του θορύβου. Για τη μείωση του θορύβου προτείνεται η χρήση ενός βαθυπερατού φίλτρου Gauss, του οποίου η κρουστική απόκριση, σε συνεχείς συντεταγμένες είναι

$$h(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}.$$

Η όλη επεξεργασία είναι ισοδύναμη με τη συνέλιξη του σήματος της φωτεινής έντασης με το φίλτρο που προκύπτει από την εφαρμογή του λαπλασιανού τελεστή στην κρουστική απόκριση Gauss, που δίνει το ακόλουθο αποτέλεσμα

$$g(r) = \frac{1}{\pi\sigma^4} \left(\frac{r^2}{2\sigma^2} - 1 \right) e^{-\frac{r^2}{2\sigma^2}} \quad (10.17)$$

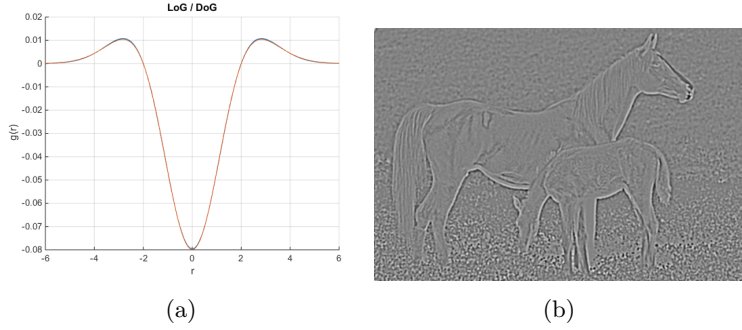
όπου $r = \sqrt{x^2 + y^2}$. Αποδεικνύεται επίσης ότι η παραπάνω κρουστική απόκριση μπορεί να προσεγγισθεί από μία διαφορά δύο συναρτήσεων Gauss,

$$g(r) \propto \frac{1}{2\pi\sigma_1^2\sigma^2} e^{-\frac{r^2}{2\sigma_1^2}} - \frac{1}{2\pi\sigma_2^2\sigma^2} e^{-\frac{r^2}{2\sigma_2^2}}, \quad (10.18)$$

όπου $\sigma_1^2 = 1,28\sigma^2$ και $\sigma_2^2 = 0,8\sigma^2$. Στο Σχήμα 10.5 δίδεται τόσο η λαπλασιανή της συνάρτησης Gauss, όσο και η προσέγγισή της με τη διαφορά δύο συναρτήσεων Gauss (διαφορετικά χρώματα). Επειδή η συνάρτηση Gauss δύο μεταβλητών είναι διαχωρίσιμη, η διαφορά δύο συναρτήσεων Gauss απαιτεί την υλοποίηση δύο διαχωρίσιμων φίλτρων. Η συνάρτηση Gauss μιας μεταβλητής προσεγγίζεται σε διακριτή μορφή, είτε από ένα φίλτρο πεπερασμένης κρουστικής απόκρισης, είτε από δύο αναδρομικά φίλτρα.

10.2 Ανάλυση ιστογράμματος

Με την ανάλυση του ιστογράμματος η τμηματοποίηση καθίσταται ένα πρόβλημα κατάταξης των σημείων της εικόνας σε κλάσεις. Από το ιστόγραμμα, μετά από τη λείανσή του, εφόσον είναι αναγκαία, προσδιορίζονται οι επικρατούσες τιμές της ολικής κατανομής. Η κατάταξη των σημείων της εικόνας γίνεται με βάση αυτές τις τιμές, ή ενδεχόμενα μέσω της κατανομής ανά κλάση που θα μπορούσε να υπολογισθεί. Η μέθοδος αυτή δεν εξασφαλίζει οπωσδήποτε τη συνεκτικότητα των αντικειμένων που εντοπίζονται, αφού η κατάταξη κάθε σημείου βασίζεται αποκλειστικά στη φωτεινή ένταση του δοσμένου μόνο σημείου. Η αξιοπιστία της μεθόδου εξαρτάται από την ομοιογένεια του φωτισμού, και από το μέγεθος των αντικειμένων, που επιπλέον ως προς το είδος θα πρέπει να είναι ολίγα τον αριθμό.



Σχήμα 10.5: Λαπλασιανή της συνάρτησης Gauss και αποτέλεσμα ανίχνευσης ακμών με εντοπισμό των διελεύσεων από το μηδέν.

Ο χωρισμός του ιστογράμματος σε διαστήματα μπορεί να βασισθεί σε κάποιο αρχικό μοντέλο για τη συνάρτηση πυκνότητας πιθανότητας της μεταβλητής που εκφράζει τη φωτεινή ένταση. Ας θεωρήσουμε την περίπτωση δύο επικρατούσων τιμών, ή ενός είδους αντικειμένου που ανιχνεύεται σε αντίθεση προς το βάθος της εικόνας. Ας υποθέσουμε ότι η τυχαία μεταβλητή του μοντέλου ακολουθεί και στις δύο περιπτώσεις την κανονική κατανομή, με την ίδια διασπορά, σ^2 , και μέση τιμή μ_0 για το βάθος της εικόνας, και μ_1 για το αντικείμενο. Με το κριτήριο της μέγιστης αληθοφάνειας το κατώφλι ανίχνευσης τοποθετείται στη μέση της απόστασης μεταξύ των θέσεων των δύο τοπικά μεγίστων τιμών του ιστογράμματος

$$\kappa = \frac{\mu_0 + \mu_1}{2}. \quad (10.19)$$

Αν επιπλέον θεωρήσουμε ότι είναι γνωστές οι a priori πιθανότητες του αντικειμένου (P_1) και συμπληρωματικά του βάθους της εικόνας ($P_0 = 1 - P_1$), τότε η μεγιστοποίηση της a posteriori πιθανότητας δίνει το ακόλουθο κατώφλι για την ανίχνευση του αντικειμένου

$$\kappa = \frac{\mu_0 + \mu_1}{2} + \frac{\sigma^2}{\mu_1 - \mu_0} \ln \frac{P_0}{P_1}. \quad (10.20)$$

Ο προσδιορισμός των παραμέτρων $\mu_0, \mu_1, \sigma^2, P_0, P_1$ μπορεί να γίνει με τη βοήθεια του ιστογράμματος.

Ο προσδιορισμός του κατωφλιού μπορεί επίσης να βασισθεί σ' ένα κριτήριο αυτόματης ομαδοποίησης. Αν μείνουμε στην περίπτωση των δύο επικρατούσων τιμών, ένα τέτοιο κριτήριο προς ελαχιστοποίηση είναι το ακόλουθο

$$\sum_{i=0}^{k-1} p_i (i - \mu_0)^2 + \sum_{i=k}^{N-1} p_i (i - \mu_1)^2,$$

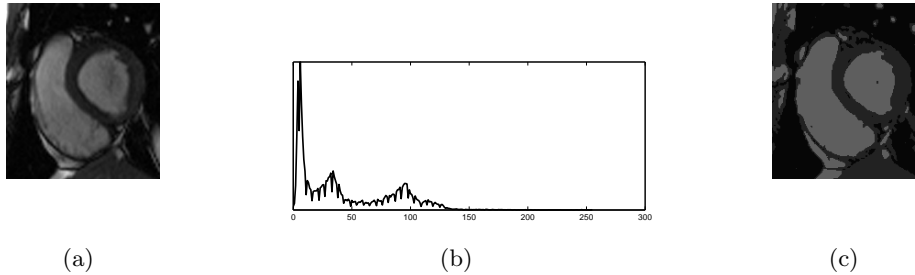
όπου p_i είναι η συχνότητα εμφάνισης, ή εμπειρική πιθανότητα, της τιμής i , για ένα σύνολο από N δυνατές τιμές. Ζητούνται οι τιμές του κατωφλιού k , και των παραμέτρων μ_0 και μ_1 που ελαχιστοποιούν το παραπάνω κριτήριο. Οι τιμές αυτές μπορούν να προσδιορισθούν χρησιμοποιώντας δύο αναγκαίες συνθήκες που ισχύουν στη θέση του ελάχιστου του κριτηρίου. Η πρώτη κατηγορία αναγκαίων συνθηκών προκύπτει για δοσμένο κατώφλι, και δίνει τις αντιπροσωπευτικές τιμές των δύο κλάσεων

$$\hat{\mu}_0 = \frac{\sum_{i=0}^{k-1} i p_i}{\sum_{i=0}^{k-1} p_i} \quad \text{και} \quad \hat{\mu}_1 = \frac{\sum_{i=k}^{N-1} i p_i}{\sum_{i=k}^{N-1} p_i} \quad (10.21)$$

Η δεύτερη αναγκαία συνθήκη δίδει το κατώφλι για δοσμένες αντιπροσωπευτικές τιμές

$$k = \lfloor \frac{\mu_0 + \mu_1}{2} + 1 \rfloor \quad (10.22)$$

Η διαδοχική χρήση των παραπάνω συνθηκών σ' ένα επαναληπτικό αλγόριθμο, επιτρέπει την ελαχιστοποίηση του κριτηρίου και τον προσδιορισμό του κατωφλιού που οδηγεί στην τμηματοποίηση της εικόνας. Η μέθοδος αυτή μπορεί να επεκταθεί για περισσότερες των δύο τελικές αποχρώσεις. Ένα παράδειγμα εφαρμογής της με τρεις αποχρώσεις δίδεται στο Σχήμα 10.6.



Σχήμα 10.6: Το ιστογράμμο της εικόνας αριστερά και το αποτέλεσμα της τμηματοποίησης με χρήση τριών αποχρώσεων.

Κάποια από τα μειονεκτήματα της ανάλυσης του ιστογράμματος, όπως η ύπαρξη μικρών στην έκταση αντικειμένων ή η ανομοιογένεια του φωτισμού, που αναφέρθηκαν εισαγωγικά, μπορούν να αντιμετωπισθούν με την εφαρμογή της μεθόδου τοπικά κατά τμήματα της εικόνας.

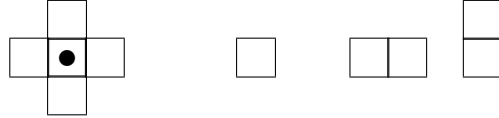
10.3 Ανάλυση ενός μοντέλου Markov ή Gibbs

Συμπληρωματικά με την απλή ανάλυση του ιστογράμματος, η ανάλυση με βάση ένα μοντέλο Markov για το πεδίο των αποφάσεων κατάταξης των σημείων σε κλάσεις, χρησιμοποιεί τις αλληλεξαρτήσεις που υπάρχουν τοπικά σε γειτονικά σημεία. Για κάθε μοντέλο Markov που δίνει τις τοπικές αλληλεξαρτήσεις ορίζεται ισοδύναμα μία κατανομή Gibbs που δίνει συνολικά όλες τις εξαρτήσεις του πεδίου των αποφάσεων. Ο ορισμός της κατανομής Gibbs δίδεται στη συνέχεια, αφού προηγουμένα δοθούν οι σχέσεις γειτνίασης που είναι αναγκαίες για την περιγραφή των εξαρτήσεων του πεδίου.

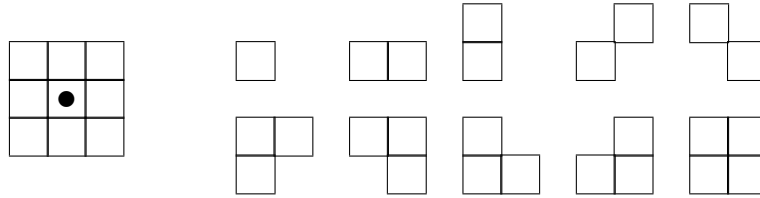
Ας είναι $L = \{(m, n) : 0 \leq m < M, 0 \leq n < N\}$ το σύνολο των σημείων της εικόνας. Η γειτονιά ενός σημείου (m, n) ορίζεται ως ένα υποσύνολο $\gamma(m, n)$ του L , που δεν περιλαμβάνει το σημείο (m, n) , με τον όρο ότι, εάν $(k, l) \in \gamma(m, n)$, τότε $(m, n) \in \gamma(k, l)$. Το σύστημα γειτνίασης στο L ορίζεται ως το σύνολο των γειτονιών όλων των σημείων του L . Με τη βοήθεια του ορισμού της γειτονιάς ορίζεται η παρέα, c , ως ένα σύνολο, είτε ενός μόνο σημείου, είτε περισσότερων του ενός σημείων, αρκεί για κάθε σημείο μιας παρέας τα υπόλοιπα σημεία της παρέας να ανήκουν στη γειτονιά του. Στο Σχήμα 10.7 δίδεται ο ορισμός μιας γειτονιάς 4 σημείων και οι αντίστοιχες σ' αυτή παρέες. Παρόμοια δίδονται οι παρέες για μια γειτονιά 8 σημείων στο Σχήμα 10.8.

Ένα πεδίο Gibbs περιγράφει συνολικά το πεδίο των αποφάσεων σε όλη την εικόνα

$$Pr\{X(m, n) = x(m, n) : (m, n) \in L\} = \frac{e^{-U(x)}}{\Psi}, \quad (10.23)$$



Σχήμα 10.7: Γειτονιά 4 σημείων και αντίστοιχες παρέες.



Σχήμα 10.8: Γειτονιά 8 σημείων και αντίστοιχες παρέες.

όπου $U(x)$ ονομάζεται συνάρτηση ενέργειας, και είναι συνάρτηση MN μεταβλητών, όσα δηλαδή είναι τα σημεία της εικόνας. Ψ είναι μία σταθερά, $\Psi = \sum e^{-U(x)}$. Ιδιαίτερα ενδιαφέρονσα είναι η περίπτωση ενός πεδίου Gibbs, όπου η συνάρτηση ενέργειας ορίζεται με τη βοήθεια του συνόλου των παρεών, C , των σημείων της εικόνας

$$U(x) = \sum_{c \in C} \phi_c(x). \quad (10.24)$$

Το $\phi_c(x)$ ονομάζεται δυναμικό της παρέας c . Συχνά για τη γειτονιά 8 σημείων μόνο ζεύγη σημείων θεωρούνται ότι έχουν μη μηδενικό δυναμικό. Ένα παράδειγμα δυναμικού για διμελείς παρέες που μπορεί να χρησιμοποιηθεί για την τμηματοποίηση των εικόνων ακολουθεί

$$\phi_c(x) = \begin{cases} 0 & \text{εάν τα δύο σημεία ανήκουν στην ίδια κλάση} \\ \zeta > 0 & \text{διαφορετικά} \end{cases} \quad (10.25)$$

Ας θεωρήσουμε τώρα το πρόβλημα της τμηματοποίησης της εικόνας κατατάσσοντας τα σημεία της σε K κλάσεις. Υποθέτουμε ότι η παρατηρούμενη εικόνα για κάθε σημείο εξαρτάται μόνο από την κλάση που ανήκει αυτό το σημείο, με μία παρενόχληση από αθροιστικό θόρυβο, που παρακάτω θα θεωρήσουμε ότι ακολουθεί την κανονική κατανομή με διασπορά σ^2

$$p(y(m, n) | X(k, l) = x(k, l), (k, l) \in L) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y(m, n) - f(x(m, n)))^2}{2\sigma^2}}. \quad (10.26)$$

Η συνάρτηση $f(\cdot)$ δίδει τη μέση αντιπροσωπευτική τιμή για κάθε κλάση S_i

$$f(x(m, n)) = \mu_i, \quad \text{εάν } (m, n) \in S_i, \quad 0 \leq i \leq K - 1. \quad (10.27)$$

Η τμηματοποίηση της εικόνας μπορεί να γίνει μεγιστοποιώντας την ύστερη πιθανότητα του πεδίου κατάταξης των σημείων σε κλάσεις

$$\begin{aligned} & Pr\{X(m, n) = x(m, n) | Y(m, n) = y(m, n); (m, n) \in L\} \\ & \propto Pr\{X(m, n) = x(m, n) : (m, n) \in L\} \prod_{(m, n) \in L} p(y(m, n) | X(m, n) = x(m, n)) \end{aligned} \quad (10.28)$$

Αυτό ισοδυναμεί με την ελαχιστοποίηση της ακόλουθης ποσότητας

$$\sum_{(m,n) \in L} \phi_c(x(m,n)) + \frac{1}{2\sigma^2} \sum_{i=0}^{K-1} \sum_{(m,n) \in S_i} (y(m,n) - \mu_i)^2.$$

Ωστόσο το κόστος εύρεσης της βέλτιστης λύσης με διεξοδική αναζήτηση είναι εξαιρετικά υψηλό, γιατί το πλήθος των δυνατών καταστάσεων είναι K^{MN} , ενώ υπάρχουν πολλά τοπικά ελάχιστα. Η ελαχιστοποίηση μπορεί να επιτευχθεί με μία πιθανοκρατική αναζήτηση του ελάχιστου προσομοιώνοντας το πεδίο Gibbs. Ωστόσο και σε αυτή την περίπτωση το υπολογιστικό κόστος είναι απαγορευτικό για πολλές εφαρμογές.

Θα πρέπει όμως να διακρίνουμε την περίπτωση όπου $K = 2$, γιατί τότε υπάρχει αλγόριθμος που ευρίσκει το ελάχιστο σε πολυωνυμικό χρόνο. Θεωρούμε ότι το πεδίο $x(m,n)$ λαμβάνει τιμές 0 με $f(0) = \mu_0$ και 1 με $f(1) = \mu_1$. Οπότε θέτοντας

$$\lambda(m,n) = \frac{1}{\sigma^2} ((y(m,n) - \mu_0)^2 - (y(m,n) - \mu_1)^2) = \frac{1}{\sigma^2} (2y(m,n) - \mu_0 - \mu_1)(\mu_1 - \mu_0), \quad (10.29)$$

μπορούμε να γράψουμε ότι η προς ελαχιστοποίηση ποσότητα ισούται με

$$- \sum_{(m,n) \in L} x(m,n)\lambda(m,n) + \zeta \sum_{\substack{(m,n) \in L \\ (k,l) \in \gamma(m,n)}} (x(m,n) - x(k,l))^2.$$

Αποδεικνύεται ότι το ανωτέρω πρόβλημα ελαχίστου ταυτίζεται με αυτό της εύρεσης της ελάχιστης τομής σε διμερή γράφο, και ισοδύναμα της επίλυσης του προβλήματος της μέγιστης ροής. Πράγματι η ανωτέρω ποσότητα διαφέρει κατά μία σταθερή ποσότητα από την ακόλουθη

$$\begin{aligned} C(x) &= \sum_{(m,n) \in L} x(m,n) \max(0, -\lambda(m,n)) + \sum_{(m,n) \in L} (1 - x(m,n)) \max(0, \lambda(m,n)) \\ &+ \zeta \sum_{\substack{(m,n) \in L \\ (k,l) \in \gamma(m,n)}} (x(m,n) - x(k,l))^2. \end{aligned} \quad (10.30)$$

Άρα μπορούμε να θεωρήσουμε ένα γράφο με κόμβους τα σημεία της εικόνας και επιπλέον ένα κόμβο 'πηγή' (s) και ένα κόμβο 'δεξαμενή' (t). Τα σύνολα $C_1 = \{s\} \cup \{(m,n) : x(m,n) = 1\}$ και $C_0 = \{t\} \cup \{(m,n) : x(m,n) = 0\}$ ορίζουν μια διαμέριση του συνόλου των κόμβων του γράφου που θεωρείται ότι είναι δίκτυο με το βάρος των ακμών να σημαίνουν τη χωρητικότητα της σύνδεσης. Ως τομή ορίζεται το σύνολο των ακμών που συνδέουν τους κόμβους των δύο μερών. Η χωρητικότητα της τομής προκύπτει επομένως από το βάρος των ακμών, ή αλλιώς τη χωρητικότητα των συνδέσεων. Ένα οποιοδήποτε σημείο της εικόνας είναι συνδεδεμένο είτε με την 'πηγή' είτε με τη 'δεξαμενή'. Εάν $\lambda(m,n) > 0$, η χωρητικότητα της σύνδεσης του σημείου (m,n) με την 'πηγή' είναι $\lambda(m,n)$. Ενώ εάν $\lambda(m,n) < 0$, η χωρητικότητα της σύνδεσης του σημείου (m,n) με τη 'δεξαμενή' είναι $-\lambda(m,n)$. Ζητούμενο τελικά είναι η διαμέριση του συνόλου των σημείων της εικόνας, ώστε η τομή να έχει την ελάχιστη χωρητικότητα $C(x)$.

Εναλλακτικά, μπορεί να χρησιμοποιηθεί, για οποιονδήποτε αριθμό κλάσεων K , η μέθοδος των επαναλαμβανόμενων επικρατουσών τιμών. Από μία αρχική κατάταξη, που συχνά είναι αυτή που προκύπτει από απλή ανάλυση του ιστογράμματος, γίνεται μία επανεκτίμηση, ώστε να μεγιστοποιείται η ακόλουθη πιθανότητα

$$Pr\{X(m,n) = x(m,n) | Y(m,n) = y(m,n), \hat{x}(k,l); (k,l) \in L - \{(m,n)\}\}.$$

Το αποτέλεσμα της μεγιστοποίησης είναι η επικρατούσα τιμή του πεδίου για δοσμένη παρατήρηση, και με δοσμένη την αρχική εκτίμηση. Αποδεικνύεται ότι αρκεί να ελαχιστοποιηθεί η ακόλουθη ποσότητα σε κάθε σημείο

$$\frac{1}{2\sigma^2}(y(m, n) - \mu_i)^2 + \sum \phi_c(x(m, n)),$$

όπου αθροίζονται μόνο τα δυναμικά των παρεών που περιλαμβάνουν το σημείο (m, n) , κι όπου οι τιμές του πεδίου για τα γειτονικά σημεία είναι αυτές της προηγούμενης εκτίμησης. Αν επομένως από τα 8 γειτονικά σημεία, κατά την προηγούμενη εκτίμηση κατατάχθηκαν $\alpha_i(m, n)$ στην κλάση i , τότε η τρέχουσα κατάταξη βασίζεται στην τοπική ελαχιστοποίηση του κριτηρίου

$$(y(m, n) - \mu_i)^2 - \sigma^2 \zeta \alpha_i(m, n).$$

Όταν ευρεθούν οι επικρατούσες τιμές σε όλα τα σημεία του πεδίου η διαδικασία επαναλαμβάνεται με τον ίδιο ακριβώς τρόπο, και σταματά όταν μειωθεί σημαντικά ο αριθμός των σημείων που από μία επανάληψη στην επόμενη αλλάζουν κλάση, ή ιδανικά όταν ουδένα σημείο αλλάζει κλάση. Ο αλγόριθμος επαναλαμβανόμενων τοπικά επικρατουσών τιμών δίδεται κατωτέρω σε ψευδοκώδικα.

Initialization:

```

for each pixel p
{
  current_label(p) := init_label(p)
  next_state(p) := active
}
change_pixels := number_of_pixels

```

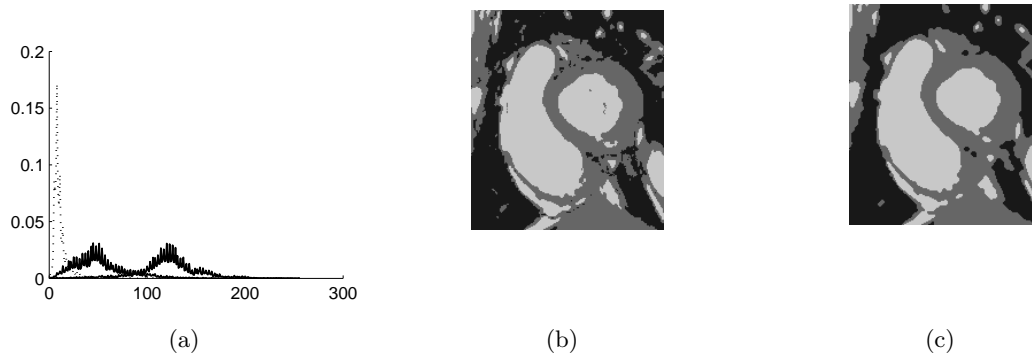
Iterations:

```

while (change_pixels != 0)
{
  change_pixels := 0
  for each pixel p
  {
    previous_label(p) := current_label(p)
    current_state(p) := next_state(p)
    next_state(p) := inactive
  }
  for each pixel p
  {
    if (current_state(p) == active)
    {
      current_label(p) := find_optimal_label(p)
      if (current_label(p) != previous_label(p))
      {
        change_pixels++
        next_state(neighbour(p)) := active
      }
    }
  }
}

```

Στο Σχήμα 10.9 δίδονται εκτιμήσεις των πυκνοτήτων πιθανότητας για τρεις κλάσεις που προέκυψαν από ανάλυση της εικόνας του Σχήματος 10.6, η αρχική κατάταξη με βάση τις πιθανοφάνειες των τριών κλάσεων και το αποτέλεσμα της τμηματοποίησης με τη μέθοδο των επαναλαμβανόμενων επικρατουσών τιμών.



Σχήμα 10.9: Τμηματοποίηση με τον αλγόριθμο επαναλαμβανόμενων επικρατουσών τιμών.

10.4 Τμηματοποίηση σε τοπολογικά συνεκτικές περιοχές

Σ' αυτή τη μέθοδο η συνεκτικότητα και η ομοιογένεια των περιοχών απαιτείται με άμεσο τρόπο. Ο αριθμός των κατηγοριών που κατατάσσονται τα σημεία της εικόνας θα μπορούσε να είναι άγνωστος και να προκύψει ως αποτέλεσμα της τμηματοποίησης, σε αντίθεση με τις προηγούμενες μεθόδους, όπου ο αριθμός αυτός ήταν δοσμένος.

Αν R είναι όλη η εικόνα σα σύνολο σημείων, ζητείται η τμηματοποίηση σε τοπολογικά συνεκτικές περιοχές R_1, R_2, \dots, R_K , ώστε

- $\bigcup_{k=1}^K R_k = R$
- $R_i \cap R_j = \emptyset, i \neq j$

Η τμηματοποίηση βασίζεται σ' ένα κριτήριο ομοιογένειας, που θα πρέπει να ικανοποιείται από όλα τα τμήματα της εικόνας, και να μην ικανοποιείται από την ένωση δύο γειτονικών τμημάτων, δηλαδή τμημάτων με κοινό σύνορο. Σαν κριτήριο ομοιογένειας μπορεί να χρησιμοποιηθεί η αντίθεση της φωτεινής έντασης σε μια περιοχή, μετρημένη είτε με τη μέγιστη αντίθεση, είτε με την τυπική απόκλιση της φωτεινής έντασης.

Με συσσωρεύσεις γύρω από κάποιο ή κάποια αφετηριακά σημεία μπορεί να υλοποιηθεί μια τμηματοποίηση που να ικανοποιεί τις παραπάνω απαιτήσεις. Δίδεται κατωτέρω ένας αλγόριθμος διαδοχικού προσεταιρισμού των σημείων της εικόνας με βάση ένα προκαθορισμένο κριτήριο ομοιότητας, $\rho(\cdot, \cdot)$, μέσω του οποίου προσδιορίζεται και η σειρά των κατατάξεων.

B1 Ονοματοθεσία των σημείων των αρχικών συνόλων.

B2 Εκτίμηση των χαρακτηριστικών των συνόλων (π.χ. μέση φωτεινότητα).

B3 Εισαγωγή στη διαταγμένη λίστα όλων των γειτονικών σημείων των αρχικών συνόλων.

B4 Όσο η λίστα δεν είναι κενή:

B4.1 Κατάταξη του πρώτου σημείου της λίστας α και αφαίρεσή του απ' αυτή.

B4.2 Έλεγχος των γειτονικών σημείων του α και ενημέρωση της λίστας:

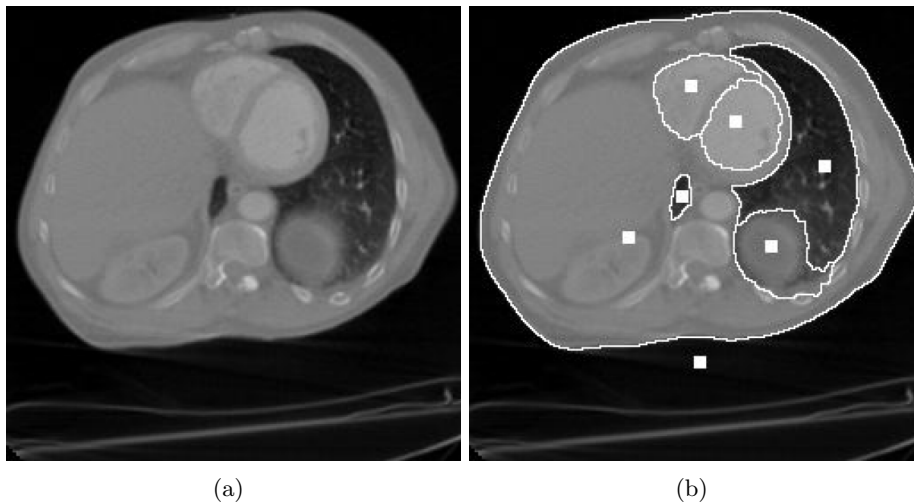
B4.2.1 Πρόσθεση των γειτόνων του α που

- δεν έχουν ήδη όνομα
- δεν είναι ήδη στη λίστα

σύμφωνα με την ομοιότητα $\rho(\cdot, \cdot)$.

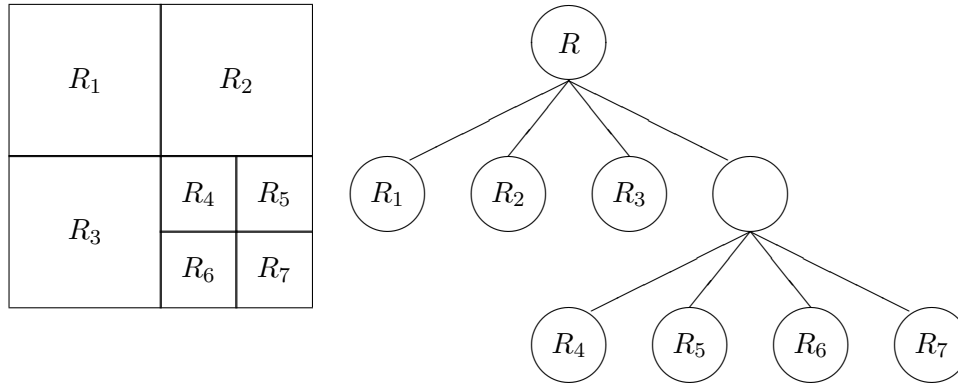
B4.2.2 Έλεγχος για γείτονες που είναι ήδη στη λίστα και που λόγω της κατάταξης του α συνορεύουν με ένα νέο σύνολο. Αυτά σημαδεύονται ως σημεία συνόρου. Επιπλέον, αν η ομοιότητα μειώνεται λόγω της νέας γειτνίασης, ανελλίσσονται αντίστοιχα στη λίστα.

Ένα παράδειγμα εφαρμογής αυτής της μεθόδου δίδεται στο Σχήμα 10.10, όπου ορίζονται αφετηριακά μπλοκ σημείων και τηρείται σειρά προτεραιότητας για τις συνενώσεις, που γίνονται με κριτήριο την κατ' απόλυτο τιμή απόκλιση της φωτεινότητας του δοσμένου σημείου από τη μέση τιμή της περιοχής στην οποία γειτονεύει.



Σχήμα 10.10: Τομογραφική εικόνα της καρδιάς και τμηματοποίησή της με συσσωρεύσεις γύρω από τα αφετηριακά μπλοκ σημείων που ορίστηκαν από το χρήστη.

Ωστόσο το αποτέλεσμα επηρεάζεται από την επιλογή των αφετηριακών σημείων και επίσης απαιτείται μεγάλος αριθμός ελέγχων για την υπόθεση της ομοιογένειας. Το αποτέλεσμα είναι πιο σταθερό, αν η τμηματοποίηση επιτυγχάνεται με ένα συνδυασμό διαιρέσεων ανομοιογενών τμημάτων και συγχωνεύσεων γειτονικών ομοιογενών τμημάτων. Η υλοποίηση ενός τέτοιου αλγορίθμου διευκολύνεται με τη χρήση μιας πυραμιδωτής δομής για την παράσταση των τμημάτων της εικόνας. Στην κορυφή της πυραμίδας βρίσκεται ολόκληρη η εικόνα, και στη βάση της τα σημεία της εικόνας. Ισοδύναμα η τμηματοποίηση μπορεί να παρασταθεί μ' ένα δένδρο, του οποίου η ρίζα είναι ολόκληρη η εικόνα, και τα φύλλα του τα σημεία της εικόνας. Πιο συχνά το δένδρο είναι



Σχήμα 10.11: Τμήματα της εικόνας και αντίστοιχο τετράκλωνο δένδρο.

τετράκλωνο, που αντιστοιχεί σε κατάτμηση κάθε τμήματος της εικόνας σε τέσσερα ίσα τμήματα. Στο Σχήμα 10.11 δίδεται ένα παράδειγμα τέτοιας παράστασης των τμημάτων της εικόνας. Το αποτέλεσμα είναι πιο ενδιαφέρον, και η πολυπλοκότητα περιορισμένη, αν η εκκίνηση του αλγορίθμου γίνει σ' ένα ενδιάμεσο επίπεδο της πυραμίδας.

Κεφάλαιο 11

Μαθηματική μορφολογία

11.1 Μορφολογική επεξεργασία δυαδικών εικόνων

Η μαθηματική μορφολογία μπορεί να χρησιμοποιηθεί για την ανάλυση δυαδικών εικόνων, που προκύπτουν μετά από τμηματοποίηση εικόνων. Σε αυτή την περίπτωση οι πράξεις και οι τελεστές της μαθηματικής μορφολογίας μετασχηματίζουν τη μορφή ενός αντικειμένου ή εξάγουν κάποια χαρακτηριστικά της μορφής του αντικειμένου. Στη συνέχεια θα αναφερθούμε συνοπτικά στους βασικούς μορφολογικούς τελεστές για δυαδικές εικόνες. Ένα αντικείμενο ορίζεται σαν ένα σύνολο σημείων, κι οι τελεστές που θα ορισθούν χρησιμοποιούν ένα αντικείμενο αντίστοιχο του επιδιωκόμενου στόχου, που ονομάζεται δομικό στοιχείο. Τελικά οι μορφολογικές πράξεις για δυαδικές εικόνες ορίζονται με τη βοήθεια των κλασικών πράξεων και σχέσεων μεταξύ συνόλων.

Δίδονται στη συνέχεια μερικοί από τους βασικότερους μορφολογικούς τελεστές.

- **Διάβρωση**

Το αποτέλεσμα της διάβρωσης του αντικειμένου X από το στοιχείο B είναι το σύνολο

$$X \ominus B = \{x : B_x \subset X\}, \quad (11.1)$$

όπου B_x είναι το στοιχείο B μετατοπισμένο στη θέση x . Σημειώνεται ότι ο ανωτέρω ορισμός και όσοι ακολουθούν είναι κατά την έννοια των πράξεων πρόσθεσης και αφαίρεσης συνόλων του Minkowski. Επικρατέστερος είναι ο ορισμός κατά έννοια αντίστοιχη της συνέλιξης, δηλαδή η ίδια πράξη με το συμμετρικό στοιχείο, που συμβολίζεται \check{B} .

Η διάβρωση είναι πράξη αναλλοίωτη στη μετατόπιση, είναι αύξουσα,

$$\text{αν } X \subset Y, \text{ τότε } X \ominus B \subset Y \ominus B$$

και συρρικνωτική,

$$\text{αν } (0,0) \in B, \text{ τότε } X \ominus B \subset X$$

Αν ένα αντικείμενο είναι μερικά γνωστό, τότε και το αποτέλεσμα της διάβρωσης είναι μερικά γνωστό

$$(X \cap Z) \ominus B = (X \ominus B) \cap (Z \ominus B).$$

Η διάβρωση διανέμεται στην ένωση δύο στοιχείων, κατά τον ακόλουθο τρόπο

$$X \ominus (B \cup B') = (X \ominus B) \cap (X \ominus B').$$

- **Διαστολή**

Το αποτέλεσμα της διαστολής του αντικειμένου X από το στοιχείο B είναι το σύνολο

$$X \oplus B = \{x : B_x \cap X \neq \emptyset\}. \quad (11.2)$$

Η διαστολή είναι πράξη αναλλοίωτη στη μετατόπιση, είναι αύξουσα,

$$\text{αν } X \subset Y, \text{ τότε } X \oplus B \subset Y \oplus B$$

και επεκτατική,

$$\text{αν } (0,0) \in B, \text{ τότε } X \subset X \oplus B.$$

Η διαστολή διατηρεί τις συνδέσεις μεταξύ σημείων ενός αντικειμένου. Η διαστολή διανέμεται στην ένωση δύο στοιχείων, κατά τον ακόλουθο τρόπο

$$X \oplus (B \cup B') = (X \oplus B) \cup (X \oplus B').$$

Η επανάληψη της διαστολής είναι ισοδύναμη με τη διαστολή από ένα στοιχείο που προκύπτει από το αποτέλεσμα της διαστολής των δύο στοιχείων,

$$(X \oplus B) \oplus B' = X \oplus (B \oplus B').$$

Το ίδιο στοιχείο προσδιορίζει και το αποτέλεσμα της επανάληψης της διάβρωσης

$$(X \ominus B) \ominus B' = X \ominus (B \oplus B').$$

Οι πράξεις της διάβρωσης και της διαστολής είναι μεταξύ τους συμπληρωματικές, με την έννοια ότι

$$X \oplus B = (X^c \ominus \check{B})^c,$$

όπου X^c είναι το συμπλήρωμα του συνόλου X , και \check{B} είναι το συμμετρικό του B ως προς το σημείο $(0,0)$.

Η διαστολή επιτρέπει την εύρεση των συνιστωσών ενός αντικειμένου, δηλαδή των υπο-συνόλων του αντικειμένου που έχουν την ιδιότητα σύνδεσης των σημείων τους. Αν x είναι ένα σημείο του αντικειμένου X , η επαναληπτική εφαρμογή της διαστολής δίδει όλα τα συνδεόμενα με το x σημεία του αντικειμένου,

$$X_{i+1} = (X_i \oplus H) \cap X, \quad X_1 = x$$

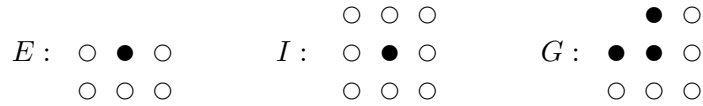
όπου το στοιχείο H εκφράζει την έννοια της σύνδεσης των σημείων (γειτονιά 4 ή 8 σημείων), και όπου οι επαναλήψεις σταματούν όταν επιτευχθεί σύγκλιση ($X_{i+1} = X_i$).

- **Εύρεση χαρακτηριστικών μορφών**

Αυτή η μορφολογική πράξη ορίζεται με τη βοήθεια δύο δομικών στοιχείων, όπου το ένα, B^1 , αφορά το αντικείμενο, και το άλλο, B^2 , το συμπλήρωμά του. Αρκεί η πράξη της διάβρωσης για τον ορισμό του τελεστή εύρεσης χαρακτηριστικών μορφών

$$X \circledast B = (X \ominus B^1) \cap (X^c \ominus B^2). \quad (11.3)$$

Στο Σχήμα 11.1 δίδονται τα σύνθετα δομικά στοιχεία: E , για την αναζήτηση τερματικών σημείων ενός αντικειμένου, I , για την αναζήτηση μεμονωμένων σημείων, και G , για την αναζήτηση των γωνιών ενός αντικειμένου.



Σχήμα 11.1: Δομικά στοιχεία εύρεσης χαρακτηριστικών μορφών

• **Άνοιγμα**

Το άνοιγμα συνίσταται στη διάβρωση ενός αντικειμένου από ένα δομικό στοιχείο, B , ακολουθούμενη από τη διαστολή από το συμμετρικό του B ,

$$X_B = (X \ominus B) \oplus \check{B}. \quad (11.4)$$

Το άνοιγμα είναι πράξη αναλλοίωτη στη μετατόπιση, αύξουσα, συρρικνωτική και αδύναμη, με την έννοια ότι η επανάληψη του ανοίγματος με το ίδιο στοιχείο δεν αλλάζει το αποτέλεσμα. Το άνοιγμα λειαίνει το περίγραμμα των αντικειμένων, κόβει στενούς ισθμούς και εξαφανίζει μικρά αντικείμενα.

• **Κλείσιμο**

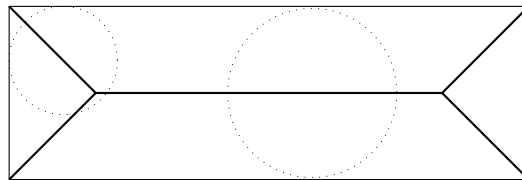
Το κλείσιμο ενός αντικειμένου συνίσταται στη διαστολή του από ένα δομικό στοιχείο, B , ακολουθούμενη από τη διάβρωση από το συμμετρικό του B ,

$$X^B = (X \oplus B) \ominus \check{B}. \quad (11.5)$$

Το κλείσιμο είναι πράξη αναλλοίωτη στη μετατόπιση, αύξουσα, επεκτατική και αδύναμη. Το κλείσιμο εξαφανίζει μικρές οπές ενός αντικειμένου και φράσσει στενά κανάλια και λεπτούς κόλπους.

• **Σκελετός**

Ο σκελετός ενός αντικειμένου μπορεί να ορισθεί ως το σύνολο των κέντρων όλων των κύκλων που εγγράφονται στο αντικείμενο και εφάπτονται σε δύο τουλάχιστον σημεία του περιγράμματός του (Σχήμα 11.2). Ο σκελετός μπορεί να υπολογισθεί χρησιμοποιώντας τη



Σχήμα 11.2: Ο σκελετός ενός ορθογώνιου σχήματος

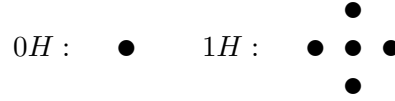
διάβρωση και το άνοιγμα

$$S(X) = \bigcup_{n=0}^{\infty} s_n(X) = \bigcup_{n=0}^{\infty} ((X \ominus nH) - (X \ominus nH)_H). \quad (11.6)$$

Το δομικό στοιχείο H αποτελεί διακριτή παράσταση ενός μοναδιαίου δίσκου, ενώ ο δίσκος με μηδενική ακτίνα είναι το ουδέτερο στοιχείο της πράξης της διάβρωσης, και δίσκοι με μεγαλύτερη ακτίνα μπορούν να ορισθούν προοδευτικά (Σχήμα 11.3). Με βάση τον παραπάνω ορισμό ο σκελετός δεν είναι κατ' ανάγκη ομοιοτοπικός του αντικειμένου, με την έννοια ότι

δεν έχει τον ίδιο αριθμό τμημάτων και οπών όπως το αντικείμενο. Το αντικείμενο μπορεί να αποκατασταθεί από τα ενδιάμεσα αποτελέσματα της σκελετοποίησης με διαστολή

$$X = \bigcup_{n=0}^{\infty} (s_n(X) \oplus nH). \quad (11.7)$$



Σχήμα 11.3: Δομικά στοιχεία παράστασης δίσκων ακτίνας 0 και 1

- **Λέπτυνση**

Η λέπτυνση συνίσταται στην επαναληπτική αφαίρεση χαρακτηριστικών του αντικειμένου

$$X \ominus B = X - X \otimes B. \quad (11.8)$$

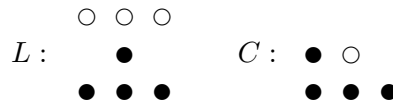
Αν B είναι δομικό στοιχείο 3×3 , που ορίζει εσωτερικά σημεία με συνδέσεις 8 σημείων, τότε η λέπτυνση δίνει το σύνορο του αντικειμένου, με συνδέσεις γειτονιάς 4 σημείων. Ενώ το στοιχείο $B = 1H$ (Σχήμα 11.3) δίνει το σύνορο με συνδέσεις γειτονιάς 8 σημείων. Το στοιχείο $B = L$ (Σχήμα 11.4), μαζί με τα στοιχεία που προκύπτουν με περιστροφή απ' αυτό, δίνει ένα σκελετό ομοιοτοπικό του αντικειμένου.

- **Πάχυνση**

Συμπληρωματική πράξη της λέπτυνσης είναι η πάχυνση

$$X \odot B = X \cup (X \otimes B). \quad (11.9)$$

Το στοιχείο $B = C$ (Σχήμα 11.4), μαζί με τα στοιχεία που προκύπτουν με περιστροφή απ' αυτό, δίνει ένα κυρτό περίγραμμα του αντικειμένου.



Σχήμα 11.4: Δομικά στοιχεία ομοιοτοπικού σκελετού και κυρτού περιγράμματος

11.2 Μορφολογική επεξεργασία εικόνων φωτεινότητας

Στις εικόνες φωτεινότητας οι μορφολογικές πράξεις συνιστούν μη γραμμικά φίλτρα επεξεργασίας. Σε αυτή την περίπτωση το δομικό στοιχείο, $b(m, n)$, είναι μια υποεικόνα, μικρής συνήθως έκτασης, που προσδιορίζεται επί ενός συνόλου σημείων D_b . Για παράδειγμα μπορεί να είναι

$$b(m, n) = 1, D_b = \{(m, n) : m^2 + n^2 \leq 8\}.$$

Η τιμή θα μπορούσε να είναι και $b(m, n) = 0$ για όλα τα σημεία του εκάστοτε D_b .

- **Διάβρωση**

Εάν $f(m, n)$ είναι οι τιμές της εικόνας η διάβρωση με το δομικό στοιχείο $b(m, n)$ δίδεται μέσω μιας ελάχιστης τιμής ως εξής

$$(f \ominus b)(m, n) = \min\{f(m+k, n+l) - b(k, l) | (k, l) \in D_b\}, \quad (11.10)$$

για όσα σημεία μπορεί να ορισθεί η τιμή $f(m+k, n+l)$. Οπότε αν το στοιχείο $b(m, n)$ λαμβάνει τη μηδενική τιμή η διάβρωση είναι απλά η ελάχιστη τιμή της αρχικής εικόνας σε μια γειτονιά που ορίζεται από το σύνολο D_b . Αν όλες οι τιμές του δομικού στοιχείου είναι μη αρνητικές, τότε η εικόνα γίνεται πιο σκούρα. Επίσης φωτεινά στίγματα, ανάλογα με το μέγεθός τους, περιορίζονται.

- **Διαστολή**

Η διαστολή δίδεται μέσω της μέγιστης τιμής ως εξής

$$(f \oplus b)(m, n) = \max\{f(m+k, n+l) + b(k, l) | (k, l) \in D_b\}, \quad (11.11)$$

για όσα σημεία μπορεί να ορισθεί η τιμή $f(m+k, n+l)$. Σε αντίθεση με τη διάβρωση, αν όλες οι τιμές του δομικού στοιχείου είναι μη αρνητικές, τότε η εικόνα γίνεται πιο φωτεινή. Επίσης σκοτεινά στίγματα, ανάλογα με το μέγεθός τους, περιορίζονται. Η διαφορά διαστολής και διάβρωσης δίδει τη μορφολογική κλίση που εντοπίζει τις ακμές,

$$g = (f \oplus b) - (f \ominus b). \quad (11.12)$$

Στην απλούστερη περίπτωση θα είναι

$$g(m, n) = \max\{f(m+k, n+l) | k^2 + l^2 \leq 2\} - \min\{f(m+k, n+l) | k^2 + l^2 \leq 2\}. \quad (11.13)$$

Μπορεί επίσης να ορισθεί ο Laplacian τελεστής ως ακολούθως

$$g(m, n) = \max\{f(m+k, n+l) | k^2 + l^2 \leq 2\} + \min\{f(m+k, n+l) | k^2 + l^2 \leq 2\} - 2f(m, n). \quad (11.14)$$

- **Άνοιγμα και κλείσιμο**

Το άνοιγμα ορίζεται κατά τρόπο αντίστοιχο με την περίπτωση των δυαδικών εικόνων με χρήση ανοίγματος και κλεισίματος,

$$f_b = (f \ominus b) \oplus \check{b}, \quad (11.15)$$

όπου το στοιχείο \check{b} αντιστοιχεί στο συμμετρικό του b ως προς την αρχή $(0, 0)$. Παρόμοια ορίζεται το κλείσιμο

$$f^b = (f \oplus b) \ominus \check{b}. \quad (11.16)$$

Το άνοιγμα λειαίνει τις κορυφές της φωτεινότητας, ενώ το κλείσιμο λειαίνει τις *χαράδρες*.

Η διαδοχική χρήση του ανοίγματος και του κλεισίματος σε μια εικόνα οδηγεί σε εξομάλυνση των τιμών της εικόνας παρόμοια με αυτή που δίδει το φίλτρο μεσαίας τιμής. Ακριβέστερα ένα φίλτρο μεσαίας τιμής σε ένα δίσκο με ακτίνα διπλάσια αυτής του δομικού στοιχείου των μορφολογικών πράξεων δίδει παραπλήσιο αποτέλεσμα.

Κεφάλαιο 12

Αποκατάσταση εικόνων

Πολλές φορές οι εικόνες κατά τη λήψη υφίστανται παραμόρφωση. Η επεξεργασία που σκοπεύει στον περιορισμό της παραμόρφωσης ονομάζεται αποκατάσταση. Πιθανές αιτίες παραμόρφωσης είναι η κίνηση κατά τη διάρκεια της λήψης της εικόνας, η κακή εστίαση, ή, στην περίπτωση δορυφορικών εικόνων, οι ατμοσφαιρικές διαταραχές.

Συχνά σε μοντέλο της παραμόρφωσης της αρχικής ιδανικής εικόνας, $f(x, y)$, χρησιμοποιείται ένα γραμμικό φίλτρο $h(x, y)$, στην έξοδο του οποίου προστίθεται θόρυβος

$$g(x, y) = \iint h(x - x', y - y') f(x', y') dx' dy' + w(x, y) \quad (12.1)$$

με $\iint h(x, y) dx dy = 1$, και θεωρώντας ότι η μέση τιμή του θορύβου, $w(x, y)$, είναι μηδέν και η διασπορά του σ^2 . Η χροστική απόκριση του φίλτρου ονομάζεται και συνάρτηση διασποράς σημείου. Παραδείγματα συναρτήσεων διασποράς σημείου δίδονται στη συνέχεια:

- Μετακίνηση στην κατεύθυνση θ κατά L

$$h(x, y) = \frac{1}{L} \Pi_1\left(\frac{x \cos \theta + y \sin \theta}{L}\right) \delta(-x \sin \theta + y \cos \theta)$$

με μετασχηματισμό Fourier,

$$H(u, v) = \text{sinc}(\pi L(u \cos \theta + v \sin \theta))$$

Η γραφική παράσταση της συνάρτησης sinc δίδεται στο Σχήμα 4.3.

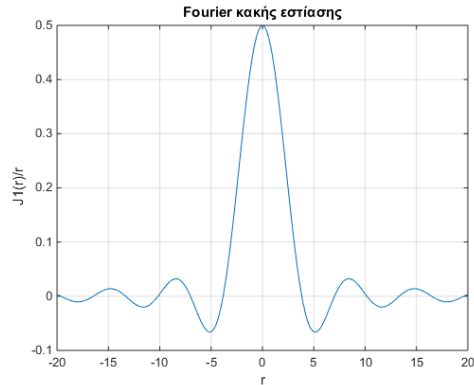
- Κακή εστίαση με ακτίνα r

$$h(x, y) = \frac{1}{\pi r^2}, \quad x^2 + y^2 \leq r^2$$

με μετασχηματισμό Fourier,

$$H(u, v) = \frac{\mathcal{J}_1(r\rho)}{r\rho}, \quad \rho = \sqrt{u^2 + v^2}.$$

όπου $\mathcal{J}_1(\cdot)$ είναι μία συνάρτηση Bessel. Η $\frac{\mathcal{J}_1(r\rho)}{r\rho}$ δίδεται στο Σχήμα 12.1. Μια εικόνα με κακή εστίαση κατά τη λήψη δίδεται στο Σχήμα 12.2.



Σχήμα 12.1: Μετασχηματισμός Fourier της απόκρισης στην κακή εστίαση



Σχήμα 12.2: Κακά εστιασμένη εικόνα

- Ατμοσφαιρικές διαταραχές διασποράς α^2

$$h(x, y) = \frac{1}{2\pi\alpha^2} e^{-\frac{x^2+y^2}{2\alpha^2}}$$

με μετασχηματισμό Fourier,

$$H(u, v) = e^{-2\pi^2\alpha^2(u^2+v^2)}.$$

Σε διακριτή μορφή η εξίσωση που δίνει την παρατηρούμενη εικόνα είναι η εξής

$$g(m, n) = \sum_{m'} \sum_{n'} h(m - m', n - n') f(m', n') + w(m, n), \quad (12.2)$$

που μπορεί να γραφεί, θεωρώντας τα διανύσματα f και g που κατασκευάζονται από τις τιμές των $f(.,.)$ και $g(.,.)$ αντίστοιχα,

$$g = Hf + w, \quad (12.3)$$

όπου ο πίνακας H κατασκευάζεται από τα δείγματα της συνάρτησης διασποράς σημείου $h(.,.)$. Η αποκατάσταση συνίσταται στην αντιστροφή του γνωστού τελεστή H της παραπάνω σχέσης για την εύρεση του f από το παρατηρούμενο g . Ο πίνακας H όπως έχει ορισθεί ανωτέρω είναι μπλοκ

Toeplitz. Αυτό σημαίνει ότι συνίσταται από πίνακες Toeplitz με την ιδιότητα της εξάρτησης των στοιχείων τους μόνο από τη διαφορά των δεικτών. Χωρίς καμία αλλοίωση των παραπάνω εξισώσεων, και επεκτείνοντας κατάλληλα τους πίνακες με μηδενικά, προκύπτουν πίνακες κυκλικοί, ενώ ο πίνακας H μετατρέπεται σε μπλοκ κυκλικό. Η δομή του πίνακα H έχει ως εξής, αν $N^2 \times N^2$ είναι η διάσταση του πίνακα,

$$H = \begin{bmatrix} H(0) & H(N-1) & \cdots & H(1) \\ H(1) & H(0) & \cdots & H(2) \\ \vdots & \vdots & \ddots & \vdots \\ H(N-1) & H(N-2) & \cdots & H(0) \end{bmatrix},$$

όπου οι πίνακες $H(m)$ έχουν τη μορφή

$$H(m) = \begin{bmatrix} h(m,0) & h(m,N-1) & \cdots & h(m,1) \\ h(m,1) & h(m,0) & \cdots & h(m,2) \\ \vdots & \vdots & \ddots & \vdots \\ h(m,N-1) & h(m,N-2) & \cdots & h(m,0) \end{bmatrix}.$$

Η αλλαγή αυτή είναι ιδιαίτερα ενδιαφέρουσα γιατί επιτρέπει τη χρήση του διακριτού μετασχηματισμού Fourier για τη λύση του προβλήματος αποκατάστασης.

Δίδονται στη συνέχεια λύσεις του αντίστροφου προβλήματος για την αποκατάσταση των εικόνων.

12.1 Αντίστροφο φίλτρο

Το αντίστροφο φίλτρο $h^{(-1)}$ ορίζεται ως ακολούθως

$$\sum_{m'} \sum_{n'} h^{(-1)}(m-m', n-n') h(m', n') = \delta(m, n) \quad (12.4)$$

και για τις αποκρίσεις στις συχνότητες

$$H^{(-1)}(u, v) H(u, v) = 1 \quad (12.5)$$

Οπότε έχοντας μετατρέψει τους αντίστοιχους πίνακες σε κυκλικούς η λύση μπορεί να δοθεί μέσω του διακριτού μετασχηματισμού Fourier

$$\hat{F}(u, v) = \frac{G(u, v)}{H(u, v)}. \quad (12.6)$$

Με δοσμένο ότι για κάποιες συχνότητες μπορεί η απόκριση $H(u, v)$ να μηδενίζεται τίθεται κατ' αρχήν ένα θέμα ύπαρξης για το αντίστροφο φίλτρο, κι οπωσδήποτε ένα ζήτημα ευστάθειας, με την έννοια της αντοχής στο θόρυβο. Εφόσον το αντίστροφο φίλτρο υπάρχει, η προσδοκητή τιμή της αντιστροφής της παρατήρησης αποκαθιστά τέλεια την εικόνα, με πολλαπλασιασμένη όμως τη διασπορά του θορύβου στις υψηλές συχνότητες. Το φαινόμενο αυτό μπορεί να περιορισθεί με τη χρήση ενός βαθυπερατού φίλτρου, που όμως θα μειώσει την ευκρίνεια της εικόνας.

12.2 Ψευδο-αντίστροφο φίλτρο

Η αντιστροφή του πίνακα H (12.3) προϋποθέτει ότι είναι τετραγωνικός. Πολύ συχνά η διάσταση της παρατηρούμενης εικόνας είναι μεγαλύτερη από αυτή της αποκαθιστόμενης. Τότε η αντιστροφή ζητείται με την έννοια της ελαχιστοποίησης της απόστασης $\|Hf - g\|^2$. Εάν ο πίνακας $H^T H$ είναι αντιστρέψιμος, η λύση είναι

$$\hat{f} = (H^T H)^{-1} H^T g \quad (12.7)$$

και, εάν χρησιμοποιηθεί ο διακριτός μετασχηματισμός Fourier,

$$\hat{F}(u, v) = \frac{H^*(u, v)G(u, v)}{|H(u, v)|^2}, \quad H(u, v) \neq 0. \quad (12.8)$$

Ωστόσο το πρόβλημα σχετικά με την ευστάθεια παραμένει. Η χρήση ενός επαναληπτικού αλγόριθμου επιτρέπει την εύρεση της λύσης χωρίς την άμεση αντιστροφή του πίνακα $H^T H$, και ταυτόχρονα δίδει ένα τρόπο για την αντιμετώπιση του ζητήματος της ευστάθειας. Σε Παραρτήματα δίδονται δύο επαναληπτικές μέθοδοι επίλυσης με κατάβαση στην κατεύθυνση της κλίσης η μία, και της συζυγούς κλίσης η άλλη. Η πρώτη δεν απαιτεί διαιρέσεις και είναι πιο αργή από τη δεύτερη. Και στις δύο περιπτώσεις κάθε νέα επανάληψη βελτιώνει την αποκαθιστόμενη εικόνα, αυξάνοντας όμως ταυτόχρονα το θόρυβο. Η βέλτιστη συνολικά λύση προκύπτει όταν οι επαναλήψεις τερματισθούν πριν την τέλεια αντιστροφή του πίνακα $H^T H$.

12.3 Γενικευμένη αντιστροφή

Στην περίπτωση που ο πίνακας $H^T H$ δεν είναι αντιστρέψιμος, μπορεί να ζητηθεί η λύση της οποίας το μέτρο είναι ελάχιστο. Αναφερόμαστε τότε στη γενικευμένη αντιστροφή, που συμβολίζεται ως εξής

$$f^\# = H^\# g.$$

Ο γενικευμένος αντίστροφος ικανοποιεί τις συνθήκες Moore-Penrose

$$\begin{aligned} H H^\# H &= H \\ H^\# H H^\# &= H^\# \\ (H H^\#)^T &= H H^\# \\ (H^\# H)^T &= H^\# H \end{aligned}$$

Ο γενικευμένος αντίστροφος μπορεί να εκφρασθεί με τη βοήθεια των ιδιάζουσών τιμών και των αντίστοιχων διανυσμάτων του πίνακα H . Ας είναι

$$H = U \Lambda V^T, \quad \text{με } U^T U = \mathbf{I} \text{ και } V^T V = \mathbf{I}$$

όπου ο πίνακας Λ περιλαμβάνει τις ιδιάζουσες τιμές του H : $\{\lambda_1, \dots, \lambda_R\}$. Θα ισχύει τότε

$$H^\# = \sum_{r=1}^R \frac{1}{\lambda_r} v_r u_r^T,$$

όπου u_r (αντίστοιχα v_r) είναι οι στήλες του U (αντίστοιχα V). Η λύση είναι

$$f^\# = \sum_{r=1}^R \frac{1}{\lambda_r} \langle g, u_r \rangle v_r.$$

Το πρόβλημα της αστάθειας παραμένει, κι αντιμετωπίζεται αγνοώντας στο παραπάνω άθροισμα τους όρους με χαμηλές ιδιάζουσες τιμές.

12.4 Εξομάλυνση

Ένα αντίστροφο πρόβλημα θεωρείται ότι έχει τεθεί κάτω από καλές συνθήκες, εάν υπάρχει λύση, και μία μόνη, κι επιπλέον αυτή η λύση είναι σταθερή. Η μέθοδος της εξομάλυνσης μετατρέπει ένα πρόβλημα που τίθεται κάτω από δυσμενείς όρους σ' ένα πρόβλημα που μπορεί να λυθεί κάτω από καλές συνθήκες. Προς τούτο προστίθεται μία συνθήκη ευστάθειας στο κριτήριο ελαχιστοποίησης που εκφράζει το αντίστροφο πρόβλημα,

$$\|Hf - g\|^2 + \mu\Omega(f)$$

Η παράμετρος μ ονομάζεται συντελεστής εξομάλυνσης.

Κατάλληλες τέτοιες συνθήκες ευστάθειας ορίζονται με τη βοήθεια των τελεστών παραγωγίσισης. Η χρήση των πρώτων μερικών παραγώγων δίδει την ακόλουθη λύση

$$\hat{f} = (H^T H + \mu(D_x^T D_x + D_y^T D_y))^{-1} H^T g.$$

Η χρήση του λαπλασιανού τελεστή δίδει την εξής λύση

$$\hat{f} = (H^T H + \mu L^T L)^{-1} H^T g.$$

Η επίλυση μπορεί να γίνει με μία από τις επαναληπτικές μεθόδους που δίδονται στα Παραρτήματα, ή με χρήση του διακριτού μετασχηματισμού Fourier.

Εάν είναι γνωστά τα φάσματα ισχύος τόσο της αρχικής εικόνας ($P_f(u, v)$), όσο και του θορύβου ($P_w(u, v)$), τότε μπορούν να χρησιμοποιηθούν για την προσαρμογή της εξομάλυνσης, δίδοντας το φίλτρο Wiener

$$\hat{F}(u, v) = \frac{P_f(u, v)H^*(u, v)G(u, v)}{P_f(u, v)|H(u, v)|^2 + P_w(u, v)}.$$

Η σχέση αυτή απλουστεύεται αν θεωρηθεί ότι ο λόγος των δύο φασμάτων ισχύος είναι σταθερός για όλες τις συχνότητες

$$\hat{F}(u, v) = \frac{H^*(u, v)G(u, v)}{|H(u, v)|^2 + \lambda}.$$

12.5 Παραρτήματα

12.5.1 Κατάβαση στην κατεύθυνση της κλίσης

Κατά την επανάληψη $k+1$ στην περίπτωση του ψευδοαντίστροφου η λύση δίδεται από την εξίσωση

$$\hat{f}^{(k+1)} = \hat{f}^{(k)} - \beta H^T (H \hat{f}^{(k)} - g)$$

με $\hat{f}^{(0)} = \beta H^T g$. Η επαναληπτική αυτή σχέση δίδει

$$\hat{f}^{(k)} = \beta \sum_{l=0}^k (\mathbf{I} - \beta H^T H)^l H^T g$$

και συγκλίνει με την προϋπόθεση ότι

$$0 < \beta < \frac{2}{\lambda_{max}^2},$$

όπου λ_{max}^2 είναι η μέγιστη ιδιοτιμή του πίνακα $H^T H$.

12.5.2 Αλγόριθμος των συζυγών κλίσεων

Για απλοποίηση θέτουμε: $Q = H^T H$ και $b = H^T g$. Ο πίνακας Q είναι συμμετρικός και θετικά ορισμένος, διαστάσεων $K \times K$. Δύο διανύσματα d_1 και d_2 ονομάζονται Q -ορθογώνια, αν $d_1^T Q d_2 = 0$. Το σύνολο K Q -ορθογώνιων διανυσμάτων αποτελεί μία πλήρη βάση στο χώρο K διαστάσεων. Θα είναι επομένως

$$f = \sum_{k=0}^{K-1} \alpha_k d_k.$$

Ο αλγόριθμος των συζυγών κλίσεων κατασκευάζει ταυτόχρονα τις συζυγείς κλίσεις και τη λύση του συστήματος γραμμικών εξισώσεων, και συγκλίνει σε K ακριβώς βήματα. Για μια οποιαδήποτε αρχική λύση $\hat{f}^{(0)}$, υπολογίζεται η κλίση $d_0 = b - Q\hat{f}^{(0)}$. Για $k = 0, \dots, K - 1$, υπολογίζεται η νέα λύση του συστήματος με μία διόρθωση στην κατεύθυνση της κλίσης

$$\hat{f}^{(k+1)} = \hat{f}^{(k)} + \alpha_k d_k,$$

όπου $\alpha_k = \frac{(b - Q\hat{f}^{(k)})^T d_k}{d_k^T Q d_k}$. Η νέα κλίση, Q -ορθογώνια στις προηγούμενες, θα είναι

$$d_{k+1} = b - Q\hat{f}^{(k+1)} + \beta_k d_k,$$

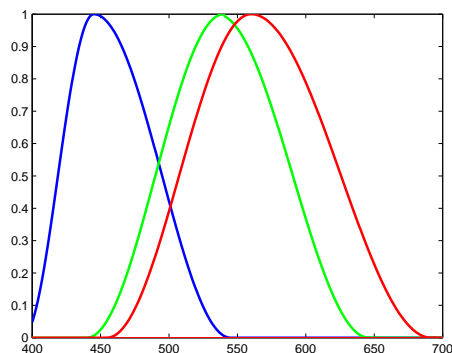
όπου $\beta_k = \frac{(Q\hat{f}^{(k+1)} - b)^T Q d_k}{d_k^T Q d_k}$.

Κεφάλαιο 13

Επεξεργασία έγχρωμων εικόνων

Το χρώμα αποτελεί πρόσθετη πηγή πληροφορίας για το περιεχόμενο των εικόνων κι επομένως υπάρχει ιδιαίτερο ενδιαφέρον, όχι μόνο για την απόδοση των χρωμάτων, αλλά και για την επεξεργασία των έγχρωμων εικόνων.

Στο Κεφάλαιο 2.2 αναφέρθηκε ότι το ανθρώπινο μάτι διαθέτει χρωματικούς αισθητήρες ευαίσθητους σε μήκη κύματος στα ονομαζόμενα κύρια χρώματα: κόκκινο R , πράσινο G και μπλε B . Στο Σχήμα 13.1 δίδονται προσεγγιστικά οι συναρτήσεις ευαισθησίας των τριών αισθητήρων ανάλογα με το μήκος κύματος σε nm.



Σχήμα 13.1: Οι συναρτήσεις ευαισθησίας.

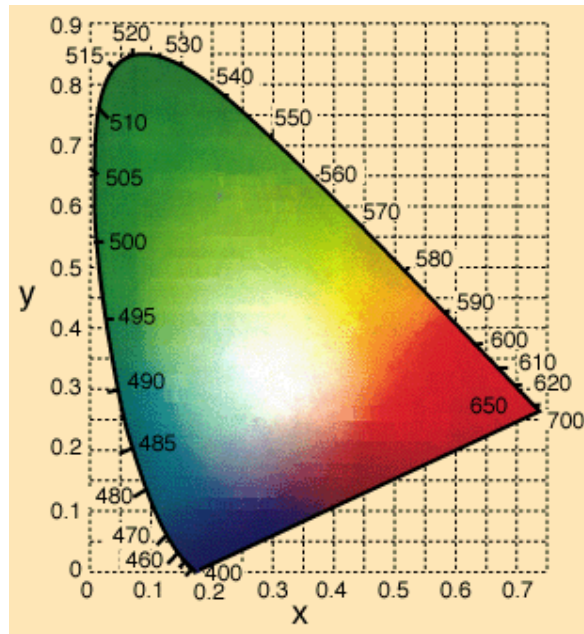
Αντίστοιχα υπάρχουν τεχνητοί αισθητήρες που δίνουν τρία κύρια χρώματα και όπου το ζητούμενο είναι να συντεθούν όλα τα δυνατά χρώματα και να ταιριάζουν στην ανθρώπινη αντίληψη των χρωμάτων και των χρωματικών διαφορών. Η Διεθνής Επιτροπή Φωτομετρίας (*Commission Internationale de l'Eclairage, C.I.E.*) έχει προσδιορίσει ένα διάγραμμα επιχρωμάτων. Ορίστηκαν αρχικά οι τριχρωματικές διεγέρσεις (X, Y, Z) , ώστε οι συναρτήσεις ταιριάσματος των χρωμάτων να είναι θετικές για όλα τα μήκη κύματος και η συνάρτηση για τη συνιστώσα Y να ανταποκρίνεται στην ευαισθησία του ανθρώπινου οφθαλμού στη φωτεινότητα. Η σχέση των (X, Y, Z) με τις συνιστώσες (R, G, B) , όπως είχαν ορισθεί επίσης από την *C.I.E.* είναι η ακόλουθη

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4900 & 0.3100 & 0.2000 \\ 0.1770 & 0.8124 & 0.0106 \\ 0.0000 & 0.0100 & 0.9900 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Η παραπάνω μετατροπή γίνεται με την παραδοχή ότι $X = Y = Z = 1$, όταν $R = G = B = 1$. Από τις τριχρωματικές διεγέρσεις ορίζονται οι συντελεστές

$$x = \frac{X}{X+Y+Z}, y = \frac{Y}{X+Y+Z}, z = \frac{Z}{X+Y+Z}.$$

Το διάγραμμα ορίζεται ως συνάρτηση των (x, y) και τα εφικτά χρώματα προσδιορίζονται από ένα τόπο σε μορφή γλώσσας. Στην καμπύλη περίμετρο ευρίσκονται τα καθαρά χρώματα του ορατού φάσματος. Η γραμμική παρεμβολή μεταξύ δύο χρωμάτων δίνει τα ενδιάμεσα χρώματα.



Σχήμα 13.2: Το χρωματικό διάγραμμα.

Με τον ορισμό των τριών κύριων χρωμάτων ορίζονται τρεις συνιστώσες. Ανάλογα όμως με τη χρήση σε συσκευές ή σε υπολογισμούς έχουν ορισθεί διαφορετικά συστήματα αναπαράστασης των χρωμάτων. Το σύστημα *RGB* χρησιμοποιείται στις οθόνες και τους προβολείς και βέβαια στις συσκευές λήψης. Στους εκτυπωτές χρησιμοποιείται το σύστημα *CMYK*, που χρησιμοποιεί τα δευτερεύοντα χρώματα: κυανό, ιώδες και κίτρινο. Για τον ποιοτικό χαρακτηρισμό και την περιγραφή των χρωμάτων είναι προτιμότερο το σύστημα *HSI*. Το *H (hue)* ορίζει το χρώμα, το *S (saturation)* ορίζει τον κορεσμό του χρώματος, δηλαδή την καθαρότητά του, και το *I (intensity)* ορίζει τη φωτεινή ένταση. Για μετρήσεις χρωματικών διαφορών καταλληλότερο σύστημα είναι το *CIE Lab*. Στα πρότυπα συμπίεσης εικόνων και βίντεο χρησιμοποιείται το σύστημα *YCbCr*. Ακολουθεί η παρουσίαση των συστημάτων χρωμάτων.

13.1 Συστήματα χρωμάτων

13.1.1 HSI

Η πρώτη συνιστώσα περιγράφει το κυρίως καθαρό χρώμα, η δεύτερη συνιστώσα το βαθμό καθαρότητας και η τρίτη την ένταση. Το σύστημα αυτό αφενός διαχωρίζει την ένταση από το χρωματικό περιεχόμενο κι αφετέρου οι δύο χρωματικές συνιστώσες συνδέονται με τον τρόπο που οι άνθρωποι αντιλαμβάνονται τα χρώματα.

Για τη μετατροπή των συνιστωσών από το σύστημα RGB στο σύστημα HSI θεωρούμε το χρωματικό σημείο στο χώρο ενός μοναδιαίου κύβου. Με αναφορά το σημείο $(1/3, 1/3, 1/3)$ ορίζουμε το χρώμα H ως τη γωνία ανάμεσα στο διάνυσμα που ορίζεται από το χρωματικό σημείο και στο διάνυσμα του κόκκινου χρώματος. Επομένως 0° αντιστοιχούν στο κόκκινο χρώμα, 60° στο κίτρινο, 120° στο πράσινο, 180° στο κυανό, 240° στο μπλε και 300° στο ιώδες. Προκύπτει τότε η ακόλουθη σχέση σε μοίρες

$$H = \arccos \frac{R - 0.5(G + B)}{\sqrt{(R - G)^2 + (R - B)(G - B)}}, \text{ για } G \geq B$$

και

$$H = 360^\circ - \arccos \frac{R - 0.5(G + B)}{\sqrt{(R - G)^2 + (R - B)(G - B)}}, \text{ για } G \leq B$$

Η ένταση ορίζεται

$$I = \frac{1}{3}(R + G + B).$$

Ο κορεσμός ορίζεται από τη σχέση

$$S = 1 - \frac{\min(R, G, B)}{I}.$$

Αν ο κορεσμός είναι μηδέν δεν ορίζεται το χρώμα. Επίσης ο κορεσμός δεν ορίζεται, αν η ένταση είναι μηδενική (μαύρο χρώμα). Ο αρχικός κύβος των χρωμάτων μετατράπηκε σε κύλινδρο, του οποίου ο άξονας αντιστοιχεί στην ένταση, η ακτίνα στον κορεσμό και η γωνία στο χρώμα.

13.1.2 C.I.E. Lab

Το σύστημα Lab έχει σχεδιασθεί από την $C.I.E.$ για να προσεγγίζει την ανθρώπινη αντίληψη των χρωμάτων και των χρωματικών διαφορών. Ο χρωματικός χώρος είναι ομοιόμορφος ως προς τις χρωματικές διαφορές, με διαχωρισμό της φωτεινότητας L από τις χρωματικές συνιστώσες (a, b) . Ο ορισμός χρησιμοποιεί το διάνυσμα τριχρωματικών διεγέρσεων (X, Y, Z) που παρουσιάστηκε προηγούμενα.

$$\begin{aligned} L &= 116f\left(\frac{Y}{Y_n}\right) - 16 \\ a &= 500\left(f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right)\right) \\ b &= 200\left(f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right)\right) \end{aligned}$$

όπου

$$f(\tau) = \begin{cases} \tau^{1/3}, & \tau > (6/29)^3 \\ (841/108)\tau + 4/29, & \tau \leq (6/29)^3 \end{cases}$$

Το (X_n, Y_n, Z_n) αντιστοιχεί στο λευκό χρώμα. Η ευκλείδεια απόσταση δίδει τη χρωματική διαφορά. Το χρώμα και ο κορεσμός του μπορούν να μετρηθούν ως εξής

$$H = \arctan \frac{b}{a}, S = \sqrt{a^2 + b^2}.$$

13.1.3 YCbCr

Η μετατροπή, όπως χρησιμοποιείται στο πρότυπο συμπίεσης *JPEG*, δίδεται ως ακολούθως

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.500 \\ 0.500 & -0.419 & -0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Το σύστημα *RGB* στις ανωτέρω σχέσεις είναι αυτό που έχει ορισθεί από την Εταιρεία Μηχανικών Τηλεόρασης (Society of Motion Picture and Television Engineers, SMPTE). Και στο σύστημα αυτό διαχωρίζεται η ένταση από τις χρωματικές συνιστώσες. Το σύστημα *YCbCr* χρησιμοποιείται στα πρότυπα συμπίεσης εικόνων και βίντεο.

13.2 Κβαντισμός χρώματος

Ο σκοπός του κβαντισμού είναι η συμπίεση ή περιγραφή του περιεχομένου, ή ακόμα η τμηματοποίηση. Ο κβαντιστής συνίσταται από ένα σύνολο αντιπροσωπευτικών διανυσμάτων $\{c_k : k = 1, \dots, K\}$. Το σύνολο των αντιπροσωπευτικών χρωμάτων κατασκευάζεται από ένα σύνολο διανυσμάτων εκμάθησης $\{x(s) : s \in \mathcal{L}\}$ και βασίζεται σ' ένα κριτήριο ελάχιστης παραμόρφωσης, όπως τετραγωνικής, που ορίζεται

$$D = \sum_{k=1}^K \sum_{x(s) \in \mathcal{S}_k} \|x(s) - c_k\|^2 \quad (13.1)$$

για K κλάσεις \mathcal{S}_k .

Ο βέλτιστος κβαντιστής ικανοποιεί δύο αναγκαίες συνθήκες για την ελαχιστοποίηση της D . Για δοσμένη κλάση ο καλύτερος αντιπρόσωπος είναι το κέντρο βάρους

$$c_k = \frac{1}{\text{card}[\mathcal{S}_k]} \sum_{x(s) \in \mathcal{S}_k} x(s) \quad (13.2)$$

Για δοσμένο σύνολο αντιπροσωπευτικών χρωμάτων η καλύτερη τιμή κβαντισμού ενός διανύσματος x συνίσταται στην επιλογή του πλησιέστερου αντιπρόσωπου

$$\|x - c_k\| < \|x - c_l\| \quad \forall l \neq k \quad \Rightarrow \quad x \in \mathcal{S}_k \quad (13.3)$$

Η χρησιμοποίηση των δύο αυτών αναγκαίων συνθηκών δίδει έναν επαναληπτικό αλγόριθμο κατασκευής ενός συνόλου αντιπροσωπευτικών χρωμάτων.

- Αρχικό βήμα: Αρχικά αντιπροσωπευτικά χρώματα, $i = 1$, και αρχική μεγάλη τιμή για την παραμόρφωση $D^{(0)}$
- Βήμα 1: Εύρεση των κλάσεων (Εξίσωση (13.3))
- Βήμα 2: Υπολογισμός της παραμόρφωσης $D^{(i)}$
- Βήμα 3: Έλεγχος σύγκλισης

$$\text{Αν } \frac{D^{(i-1)} - D^{(i)}}{D^{(i-1)}} \leq \epsilon, \quad \text{τέλος}$$

Διαφορετικά, προσαύξηση του i , και συνέχιση των επαναλήψεων

- Βήμα 4: Εύρεση του καλύτερου αντιπρόσωπου για κάθε κλάση (Εξίσωση (13.2)), κι επιστροφή στο Βήμα 1.

13.3 Μείωση θορύβου

Ο θόρυβος παράγεται είτε κατά τη διαδικασία της λήψης, είτε κατά τη μετάδοση των εικόνων. Πολύ συχνά μπορεί να θεωρηθεί αθροιστικός και ανεξάρτητος από σημείο σε σημείο. Τα γραμμικά φίλτρα μείωσης θορύβου που παρουσιάστηκαν στο Κεφάλαιο 9 μπορούν να χρησιμοποιηθούν ξεχωριστά για κάθε συνιστώσα των έγχρωμων εικόνων. Και το φίλτρο μεσαίας τιμής θα μπορούσε να χρησιμοποιηθεί ανεξάρτητα στις ξεχωριστές συνιστώσες. Ωστόσο είναι προτιμότερο να ληφθεί υπόψη ότι τα δεδομένα είναι τρισδιάστατα.

Το φίλτρο μεσαίου διάνυσματος εφαρμόζεται σε μια περιοχή W γύρω από το σημείο που λαμβάνεται η απόκριση και συνήθως είναι ένα τετράγωνο μπλοκ. Ας είναι

$$\{x_n, n = 1, 2, \dots, N\}$$

τα διάνυσμα χρώματος στην περιοχή W . Το μεσαίο διάνυσμα είναι εκείνο (x_i) μεταξύ των παραπάνω που ελαχιστοποιεί την απόσταση¹

$$\sum_{n=1}^N \|x_i - x_n\|_p.$$

Το φίλτρο μεσαίου διάνυσματος αφαιρεί τον χρουστικό θόρυβο και μειώνει τον αθροιστικό θόρυβο, ιδίως αν συνδυασθεί και με γραμμικά φίλτρα.

Το διάνυσμα μεσαίας κατεύθυνσης αποκρίνεται x_i , εφόσον αυτό το διάνυσμα ελαχιστοποιεί την απόσταση

$$\sum_{n=1}^N \theta(x_i, x_n),$$

που ορίζεται με βάση τις γωνιακές αποκλίσεις. Υπάρχει επίσης η δυνατότητα χρήσης φίλτρων και ως προς την κατεύθυνση και ως προς το μέτρο των διανυσμάτων της θεωρούμενης περιοχής.

Ορίζεται επίσης το φίλτρο τοπικά σταθμισμένης μέσης διανυσματικής τιμής ως ακολούθως

$$f \left(\frac{\sum_{i=1}^N w_i x_i}{\sum_{i=1}^N w_i} \right), w_i = e^{-\frac{(\sum_{n=1}^N \|x_i - x_n\|_p)^r}{\beta}}$$

με ενδεχόμενη χρήση μιας μη γραμμικής συνάρτησης $f(\cdot)$.

¹ $\|x\|_p = (\sum_{k=1}^K |x_k|^p)^{1/p}$

Κεφάλαιο 14

Ανάλυση σε ζώνες συχνοτήτων

Με την ανάλυση η εικόνα διαχωρίζεται σε μικρότερες εικόνες που η καθεμία αντιστοιχεί σε μία διαφορετική ζώνη συχνοτήτων. Επειδή κατά κανόνα για τα 2-Δ σήματα χρησιμοποιούνται διαχωρίσιμα φίλτρα για την ανάλυση, είναι αρκετό να αναπτύξουμε το ζήτημα για 1-Δ σήματα.

Ας ξεκινήσουμε με κάποιους χρήσιμους ορισμούς. Ονομάζεται αποδεκατισμός τάξης p (όπου $p > 1$ φυσικός αριθμός) η περιοδική δειγματοληψία ενός από κάθε p δείγματα ενός σήματος. Το αποδεκατισμένο σήμα που προκύπτει από το σήμα $x(n)$ είναι

$$y(n) = x(pn) \quad (14.1)$$

Ας ονομάσουμε $X(z)$ (αντίστοιχα $Y(z)$) το μετασχηματισμό Z του $x(n)$ (αντίστοιχα $y(n)$). Η σχέση μεταξύ των δύο μετασχηματισμών Z προκύπτει ως εξής

$$\begin{aligned} Y(z) &= \sum_{n=-\infty}^{\infty} x(pn)z^{-n} = \sum_{n=-\infty}^{\infty} \frac{z^{-n}}{2\pi j} \oint X(\zeta)\zeta^{pn-1}d\zeta \\ &= \frac{1}{2\pi j} \oint X(\zeta) \sum_{n=-\infty}^{\infty} \left(\frac{z}{\zeta^p}\right)^n \frac{d\zeta}{\zeta} = \frac{1}{2\pi j} \oint X(\zeta) \frac{\zeta^{p-1}}{\zeta^p - z} d\zeta \\ &= \frac{1}{2\pi j} \oint X(\zeta) \frac{1}{p} \sum_{k=0}^{p-1} \frac{1}{\zeta - (ze^{-j2\pi k})^{1/p}} d\zeta \\ &= \frac{1}{p} \sum_{k=0}^{p-1} X(z^{1/p} e^{-\frac{j2\pi k}{p}}) \end{aligned} \quad (14.2)$$

Απ' αυτή τη σχέση μπορεί να εξαχθεί επίσης η σχέση μεταξύ των δύο μετασχηματισμών Fourier

$$\mathcal{Y}(u) = \frac{1}{p} \sum_{k=0}^{p-1} \mathcal{X}\left(\frac{u-k}{p}\right) \quad (14.3)$$

Είναι φανερό ότι με τον αποδεκατισμό δημιουργείται ένα πρόβλημα ψευδώνυμων συχνοτήτων. Το φαινόμενο αυτό δεν είναι επιθυμητό, αντίθετα επιζητείται ο περιορισμός του αποδεκατισμένου σήματος σε μια ζώνη συχνοτήτων εύρους $1/p$. Γι' αυτό απαιτείται η χρήση ενός βαθυπερατού φίλτρου πριν τον αποδεκατισμό. Το ιδανικό φίλτρο, εάν υπήρχε, θα είχε την ακόλουθη απόκριση στις συχνοτήτες

$$\mathcal{H}(u) = \begin{cases} 1 & |u| \leq \frac{1}{2p} \\ 0 & \frac{1}{2p} < |u| \leq \frac{1}{2} \end{cases} \quad (14.4)$$

Στην πράξη για ένα οποιοδήποτε φίλτρο και μετά τον αποδεκατισμό θα ομιλούμε για προσέγγιση του σήματος σε κλίμακα $1/p$ και θα έχουμε τις ακόλουθες σχέσεις

$$x_a(n : p) = \sum_{k=-\infty}^{\infty} h(pn - k)x(k) \quad (14.5)$$

$$X_a(z : p) = \frac{1}{p} \sum_{k=0}^{p-1} H(z^{1/p} e^{-j\frac{2\pi k}{p}}) X(z^{1/p} e^{-j\frac{2\pi k}{p}}) \quad (14.6)$$

$$\mathcal{X}_a(z : p) = \frac{1}{p} \sum_{k=0}^{p-1} \mathcal{H}\left(\frac{u-k}{p}\right) \mathcal{X}\left(\frac{u-k}{p}\right) \quad (14.7)$$

Στην αντίθετη κατεύθυνση σε σχέση με τον αποδεκατισμό ευρίσκεται η παρεμβολή δειγμάτων με μηδενική τιμή για την επέκταση του σήματος κατά ένα παράγοντα p . Το σήμα που προκύπτει είναι το ακόλουθο

$$y(n) = \begin{cases} x\left(\frac{n}{p}\right) & n \% p = 0 \\ 0 & n \% p \neq 0 \end{cases} \quad (14.8)$$

όπου $n \% p$ δηλώνει το υπόλοιπο της διαίρεσης του n δια p . Αποδεικνύεται εύκολα ότι μεταξύ των μετασχηματισμών Z και Fourier ισχύουν οι ακόλουθες σχέσεις

$$Y(z) = X(z^p) \quad (14.9)$$

$$\mathcal{Y}(u) = \mathcal{X}(pu) \quad (14.10)$$

Η σταθμισμένη παρεμβολή $p-1$ τιμών μεταξύ δύο τιμών του αρχικού σήματος $x(n)$ προκύπτει με τη χρήση ενός βαθυπερατού φίλτρου μετά την παρεμβολή των μηδενικών τιμών. Οπότε θα έχουμε

$$x_i(n : p) = \sum_{k=-\infty}^{\infty} h(n-k)y(k) = \sum_{l=-\infty}^{\infty} h(n-pl)x(l) \quad (14.11)$$

Μεταξύ των μετασχηματισμών Z και Fourier ισχύουν οι ακόλουθες σχέσεις

$$X_i(z : p) = H(z)X(z^p) \quad (14.12)$$

$$\mathcal{X}_i(u : p) = \mathcal{H}(u)\mathcal{X}(pu) \quad (14.13)$$

Στη συνέχεια θα περιορισθούμε μόνο στην περίπτωση της ανάλυσης σε δύο ζώνες συχνοτήτων ($p=2$) με βάση τους παραπάνω ορισμούς. Η προσέγγιση του σήματος σε κλίμακα $1/2$ αντιστοιχεί στη ζώνη χαμηλών συχνοτήτων. Συνήθως το φίλτρο $h(n)$ επιλέγεται ώστε να ικανοποιεί τη συνθήκη

$$\sum_{n=-\infty}^{\infty} h(2n) = \sum_{n=-\infty}^{\infty} h(2n+1)$$

Με παρόμοιο τρόπο και στην ίδια κλίμακα μπορούν να εξαχθούν οι λεπτομέρειες του σήματος που αντιστοιχούν στις υψηλές συχνότητες

$$x_d(n) = \sum_{k=-\infty}^{\infty} g(2n-k)x(k) \quad (14.14)$$

Οπότε θα έχουμε, χρησιμοποιώντας την Εξίσωση (14.6)

$$X_d(z) = \frac{1}{2} \left(G(z^{1/2})X(z^{1/2}) + G(-z^{1/2})X(-z^{1/2}) \right) \quad (14.15)$$

Το φίλτρο για την εξαγωγή των υψηλών συχνοτήτων επιλέγεται με βάση το βαθυπερατό φίλτρο που χρησιμοποιείται για την προσέγγιση. Συνήθης επιλογή είναι η ακόλουθη

$$\mathcal{G}(u) = e^{-j2\pi u} \mathcal{H}^*(u - \frac{1}{2}) \quad (14.16)$$

$$G(z) = z^{-1} H(-\frac{1}{z}) \quad (14.17)$$

$$g(n) = (-1)^{1-n} h(1-n) \quad (14.18)$$

Από την προσέγγιση και από τις λεπτομέρειες του σήματος, ακόμα κι αν το φίλτρο h δεν είναι ιδανικό μπορεί να αποκατασταθεί τέλεια το αρχικό σήμα αρκεί να χρησιμοποιηθεί σταθμισμένη παρεμβολή με το φίλτρο $H(\frac{1}{z})$ για τις χαμηλές συχνότητες και με το $G(\frac{1}{z})$ για τις υψηλές συχνότητες, και τα δύο προκύπτοντα σήματα να προστεθούν. Ας είναι $\hat{X}(z)$ ο μετασχηματισμός Z που προκύπτει μ'αυτό τον τρόπο. Θα έχουμε

$$\begin{aligned} \hat{X}(z) &= \frac{1}{2} (H(z)X(z) + H(-z)X(-z)) H(\frac{1}{z}) + \frac{1}{2} (G(z)X(z) + G(-z)X(-z)) G(\frac{1}{z}) \\ &= \frac{1}{2} \left(H(z)H(\frac{1}{z}) + H(-z)H(-\frac{1}{z}) \right) X(z) + \frac{1}{2} \left(H(-z)H(\frac{1}{z}) - H(\frac{1}{z})H(-z) \right) X(-z) \end{aligned}$$

Αν επομένως $H(z)H(\frac{1}{z}) + H(-z)H(-\frac{1}{z}) = 2$, τότε $\hat{X}(z) = X(z)$. Η παραπάνω συνθήκη είναι ισοδύναμη με

$$|\mathcal{H}(u)|^2 + |\mathcal{H}(u - \frac{1}{2})|^2 = 2 \quad (14.19)$$

Επιπλέον μεταξύ των φίλτρων H και G απαιτείται μία σχέση ορθογωνιότητας

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} \mathcal{H}(u) \mathcal{G}^*(u) du = 0$$

Φίλτρα που επαληθεύουν αυτές τις σχέσεις ανήκουν στην κατηγορία των συζυγών ορθογώνιων φίλτρων.

Ένα ζεύγος φίλτρων που ικανοποιεί όλες τις παραπάνω απαιτήσεις είναι το ακόλουθο

$$H(z) = \frac{1}{\sqrt{2}}(1 + z^{-1}) \quad \text{και} \quad G(z) = \frac{1}{\sqrt{2}}(-1 + z^{-1})$$

που όμως έχει πτωχές επιδόσεις ως προς την ανάλυση σε ζώνες συχνοτήτων. Κατά κανόνα χρησιμοποιούνται φίλτρα με πεπερασμένη, αλλά μεγαλύτερη σε έκταση, χροστική απόκριση, που ικανοποιούν προσεγγιστικά τις παραπάνω απαιτήσεις. Τέτοια παραδείγματα φίλτρων δίδονται στον Πίνακα 14.1.

Στην περίπτωση 2- Δ σημάτων τα χρησιμοποιούμενα 2- Δ φίλτρα προκύπτουν με βάση την αρχή της διαχωρισιμότητας. Έτσι για ανάλυση σε τέσσερις ζώνες συχνοτήτων χρησιμοποιούνται τα εξής φίλτρα

$$h_{LL}(m, n) = h(m)h(n)$$

$$h_{LH}(m, n) = h(m)g(n)$$

$$h_{HL}(m, n) = g(m)h(n)$$

$$h_{HH}(m, n) = g(m)g(n)$$

Η ανάλυση μπορεί να επαναληφθεί για οποιαδήποτε από τις ζώνες συχνοτήτων, αλλά κυρίως έχει ενδιαφέρον η περαιτέρω ανάλυση της ζώνης χαμηλών συχνοτήτων για προσέγγιση σε μικρότερη κλίμακα. Συνολικά το 2- Δ σήμα μπορεί να αναλυθεί σε K σήματα I_k , με διαφορετική έκταση το καθένα στη γενική περίπτωση.

n	$h_4(n)$	$h_6(n)$	$h_8(n)$	$h_{10}(n)$	$h_{12}(n)$
0	.483	.333	.230	.160	.112
1	.837	.807	.715	.604	.495
2	.224	.460	.631	.724	.751
3	-.129	-.135	-.028	.138	.315
4		-.085	-.187	-.242	-.226
5		.035	.031	-.032	-.130
6			.033	.078	.098
7			-.011	-.006	.028
8				-.013	-.032
9				.003	.001
10					.005
11					-.001

Πίνακας 14.1: Φίλτρα προσέγγισης σε κλίμακα 1/2

Κεφάλαιο 15

Συμπίεση εικόνων

Η ανάγκη για συμπίεση των εικόνων προκύπτει από το μεγάλο όγκο των δεδομένων προς μετάδοση ή αποθήκευση. Ζητούμενο της συμπίεσης είναι η οικονομικότερη μετάδοση ή αποθήκευση των εικόνων. Η συμπίεση των εικόνων βασίζεται στον περιορισμό της πλεονάζουσας πληροφορίας, χωρίς καμία αλλοίωση των δεδομένων της εικόνας, και στον περιορισμό της περιττής πληροφορίας, χωρίς αντιληπτή αλλοίωση των δεδομένων της εικόνας, ή έστω ως τα όρια της ανεκτής από το ανθρώπινο σύστημα όρασης αλλοίωσης. Κρίτηρια επίδοσης ενός συστήματος συμπίεσης είναι: ο βαθμός συμπίεσης, η ποιότητα των εικόνων που μπορούν να ανασυντεθούν από τη συμπιεσμένη ποσότητα πληροφορίας, η πολυπλοκότητα του συμπιεστή/αποσυμπιεστή, και, στην περίπτωση ύπαρξης λαθών μεταξύ συμπιεστή και αποσυμπιεστή, η αντοχή του συστήματος συμπίεσης σ'αυτά τα λάθη.

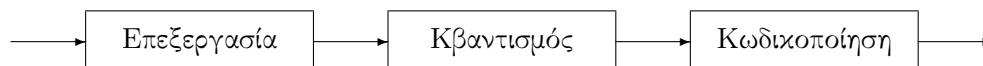
Ο βαθμός συμπίεσης ορίζεται ως ο λόγος της ποσότητας πληροφορίας πριν και μετά τη συμπίεση. Καταλληλότερος κριτής της ποιότητας των συμπιεσμένων εικόνων είναι ο χρήστης των εικόνων. Ωστόσο για τη σχεδίαση συστημάτων συμπίεσης εικόνων μπορούν να χρησιμοποιηθούν στατιστικά κριτήρια μέτρησης της επίδοσής τους, όπως η σηματοθυροβική σχέση σε λογαριθμική κλίμακα (decibel)

$$\text{SNR} = 10 \log_{10} \frac{\sum_{m=1}^M \sum_{n=1}^N I^2(m, n)}{\sum_{m=1}^M \sum_{n=1}^N (I(m, n) - \bar{I}(m, n))^2},$$

που επίσης μπορεί να ορισθεί με βάση τη μέγιστη τιμή του σήματος

$$\text{PSNR} = 10 \log_{10} \frac{MN(I_{max} - I_{min})^2}{\sum_{m=1}^M \sum_{n=1}^N (I(m, n) - \bar{I}(m, n))^2}.$$

Κατά γενικό κανόνα ένας συμπιεστής συνίσταται από τρία μέρη: στο πρώτο μέρος η εικόνα υφίσταται μία επεξεργασία, ανάλυση ή μετασχηματισμό, χωρίς απώλεια πληροφορίας, στο δεύτερο μέρος το αποτέλεσμα χβαντίζεται, με απώλεια πληροφορίας, και στο τρίτο μέρος τα χβαντισμένα μεγέθη κωδικοποιούνται (Σχήμα 15.1). Συμμετρικά ο αποσυμπιεστής αποκωδικοποιεί, αποκα-



Σχήμα 15.1: Τυπικό σχήμα συμπίεσης εικόνων

θιστά την αντιστοιχία με τα χβαντισμένα μεγέθη και ανασυνθέτει την εικόνα.

Η χρήση ενός εντροπικού κωδικοποιητή, όπως αυτός που δίδει ο αλγόριθμος του Huffman (βλέπε Παράρτημα), επιτρέπει τη μέγιστη δυνατή συμπίεση στην έξοδο του κβαντιστή χωρίς απώλεια πληροφορίας.

Ο κβαντισμός οδηγεί σε σημαντική συμπίεση, εισάγοντας παραμόρφωση, που επιδιώκεται να περιορίζεται σε όρια μη αντιληπτά. Ο διανυσματικός κβαντισμός μπορεί να αποτελέσει αυτοτελή μέθοδο συμπίεσης των δεδομένων της εικόνας. Δίδονται στη συνέχεια τα βασικά χαρακτηριστικά ενός διανυσματικού κβαντιστή, που ισχύουν και για την περίπτωση του βαθμωτού κβαντιστή, αρκεί να περιορισθεί ανάλογα η διάσταση του διανύσματος.

15.1 Διανυσματικός κβαντισμός

Τα διανύσματα κατασκευάζονται από στοιχεία που ανήκουν σε μπλοκ από την εικόνα διάστασης $M_b \times N_b$ (τυπική διάσταση του μπλοκ είναι 4×4 , οπότε η διάσταση του διανύσματος είναι 16). Ο συμπίεστος διαθέτει ένα σύνολο από αντιπροσωπευτικά διανύσματα $\{r_k : k = 1, \dots, K\}$ που ονομάζεται λεξικό. Για κάθε διάνυσμα προς κωδικοποίηση επιλέγεται το πλησιέστερο σ αυτό αντιπροσωπευτικό διάνυσμα, και κωδικοποιείται ο δείκτης αυτού του διανύσματος. Ο αποσυμπίεστος διαθέτει το λεξικό και με την αποκωδικοποίηση προσδιορίζει το αντίστοιχο διάνυσμα από το λεξικό, το οποίο και αποκαθιστά. Ο αριθμός των βιτς ανά σημείο της εικόνας θα είναι

$$B = \frac{\log_2 K}{M_b N_b}.$$

Το λεξικό κατασκευάζεται από ένα σύνολο διανυσμάτων εκμάθησης $\{x_i : i = 1, \dots, N_i\}$ και βασίζεται σ'ένα κριτήριο ελάχιστης τετραγωνικής παραμόρφωσης

$$D = \sum_{k=1}^K \sum_{x_i \in C_k} \|x_i - r_k\|^2 \quad (15.1)$$

για K κλάσεις C_k .

Ο βέλτιστος διανυσματικός κβαντιστής ικανοποιεί δύο αναγκαίες συνθήκες για την ελαχιστοποίηση της D . Για δοσμένη κλάση ο καλύτερος αντιπρόσωπος είναι το κέντρο βάρους

$$r_k = \frac{1}{\text{card}[C_k]} \sum_{x_i \in C_k} x_i \quad (15.2)$$

Για δοσμένο λεξικό η καλύτερη τιμή κβαντισμού ενός διανύσματος x συνίσταται στην επιλογή του πλησιέστερου αντιπρόσωπου

$$\|x - r_k\| < \|x - r_l\| \quad \forall l \neq k \quad \Rightarrow \quad x \in C_k \quad (15.3)$$

Η χρησιμοποίηση των δύο αυτών αναγκαίων συνθηκών δίδει έναν επαναληπτικό αλγόριθμο κατασκευής ενός λεξικού από ένα σύνολο διανυσμάτων εκμάθησης.

- Αρχικό βήμα: Αρχικό λεξικό, $i = 1$, και αρχική μεγάλη τιμή για την παραμόρφωση $D^{(0)}$
- Βήμα 1: Εύρεση των κλάσεων (Εξίσωση (15.3))
- Βήμα 2: Υπολογισμός της παραμόρφωσης $D^{(i)}$

- Βήμα 3: Έλεγχος σύγκλισης

$$\text{Αν } \frac{D^{(i-1)} - D^{(i)}}{D^{(i-1)}} \leq \epsilon, \quad \text{τέλος}$$

Διαφορετικά, προσαύξηση του i , και συνέχιση των επαναλήψεων

- Βήμα 4: Εύρεση του καλύτερου αντιπρόσωπου για κάθε κλάση (Εξίσωση (15.2)), και επιστροφή στο Βήμα 1.

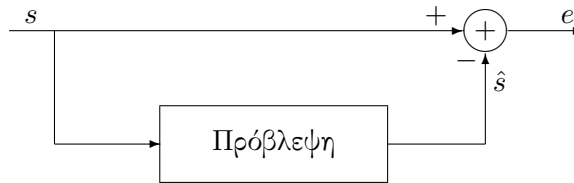
Το υπολογιστικό κόστος για την εύρεση του πλησιέστερου αντιπρόσωπου από το λεξικό είναι $O(M_b N_b K)$. Το κόστος αποθήκευσης του λεξικού είναι $M_b N_b K$. Επειδή το υπολογιστικό κόστος είναι υψηλό, κι ενδεχόμενα απαγορευτικό για μετάδοση σε πραγματικό χρόνο, μπορεί να επιβληθεί μία δομή δένδρου στο λεξικό. Αν το δένδρο περιλαμβάνει ν κλάδους, τότε το υπολογιστικό κόστος περιορίζεται σε $O(M_b N_b \nu \log_2 K)$, ενώ το κόστος αποθήκευσης του λεξικού αυξάνει σε $M_b N_b (K - 1) \frac{\nu}{\nu - 1}$. Αν για παράδειγμα $\nu = 2$, το υπολογιστικό κόστος είναι $O(2M_b N_b \log_2 K)$ και το κόστος αποθήκευσης $2M_b N_b (K - 1)$. Η κατασκευή του λεξικού γίνεται με τη χρήση του παραπάνω αλγόριθμου για ν κλάσεις και για όλους τους ενδιάμεσους κόμβους του δένδρου από τη ρίζα προς τα φύλλα.

15.2 Συμπίεση με πρόβλεψη

Η βασική αρχή ενός συστήματος συμπίεσης με πρόβλεψη είναι ‘ό,τι είναι προβλέψιμο, πλεονάζει’. Στο Σχήμα 15.2 δίδεται ένα σύστημα εξαγωγής της πλεονάζουσας πληροφορίας με πρόβλεψη. Το σύστημα αυτό χαρακτηρίζεται από το κέρδος της πρόβλεψης

$$G_p = \frac{\sigma_s^2}{\sigma_e^2}. \quad (15.4)$$

Για δοσμένο ρυθμό πληροφορίας, και εάν χρησιμοποιηθεί χβαντισμός, που συνεπάγεται παρα-



Σχήμα 15.2: Εξαγωγή του σφάλματος της πρόβλεψης

μόρφωση, η πρόβλεψη οδηγεί σε μείωση της παραμόρφωσης πρακτικά ίση με το κέρδος της πρόβλεψης

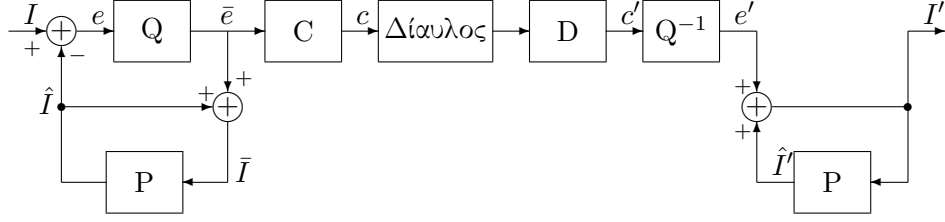
$$\frac{D_s}{D_e} = G_p.$$

Ενώ για δοσμένη παραμόρφωση, ο αριθμός βιτς ανά δείγμα του σήματος, μειώνεται

$$B_s - B_e = \frac{1}{2} \log_2 G_p.$$

Το σύστημα συμπίεσης που περιλαμβάνει και χβαντιστή δίδεται στο Σχήμα 15.3. Απ'αυτό το σχήμα προκύπτει ότι η παραμόρφωση είναι ίση με το σφάλμα του χβαντισμού

$$I - \bar{I} = e - \bar{e}.$$



Σχήμα 15.3: Συμπίεστης/αποσυμπίεστης με πρόβλεψη (P: πρόβλεψη, Q: κβαντισμός, C: κωδικοποίηση, D: αποκωδικοποίηση, Q^{-1} : αντίστροφος κβαντισμός).

Κατά κανόνα για την πρόβλεψη χρησιμοποιείται ένα γραμμικό αναδρομικό φίλτρο. Ένας μικρός αριθμός συντελεστών είναι αρκετός για τον προσδιορισμό του φίλτρου πρόβλεψης, όπως για παράδειγμα

$$\hat{I}(m, n) = a(0, 1)\bar{I}(m, n - 1) + a(1, 0)\bar{I}(m - 1, n) + a(1, 1)\bar{I}(m - 1, n - 1) \quad (15.5)$$

Οι συντελεστές του φίλτρου μπορούν να εκτιμηθούν μέσω της συνάρτησης συμμεταβλητότητας του 2- Δ σήματος.

Ας θεωρήσουμε ένα γραμμικό μοντέλο πρόβλεψης για το 2- Δ σήμα με μηδενική μέση τιμή

$$\hat{I}(m, n) = a(0, 1)I(m, n - 1) + a(1, 0)I(m - 1, n) + a(1, 1)I(m - 1, n - 1). \quad (15.6)$$

Οι συντελεστές μπορούν να προσδιορισθούν χρησιμοποιώντας το κριτήριο ελάχιστου τετραγωνικού σφάλματος

$$E\{(I(m, n) - (a(0, 1)I(m, n - 1) + a(1, 0)I(m - 1, n) + a(1, 1)I(m - 1, n - 1)))^2\}.$$

Η ελαχιστοποίηση του παραπάνω σφάλματος δίδει τις ακόλουθες εξισώσεις

$$E\{(I(m, n) - (a(0, 1)I(m, n - 1) + a(1, 0)I(m - 1, n) + a(1, 1)I(m - 1, n - 1)))I(m - k, n - l)\} = 0,$$

για $(k, l) \in \{(0, 1), (1, 0), (1, 1)\}$, που γράφονται ως ακολούθως

$$\begin{aligned} \gamma(0, 1) &= a(0, 1)\gamma(0, 0) + a(1, 0)\gamma(1, 1) + a(1, 1)\gamma(1, 0) \\ \gamma(1, 0) &= a(0, 1)\gamma(1, 1) + a(1, 0)\gamma(0, 0) + a(1, 1)\gamma(0, 1) \\ \gamma(1, 1) &= a(0, 1)\gamma(1, 0) + a(1, 0)\gamma(0, 1) + a(1, 1)\gamma(0, 0) \end{aligned} \quad (15.7)$$

Οι τιμές της συνάρτησης $\gamma(m, n)$ μπορούν να προσδιορισθούν είτε εμπειρικά είτε μέσω ενός μοντέλου. Παράδειγμα διαχωρίσιμου μοντέλου είναι το εξής

$$\gamma(m, n) = \gamma(0, 0)\rho_v^{|m|}\rho_h^{|n|}.$$

Η λύση του παραπάνω συστήματος σάυτην την περίπτωση δίδει

$$a(0, 1) = \rho_h, a(1, 0) = \rho_v, a(1, 1) = -\rho_h\rho_v$$

Το κέρδος της πρόβλεψης είναι τόσο μεγαλύτερο, όσο οι συντελεστές συσχέτισης είναι υψηλότεροι

$$G_p = \frac{1}{(1 - \rho_v^2)(1 - \rho_h^2)}.$$

Ο χβαντιστής μπορεί να σχεδιασθεί με βάση τη μέθοδο που περιγράφηκε στο Τμήμα 6.2. Ωστόσο το σφάλμα της πρόβλεψης διαφέρει σημαντικά ανάλογα με τα τοπικά χαρακτηριστικά του σημείου. Όπως είναι αναμενόμενο, το σφάλμα της πρόβλεψης είναι ισχυρότερο κοντά στις ακμές, κι ασθενέστερο σε ομοιογενείς περιοχές. Αν επιπλέον ληφθεί υπόψη ότι το ανθρώπινο σύστημα όρασης είναι πιο ευαίσθητο σε παραμορφώσεις της έντασης σε ομοιογενείς περιοχές, είναι προτιμότερο ο χβαντιστής να προσαρμόζεται τοπικά στο περιεχόμενο της εικόνας, και το βήμα του χβαντισμού να είναι μικρότερο στις ομοιογενείς περιοχές, και μεγαλύτερο κοντά στις ακμές.

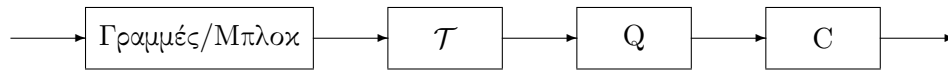
Χωρίς χβαντισμό πρόβλεψη χρησιμοποιείται στο σύστημα κωδικοποίησης JPEG για συμπίεση χωρίς απώλειες. Πρόβλεψη από εικόνα σε εικόνα μετά και από εκτίμηση της κίνησης χρησιμοποιείται στα συστήματα κωδικοποίησης βίντεο MPEG.

15.3 Συμπίεση με χρήση 2-Δ μετασχηματισμού

Η χρήση ενός 2-Δ ορθομοναδιαίου μετασχηματισμού επιτυγχάνει τη συμπίεση, χάρη στη σημαντική αποσυσχέτιση των δεδομένων και στη δυνατότητα ο χβαντισμός να γίνει για ένα ολόκληρο μπλοκ διάστασης $M_b \times N_b$. Τυπική διάσταση των μπλοκ που χρησιμοποιούνται για το μετασχηματισμό είναι 8×8 . Μεταξύ των μετασχηματισμών που παρουσιάστηκαν στο Κεφάλαιο 8 ο πλέον χρησιμοποιούμενος με σκοπό τη συμπίεση είναι ο διακριτός μετασχηματισμός συνημιτόνου (βλέπε σύστημα JPEG). Το κέρδος που επιτυγχάνεται με τη χρήση μετασχηματισμού, αν σ^2 είναι η διασπορά του αρχικού σήματος και σ_{mn}^2 είναι η διασπορά του συντελεστή του μετασχηματισμού στη θέση (m, n) , είναι

$$G_t = \frac{\sigma^2}{\left(\prod_{m=0}^{M_b-1} \prod_{n=0}^{N_b-1} \sigma_{mn}^2 \right)^{\frac{1}{M_b N_b}}} \quad (15.8)$$

πάντοτε μεγαλύτερο της μονάδας, αφού $\sum_{m=0}^{M_b-1} \sum_{n=0}^{N_b-1} \sigma_{mn}^2 = M_b N_b \sigma^2$.



Σχήμα 15.4: Συμπίεση με χρήση μετασχηματισμού (T : 2-Δ ορθομοναδιαίος μετασχηματισμός, Q : χβαντισμός, C : κωδικοποίηση).

Μπορεί κανείς να διακρίνει δύο τεχνικές για τον χβαντισμό και την κωδικοποίηση. Κατά την πρώτη χρησιμοποιείται διαφορετικός χβαντιστής για κάθε συντελεστή του μετασχηματισμού, και σταθερός αριθμός bits για κάθε επίπεδο χβαντισμού. Κατά τη δεύτερη κατά κανόνα χρησιμοποιείται ο ίδιος ομοιόμορφος χβαντιστής για όλους τους συντελεστές του μετασχηματισμού και μεταβλητός αριθμός bits ανάλογα με το επίπεδο χβαντισμού. Οι δύο αυτές τεχνικές παρουσιάζονται στη συνέχεια.

15.3.1 Σταθερός ρυθμός πληροφορίας

Έστω ότι ο ρυθμός μετάδοσης της πληροφορίας είναι σταθερός και καθορισμένος

$$R = \frac{1}{M_b N_b} \sum_{m=0}^{M_b-1} \sum_{n=0}^{N_b-1} B_{mn}, \quad (15.9)$$

όπου B_{mn} είναι ο αριθμός των bits που χρησιμοποιούνται για το συντελεστή του μετασχηματισμού στη θέση (m, n) . Ζητείται η βέλτιστη κατανομή των bits ανά συντελεστή του μετασχηματισμού, με την έννοια της ελαχιστοποίησης της μέσης παραμόρφωσης λόγω χβαντισμού

$$D = \frac{1}{M_b N_b} \sum_{m=0}^{M_b-1} \sum_{n=0}^{N_b-1} \sigma_{mn}^2 f(B_{mn}) \quad (15.10)$$

Η συνάρτηση $f(\cdot)$ χαρακτηρίζει τον χβαντιστή. Για παράδειγμα για τον ιδανικό χβαντιστή του Shannon είναι

$$f(x) = 2^{-2x}. \quad (15.11)$$

Η λύση του παραπάνω προβλήματος είναι γνωστή. Αν $h(\cdot)$ είναι η αντίστροφη συνάρτηση της $f'(\cdot)$, τότε ο αριθμός bits είναι

$$B_{mn} = \begin{cases} h\left(\frac{\theta f'(\theta)}{\sigma_{mn}^2}\right) & 0 < \theta < \sigma_{mn}^2 \\ 0 & \theta \geq \sigma_{mn}^2 \end{cases}, \quad (15.12)$$

όπου θ είναι η λύση της εξίσωσης

$$R = \frac{1}{M_b N_b} \sum_{\sigma_{mn}^2 \geq \theta} h\left(\frac{\theta f'(\theta)}{\sigma_{mn}^2}\right). \quad (15.13)$$

Για την περίπτωση του χβαντιστή του Shannon (15.11), το αποτέλεσμα είναι

$$B_{mn} = \max\left(0, \frac{1}{2} \log_2 \frac{\sigma_{mn}^2}{\theta}\right),$$

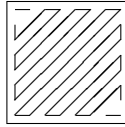
οπότε

$$D = \frac{1}{M_b N_b} \sum_{m=0}^{M_b-1} \sum_{n=0}^{N_b-1} \min(\theta, \sigma_{mn}^2).$$

Η Εξίσωση (15.12) δεν εξασφαλίζει ακέραιο αριθμό bits. Απαιτείται επομένως η στρογγύλευση του αποτελέσματος, για να έχουμε ακέραιο αριθμό bits ανά συντελεστή του μετασχηματισμού, και το δοσμένο ρυθμό μετάδοσης της πληροφορίας.

15.3.2 Μεταβλητός και προσαρμοζόμενος ρυθμός πληροφορίας

Η χρήση μεταβλητού αριθμού bits ανά μπλοκ δίνει τη δυνατότητα προσαρμογής του ρυθμού μετάδοσης της πληροφορίας στο περιεχόμενο του κάθε μπλοκ. Προς τούτο ο χβαντιστής είναι ομοιόμορφος και συμμετρικός με περιττό αριθμό επιπέδων χβαντισμού. Παρατηρείται ότι ένας μεγάλος αριθμός συντελεστών του μετασχηματισμού αντιστοιχεί στο ενδιάμεσο επίπεδο χβαντισμού με τιμή 0. Κατά κύριο λόγο οι συντελεστές αυτοί αντιστοιχούν σε υψηλούς δείκτες, που με τη σειρά τους αντιστοιχούν σε υψηλές συχνότητες για τους χρησιμοποιούμενους μετασχηματισμούς. Αρκεί να δοθεί το ορθογώνιο παραλληλόγραμμο που περιλαμβάνει όλους τους μη 'μηδενικούς' συντελεστές, για να καθορισθεί η τιμή όλων των υπόλοιπων, ή να χρησιμοποιηθεί μία αντι-διαγώνια σάρωση του μπλοκ, που να διέρχεται πρώτα από τους συντελεστές με χαμηλούς δείκτες, και να δοθεί η διεύθυνση του τελευταίου μη 'μηδενικού' συντελεστή (Σχήμα 15.5). Επιπλέον στη ζώνη που περιλαμβάνει τους μη 'μηδενικούς' συντελεστές μπορεί να υπάρχει ακόμη ικανός αριθμός 'μηδενικών' συντελεστών. Ο βαθμός συμπίεσης είναι υψηλότερος, εάν κωδικοποιηθούν τα διαστήματα 'μηδενικών' συντελεστών, αντί να κωδικοποιηθεί κάθε τιμή ξεχωριστά.



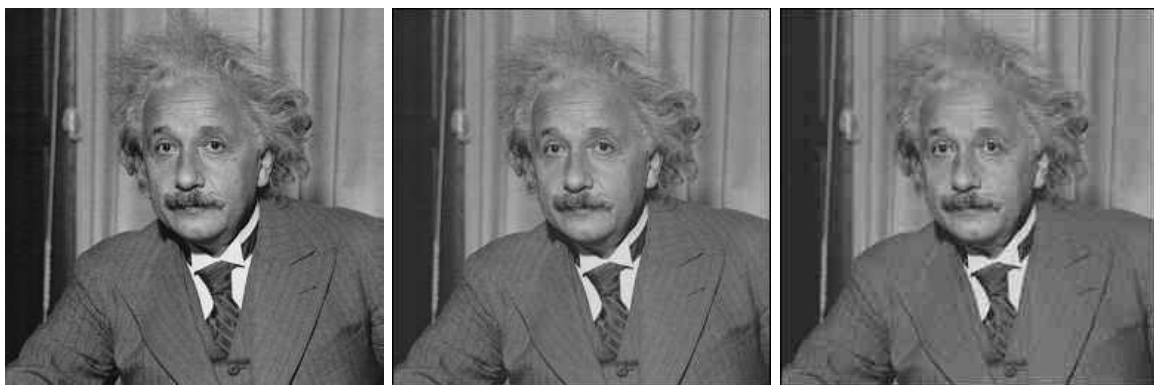
Σχήμα 15.5: Αντι-διαγώνια σάρωση του μπλοκ

	0	1	2	3	4	5	6	7
0	16	11	10	16	24	40	51	61
1	12	12	14	19	26	58	60	55
2	14	13	16	24	40	57	69	56
3	14	17	22	29	51	87	80	62
4	18	22	37	56	68	109	103	77
5	24	35	55	64	81	104	113	92
6	49	64	78	87	103	121	120	101
7	72	92	95	98	112	100	103	99

Πίνακας 15.1: Πίνακας κβαντισμού για ένα μπλοκ συντελεστών DCT 8×8 συνιστόμενος στο σύστημα JPEG

Η χρήση μεταβλητού αριθμού bits ανάλογα με το περιεχόμενο του κάθε μπλοκ αξιοποιεί επίσης τη χρήση ενός πίνακα κβαντισμού, που λαμβάνει υπόψη την απόκριση του ανθρώπινου συστήματος όρασης ανάλογα με τη συχνότητα, δηλαδή με τη θέση του συντελεστή. Ένα παράδειγμα τέτοιου πίνακα με διαφορετικό βήμα κβαντισμού ανάλογα με τη θέση του συντελεστή δίδεται στον Πίνακα 15.1.

Η μέθοδος αυτή κωδικοποίησης με μεταβλητό αριθμό bits ανά μπλοκ χρησιμοποιείται στο σύστημα συμπίεσης JPEG. Ένα παράδειγμα αποτελεσμάτων δίδεται στο Σχήμα 15.6.



Σχήμα 15.6: Αποτελέσματα συμπίεσης με το σύστημα JPEG. Αριστερά: η αρχική εικόνα, Κέντρο: συμπιεσμένη εικόνα με συντελεστή συμπίεσης 10, Δεξιά: συμπιεσμένη εικόνα με συντελεστή συμπίεσης 16.

15.4 Παράρτημα: Κωδικοποίηση Huffman

Ας είναι X μία διακριτή τυχαία μεταβλητή με πιθανότητες $\{p_1, \dots, p_N\}$ για τις τιμές της $\{x_1, \dots, x_N\}$. Το ζητούμενο είναι να ευρεθούν δυαδικοί κώδικες μεταβλητού μήκους $\{l_1, \dots, l_N\}$, για τις τιμές της τυχαίας μεταβλητής, ώστε το μέσο μήκος κωδικής λέξης

$$\bar{l} = \sum_{i=1}^N p_i l_i$$

να είναι ελάχιστο.

Η λύση δίδεται από τον αλγόριθμο του Huffman, που κατασκευάζει το ομώνυμο δυαδικό δένδρο για την παράσταση του κώδικα. Χωρίς βλάβη της γενικότητας μπορούμε να υποθέσουμε ότι οι πιθανότητες είναι διαταγμένες, $p_1 \geq p_2 \geq \dots \geq p_N$. Ο αλγόριθμος περιλαμβάνει δύο βήματα:

1. Λαμβάνονται οι δύο μικρότερες πιθανότητες, δημιουργείται ένας κόμβος γονέας, και οι δύο δημιουργούμενοι κλάδοι διακρίνονται με τον κώδικα '0' για τον ένα, και '1' για τον άλλο.
2. Αντικαθίστανται οι δύο πιθανότητες με το άθροισμά τους. Αν η λίστα των πιθανοτήτων περιλαμβάνει ένα μόνο στοιχείο, τέλος. Διαφορετικά, επιστροφή στο βήμα 1.