



Εισαγωγή στην ψηφιακή εικόνα

Σκοπός

Στο κεφάλαιο αυτό γίνεται περιγραφή των βασικών χαρακτηριστικών των ψηφιακών εικόνων. Τις ψηφιακές εικόνες θα τις θεωρήσουμε ως ψηφιακά σήματα στις δύο διαστάσεις. Θα γνωρίσουμε επίσης στοιχεία από τη λειτουργία του ανθρώπινου συστήματος όρασης, τα οποία θα βοηθήσουν στην καλύτερη κατανόηση της πληροφορίας που περιέχει η ψηφιακή εικόνα. Έμφαση θα δώσουμε στα είδη των πράξεων που μπορούν να γίνουν σε μία ψηφιακή εικόνα. Θα δούμε επίσης τι είναι φάσμα μιας εικόνας, τον τρόπο που υπολογίζεται και τη σημασία της φάσης στη σωστή αναπαράσταση της πληροφορίας που περιέχει η εικόνα.

Προσδοκώμενα αποτελέσματα

Όταν ολοκληρώσετε τη μελέτη του κεφαλαίου αυτού θα είστε σε θέση να:

- Εκτιμάτε το μέγεθος μιας ψηφιακής εικόνας
- Αναγνωρίζετε τον τύπο (format) με τον οποίο είναι αποθηκευμένη σε αρχείο μία εικόνα
- Κατανοείτε τη συμπεριφορά του ανθρώπινου συστήματος όρασης στις διαφορετικές αποχρώσεις της αμαύρωσης
- Εκτελείτε απλές πράξεις πάνω σε ψηφιακές εικόνες
- Εκτιμάτε το πλήθος των υπολογισμών που απαιτούνται για μία συγκεκριμένη διαδικασία επεξεργασίας της ψηφιακής εικόνας
- Κατανοείτε το φάσμα της εικόνας και τον σημαντικό ρόλο που παίζει η φάση του

Έννοιες κλειδιά

- Εικονοστοιχείο
- Διαβαθμίσεις της αμαύρωσης
- Δισδιάστατη μοναδιαία κρούση
- Δισδιάστατη κρουστική γραμμή
- Δισδιάστατη βαθμίδα
- Απόκτηση της εικόνας

- Αποκατάσταση της εικόνας
- Βελτίωση της εικόνας
- Ανάλυση της εικόνας
- Τεχνητή όραση
- Φωτεινότητα
- Λαμπρότητα
- Συνάρτηση μεταφοράς διαμόρφωσης
- Τοπικές πράξεις
- Καθολικές πράξεις
- Πράξεις σημείου
- Γεωμετρικές πράξεις
- Χωρικές συχνότητες
- Φάσμα εικόνας
- Φάση του Φάσματος της εικόνας

Εισαγωγικές Παρατηρήσεις

Η ψηφιακή εικόνα είναι ένα από τα δισδιάστατα ψηφιακά σήματα που παρουσιάζουν σημαντικό ενδιαφέρον στη σύγχρονη κοινωνία της πληροφορίας. Στο κεφάλαιο αυτό θα γνωρίσουμε τα στοιχειώδη συστατικά της ψηφιακής εικόνας, όπως είναι το οικονομικό, το μέγεθος της εικόνας, οι πράξεις που μπορούμε να κάνουμε πάνω σε αυτή καθώς και το πλήθος των υπολογισμών που απαιτούνται. Θα γνωρίσουμε επίσης πώς είναι το φάσμα μιας εικόνας και τον τρόπο που αυτό μπορεί να υπολογιστεί. Τέλος, θα γνωρίσουμε στοιχεία από τον τρόπο λειτουργίας του ανθρώπινου οπτικού συστήματος. Περισσότερες πληροφορίες ο αναγνώστης μπορεί να βρει στις αναφορές της ελληνικής και της ξενόγλωσσης βιβλιογραφίας.

Το κεφάλαιο αυτό περιλαμβάνει τις ακόλουθες ενότητες:

- 6.1. Η ψηφιακή εικόνα.*
- 6.2. Η φυσιολογία της όρασης.*
- 6.3. Είδη πράξεων στην επεξεργασία της ψηφιακής εικόνας.*
- 6.4. Το φάσμα της εικόνας.*

6.1 Η ψηφιακή εικόνα

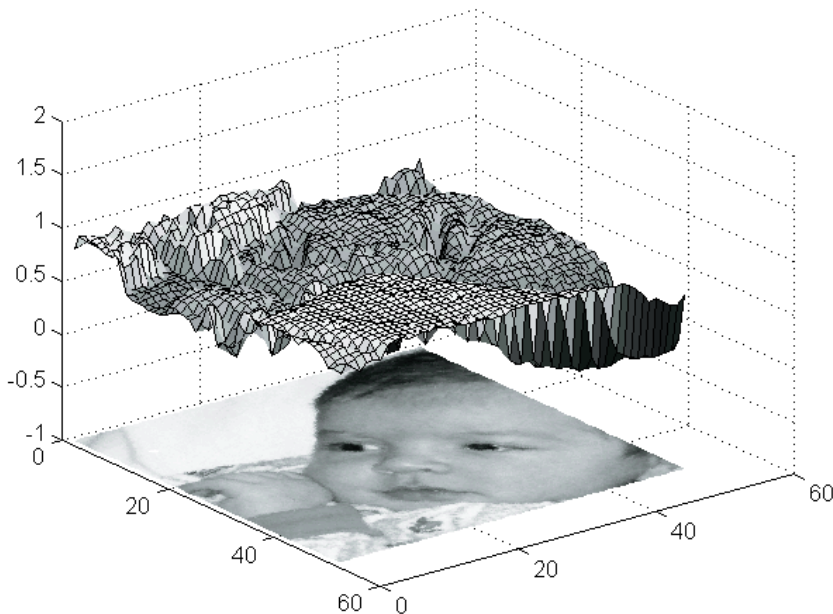
Η εικόνα αποτελεί στην εποχή μας μία από τις σημαντικότερες πηγές πληροφορίας. Τη συναντούμε ως εικόνα ακίνητη (φωτογραφία) ή κινούμενη (τηλεόραση), ασπρόμαυρη ή έγχρωμη. Η ψηφιακή εικόνα αποτελεί ό,τι πιο σύγχρονο, τόσο στο χώρο της ενημέρωσης (Internet) και της εκπαίδευσης (multimedia), όσο και στο χώρο του θεάματος και της ψυχαγωγίας (ψηφιακή τηλεόραση, DVD κτλ.). Στην παράγραφο αυτή θα μελετήσουμε εισαγωγικά θέματα που αφορούν την ψηφιακή εικόνα.

6.1.1 Η εικόνα ως δισδιάστατο σήμα

Μέχρι το προηγούμενο κεφάλαιο είχαμε θεωρήσει τα σήματά μας ως συναρτήσεις μιας μόνο ανεξάρτητης μεταβλητής, του χρόνου. Η ανάγκη για επεξεργασία της πληροφορίας της εικόνας οδηγεί κατ' αρχήν στη μαθηματική περιγραφή του περιεχομένου της. Έτσι, την ασπρόμαυρη εικόνα μπορούμε να τη δούμε ως μία δισδιάστατη συνάρτηση των συντεταγμένων του χώρου x και y :

$$f \rightarrow f(x, y) \quad (6.1)$$

Αυτό σημαίνει ότι η f περιγράφει μία επιφάνεια, η οποία έχει μεγάλη τιμή όπου η εικόνα είναι πιο λευκή, ενώ για μικρές τιμές της f η εικόνα πλησιάζει το μαύρο. Στην ουσία η f περιγράφει την αμαύρωση της εικόνας σε κάθε θέση (x, y) . Η αναπαράσταση αυτή φαίνεται στο Σχήμα 6.1, όπου παρουσιάζεται η εικόνα τμήματος προσώπου και η αντίστοιχη συνάρτηση της αμαύρωσης f .



Σχήμα 6.1

Η εικόνα προσώπου και η αντίστοιχη δισδιάστατη συνάρτηση της αμαύρωσης.

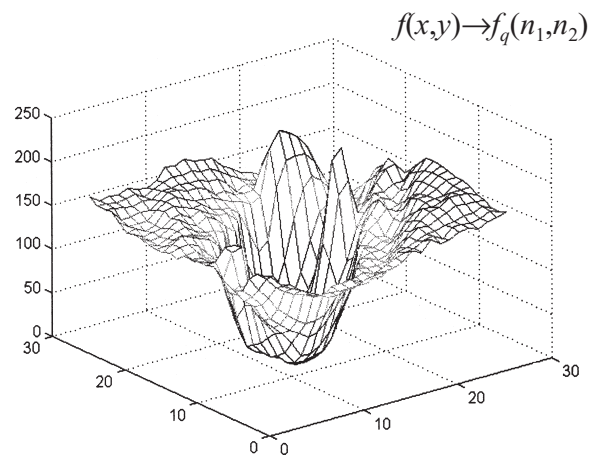
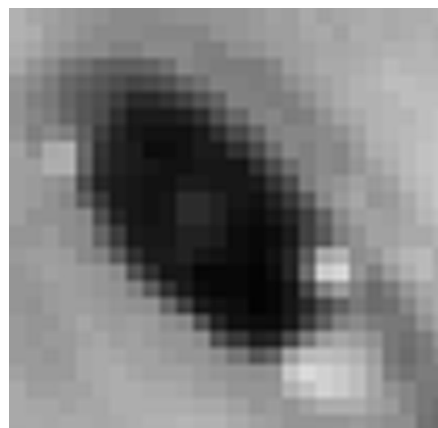
Στην πράξη, κάθε εικόνα για να υποστεί ψηφιακή επεξεργασία θα πρέπει κατ' αρχήν να μετατραπεί σε ψηφιακή. Η διαδικασία αυτή είναι περίπου η ίδια όπως και στην περίπτωση του μονοδιάστατου σήματος. Έτσι, θα πρέπει να λάβουμε ισαπέχοντα δείγματα της συνάρτησης $f(x,y)$ στις διευθύνσεις x και y . Η πυκνότητα με την οποία θα ληφθούν τα δείγματα καθορίζεται και στην περίπτωση αυτή (δισδιάστατη συνάρτηση στο χώρο) από το θεώρημα δειγματοληψίας:

Η απόσταση δύο διαδοχικών δειγμάτων στο επίπεδο (x,y) θα πρέπει να είναι μικρότερη από την ημιπερίοδο των ταχύτερων εναλλαγών της συνάρτησης $f(x,y)$.

Με άλλα λόγια, θα πρέπει να λαμβάνουμε τα δείγματα αρκετά πυκνά, ώστε να προλαβαίνουμε τις γρήγορες εναλλαγές της αμαύρωσης της εικόνας. Στη συνέχεια τα δείγματα θα πρέπει να κβαντιστούν σε πεπερασμένο αριθμό σταθμών. Συνήθως οι στάθμες που χρησιμοποιούνται είναι 256 (2^8). Στη στάθμη 255 αντιστοιχούμε το λευκό ενώ στη στάθμη 0 το μαύρο. Έτσι, ένας δυαδικός αριθμός των 8 bits (1 byte) επαρκεί για να περιγράψουμε την τιμή της αμαύρωσης ενός δείγματος της εικόνας, που καλείται και «εικονοστοιχείο» (picture element – pixel). Στην τελική της μορφή μία ασπρόμαυρη ψηφιακή εικόνα θα είναι όπως φαίνεται στο Σχήμα 6.2.

Σχήμα 6.2

Η ψηφιακή εικόνα αποτελείται από μικρές κουκίδες τα εικονοστοιχεία (pixels). Τα λευκά εικονοστοιχεία αντιστοιχούν στο 255 ενώ τα μαύρα στο 0. Καθένα από τα εικονοστοιχεία είναι ένα δείγμα από τη συνάρτηση $f(x,y)$ που αντιστοιχεί στην αναλογική εικόνα.



Η ψηφιακή εικόνα της μορφής του σχήματος 6.2 παριστάνεται μαθηματικά ως η κβαντισμένη σε πλάτος συνάρτηση

$$f \rightarrow f_q(n_1, n_2) \quad (6.2)$$

όπου οι διακριτές χωρικές μεταβλητές n_1 και n_2 αντιστοιχούν στις συνεχείς χωρικές μεταβλητές x και y . Στη γενική περίπτωση, η μέγιστη τιμή του n_1 είναι M και του n_2 είναι N . Έτσι, μπορούμε να θεωρήσουμε την ψηφιακή εικόνα ως μία μήτρα αριθμών διαστάσεων $M \times N$, που καθένας από αυτούς εκπροσωπεί την τιμή του αντίστοιχου εικονοστοιχείου. Για παράδειγμα, η εικόνα του Σχήματος 6.2 αποτελείται από $M \times N = 25 \times 25 = 625$ εικονοστοιχεία.

Σε ειδικού τύπου εικόνες, όπως είναι το κείμενο ή το σχέδιο, αρκούν δύο μόνο στάθμες, οι 0 και 1, που αντιστοιχούν στο μαύρο και το λευκό. Στην περίπτωση αυτή ένα μόνο δυαδικό ψηφίο είναι αρκετό για να αναπαραστήσει την πληροφορία. Μία τέτοιου τύπου εικόνα απεικονίζεται στο Σχήμα 6.3.

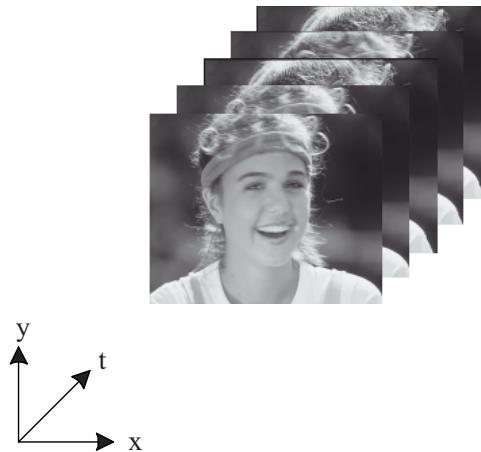
Σχήμα 6.3

Χειρόγραφη λέξη όπου η πληροφορία έχει καταγραφεί μόνο με μαύρα και λευκά εικονοστοιχεία (0 και 1).

Σε περίπτωση που η εικόνα είναι έγχρωμη τότε σε κάθε θέση (n_1, n_2) η f έχει τρεις τιμές (άνυσμα), οι οποίες αντιστοιχούν στις τιμές των χρωμάτων κόκκινο, πράσινο και μπλε (Red, Green and Blue – RGB). Όταν, τέλος, έχουμε διαδοχή εικόνων η συνάρτηση f έχει μία επιπλέον μεταβλητή, το χρόνο t . Το σήμα είναι διακριτό και ως προς το χρόνο ($t \rightarrow n_3$):

$$f \rightarrow f_q(n_1, n_2, n_3) \quad (6.3)$$

Αυτό σημαίνει ότι κάθε εικόνα απέχει χρονικά από την άλλη κάποιο χρονικό διάστημα. Το διάστημα αυτό θα πρέπει να είναι αρκετά μικρό, περίπου 1/30 sec, για να γίνει δυνατή η παρουσίαση της κίνησης. Η συνάρτηση αυτή στον τρισδιάστατο χώρο (x, y, t) φαίνεται στο Σχήμα 6.4.



Σχήμα 6.4

Η διαδοχική παρουσίαση εικόνων που δίνει την αίσθηση της κίνησης μπορεί να παρασταθεί μαθηματικά ως συνάρτηση τριών μεταβλητών.

6.1.2 Στοιχειώδη δισδιάστατα ψηφιακά σήματα

Ένα δισδιάστατο ψηφιακό σήμα μπορεί να παρασταθεί γραφικά όπως φαίνεται στο Σχήμα 6.5α. Το σήμα έχει τιμή μόνο για τα σημεία του πλέγματος (κβάντιση των x και y) καθώς επίσης και κβαντισμένο πλάτος. Τα πλέον στοιχειώδη δισδιάστατα ψηφιακά σήματα είναι η δισδιάστατη κρουστική $\delta(n_1, n_2)$ (Σχήμα 6.5β):

$$\delta(n_1, n_2) = \begin{cases} 1 & n_1 = n_2 = 0 \\ 0 & \text{οπουδήποτε αλλού} \end{cases} \quad (6.4)$$

η δισδιάστατη κρουστική γραμμή $\delta(n_1)$ ή $\delta(n_2)$ (Σχήμα 6.5γ και 6.5δ)

$$\delta(n_1) = \begin{cases} 1 & n_1 = 0 \\ 0 & \text{οπουδήποτε αλλού} \end{cases} \quad \delta(n_2) = \begin{cases} 1 & n_2 = 0 \\ 0 & \text{οπουδήποτε αλλού} \end{cases} \quad (6.5)$$

καθώς και η δισδιάστατη βηματική $u(n_1, n_2)$ (Σχήμα 6.5ε):

$$u(n_1, n_2) = \begin{cases} 1 & n_1 \geq 0 \text{ και } n_2 \geq 0 \\ 0 & \text{οπουδήποτε αλλού} \end{cases} \quad (6.6)$$

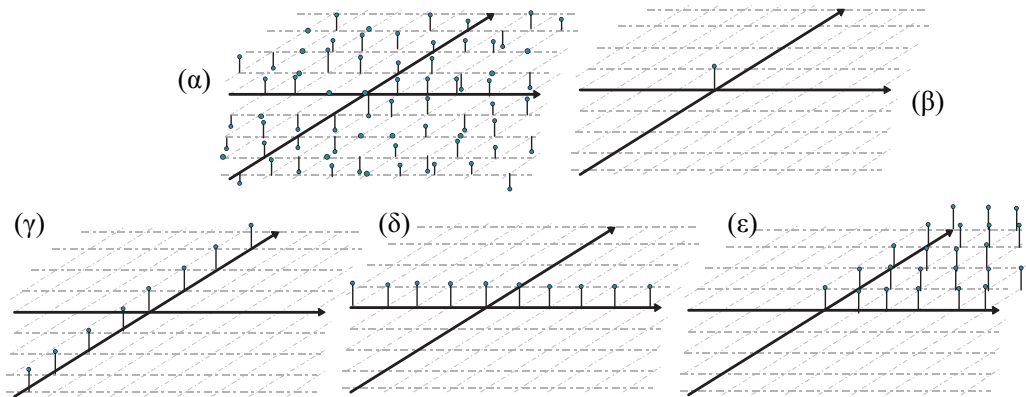
Χρησιμοποιώντας τα στοιχειώδη αυτά σήματα μπορούμε να εκφράσουμε κάθε δισδιάστατο ψηφιακό σήμα ως γραμμικό συνδυασμό τους. Έτσι, για το σήμα $f(n_1, n_2)$ ισχύει:

$$f(n_1, n_2) = \sum_{r_1=-\infty}^{\infty} \sum_{r_2=-\infty}^{\infty} f(r_1, r_2) \delta(n_1 - r_1, n_2 - r_2) \quad (6.7)$$

Το σήμα $f(n_1, n_2)$, δηλαδή, είναι γραμμικός συνδυασμός δισδιάστατων κρουστικών. Μία πολύ σημαντική ιδιότητα που είναι επιθυμητή για τα δισδιάστατα ψηφιακά σήματα είναι η διαχωρισιμότητα. Η ιδιότητα αυτή συνεπάγεται τη μαθηματική έκφραση ενός δισδιάστατου σήματος $f(n_1, n_2)$ ως γινομένου δύο μονοδιάστατων σημάτων:

$$f(n_1, n_2) = f_1(n_1) f_2(n_2) \quad (6.8)$$

Την ιδιότητα αυτή την έχει τόσο η $\delta(n_1, n_2)$, όσο και η $u(n_1, n_2)$.

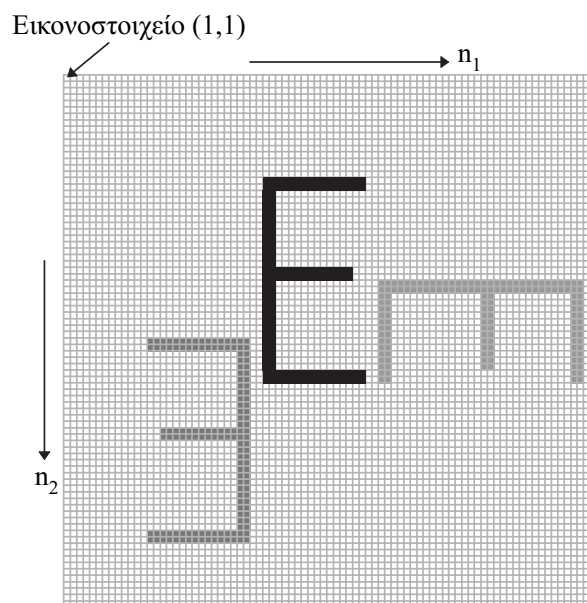
**Σχήμα 6.5**

α) Η μορφή ενός δισδιάστατου ψηφιακού σήματος. β) Δισδιάστατη κρουστική. γ) και δ) δισδιάστατη κρουστική γραμμή. ε) Δισδιάστατη βηματική.

6.1.3 Η γεωμετρία της εικόνας

Όπως προαναφέρθηκε, η εικόνα θεωρείται ως η κατανομή της πληροφορίας (βαθμός αμαύρωσης ή χρώμα) στο επίπεδο (x,y) . Τα περισσότερα οπτικά συστήματα (φωτογραφικές μηχανές, βιντεοκάμερες) αποδίδουν αναλογικές εικόνες, δηλαδή, εικόνες όπου οι συντεταγμένες (x,y) αλλά και ο βαθμός αμαύρωσης f είναι συνεχείς ποσότητες. Αντίθετα, οι ψηφιακές εικόνες, που πρόκειται να αποθηκευτούν στη μνήμη του υπολογιστή, ή να τύχουν επεξεργασίας από αυτόν, έχουν τιμή μόνο σε διακριτό αριθμό συντεταγμένων (n_1, n_2) , που αποτελούν και τις θέσεις των εικονοστοιχείων. Επίσης, κάθε εικονοστοιχείο λαμβάνει πεπερασμένο αριθμό τιμών αμαύρωσης.

Ο προσανατολισμός των συντεταγμένων σε μία ψηφιακή εικόνα μπορεί να είναι αυθαίρετος. Στις περισσότερες όμως περιπτώσεις και κυρίως στις ποικίλες εφαρμογές λογισμικού που κυκλοφορούν στο εμπόριο, η αρχή των συντεταγμένων $(n_1, n_2) = (1,1)$ λαμβάνεται στην επάνω αριστερή γωνία της εικόνας (Σχήμα 6.6).

**Σχήμα 6.6**

Συνήθως τα εικονοστοιχεία σε μία ψηφιακή εικόνα είναι διατεταγμένα σε σειρές και στήλες με πρώτο αυτό της επάνω αριστερής γωνίας.

Καθένα εικονοστοιχείο έχει τις δικές του συντεταγμένες. Συνεπώς, μία εικόνα που έχει M εικονοστοιχεία στη διεύθυνση n_1 και N στη διεύθυνση n_2 , περιέχει συνολικά $N \times M$ εικονοστοιχεία. Το μέγεθος της πληροφορίας της εικόνας είναι γενικά μεγάλο. Ακόμα και μικρές σε διαστάσεις εικόνες περιέχουν μεγάλο αριθμό εικονοστοιχείων, γεγονός που καθιστά χρονοβόρα την ψηφιακή τους επεξεργασία. Για το λόγο αυτό, ο έλεγχος αλγορίθμων επεξεργασίας εικόνας γίνεται αρχικά σε ένα μικρό τμήμα της εικόνας.

Παράδειγμα 6.1

Μία εικόνα με διαστάσεις 1.000×1.000 έχει συνολικά 1.000.000 εικονοστοιχεία. Η επεξεργασία του πλήθους των εικονοστοιχείων αυτών με πολύπλοκους αλγόριθμους απαιτεί μεγάλο πλήθος υπολογισμών. Πόσο λιγότεροι υπολογισμοί θα χρειαστούν αν δοκιμάσει κανείς τους αλγορίθμους αυτούς σε ένα τμήμα της εικόνας μεγέθους 50×50 εικονοστοιχείων;

Λύση

Το τμήμα αυτό της εικόνας περιέχει $50 \times 50 = 2.500$ εικονοστοιχεία, δηλαδή τετρακόσιες φορές λιγότερα απ' ότι σε ολόκληρη την εικόνα. Έτσι, οι απαιτούμενοι υπολογισμοί είναι τετρακόσιες φορές λιγότεροι. Από άποψη χρόνου, αν ολόκληρη η εικόνα απαιτούσε μία ολόκληρη ώρα για τη συγκεκριμένη ψηφιακή επεξεργασία, η δοκιμή των αλγορίθμων μόνο στο συγκεκριμένο τμήμα της χρειάζεται εννέα μόλις δευτερόλεπτα.

Δραστηριότητα 6.1

Ποιος είναι ο αριθμός των εικονοστοιχείων μιας εικόνας μεσαίων διαστάσεων; Πόσο χώρο στη μνήμη του υπολογιστή απαιτεί η εικόνα αυτή αν η πληροφορία κάθε εικονοστοιχείου είναι κωδικοποιημένη με δυαδικό αριθμό του ενός byte; Πόσος είναι ο χώρος αυτός αν η ίδια η εικόνα είναι έγχρωμη;

Μία τυπική εικόνα μεσαίων διαστάσεων έχει πλήθος εικονοστοιχείων $512 \times 512 = 262144$. Αν κάθε εικονοστοιχείο χρειάζεται ένα byte μνήμης για αποθήκευση (με ένα byte μπορούν να αναπαρασταθούν 256 διαβαθμίσεις του γκριζου), τότε η συνολική εικόνα απαιτεί 262.144 bytes ή 0,25 Mbytes μνήμης. Στην περίπτωση που η εικόνα είναι έγχρωμη, κάθε εικονοστοιχείο θέλει 3 bytes για αποθήκευση (1 για κάθε χρώμα από τα κόκκινο, πράσινο και μπλε). Για ολόκληρη την εικόνα χρειάζονται συνολικά 0,75 Mbytes μνήμης.

Άσκηση Αυτοαξιολόγησης 6.1

Μία ασπρόμαυρη εικόνα (1 byte/εικονοστοιχείο) 128×128 εικονοστοιχείων συγκριτικά με μία έγχρωμη 512×512 εικονοστοιχείων, απαιτεί για αποθήκευση:

- α. 4 φορές λιγότερη μνήμη.
- β. 16 φορές λιγότερη μνήμη.
- γ. 48 φορές λιγότερη μνήμη.
- δ. 64 φορές λιγότερη μνήμη.

Ποια από τις απαντήσεις είναι η σωστή.

Οι σύγχρονες οθόνες των υπολογιστών έχουν τη δυνατότητα να επιδείξουν ψηφιακές εικόνες μεγάλων διαστάσεων. Η δυνατότητα αυτή συνδέεται άμεσα με την ανάλυση της οθόνης, ή αλλιώς με τον αριθμό των εικονοστοιχείων που μπορούν να εμφανίσουν. Η απλούστερη έγχρωμη οθόνη, με τη συνοδεία των αντίστοιχων ηλεκτρονικών κυκλωμάτων (κάρτα Video Graphics Adapter – VGA), μπορεί να παρουσιάσει $800 \times 600 = 480.000$ έγχρωμα εικονοστοιχεία. Με διάφορες τεχνικές είναι δυνατό να παρουσιάσουμε σε μία τέτοια οθόνη πολύ μεγαλύτερες εικόνες.

Τέλος, πρέπει να σημειώσουμε ότι ο απαιτούμενος χώρος σε bytes για να αποθηκευτεί η εικόνα εξαρτάται τόσο από το πλήθος των εικονοστοιχείων, όσο και από την πολυπλοκότητα στη διάταξή τους. Έτσι, αν πρόκειται για εικόνα που περιέχει σχέδιο ή κείμενο κάθε εικονοστοιχείο μπορεί να είναι μαύρο ή άσπρο, οπότε χρειάζεται μόνο 1 bit για την πληροφορία αυτή. Αν η εικόνα περιέχει ασπρόμαυρη φωτογραφία με 256 διαβαθμίσεις στην αμαύρωση, τότε χρειάζεται ένα byte για την αποθήκευση κάθε εικονοστοιχείου. Στις έγχρωμες εικόνες θέλουμε 3 bytes για κάθε εικονοστοιχείο. Με σύγχρονες μεθόδους συμπίεσης της πληροφορίας (image compression) είναι δυνατόν ο απαιτούμενος χώρος αποθήκευσης να είναι αρκετά μικρότερος.

Παράδειγμα 6.2

Πόσο χώρο σε bytes απαιτεί για την αποθήκευσή της μία εικόνα διαστάσεων 1024×1024 εικονοστοιχείων, όταν περιέχει :

- α. Κείμενο
- β. Ασπρόμαυρη εικόνα με 256 διαβαθμίσεις του γκριζου.
- γ. Έγχρωμη εικόνα

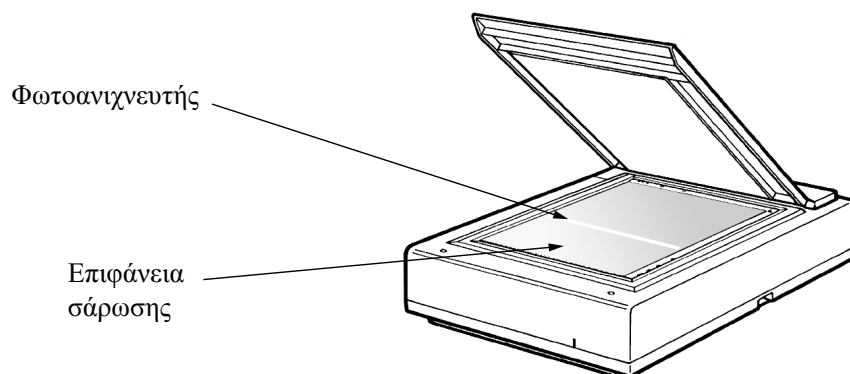
Λύση

- α. Στην περίπτωση κειμένου κάθε εικονοστοιχείο χρειάζεται 1 bit για να αποθηκευτεί. Έτσι, τα bits που απαιτούνται είναι $1.024 \times 1.024 \times 1 = 1.048.576$ ή 131.072 bytes που ισοδυναμούν με 128 kbytes.
- β. Στην περίπτωση της ασπρόμαυρης εικόνας χρειάζεται 1 byte για κάθε εικονοστοιχείο, οπότε 1.024×1.024 bytes για όλη την ασπρόμαυρη εικόνα που ισοδυναμούν με 1 Mbyte μνήμης.
- γ. Για την έγχρωμη εικόνα το ποσό αυτό είναι τριπλάσιο.

6.1.4 Τρόποι απόκτησης της ψηφιακής εικόνας

Η ψηφιακή εικόνα που, όπως αναφέραμε, αποτελεί ένα δισδιάστατο διακριτό σήμα στο χώρο (x,y) με κβαντισμένες τιμές πλάτους, μπορεί να ληφθεί είτε από αναλογικές εικόνες είτε απευθείας από συστήματα λήψης ψηφιακών εικόνων. Στην πρώτη περίπτωση, η αναλογική εικόνα θα μετατραπεί σε ψηφιακή με τη βοήθεια (το πιθανότερο) ενός σαρωτή εικόνας (scanner). Η συσκευή αυτή σαρώνει την εικόνα από πάνω προς τα κάτω σε γραμμές με τη βοήθεια ενός επιμήκους φωτο-ανιχνευτή (Σχήμα 6.7). Ο φωτο-ανιχνευτής διαθέτει λαμπτήρα λευκού φωτός, που φωτίζει μία στενή λωρίδα της εικόνας. Το ανακλώμενο φως καταγράφεται από τα φωτο-ευαίσθητα κύτταρα του ανιχνευτή που βρίσκονται σε γραμμική διάταξη. Οι σύγχρονοι σαρωτές εικόνας αποδίδουν πολύ καλή ανάλυση στην εικόνα. Η ανάλυση αυτή εκφράζεται σε εικονοστοιχεία ανά ίντσα (pixels per inch – ppi). Τυπικές τιμές αυτής της ανάλυσης είναι 75, 150, 300, 600 ή και 1.440 ppi.

Σχήμα 6.7
Σαρωτής εικόνας
της εταιρίας
Hewlett Packard
(μοντέλο HP
9195A).



Αν πρόκειται για αναλογική εικόνα από video (διαδοχή εικόνων), τότε η μετατροπή σε ψηφιακή εικόνα γίνεται με ταχύτατα ηλεκτρονικά κυκλώματα, τους frame grabbers. Τα κυκλώματα αυτά εκτελούν τη μετατροπή κάθε αναλογικής εικόνας σε

ψηφιακή σε πραγματικό χρόνο (30 στιγμιότυπα/δευτερόλεπτο). Παρότι η ανάλυσή τους σε εικονοστοιχεία δε φθάνει αυτήν που αποδίδει ένας σαρωτής, σε γενικές γραμμές θεωρείται πολύ ικανοποιητική και δίνει εικόνες πολύ καλής ποιότητας.

Τα συστήματα απευθείας λήψης ψηφιακής εικόνας αναπτύσσονται τον τελευταίο καιρό, και η τεχνολογία τους βρίσκεται σε εξέλιξη. Έτσι, οι ψηφιακές φωτογραφικές μηχανές, που ήδη κυκλοφορούν, έχουν ενσωματωμένη μνήμη RAM και δισκέτα για καταχώρηση της πληροφορίας. Το είδωλο της εικόνας δεν καταγράφεται πάνω σε φιλμ, αλλά σε μία μήτρα στοιχειωδών φωτο-ανιχνευτών (λ.χ. 600×350), το σήμα των οποίων κβαντίζεται και καταχωρείται στη μνήμη της μηχανής. Οι εικόνες αυτές έχουν φθάσει ήδη την ποιότητα της κλασικής φωτογραφίας, αφού η ανάλυσή τους έχει υπερβεί τα 1.024×1.024 εικονοστοιχεία. Το κυριότερο πλεονέκτημα των ψηφιακών συστημάτων είναι ότι παρέχουν την πληροφορία σε ψηφιακή μορφή, που μπορεί εύκολα να αντιγραφεί, να διαμορφωθεί και κυρίως να παραμένει αναλλοίωτη στο χρόνο λόγω της ψηφιακής της υφής.

Τα συστήματα λήψης ψηφιακής εικόνας παρέχουν τη δυνατότητα στο χρήστη να λάβει την ψηφιακή εικόνα με διαφορετικές αναλύσεις. Έτσι, του δίνεται η δυνατότητα να δοκιμάσει το ανεκτό όριο τόσο στην ανάλυση, όσο και στο πλήθος των διαβαθμίσεων της αμαύρωσης. Στο Σχήμα 6.8 δίνεται διαδοχικά η εικόνα του προσώπου η οποία έχει ληφθεί με διαφορετική κάθε φορά ανάλυση και πλήθος διαβαθμίσεων του γκριζου. Η ποιότητα της εικόνας υποβαθμίζεται με την ελάττωση της ανάλυσης, αλλά ταυτόχρονα μειώνεται και η απαιτούμενη μνήμη για αποθήκευσή της.



α



β



γ



δ



ε

Σχήμα 6.8

Οι πέντε εικόνες του προσώπου έχουν ληφθεί με διαφορετικές αναλύσεις και απαιτούν διαφορετική μνήμη.

(α) 300 εικονοστοιχεία/ίντσα με 256 διαβαθμίσεις της αμαύρωσης απαιτούν για αποθήκευση 186 kbytes μνήμης

(β) 75 εικονοστοιχεία/ίντσα με 256 διαβαθμίσεις της αμαύρωσης απαιτούν για αποθήκευση 12 kbytes μνήμης.

(γ) 25 εικονοστοιχεία/ίντσα με 256 διαβαθμίσεις της αμαύρωσης απαιτούν για αποθήκευση 2 kbytes μνήμης.

(δ) 75 εικονοστοιχεία/ίντσα με 16 διαβαθμίσεις της αμαύρωσης απαιτούν για αποθήκευση 6 kbytes μνήμης.

(ε) 25 εικονοστοιχεία/ίντσα με 2 διαβαθμίσεις της αμαύρωσης απαιτούν για αποθήκευση 0,25 kbytes μνήμης.

Δραστηριότητα 6.2

Πόση μνήμη απαιτείται για να φυλαχτεί ως εικόνα η παρούσα σελίδα του βιβλίου. Εκτιμήστε μόνοι σας την απαιτούμενη ανάλυση σε ppi του σαρωτή εικόνας.

Για μία αποδεκτή ποιότητα εικόνας είναι αρκετά 100 ppi. Η ανάλυση αυτή αντιστοιχεί σε 100 εικονοστοιχεία/ίντσα σε κάθε κατεύθυνση. Αυτό σημαίνει ότι σε μία τετραγωνική ίντσα περιέχονται 10.000 εικονοστοιχεία. Μία σελίδα χαρτιού A4 είναι περίπου 90 τετραγωνικές ίντσες (8×11), επομένως θα έχουμε συνολικά 900.000 εικονοστοιχεία. Κάθε ένα εικονοστοιχείο απαιτεί μόνο ένα bit για την αντιπροσωπευσή του (μαύρο ή άσπρο). Επομένως η συνολική μνήμη θα είναι 900.000 bits ή 112.500 bytes.

Φυσικά, διαφορετικές απαιτήσεις στην ανάλυση της εικόνας θα οδηγήσουν σε διαφορετικά μεγέθη.

Οι ψηφιακές εικόνες είναι οργανωμένες και αποθηκεύονται σε αρχεία (files) διαφόρων ειδών ανάλογα με τον τρόπο και τη συσκευή απόκτησής τους, αλλά και την ανάγκη για γρήγορη μετάδοση ή οικονομική αποθήκευση. Έτσι, στην αρχή του αρχείου υπάρχει συνήθως η πληροφορία (header) για τον τρόπο που έχει καταγραφεί η εικόνα στο αρχείο. Οι κυριότεροι τύποι αρχείων που περιέχουν ψηφιακές εικόνες είναι οι ακόλουθοι:

im名称.bmp

Τυποποίηση για έγχρωμες εικόνες σε Windows. Τα γράμματα της κατάληξης του ονόματος του αρχείου προέρχονται από τον όρο «Bit Mapped», που σημαίνει αντιστοιχία εικονοστοιχείου και byte. Είναι αρχεία ανεξάρτητα της συσκευής (device independent), μπορούν δηλαδή να «ανοιχτούν» χωρίς πρόβλημα από διαφορετικού τύπου λογισμικό, χωρίς αλλαγές.

im名称.tif

Tagget Image File Format. Είναι και αυτή τυποποίηση bit mapped που χρησιμοποιείται από όλα σχεδόν τα πακέτα λογισμικού για εικόνες, καθώς και όλα τα λειτουργικά συστήματα. Η διαφορά με την προηγούμενη τυποποίηση είναι ότι εδώ χρησιμοποιείται πίνακας αναφοράς για την επιλογή των χρωμάτων.

im名称.jpg

Συμπίεσμένες εικόνες σύμφωνα με το πρότυπο JPEG (Joint Photographic Experts Group). Οι συμπίεσμένες εικόνες καταλαμβάνουν λιγότερο χώρο μνήμης και επο-

μένως, μεταδίδονται πιο γρήγορα μέσα στο δίκτυο. Κατά τη συμπίεση ένα μέρος της πληροφορίας στην εικόνα χάνεται, επομένως η ανάκτησή της δεν γίνεται με την αρχική ακρίβεια.

image_name.gif

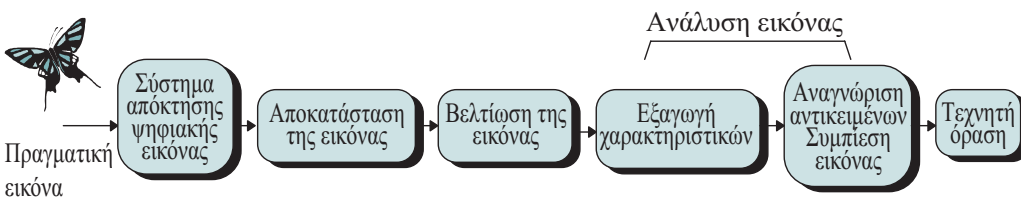
Ο τύπος αυτός αφορά τις εικόνες που είναι κωδικοποιημένες σύμφωνα με το Graphics Interchange Format, μία τυποποίηση για έγχρωμες εικόνες που παρουσιάζονται ή μεταφέρονται σε δικτυακές διασυνδέσεις και στο Διαδίκτυο (Internet).

6.1.5 Η επεξεργασία της ψηφιακής εικόνας

Η επεξεργασία της ψηφιακής εικόνας περιλαμβάνει στη γενικότερη περίπτωση πολλά και διαφορετικά στάδια ανάλογα με

- τη μορφή και την αρχική κατάσταση της ψηφιακής εικόνας
- το επιθυμητό τελικό αποτέλεσμα.

Στο Σχήμα 6.9 δίνεται ένα γενικό διάγραμμα που περιλαμβάνει πολλά από τα στάδια αυτά καθώς και τη μεταξύ τους σχέση. Έτσι, η πραγματική εικόνα θα καταγραφεί από το σύστημα απόκτησης της ψηφιακής εικόνας (image acquisition system). Θα ακολουθήσει το στάδιο της εξαγωγής του θορύβου (filtering) και της διόρθωσης (αποκατάστασης – restoration) σφαλμάτων που προκλήθηκαν είτε από το μέσο μεταξύ εικόνας και συστήματος απόκτησης είτε και από το ίδιο το σύστημα απόκτησης. Τα σφάλματα αυτά θα πρέπει να διορθωθούν πριν προχωρήσουμε σε επόμενο στάδιο στην επεξεργασία της εικόνας. Σφάλματα που παρουσιάζονται στην εικόνα κατά τη διάρκεια της απόκτησής της είναι ο θερμικός θόρυβος ή/και ο θόρυβος αιχμών (από το σύστημα απόκτησης) και η θόλωση της εικόνας. Η θόλωση προκαλείται συνήθως από την οπτική ανομοιογένεια του μέσου (αέρας) ή τη σχετική κίνηση της εικόνας ως προς το σύστημα απόκτησης.



Σχήμα 6.9

Γενικό διάγραμμα των διαφόρων σταδίων στην επεξεργασία της ψηφιακής εικόνας.

Οι αιτίες δημιουργίας σφαλμάτων κατά την απόκτηση της εικόνας έχουν καταγραφεί στη βιβλιογραφία εκτενώς. Καλύτερος τρόπος για τη διόρθωση των σφαλμάτων αυτών είναι να περιγραφεί μαθηματικά η αιτία που τα προκαλεί και να υλοποιηθεί

με το σύστημα αποκατάστασης η αντίστροφη διαδικασία. Υπάρχουν φυσικά περιπτώσεις όπου η διαδικασία απόκτησης είναι αρκετά καλή, με αποτέλεσμα η εικόνα να μη χρειάζεται αποκατάσταση. Σε άλλες περιπτώσεις το στάδιο της αποκατάστασης είναι το τελευταίο πριν από την επίδειξη της εικόνας (τηλεόραση).

Η βελτίωση της εικόνας αφορά κυρίως στον τονισμό συγκεκριμένων στοιχείων της εικόνας με σκοπό τη διευκόλυνση των διαδικασιών που θα ακολουθήσουν. Τα επόμενα στάδια εντάσσονται σε ό, τι καλούμε «ανάλυση της εικόνας». Σημαντικά θέματα εδώ είναι:

- Η εξαγωγή των κατάλληλων χαρακτηριστικών για αναγνώριση και ταξινόμηση αντικειμένων που υπάρχουν μέσα στην εικόνα.
- Η συμπίεση της εικόνας για γρήγορη μετάδοσή της ή οικονομική αποθήκευση.
- Ο μετασχηματισμός της σε άλλους χώρους, όπως αυτός των συχνοτήτων, για την αποκάλυψη πληροφορίας που στον αρχικό χώρο δεν είναι εμφανής.
- Οι γεωμετρικοί μετασχηματισμοί που είναι απαραίτητοι για τη σύγκριση εικόνων, τη σύνθεση της πληροφορίας σκηνών που έχουν ληφθεί από διαφορετική γωνία καθώς και τη σύνθεση εγκάρσιων τομών αντικειμένων από προβολικές εικόνες (τομοσύνθεση και αξονική τομογραφία).

Το τελευταίο στάδιο είναι η τεχνητή όραση. Εδώ, περιλαμβάνονται όλα τα αυτοματοποιημένα συστήματα επεξεργασίας εικόνας, αναγνώρισης αντικειμένων και λήψης αποφάσεων για συγκεκριμένες ενέργειες (robot) (Βλέπε «Ανάλυση εικόνας και Αναγνώριση Προτύπων», Ε. Σαγκριώτη και Σ. Θεοδωρίδη) .

Σύνοψη ενότητας

Στην ενότητα αυτή γνωρίσαμε την εικόνα ως δισδιάστατο σήμα, τη γεωμετρία της, καθώς και τους τρόπους που μπορεί να αποκτηθεί. Είδαμε επίσης αναλυτικά πώς μπορούμε να υπολογίσουμε το μέγεθος μιας εικόνας και το χώρο μνήμης που απαιτείται για την αποθήκευσή της. Τέλος, γνωρίσαμε μερικά από τα διαφορετικά στάδια από τα οποία μπορεί να αποτελείται η διαδικασία της επεξεργασίας της ψηφιακής εικόνας.

6.2 Η φυσιολογία της όρασης

Η όραση είναι ίσως η πλέον σημαντική από τις ανθρώπινες αισθήσεις, αφού δίνει τη δυνατότητα στον άνθρωπο να αντιληφθεί και να δράσει στο χώρο που τον περιβάλλει. Το ανθρώπινο σύστημα όρασης αποτελείται από τα ακόλουθα μέρη:

- Τον οφθαλμό (οπτικός αισθητήρας) που χρησιμεύει στη δημιουργία του ειδώλου.
- Το οπτικό νεύρο που μεταδίδει το ερέθισμα στον εγκέφαλο.
- Το τμήμα του εγκεφάλου που εκτελεί την επεξεργασία και την τελική κατανόηση της πληροφορίας

Η διαδικασία της όρασης είναι πολύπλοκη και δεν έχει ακόμη κατανοηθεί πλήρως. Η καλύτερη κατανόηση του τρόπου με τον οποίο ο άνθρωπος βλέπει γύρω του, θα συμβάλει στη βελτίωση των τεχνητών οπτικών συστημάτων απόκτησης και επεξεργασίας της εικόνας. Επιπλέον, θα δώσει ώθηση στην ανάπτυξη νέων αλγορίθμων για την ανάλυση της εικόνας και τη βελτίωση των συστημάτων μηχανικής όρασης (robot).

Η καταγραφή της συμπεριφοράς του ανθρώπινου συστήματος όρασης έχει γίνει με κάθε λεπτομέρεια. Έτσι, γνωρίζουμε:

- Πόσο φως είναι απαραίτητο για να διεγερθεί ο οφθαλμός.
- Σε πόσο έντονη ακτινοβολία μπορεί να αντέξει.
- Πόσο γρήγορα μπορεί να προσαρμοστεί στις αλλαγές της φωτεινότητας των αντικειμένων.
- Πότε και πόσο καλά μπορεί να διακρίνει τα χρώματα.
- Πώς αντιλαμβάνεται τα σχήματα παρουσία ή απουσία άλλων σχημάτων.
- Πώς γίνεται δυνατή η αίσθηση του βάθους.

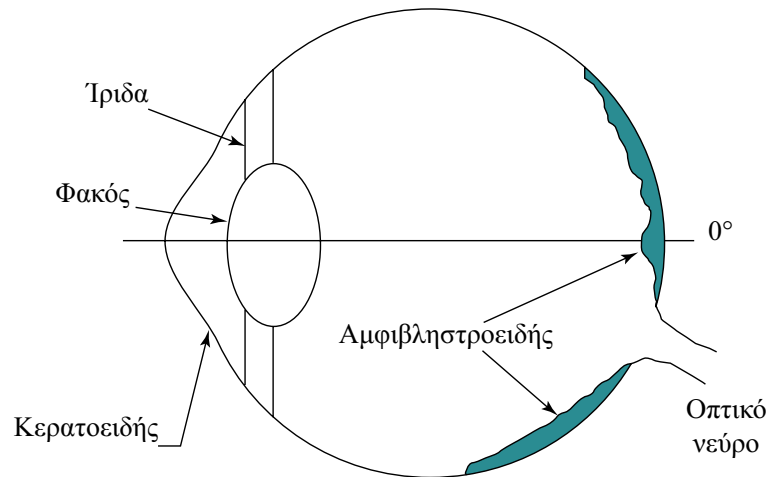
Με βάση αυτή την καταγραφή οι επιστήμονες προσπαθούν να περιγράψουν και να εξηγήσουν τη διαδικασία της όρασης. Στη συνέχεια, θα παραθέσουμε τα πλέον σημαντικά από τα στοιχεία αυτά, αφού πρώτα δώσουμε μια απλή περιγραφή της φυσιολογίας του οφθαλμού.

6.2.1 Ο ανθρώπινος οφθαλμός

Στο Σχήμα 6.10 παρουσιάζεται μια απλουστευμένη εικόνα του οφθαλμού. Στο μπροστινό μέρος βρίσκεται το σύστημα συγκέντρωσης και εστίασης των εικόνων (ίριδα, φακός). Το είδωλο της εικόνας σχηματίζεται ανεστραμμένο στο πίσω μέρος του οφθαλ-

μού πάνω στον αμφιβληστροειδή χιτώνα (retina). Στο σημείο της εστίασης υπάρχει μεγάλη συγκέντρωση των φωτοαισθητήρων. Δίπλα ακριβώς ξεκινά το οπτικό νεύρο που μεταφέρει τις ηλεκτροχημικές αντιδράσεις των φωτοαισθητήρων στον εγκέφαλο.

Σχήμα 6.10
Απλουστευμένη
εικόνα
του οφθαλμού.



Υπάρχουν δύο είδη φωτοαισθητήρων στον αμφιβληστροειδή, τα ραβδία και οι κώνοι. Τα ραβδία είναι περίπου 100 εκατομμύρια και σχετίζονται με την όραση στο χαμηλό φως (scotopic vision). Οι κώνοι είναι γύρω στα έξι εκατομμύρια και είναι λιγότερο ευαίσθητοι από τα ραβδία. Σε φωτεινό περιβάλλον το κύριο βάρος της όρασης φέρουν οι κώνοι. Οι κώνοι επίσης είναι υπεύθυνοι για το χρώμα στην όραση. Το οπτικό σύστημα του ανθρώπου είναι ευαίσθητο στην ηλεκτρομαγνητική ακτινοβολία και σε μήκη κύματος από 350 έως 780 nm (ορατό φως). Παρουσιάζει όμως μέγιστη ευαισθησία στο πράσινο φως (γύρω στα 550 nm). Επίσης, είναι ικανό να λειτουργήσει με ελάχιστο φως (ραβδία), αλλά και σε πολύ φωτεινό περιβάλλον (κώνοι). Η ικανότητα του οφθαλμού να προσαρμόζεται σε μεγάλη διαφορά φωτεινότητας φθάνει τις 10 τάξεις μεγέθους:

$$\frac{\text{Μέγιστη φωτεινή ενέργεια ανεκτή από το μάτι}}{\text{Ελάχιστη φωτεινή ενέργεια που διεγείρει το μάτι}} \approx 10^{10} \quad (6.9)$$

6.2.2 Φωτεινότητα, λαμπρότητα και αντίθεση

Δύο φυσικά μεγέθη, που σχετίζονται με τη φωτεινή ακτινοβολία και παίζουν σημαντικό ρόλο στην ανθρώπινη όραση είναι η φωτεινότητα και η λαμπρότητα.

- Η φωτεινότητα ενός αντικειμένου (luminance) σχετίζεται με τη φωτεινή ενέργεια που εκπέμπει το αντικείμενο, ανεξάρτητα από τη φωτεινότητα των γειτονικών σωμάτων.

- Η λαμπρότητα ενός αντικειμένου (brightness) είναι η λαμβανόμενη από το μάτι φωτεινή ενέργεια που προέρχεται από το αντικείμενο και η οποία εξαρτάται από τη φωτεινότητα του περιβάλλοντος χώρου.

Τον διαφορετικό τρόπο (λαμπρότητα) που αντιλαμβάνεται το μάτι τα αντικείμενα της ίδιας φωτεινότητας (luminance) μπορούμε να τον δούμε στο παράδειγμα του Σχήματος 6.11. Οι δύο μικροί δίσκοι έχουν την ίδια φωτεινότητα, αλλά στο μάτι δείχνουν διαφορετικοί. Ο πρώτος φαίνεται περισσότερο φωτεινός, αφού έχει σκοτεινότερο υπόβαθρο.



Σχήμα 6.11

Οι δύο δίσκοι έχουν την ίδια φωτεινότητα, αλλά φαίνονται διαφορετικοί λόγω της διαφορετικής φωτεινότητας του υπόβαθρου.

Πρέπει να σημειωθεί ότι σύμφωνα με το νόμο του Weber, ένα αντικείμενο με φωτεινότητα f ξεχωρίζει από ένα περιβάλλον (background) φωτεινότητας f_0 όταν η σχετική τους διαφορά (αντίθεση ή contrast) είναι το λιγότερο 2%:

$$\frac{f_0 - f}{f_0} \geq 2\% \quad (6.10)$$

Η τελευταία αυτή σχέση δηλώνει ότι μεταξύ απολύτως μαύρου και απολύτως λευκού, το ανθρώπινο μάτι μπορεί να ξεχωρίσει το πολύ 50 διαβαθμίσεις του γκριζου (βλέπε Σχήμα 6.12). Η τελευταία σχέση φανερώνει τη μη γραμμική (λογαριθμική) συμπεριφορά στην αντίληψη της αμαύρωσης από το ανθρώπινο σύστημα όρασης αφού:

$$\frac{\Delta f}{f} = a \Rightarrow d(\log f) = a \quad (6.11)$$

Δηλαδή, ίσες αυξήσεις στο λογάριθμο της φωτεινότητας εκλαμβάνονται από τον άνθρωπο έως ίσες διαφορές.

Σχήμα 6.12

Το ανθρώπινο μάτι μπορεί να διακρίνει σε ασπρόμαυρες εικόνες το πολύ μέχρι 50 διαφορετικές διαβαθμίσεις του γκριζου. Στην πρώτη λωρίδα παρουσιάζονται 64 διαβαθμίσεις (δύσκολα διακρίνονται μεταξύ τους) από το μαύρο έως το λευκό, ενώ στη δεύτερη 16.

**Παράδειγμα 6.3**

Τρία αντικείμενα α , β και γ έχουν φωτεινότητα 1, 10 και 100 αντίστοιχα, (εκφρασμένη πάντα σε κατάλληλες μονάδες). Πόσο φωτεινά αντιλαμβάνεται ο άνθρωπος τα αντικείμενα αυτά;

Λύση

Με βάση τη σχέση (6.11), έχουμε για τους λογαρίθμους των φωτεινοτήτων:

$$\log(1) = 0$$

$$\log(10) = 1$$

$$\log(100) = 2$$

Επομένως, ο άνθρωπος αντιλαμβάνεται ότι τα τρία διαφορετικά σώματα έχουν την ίδια διαφορά στη φωτεινότητα.

Μία άλλη συνάρτηση που χαρακτηρίζει την ευαισθησία του ανθρώπινου συστήματος όρασης στην αντίθεση της φωτεινότητας των σωμάτων, είναι η συνάρτηση μεταφοράς διαμόρφωσης (Modulation Transfer Function – MTF). Η συνάρτηση αυτή δείχνει ότι η μέγιστη ευαισθησία βρίσκεται στις μεσαίες χωρικές συχνότητες (βλέπε Σχήμα 6.13).

Σχήμα 6.13

Η ευαισθησία στις μεσαίες χωρικές συχνότητες φαίνεται από το γεγονός ότι το μάτι διεγείρεται περισσότερο στις μεσαίες εναλλαγές της αμαύρωσης (ούτε πολύ γρήγορες ούτε πολύ αργές).



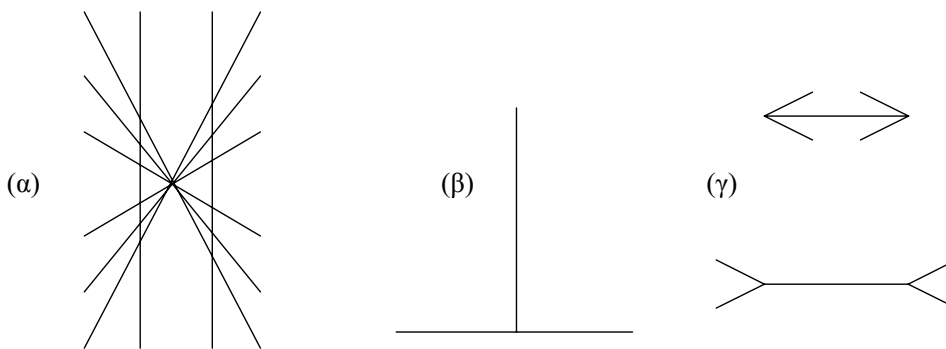
Μεσαίες εναλλαγές αμαύρωσης

Τέλος, το ανθρώπινο σύστημα όρασης δεν μπορεί να ξεχωρίσει περιοδικές αναλαμπές φωτεινής πηγής όταν η συχνότητά της ξεπερνά τα 50 Hz περίπου. Στην περίπτωση αυτή εκλαμβάνει ως συνεχή την εκπομπή της φωτεινής ακτινοβολίας. Το φαινόμενο καθορίζει την ταχύτητα με την οποία πρέπει να ανανεώνεται η πληροφορία στην οθόνη της τηλεόρασης ή του υπολογιστή. Συχνότητα ανανέωσης 60 Hz, δίνει σταθερή εικόνα χωρίς την αίσθηση της ασυνέχειας μεταξύ των διαδοχικών στιγμιότυπων.

6.2.3 Χρώμα και ψυχοφυσικά φαινόμενα

Το ανθρώπινο οπτικό σύστημα παρουσιάζει την ιδιότητα να είναι πολύ περισσότερο ευαίσθητο στις έγχρωμες απ' ό,τι στις ασπρόμαυρες εικόνες. Έτσι, ενώ μπορεί να ξεχωρίσει μόνο μερικές δεκάδες αποχρώσεις του γκριζου στις ασπρόμαυρες εικόνες, έχει τη δυνατότητα να διακρίνει χιλιάδες διαφορετικά χρώματα. Τα μεγέθη που καθορίζουν τα χρώματα είναι η λαμπρότητα (brightness), η απόχρωση (hue) και ο κορεσμός (saturation). Η αναπαράσταση της χρωματικής πληροφορίας στηρίζεται στην «κλασική» θεωρία του Thomas Young (1802), ότι κάθε χρώμα μπορεί να δημιουργηθεί με την κατάλληλη μείξη τριών χρωμάτων, όπως για παράδειγμα του κόκκινου, πράσινου και μπλε (Red–Green–Blue, RGB).

Άλλα φαινόμενα, κυρίως ψυχοφυσικά (βλέπε Σχήμα 6.14), που συνδέονται με την ψυχοφυσιολογία της όρασης περιγράφονται εκτενώς στη βιβλιογραφία. Η ανάλυση παρόμοιων φαινομένων μπορεί να βοηθήσει σημαντικά στην ανάπτυξη και τη βελτίωση των συστημάτων τεχνητής όρασης.



Σχήμα 6.14

Τα παράξενα της ψυχοφυσιολογίας της όρασης. α) Αυταπάτη Wundt, οι παράλληλες γραμμές δείχνουν παραμορφωμένες. β) Αυταπάτη οριζόντιας–κάθετης γραμμής, το κατακόρυφο τμήμα φαίνεται μεγαλύτερο. γ) Η αυταπάτη Mueller–Lyer, τα δύο οριζόντια τμήματα είναι ίσα.

6.2.4 Κριτήρια ποιότητας της εικόνας

Για να μετρήσουμε την ποιότητα των ψηφιακών εικόνων, αλλά και για να αξιολογήσουμε την απόδοση των συστημάτων απόκτησης και επεξεργασίας εικόνας, χρησιμοποιούμε διάφορα κριτήρια. Τα κριτήρια αυτά μπορούν να χωριστούν σε υποκειμενικά και ποσοτικά. Τα υποκειμενικά στηρίζονται στη γνώμη μιας μικρής ή μεγάλης ομάδας παρατηρητών που χαρακτηρίζουν την ποιότητα της εικόνας ως εξαιρετική, καλή, μέτρια, φτωχή ή μη ικανοποιητική. Τα ποσοτικά κριτήρια είναι καθιερωμένες μαθηματικές διαδικασίες που χρησιμοποιούν ποσότητες, όπως ο λόγος του σήματος προς το θόρυβο ή το μέσο τετραγωνικό σφάλμα. Πρέπει να τονιστεί ότι η επιλογή του κριτηρίου ποιότητας της εικόνας ποικίλλει ανάλογα με την εφαρμογή και δεν είναι μονοσήμαντα καθορισμένη.

Σύνοψη ενότητας

Στην ενότητα αυτή γνωρίσαμε τα βασικά χαρακτηριστικά του ανθρώπινου συστήματος όρασης. Μεταξύ αυτών επισημαίνουμε τη μεγάλη προσαρμογή του οφθαλμού σε ελάχιστο, αλλά και πολύ έντονο φως, που οφείλεται στους δύο διαφορετικούς τύπους αισθητήρων, τους κώνους και τα ραβδία. Στην ουσία, η συμπεριφορά του οφθαλμού στη φωτεινότητα είναι λογαριθμική. Επίσης τονίσαμε τη σχετικότητα στη φωτεινότητα των σωμάτων και τον τρόπο που αντιλαμβανόμαστε την αντίθεση (contrast). Τέλος, τόσο η υψηλή ευαισθησία στις μεσαίες συχνότητες, όσο και η απαίτηση για τουλάχιστον 30 στιγμιότυπα ανά δευτερόλεπτο για σχηματισμό της κίνησης είναι από τα ενδιαφέροντα χαρακτηριστικά της ανθρώπινης όρασης.

6.3 Είδη πράξεων στην επεξεργασία της ψηφιακής εικόνας

Όπως είδαμε στην υποενότητα 6.1.5, η διαδικασία της επεξεργασίας της ψηφιακής εικόνας περιλαμβάνει αρκετές βαθμίδες, ανάλογα με την ποιότητα της αρχικής εικόνας και την επιθυμητή έξοδο. Θα περίμενε κανείς πως κάθε μία από τις βαθμίδες αυτές θα έχει είσοδο μία εικόνα και έξοδο επίσης μία άλλη εικόνα. Όμως στη γενική της μορφή μία βαθμίδα επεξεργασίας μπορεί να έχει είσοδο ή έξοδο εικόνα ή μαθηματική περιγραφή (συνάρτηση ή άνυσμα). Στον Πίνακα 6.1 παρουσιάζονται οι δυνατοί συνδυασμοί εισόδου–εξόδου για μία βαθμίδα επεξεργασίας ψηφιακής εικόνας και οι πιθανές διαδικασίες που υλοποιούνται.

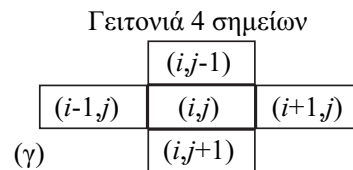
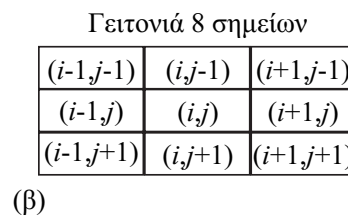
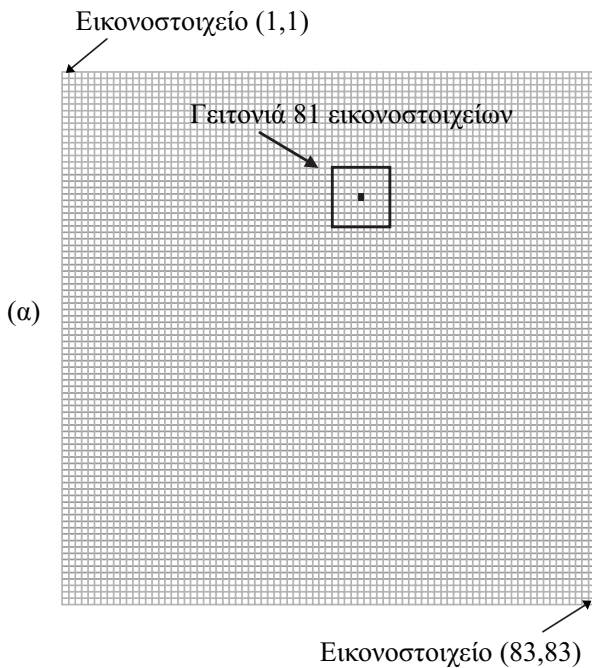
Πίνακας 6.1

Πιθανές μορφές εισόδου–εξόδου σε μία βαθμίδα ψηφιακής επεξεργασίας εικόνας.

Είσοδος	Έξοδος	Χαρακτηρισμός	Πιθανή πράξη
Εικόνα	Εικόνα	Επεξεργασία εικόνας	Εξαγωγή θορύβου
Εικόνα	Περιγραφή	Ανάλυση εικόνας	Εξαγωγή χαρακτηριστικών
Περιγραφή	Εικόνα	Γραφικά υπολογιστών	Σχεδιασμός με υπολογιστή
Περιγραφή	Περιγραφή	Μετασχηματισμοί	Μετασχηματισμός δεδομένων

Στην παράγραφο αυτή θα αναπτύξουμε μόνο τα είδη των πράξεων που μπορεί κανείς να συναντήσει στην περίπτωση που είσοδος και έξοδος της βαθμίδας είναι εικόνα. Απαραίτητη έννοια στις πράξεις αυτές είναι η έννοια της γειτονιάς ενός εικονοστοιχείου. Πρόκειται για τα εικονοστοιχεία που το περιβάλλουν (Σχήμα 6.15α). Ειδικότερα, ως γειτονιά των 8 σημείων ενός εικονοστοιχείου με συντεταγμένες $(n_1, n_2) = (i, j)$ ορίζεται η διάταξη των 8 εικονοστοιχείων γύρω από το (i, j) που οι συντεταγμέ-

νες δίνονται στο Σχήμα 6.15β. Επίσης, γειτονιά των 4 σημείων ορίζεται η διάταξη των εικονοστοιχείων γύρω από το (i,j) όπως δίνεται στο Σχήμα 6.15γ. Όλοι οι γείτονες του εικονοστοιχείου (i,j) χρησιμοποιούνται μαζί με αυτό για τον υπολογισμό του αντίστοιχου (στην ίδια θέση) εικονοστοιχείου στην εικόνα εξόδου. Ο όρος «παράθυρο με κέντρο το εικονοστοιχείο (i,j) » χρησιμοποιείται επίσης αντί του όρου «γειτονιά του εικονοστοιχείου».



Σχήμα 6.15

Η διάταξη των εικονοστοιχείων σε σειρές και στήλες, δημιουργεί για καθένα από αυτά γείτονες. Οι τιμές των 8 ή 4 γειτονικών εικονοστοιχείων χρησιμοποιούνται συχνότερα στις διάφορες πράξεις.

Οι πράξεις στην ψηφιακή επεξεργασία της εικόνας έχουν στόχο τον υπολογισμό κάθε εικονοστοιχείου της εικόνας εξόδου. Ανάλογα με τον τρόπο που υπολογίζεται η τιμή του εικονοστοιχείου αυτού, οι πράξεις διακρίνονται σε

Τοπικές (local)

Καθολικές (global)

Γεωμετρικές (geometrical)

6.3.1 Τοπικές πράξεις

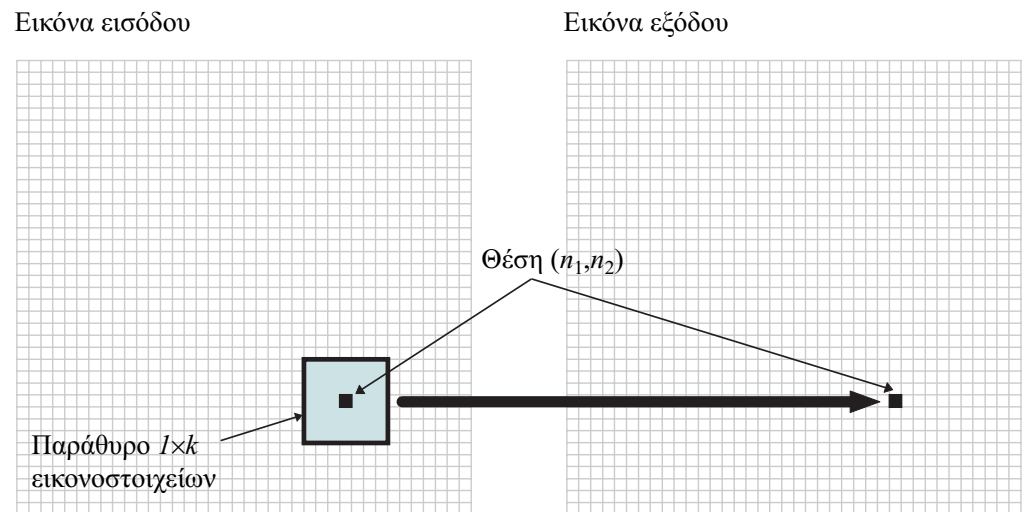
Στις πράξεις αυτές η τιμή κάθε εικονοστοιχείου στη θέση (n_1, n_2) στην εικόνα εξόδου υπολογίζεται χρησιμοποιώντας την τιμή του εικονοστοιχείου στην αντίστοιχη θέση της εικόνας εισόδου, καθώς επίσης και τις τιμές των γειτονικών εικονοστοιχείων. Η διαδικασία αυτή παρουσιάζεται στο Σχήμα 6.16. Πιο συγκεκριμένα, για τον υπολογισμό της τιμής του εικονοστοιχείου στη θέση (n_1, n_2) της εικόνας εξόδου, θεωρού-

Σχήμα 6.16

Το εικονοστοιχείο της εικόνας εξόδου υπολογίζεται από τα εικονοστοιχεία της εικόνας εισόδου που βρίσκονται σε ένα παράθυρο μεγέθους ($l \times k = 7 \times 7$) γύρω από το αντίστοιχο εικονοστοιχείο της εικόνας εισόδου. Η διαδικασία αυτή επαναλαμβάνεται για όλα τα εικονοστοιχεία της εικόνας εξόδου.

με μία γειτονιά (παράθυρο) $l \times k$ εικονοστοιχείων στην εικόνα εισόδου με κέντρο την αντίστοιχη θέση (n_1, n_2) . Οι τιμές q όλων των εικονοστοιχείων μέσα στο παράθυρο θα συμβάλουν στον υπολογισμό της τιμής $p(n_1, n_2)$ του εικονοστοιχείου εξόδου, για παράδειγμα:

$$p(n_1, n_2) = \sum_{i=-l/2}^{l/2} \sum_{j=-k/2}^{k/2} a(i, j) f(q(n_1 - i, n_2 - j)) \quad (6.12)$$



Επομένως, στη γενικότερη περίπτωση οι τιμές q μετασχηματίζονται από μία συνάρτηση f (γραμμική ή μη γραμμική), και στη συνέχεια η έξοδος $p(n_1, n_2)$ υπολογίζεται ως γραμμικός συνδυασμός των αποτελεσμάτων αυτών με τους συντελεστές βάρους $a(i, j)$. Η διαδικασία αυτή επαναλαμβάνεται για όλα τα εικονοστοιχεία της εικόνας εξόδου. Στην πράξη, ο πίνακας των συντελεστών $a(i, j)$ μοιάζει με παράθυρο που κινείται σε όλη την εικόνα και κάθε φορά συμβάλλει σύμφωνα με την (6.12) στον υπολογισμό του εικονοστοιχείου της εικόνας εξόδου, που αντιστοιχεί στη θέση του κέντρου του. Στην περίπτωση που $f(q) = q$, τότε η πράξη είναι γραμμική και δεν είναι τίποτε άλλο από τη γενίκευση της γνωστής μας συνέλιξης σε δύο διαστάσεις και ονομάζεται δισδιάστατη συνέλιξη. Τη δισδιάστατη συνέλιξη θα τη μελετήσουμε αναλυτικότερα στο επόμενο κεφάλαιο όπου εξετάζονται τα γραμμικά δισδιάστατα συστήματα. Στην περίπτωση αυτή, οι συντελεστές $a(i, j)$ χαρακτηρίζουν το γραμμικό σύστημα. Βεβαίως, η μορφή της εξίσωσης 6.12 μπορεί να είναι διαφορετική ανάλογα με τις απαιτήσεις της εφαρμογής.

Όταν το μέγεθος του παραθύρου μικραίνει, η τιμή της $p(n_1, n_2)$ εξαρτάται όλο και λιγότερο από τα γειτονικά εικονοστοιχεία της θέσης (n_1, n_2) . Στην ειδική περίπτωση

όπου $l = k = 1$ η τοπική πράξη εκφυλίζεται σε πράξη σημείου (point operation). Τότε η τιμή του εικονοστοιχείου στη θέση (n_1, n_2) της εικόνας εξόδου εξαρτάται μόνο από την τιμή του εικονοστοιχείου στην αντίστοιχη θέση της εικόνας εισόδου.

Παράδειγμα 6.4

Η παρακάτω εικόνα εισόδου (Σχήμα 6.17) έχει μέγεθος 5×5 και οι τιμές των εικονοστοιχείων κυμαίνονται μεταξύ 0 και 51. Να πολλαπλασιαστούν κατάλληλα οι τιμές των εικονοστοιχείων, ώστε αυτές να καταλαμβάνουν όλη την περιοχή από 0 ως 255 που καλύπτει δυαδικός αριθμός του ενός byte. Σε ποια κατηγορία ανήκει η πράξη αυτή;

Λύση

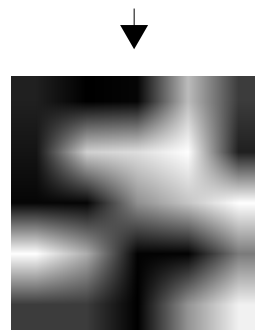
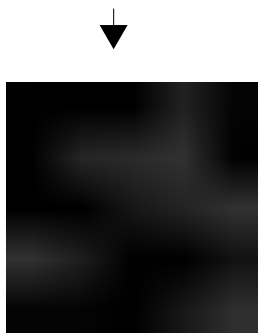
Αρκεί να πολλαπλασιάσουμε την τιμή κάθε εικονοστοιχείου με το 5. Με τον τρόπο αυτό δεν ξεπερνάμε την τιμή 255 για τα εικονοστοιχεία της εικόνας εξόδου (Σχήμα 6.17). Πρόκειται για πράξη σημείου (point operation), αφού κάθε εικονοστοιχείο της εικόνας εξόδου υπολογίζεται χρησιμοποιώντας το αντίστοιχο εικονοστοιχείο της εικόνας εισόδου. Στην ψηφιακή επεξεργασία εικόνας η πράξη αυτή καλείται βελτίωση της αντίθεσης (contrast stretching).

Εικόνα εισόδου

7	0	3	37	12
5	41	44	50	7
3	3	28	39	51
51	34	2	2	24
12	12	0	29	48

Εικόνα εξόδου

35	0	15	185	60
25	205	220	250	35
15	15	140	195	255
255	170	10	10	120
60	60	0	145	240



Σχήμα 6.17

Βελτίωση της αντίθεσης δια του πολλαπλασιασμού όλων των εικονοστοιχείων της εικόνας εισόδου με την ίδια τιμή (πράξη σημείου)

Παράδειγμα 6.5

Σε εικόνα εισόδου μεγέθους 7×7 εικονοστοιχείων (Σχήμα 6.18), πρόκειται να εκτελεστεί διαδικασία μέσης τιμής με παράθυρο 3×3 και τιμές για τους συντελεστές $a(i, j) = 1/9$. Να υπολογιστεί η εικόνα εξόδου. Σε ποια κατηγορία ανήκει η πράξη αυτή;

Λύση

Για να υπολογίσουμε κάθε ένα εικονοστοιχείο της εικόνας εξόδου, λαμβάνουμε το αντίστοιχο της εικόνας εισόδου καθώς και τους 8 γείτονές του, πολλαπλασιάζουμε με $1/9$ και αθροίζουμε τα γινόμενα. Η διαδικασία είναι ίδια με αυτή που περιγράφεται στο Σχήμα 6.15 για $l = k = 3$. Πρόκειται για τοπική πράξη στην ψηφιακή επεξεργασία εικόνας. Η διαδικασία που περιγράψαμε είναι διαδικασία μέσης τιμής και εύκολα μπορεί να ξεχωρίσει κανείς ότι η αρχική εικόνα (μαύρο τετράγωνο) έχει θολώσει.

Εικόνα εισόδου

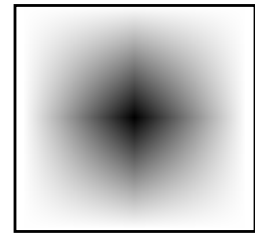
255	255	255	255	255	255	255
255	255	255	255	255	255	255
255	255	0	0	0	255	255
255	255	0	0	0	255	255
255	255	0	0	0	255	255
255	255	255	255	255	255	255
255	255	255	255	255	255	255

Εικόνα εξόδου

255	255	255	255	255	255	255
255	226	198	170	198	226	255
255	198	141	85	141	198	255
255	170	85	0	85	170	255
255	198	141	85	141	198	255
255	226	198	170	198	226	255
255	255	255	255	255	255	255

Σχήμα 6.18

Καθένα εικονοστοιχείο της εικόνας εξόδου (δεξιά) υπολογίζεται από τον μέσο όρο 9 εικονοστοιχείων της εικόνας εισόδου (μέση τιμή)



6.3.2 Καθολικές πράξεις

Στις πράξεις αυτές, για να υπολογιστεί η τιμή ενός εικονοστοιχείου της εικόνας εξόδου χρησιμοποιούνται όλα τα εικονοστοιχεία της εικόνας εισόδου. Οι πράξεις αυτές αποτελούν ουσιαστικά ειδική περίπτωση των τοπικών πράξεων όπου το μέγεθος $l \times k$ του παραθύρου γίνεται τόσο μεγάλο, όσο και η ίδια η εικόνα. Εδώ θα πρέπει να διευκρινίσουμε ότι οι τιμές των εικονοστοιχείων της εικόνας εξόδου διαφοροποιούνται μεταξύ τους από το γεγονός ότι, ανάλογα με τη θέση (n_1, n_2) του εικονοστοιχείου, οι συντελεστές $a(i, j)$ της σχέσης (6.12) αλλάζουν. Καθολικές πράξεις στην ψηφιακή επεξεργασία εικόνας είναι οι πράξεις των μετασχηματισμών που θα δούμε στη συνέχεια.

Το ολικό πλήθος των απαιτούμενων υπολογισμών για να υλοποιηθεί μία τοπική πράξη εξαρτάται από το πλήθος των εικονοστοιχείων της εικόνας και από το μέγεθος της γειτονιάς που χρησιμοποιείται. Για μία εικόνα μεγέθους $M \times N$ εικονοστοιχείων και μία γειτονιά $k \times l$ εικονοστοιχείων, με βάση τη σχέση (6.12), για το σχηματισμό της

εικόνας εξόδου απαιτούνται συνολικά $N \times M \times k \times l$ πράξεις (πολλαπλασιασμοί και προσθέσεις). Στην περίπτωση που η πράξη είναι καθολική και η εικόνα τετραγωνική ($N \times N$), το πλήθος των απαιτούμενων υπολογισμών είναι της τάξης N^4 .

Δραστηριότητα 6.3

Να εκτελεστεί μία μη γραμμική πράξη σημείου στην εικόνα εισόδου του Παραδείγματος 6.4. Να προσέξετε ώστε οι τελικές τιμές να είναι στο όριο 0 ως 255. Για το λόγο αυτό να γίνει η κατάλληλη κανονικοποίηση.

Μία από τις σπουδαίες μη γραμμικές πράξεις σημείου είναι η λογαρίθμιση (εδώ χρησιμοποιούμε δεκαδικό λογάριθμο). Με τη βοήθεια του λογαρίθμου μπορούμε να μεταβάλλουμε τη φωτεινότητα των εικονοστοιχείων ανισομερώς. Στην εικόνα εισόδου που φαίνεται στο Σχήμα 6.19, πριν λογαριθμίσουμε την τιμή του κάθε εικονοστοιχείου, θα προσθέσουμε σε όλα τη μονάδα για να αποφύγουμε το λογάριθμο του μηδενός. Μετά τη λογαρίθμιση θα πολλαπλασιάσουμε όλες τις τιμές που προκύπτουν με το $255/\log_{10}(52)$, για να τις κανονικοποιήσουμε σε μέγιστη τιμή το 255. Συνολικά θα εκτελέσουμε την πράξη:

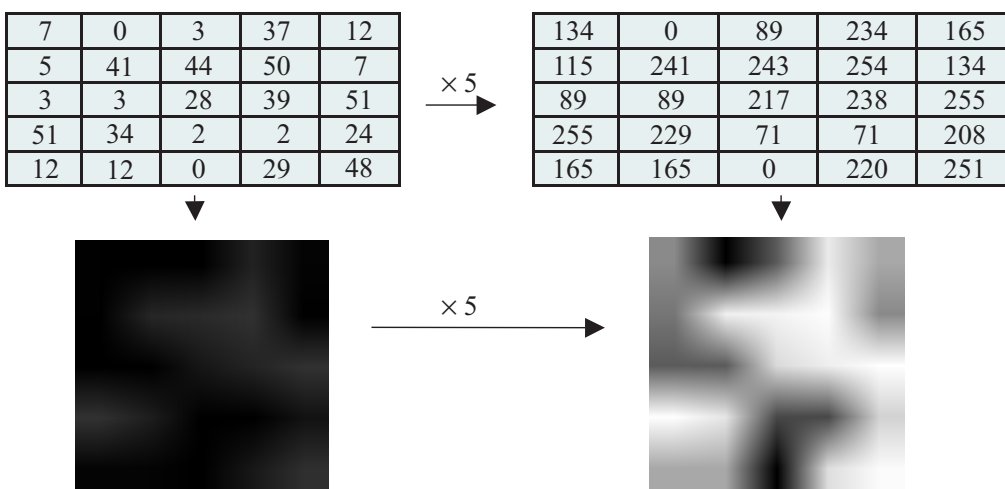
$$\log_{10}(\text{τιμή του εικονοστοιχείου} + 1) * 255 / \log_{10}(52)$$

όπου 52 είναι η μεγαλύτερη τιμή της εισόδου αυξημένη κατά 1.

Συγκρίνοντας το αποτέλεσμα (εικόνα εξόδου Σχήμα 6.19) με αυτό του Παραδείγματος 6.4, βλέπουμε ότι εδώ τα εικονοστοιχεία μικρής τιμής έχουν μεγαλώσει περισσότερο.

Εικόνα εισόδου

Εικόνα εξόδου



Σχήμα 6.19

Αύξηση της φωτεινότητας δυσανάλογα με την τιμή των εικονοστοιχείων. Τα αμυδρά εικονοστοιχεία ενισχύονται περισσότερο.

Άσκηση Αυτοαξιολόγησης 6.2

Οι πράξεις που απαιτούνται για να υπολογιστεί η εικόνα εξόδου, όταν η εικόνα εισόδου είναι 31×31 εικονοστοιχεία και η διαδικασία που θα εκτελεστεί περιλαμβάνει τοπική πράξη με παράθυρο 5×5 , είναι συνολικά: (α) 961, (β) 24,025, (γ) 155, (δ) 775.

Δραστηριότητα 6.4

Να αναφέρετε ένα πλεονέκτημα και ένα μειονέκτημα, που έχουμε, όταν το μέγεθος του παραθύρου σε μία τοπική πράξη μεγαλώνει.

Όταν το μέγεθος του παραθύρου σε μία τοπική πράξη αυξάνει, παρουσιάζεται το μειονέκτημα ότι το πλήθος των απαιτούμενων πράξεων για τον υπολογισμό της εικόνας εξόδου αυξάνεται. Το πλεονέκτημα έγκειται στο γεγονός ότι με πιο μεγάλο παράθυρο, κάθε εικονοστοιχείο της εικόνας εξόδου, περιέχει πληροφορία από ευρύτερη περιοχή της εικόνας εισόδου.

6.3.3 Γεωμετρικές πράξεις

Οι πράξεις αυτές αφορούν στο μετασχηματισμό των συντεταγμένων της αρχικής εικόνας. Από τις κυριότερες γεωμετρικές πράξεις είναι η μετατόπιση (translation), η στροφή (rotation) και ο καθρεπτισμός (mirroring). Οι πράξεις αυτές αποδίδουν συνήθως εικόνα εξόδου ίδια σε διαστάσεις με την εικόνα εισόδου (μετασχηματισμός ένα προς ένα). Υπάρχουν όμως και γεωμετρικές πράξεις που αποδίδουν εικόνες μεγαλύτερες ή μικρότερες από την αρχική. Πρέπει να επισημάνουμε ότι οι γεωμετρικοί μετασχηματισμοί βρίσκουν εφαρμογή σε πληθώρα περιπτώσεων επεξεργασίας εικόνας. Έτσι, για παράδειγμα, η στροφή είναι απαραίτητη για το ταίριασμα (matching) και σύγκριση εικόνων της ίδιας περιοχής που έχουν ληφθεί από διαφορετικές γωνίες.

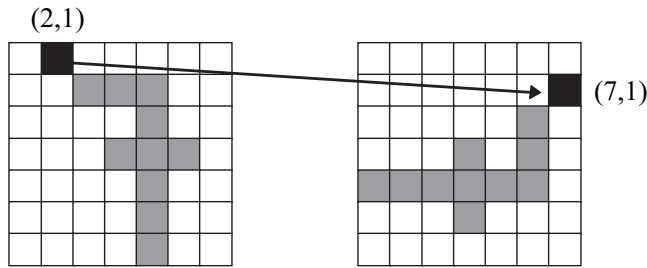
Παράδειγμα 6.6

Να δοθεί η μαθηματική έκφραση για μετασχηματισμό στροφής μιας εικόνας κατά 90° .

Λύση

Θεωρούμε την εικόνα μεγέθους $N \times N = 7 \times 7$ όπως φαίνεται στο Σχήμα 6.20. Στροφή κατά 90° σημαίνει το εικονοστοιχείο με συντεταγμένες (n_1, n_2) να πάει στη θέση $(N + 1 - n_2, n_1)$. Επομένως, πρέπει να γίνει ο μετασχηματισμός

$$T(n_1, n_2) = (N + 1 - n_2, n_1) \text{ όπου } N = 7$$

**Σχήμα 6.20**

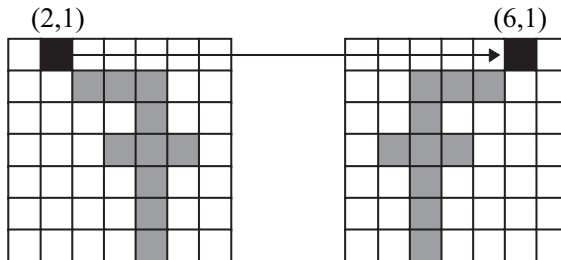
Στροφή εικόνας
κατά 90° .

Δραστηριότητα 6.5

Να δοθεί μετασχηματισμός με τον οποίο λαμβάνουμε το συμμετρικό της προηγούμενης εικόνας.

Δεδομένου ότι είναι δυνατό να έχουμε συμμετρία πολλών ειδών (γύρω από κατακόρυφο, οριζόντιο ή διαγώνιο άξονα) θα αντιμετωπίσουμε εδώ μόνο την περίπτωση συμμετρίας γύρω από κατακόρυφο άξονα. Έτσι, με βάση το Σχήμα 6.21, για να λάβουμε τη συμμετρική εικόνα γύρω από άξονα που ταυτίζεται με τη μεσαία στήλη των εικονοστοιχείων, θα πρέπει να ισχύει:

$$T(n_1, n_2) = (N - n_1 + 1, n_2)$$

**Σχήμα 6.21**

Δημιουργία του
συμμετρικού
εικόνας γύρω
από κατακόρυφο
άξονα.

Άσκηση Αυτοαξιολόγησης 6.3

Ποιος από τους παρακάτω μετασχηματισμούς δίνει μία εικόνα $N \times N$ εικονοστοιχείων αναστραμμένη;

$$T(n_1, n_2) = (n_2, n_1)$$

$$T(n_1, n_2) = (N - n_2 + 1, n_1)$$

$$T(n_1, n_2) = (n_1, N - n_2 + 1)$$

Σύνοψη ενότητας

Στην παράγραφο αυτή γνωρίσαμε τις στοιχειώδεις πράξεις που μπορούν να εκτελεστούν πάνω σε μία εικόνα. Τις πράξεις αυτές τις ταξινομήσαμε σε τοπικές, καθολικές και γεωμετρικές. Σημαντικό χαρακτηριστικό στην εκτέλεση των πράξεων αποτελεί η γειτονιά των εικονοστοιχείων που χρησιμοποιούνται στην εικόνα εισόδου, για να υπολογιστεί ένα εικονοστοιχείο στην εικόνα εξόδου. Όταν η γειτονιά αυτή (παράθυρο) εκφυλίζεται σ' ένα σημείο τότε έχουμε τις σημειακές πράξεις, ενώ όταν εκτείνεται σε ολόκληρη την εικόνα λαμβάνουμε τις καθολικές πράξεις. Γνωρίσαμε επίσης τον τρόπο υπολογισμού του απαιτούμενου πλήθους πράξεων.

6.4 Το φάσμα της εικόνας

Στο Κεφάλαιο 2 είδαμε ότι ο μετασχηματισμός ενός σήματος από το πεδίο του χρόνου στο πεδίο των συχνοτήτων προσφέρει σημαντικές δυνατότητες για τη μελέτη του σήματος, αφού μπορούμε να το εξετάσουμε από μία άλλη σκοπιά. Επισημάνουμε επίσης ότι τον διακριτό μετασχηματισμό Fourier (DFT) τον υλοποιούμε γρήγορα με τον αλγόριθμο FFT. Στα δισδιάστατα σήματα, όπως είναι οι εικόνες, υπάρχει ακριβώς η ίδια δυνατότητα μετασχηματισμού. Φυσικά, αφού η εικόνα αποτελεί σήμα που περιγράφει τον τρόπο μεταβολής της αμαύρωσης (ή του χρώματος) στο χώρο και όχι στο χρόνο, ο διακριτός μετασχηματισμός Fourier θα μας μεταφέρει στην περιοχή των χωρικών συχνοτήτων (διαστάσεις αντιστρόφου χώρου). Κατά τα άλλα, ο δισδιάστατος διακριτός μετασχηματισμός Fourier έχει όλες τις ιδιότητες του μονοδιάστατου μετασχηματισμού. Ο δισδιάστατος μετασχηματισμός Fourier $p(k_1, k_2)$, μιας εικόνας $q(n_1, n_2)$ μεγέθους $N \times N$ εικονοστοιχείων, ορίζεται ως:

$$p(k_1, k_2) = \sum_{n_1=0}^{N-1} \sum_{n_2=0}^{N-1} q(n_1, n_2) W_N^{k_1 n_1} W_N^{k_2 n_2} \quad 0 \leq k_1, k_2 \leq N-1, \quad (6.13)$$

ενώ ο αντίστροφος μετασχηματισμός είναι ο εξής:

$$q(n_1, n_2) = \frac{1}{N^2} \sum_{k_1=0}^{N-1} \sum_{k_2=0}^{N-1} p(k_1, k_2) W_N^{-k_1 n_1} W_N^{-k_2 n_2} \quad 0 \leq n_1, n_2 \leq N-1 \quad (6.14)$$

Από την πρώτη σχέση γίνεται σαφές ότι σε μία εικόνα q με $N \times N = N^2$ εικονοστοιχεία, αντιστοιχεί ένα φάσμα p με $N \times N = N^2$ όρους, που στη γενική τους περίπτωση είναι μιγαδικοί αριθμοί. Κάθε όρος $p(k_1, k_2)$ απαιτεί για τον υπολογισμό του N^2 πολλαπλασιασμούς και σχεδόν άλλες τόσες προσθέσεις. Επομένως, για τον υπολογισμό όλων των όρων του φάσματος απαιτείται ένα πλήθος πράξεων της τάξης N^4 .

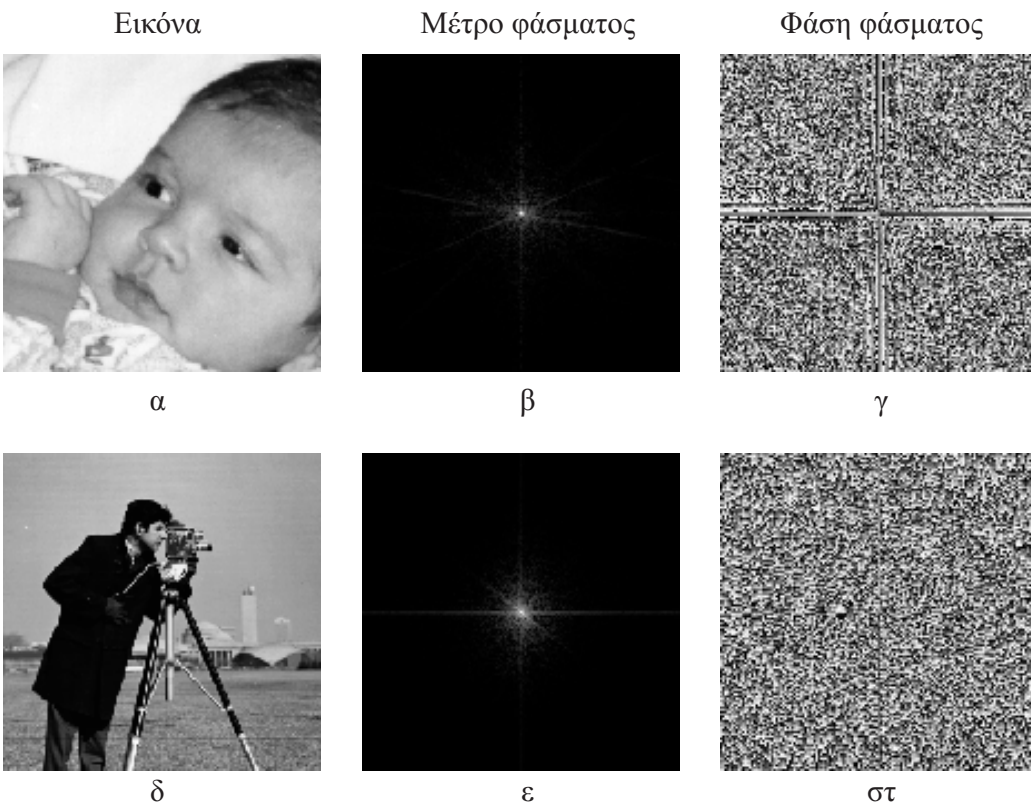
Επειδή ο διδιάστατος DFT είναι διαχωρίσιμη διαδικασία, το πλήθος των πράξεων είναι της τάξης του N^3 . Η διαχωρισιμότητα της σχέσης (6.13) απορρέει από το γεγονός ότι μπορεί να γραφεί και ως

$$p(k_1, k_2) = \sum_{n_1=0}^{N-1} W_N^{k_1 n_1} \sum_{n_2=0}^{N-1} q(n_1, n_2) W_N^{k_2 n_2} \quad 0 \leq k_1, k_2 \leq N-1 \quad (6.15)$$

που σημαίνει ότι μπορούμε να υπολογίσουμε πρώτα το εσωτερικό άθροισμα της σχέσης. Η πράξη αυτή ισοδυναμεί με το να υπολογίσουμε τον DFT κάθε μιας γραμμής της εικόνας και στο αποτέλεσμα που προκύπτει να υπολογίσουμε τον DFT της κάθε στήλης. Η χρησιμοποίηση του FFT θα μειώσει τις απαιτούμενες πράξεις για τον υπολογισμό του φάσματος σε πλήθος της τάξης του $2N^2 \log_2 N$.

Δεδομένου ότι οι όροι $p(k_1, k_2)$ του φάσματος είναι μιγαδικοί αριθμοί, για να μελετήσουμε το φάσμα μιας εικόνας, συνήθως μελετούμε το μέτρο των χωρικών συνιστωσών και τη φάση τους ξεχωριστά.

Στο Σχήμα 6.22 φαίνεται το μέτρο και η φάση του φάσματος φυσικών εικόνων. Συγκρίνοντας το μέτρο των δύο εικόνων μπορούμε να πούμε ότι παρουσιάζει τις περισσότερες αρμονικές στις χαμηλές χωρικές συχνότητες (κέντρο των εικόνων του μέτρου). Αντίθετα, οι εικόνες της φάσης παρουσιάζουν έντονη τυχαιότητα και δεν είναι δυνατό να συγκριθούν.



Σχήμα 6.22

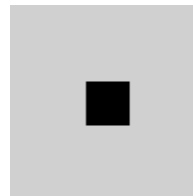
Το μέτρο και η φάση του φάσματος δύο φυσικών εικόνων.

Στο Σχήμα 6.23 παρουσιάζονται τέσσερις γεωμετρικές εικόνες και το μέτρο των αντίστοιχων φασμάτων. Από τις φασματικές εικόνες παρατηρούμε ότι το φάσμα έχει έντονο αρμονικό περιεχόμενο σε διευθύνσεις όπου στην αντίστοιχη εικόνα υπάρχει απότομη μεταβολή της αμαύρωσης. Έτσι, στην περίπτωση του τετραγώνου έντονο αρμονικό περιεχόμενο παρουσιάζεται σε διευθύνσεις κάθετες στις πλευρές του. Το ίδιο ακριβώς συμβαίνει στο τρίγωνο, ενώ στον κύκλο έχουμε αρμονικό περιεχόμενο προς όλες τις κατευθύνσεις αφού η απότομη μεταβολή της αμαύρωσης συμβαίνει κυκλικά κατά μήκος της περιφέρειάς του. Τέλος, στο σχήμα με τις οριζόντιες γραμμές έχουμε μεταβολή της αμαύρωσης και μάλιστα περιοδική μόνο κατά την κατακόρυφη διεύθυνση. Το αντίστοιχο φάσμα είναι γραμμικό κατά τη διεύθυνση αυτή.

Σχήμα 6.23

Στην πρώτη σειρά παρουσιάζονται γεωμετρικά σχήματα, ενώ στην δεύτερη σειρά το μέτρο των αντίστοιχων φασμάτων. Τα φάσματα έχουν έντονο αρμονικό περιεχόμενο σε διευθύνσεις όπου στην αντίστοιχη εικόνα υπάρχει απότομη μεταβολή της αμαύρωσης.

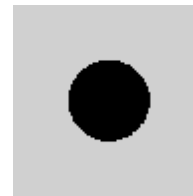
Εικόνα



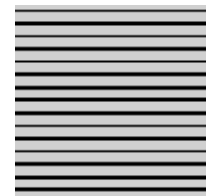
α



β

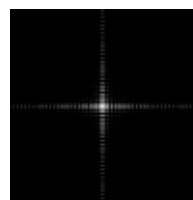


γ



δ

Μέτρο φάσματος



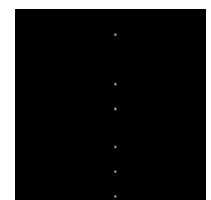
ε



στ



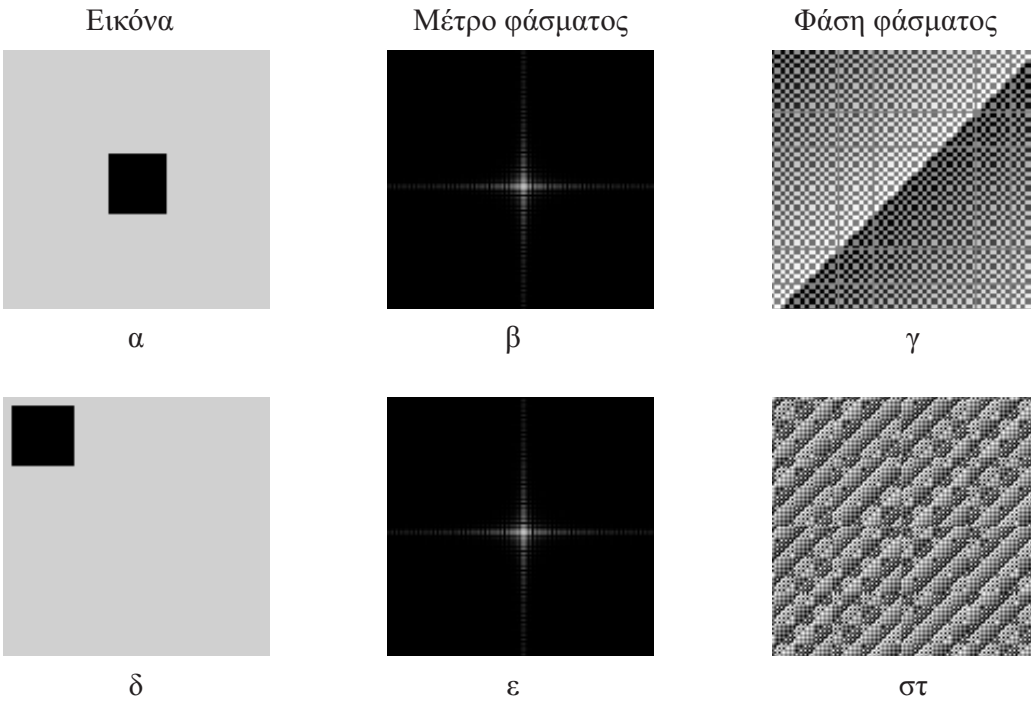
ζ



η

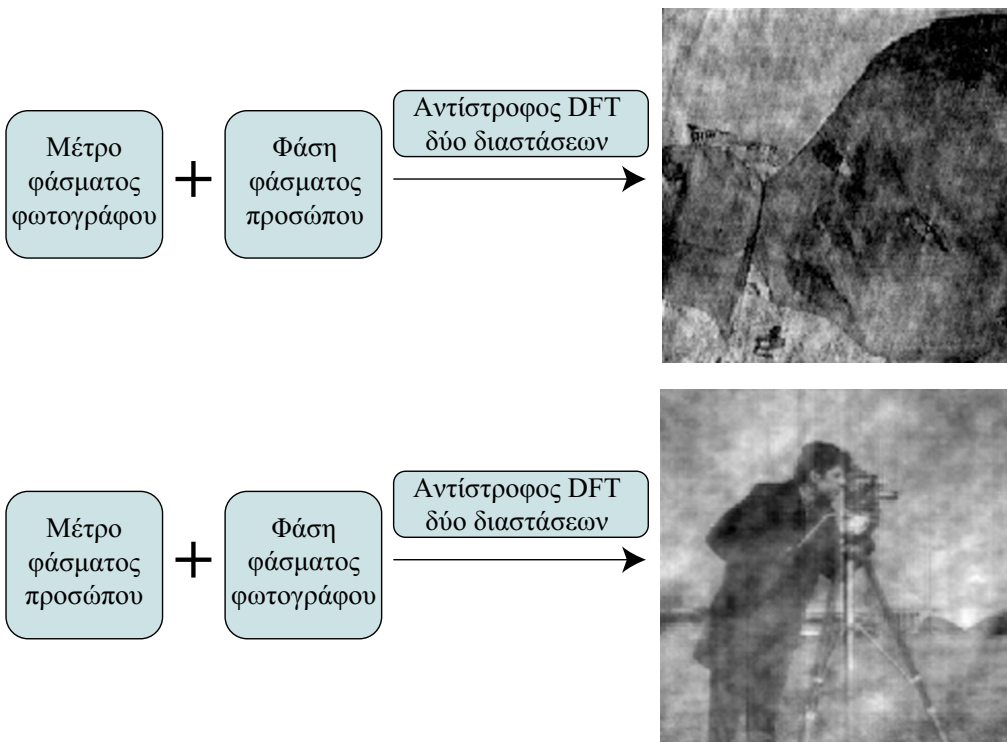
Τέλος, πρέπει να επισημάνουμε τη σημασία που έχει η φάση για την εικόνα. Στο Σχήμα 6.24 παρουσιάζονται το μέτρο και η φάση του φάσματος για το ίδιο σχήμα (ένα τετράγωνο) που είναι τοποθετημένο σε δύο διαφορετικές θέσεις της εικόνας. Το μέτρο του φάσματος και για τις δύο εικόνες είναι το ίδιο, ενώ η φάση είναι διαφορετική. Το γεγονός αυτό αποτελεί ένδειξη ότι η φάση περιέχει σημαντική πληροφορία για τη θέση ενός σχήματος μέσα στην εικόνα. Επίσης πρέπει να σημειωθεί ότι η εικόνα κρατά τα γενικά της χαρακτηριστικά αν το μέτρο του φάσματος μεταβληθεί, αλλά η φάση διατηρηθεί αναλλοίωτη. Έτσι η ανασύνθεση εικόνας, χρησιμοποιώντας την πληροφορία της φάσης του φάσματός της και μέτρο από φάσμα οποιασδήποτε φυσικής σκηνής, δίνει ικανοποιητική αναπαράσταση της εικόνας όπως φαίνεται στο Σχήμα 6.25. Συμπερασματικά, θα πρέπει να τονίσουμε ότι στην ψηφιακή επεξεργασία εικόνας με διάφορα γραμμικά φίλτρα, αυτά θα πρέπει να μην επηρεάζουν τη φασική σχέση των χωρικών αρμονικών της εικόνας. Τέτοια φίλτρα είναι

τα δισδιάστατα FIR, που θα παρουσιάσουμε στο επόμενο κεφάλαιο. Τα FIR φίλτρα ως γνωστό (Κεφάλαιο 4) είναι δυνατό να δώσουν γραμμική απόκριση φάσης με αποτέλεσμα να διατηρούν τη φασική σχέση των αρμονικών του σήματος και επομένως, τη μορφή του.



Σχήμα 6.24

Η πληροφορία για τη θέση του τετραγώνου βρίσκεται στη φάση του φάσματος της εικόνας αφού αυτή διαφοροποιείται, και όχι στο μέτρο που παραμένει το ίδιο.



Σχήμα 6.25

Η φάση του φάσματος μιας εικόνας είναι καθοριστική στην ανασύνθεση της εικόνας. Το μέτρο του φάσματος παίζει δευτερεύοντα ρόλο.

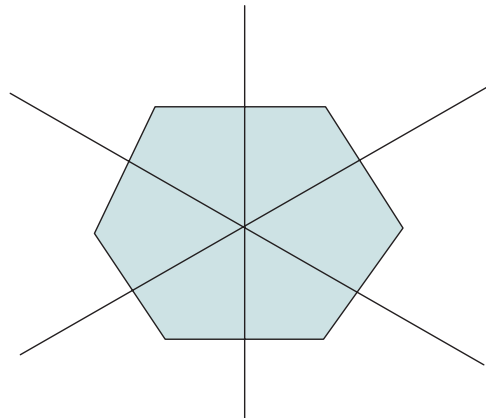
Εκτός από το δισδιάστατο μετασχηματισμό Fourier, στην επεξεργασία της ψηφιακής εικόνας χρησιμοποιούνται και πλήθος άλλων μετασχηματισμών με ενδιαφέρουσες ιδιότητες. Ο σημαντικότερος εξ αυτών είναι ο δισδιάστατος μετασχηματισμός συνημιτόνου (Discrete Cosine Transform – DCT). Οι πολύ καλές ιδιότητες του DCT τον καθιστούν πολύ χρήσιμο σε διαδικασίες συμπίεσης εικόνας. Αυτό οφείλεται στο γεγονός ότι παρουσιάζει συγκέντρωση της ενέργειας του σήματος σε λίγους μόνο συντελεστές στον μετασχηματισμένο χώρο (με προϋποθέσεις φυσικά). Ο DCT μπορεί να υλοποιηθεί ταχύτατα μέσω του FFT, ή και απευθείας με τεχνικές αποδελτατισμού (decimation) ανάλογες εκείνων του FFT που γνωρίσαμε στο Κεφάλαιο 2.

Δραστηριότητα 6.6

Να σχεδιάσετε κατά προσέγγιση τη μορφή του μέτρου του φάσματος κανονικού πολυγώνου.

Εδώ θα θεωρήσουμε ένα κανονικό εξάγωνο όπως φαίνεται στην εικόνα του Σχήματος 6.26. Περιμένουμε το μέτρο του φάσματος να έχει πλούσιο αρμονικό περιεχόμενο σε κατευθύνσεις κάθετες στις πλευρές του, δεδομένου ότι εκεί παρατηρούμε απότομη μεταβολή της αμαύρωσης.

Σχήμα 6.26
 Το φάσμα της εικόνας του εξάγωνου θα είναι πλούσιο σε αρμονικές συνιστώσες σε κατευθύνσεις κάθετες στις πλευρές του.



Σύνοψη ενότητας

Στην παράγραφο αυτή μελετήσαμε ποιοτικά το φάσμα της ψηφιακής εικόνας. Αφού έγινε εισαγωγή στο δισδιάστατο διακριτό μετασχηματισμό Fourier, στη συνέχεια είδαμε τα χαρακτηριστικά του μέτρου και της φάσης διαφόρων εικόνων και συζητήσαμε τη σπουδαιότητα της φάσης στη μορφή και τη θέση του περιεχόμενου της εικόνας.

Σύνοψη κεφαλαίου

Στο κεφάλαιο αυτό γνωρίσαμε τα απλά δισδιάστατα ψηφιακά σήματα και είδαμε ότι η ψηφιακή εικόνα αποτελεί ένα τέτοιο σήμα. Γνωρίσαμε επίσης, απλά στοιχεία από την ψηφιακή εικόνα, όπως είναι η γεωμετρία της και οι διαδικασίες που μπορούν να εκτελεστούν πάνω σε αυτή. Γνωρίσαμε τα σημαντικότερα χαρακτηριστικά της ανθρώπινης όρασης. Ταξινομήσαμε τα είδη των πράξεων που μπορούν να εκτελεστούν στην ψηφιακή εικόνα. Τέλος, μελετήσαμε τα χαρακτηριστικά του φάσματος της εικόνας, όπως αυτά εμφανίζονται τόσο στο μέτρο του, όσο και στη φάση του.

Βιβλιογραφία κεφαλαίου

ΕΛΛΗΝΟΓΛΩΣΣΗ

- Ι. Πήτας, *Ψηφιακή επεξεργασία Εικόνας*, Θεσσαλονίκη 1996.
- Ε. Σαγκριώτης και Σ. Θεοδωρίδης, *Ανάλυση Εικόνας και Αναγνώριση Προτύπων*, Ελληνικό Ανοικτό Πανεπιστήμιο (υπό έκδοση).

ΞΕΝΟΓΛΩΣΣΗ

- M.D. Levine, *Vision in Man and Machine*, McGraw–Hill, 1985.
- A.K. Jain, *Fundamentals of Digital Image Processing*, Prentice Hall, 1989.
- R.C. Gonzalez and R.E. Woods, *Digital Image Processing*, Addison–Wesley, New York, 1993.
- R. Klette and P. Zamperoni, *Handbook of Image Processing Operators*, John Wiley and Sons, Chichester, 1996.
- J.C. Russ, *The Image Processing Handbook*, CRC Press, Springer, IEEE Press, 1999.
- K.R. Castleman, *Digital Image Processing*, Prentice Hall, 1996

ΟΔΗΓΟΣ ΠΕΡΑΙΤΕΡΩ ΜΕΛΕΤΗΣ

- Ι. Πήτας, *Ψηφιακή επεξεργασία Εικόνας*, Θεσσαλονίκη 1996.

Έμφαση μπορεί να δοθεί στα κεφάλαια 1,2,4 και 6

Απόδοση αγγλικών όρων στην ελληνική

Brightness	Λαμπρότητα
Geometrical operation	Γεωμετρική πράξη
Global operation	Καθολική πράξη
Hue	Απόχρωση
Image compression	Συμπίεση εικόνας
Image acquisition	Απόκτηση εικόνας
Image restoration	Αποκατάσταση της εικόνας
Local operation	Τοπική πράξη
Luminance	Φωτεινότητα
Matching	Ταίριασμα
Mirroring	Καθρεπτισμός
MTF	Συνάρτηση Μεταφοράς Διαμόρφωσης
Picture element ή pixel	εικονοστοιχείο
Pixels per inch	Εικονοστοιχεία ανά ίντσα
Point operation	Πράξη σημείου
Retina	Αμφιβληστροειδής
RGB ή Red–Green–Blue	Κόκκινο–Πράσινο–Μπλε
Rotation	Στροφή
Saturation	Κορεσμός
Scanner	Σαρωτής εικόνας
Translation	Μεταφορά ή μετατόπιση



Δισδιάστατα ψηφιακά γραμμικά συστήματα

Σκοπός

Στο κεφάλαιο αυτό θα γνωρίσουμε τα δισδιάστατα ψηφιακά φίλτρα (*Two-dimensional digital filters*) και τις ιδιότητές τους. Η μετάβαση από τα μονοδιάστατα στα δισδιάστατα φίλτρα δεν είναι πολύ εύκολη, παρότι οι έννοιες είναι σε γενικές γραμμές οι ίδιες. Ο κυριότερος λόγος είναι ότι οι δισδιάστατες συναρτήσεις, όπως η δισδιάστατη κρουστική απόκριση και η δισδιάστατη απόκριση συχνότητας, είναι δυσκολότερες στη μαθηματική τους επεξεργασία από τις αντίστοιχες μονοδιάστατες. Σκοπός του κεφαλαίου είναι να γνωρίσουμε τα κύρια χαρακτηριστικά των δισδιάστατων φίλτρων που έχουν μεγάλη σημασία στην ψηφιακή επεξεργασία της εικόνας. Έτσι, θα γνωρίσουμε κυρίως τα δισδιάστατα *FIR* φίλτρα και τις εφαρμογές τους.

Προσδοκώμενα αποτελέσματα

Όταν ολοκληρώσετε τη μελέτη του κεφαλαίου αυτού θα είστε σε θέση να:

- Σχεδιάζετε τη μορφή δισδιάστατου σήματος ή την κρουστική απόκριση δισδιάστατου φίλτρου
- Υλοποιείτε δισδιάστατη συνέλιξη με σκοπό να υπολογίσετε την έξοδο δισδιάστατου ψηφιακού φίλτρου
- Αναγνωρίζετε αν ένα δισδιάστατο φίλτρο μπορεί να υλοποιηθεί ως συνδυασμός δύο μονοδιάστατων φίλτρων (διαχωρισσιμότητα)
- Σχεδιάζετε ένα απλό δισδιάστατο *FIR* φίλτρο
- Δείχνετε αν ένα δισδιάστατο φίλτρο έχει απόκριση μηδενικής φάσης (*zero phase response*)

Έννοιες κλειδιά

- Αμεταβλητότητα στην ολίσθηση ή μετατόπιση
- Αρχή της υπέρθεσης
- Ανιχνευτές ακμών
- Αντίστροφα προβλήματα

- Γραμμικά συστήματα
- Διακριτές χωρικές μεταβλητές
- Διαχωρίσιμες συναρτήσεις
- Διαχωρίσιμα φίλτρα
- Δισδιάστατα FIR φίλτρα
- Δισδιάστατη απόκριση συχνότητας
- Δισδιάστατη συνέλιξη
- Ευσταθή δισδιάστατα φίλτρα
- Μέθοδος των παραθύρων
- Μηδενική απόκριση φάσης
- Παράλληλη σύνδεση
- Περιοδική απόκριση συχνότητας
- Συμμετρική κρουστική απόκριση
- Σύνδεση σε σειρά
- Φίλτρο μέσης τιμής

Εισαγωγικές Παρατηρήσεις

Στο Κεφάλαιο 6 γνωρίσαμε τα βασικά χαρακτηριστικά της ψηφιακής εικόνας και τον τρόπο με τον οποίο την αντιλαμβάνεται και μπορεί να την επεξεργαστεί ο άνθρωπος. Δεδομένου ότι η ψηφιακή εικόνα είναι ένα δισδιάστατο σήμα που εκφράζει τις μεταβολές της αμαύρωσης στο επίπεδο (x,y) , η επεξεργασία της μπορεί να γίνει με τα δισδιάστατα ψηφιακά φίλτρα. Τα φίλτρα αυτά χαρακτηρίζονται από την κρουστική τους απόκριση. Η κρουστική απόκριση είναι δισδιάστατη διακριτή συνάρτηση και οι ιδιότητές της επηρεάζουν και καθορίζουν το σήμα εξόδου από το δισδιάστατο φίλτρο. Επειδή η εικόνα είναι πολύ ευαίσθητη στην αλλαγή της φασικής σχέσης μεταξύ των αρμονικών της, θα πρέπει κατά την ψηφιακή της επεξεργασία, το δισδιάστατο φίλτρο να μην αλλάζει τη σχέση αυτή. Τέτοια φίλτρα έχουν συμμετρική κρουστική απόκριση και καλούνται φίλτρα μηδενικής φάσης. Στο κεφάλαιο αυτό θα γνωρίσουμε τον τρόπο σχεδίασης απλών δισδιάστατων FIR φίλτρων που έχουν τις παραπάνω ιδιότητες. Το κεφάλαιο θα το ολοκληρώσουμε κάνοντας απλή αναφορά σε βασικές εφαρμογές των δισδιάστατων FIR φίλτρων στην ψηφιακή επεξεργασία της εικόνας. Σε όλη

την έκταση του κεφαλαίου οι έννοιες ψηφιακά φίλτρα και ψηφιακά συστήματα εναλλάσσονται. Απαραίτητη γνώση για το κεφάλαιο αυτό αποτελεί η ύλη των Κεφαλαίων 3, 4 και 6 του βιβλίου. Περισσότερες πληροφορίες ο αναγνώστης μπορεί να βρει στις αναφορές της ελληνικής και της ξενόγλωσσης βιβλιογραφίας.

Το κεφάλαιο αυτό περιλαμβάνει τις ακόλουθες ενότητες:

- 7.1. Δισδιάστατη συνέλιξη.
- 7.2. Χαρακτηριστικά των δισδιάστατων ψηφιακών συστημάτων.
- 7.3. Απόκριση στη συχνότητα των δισδιάστατων ψηφιακών συστημάτων.
- 7.4. Σχεδίαση των δισδιάστατων ψηφιακών φίλτρων.

7.1 Δισδιάστατη συνέλιξη

Ένα γραμμικό ψηφιακό σύστημα ουσιαστικά μετασχηματίζει ένα ψηφιακό σήμα $x(n)$ σε ένα νέο $y(n)$ (Σχήμα 7.1α). Η πράξη που εκτελεί είναι η συνέλιξη της κρουστικής απόκρισης $h(n)$ που χαρακτηρίζει το σύστημα με το σήμα εισόδου $x(n)$:

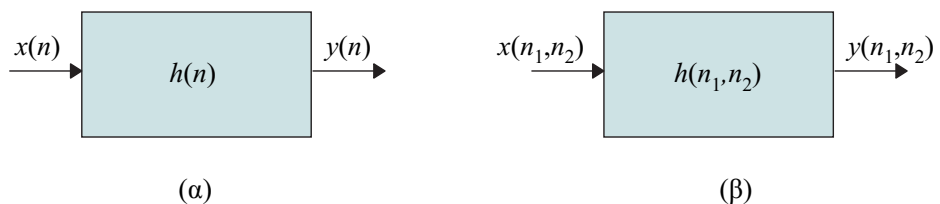
$$y(n) = x(n) * h(n) = \sum_{r=-\infty}^{\infty} x(n-r)h(r) \quad (7.1)$$

Με τον ίδιο τρόπο ένα δισδιάστατο ψηφιακό σήμα $x(n_1, n_2)$ μετασχηματίζεται από ένα δισδιάστατο γραμμικό ψηφιακό σύστημα σε ένα νέο δισδιάστατο ψηφιακό σήμα $y(n_1, n_2)$ όπως παρουσιάζεται στο Σχήμα 7.1β. Και εδώ η πράξη που εκτελεί το δισδιάστατο γραμμικό σύστημα είναι η συνέλιξη του σήματος εισόδου $x(n_1, n_2)$ με τη δισδιάστατη κρουστική του απόκριση $h(n_1, n_2)$:

$$y(n_1, n_2) = x(n_1, n_2) * h(n_1, n_2) = \sum_{r_1=-\infty}^{\infty} \sum_{r_2=-\infty}^{\infty} x(n_1 - r_1, n_2 - r_2) h(r_1, r_2) \quad (7.2)$$

Σχήμα 7.1

Σχηματική παράσταση α) Μονοδιάστατου ψηφιακού συστήματος β) Δισδιάστατου ψηφιακού συστήματος.



Στα μονοδιάστατα ψηφιακά φίλτρα η διακριτή μεταβλητή n αντιστοιχούσε κυρίως στο χρόνο, αφού τα μονοδιάστατα ψηφιακά σήματα είναι συνήθως σήματα χρόνου. Αντίθετα, στα δισδιάστατα ψηφιακά φίλτρα οι ανεξάρτητες μεταβλητές n_1 και n_2 είναι συνήθως διακριτές χωρικές μεταβλητές που αντιστοιχούν στις διευθύνσεις x και y . Ο κυριότερος λόγος για την επιλογή αυτή είναι ότι οι ψηφιακές εικόνες θεωρούνται δισδιάστατα διακριτά σήματα στο επίπεδο (x, y) .

Η σχέση (7.2) που μας περιγράφει τη δισδιάστατη συνέλιξη αποτελεί τη σχέση εισόδου–εξόδου του γραμμικού ψηφιακού συστήματος, αφού συνδέει τα δείγματα της εισόδου $x(n)$ με αυτά της εξόδου $y(n)$. Η δισδιάστατη συνέλιξη είναι άμεσα συνδεδεμένη με την ιδιότητα της γραμμικότητας και αυτή της αμεταβλητότητας στο χρόνο. Στην συνέχεια θα εξηγήσουμε τον τρόπο με τον οποίο οι δύο αυτές ιδιότητες μας οδηγούν στη δισδιάστατη συνέλιξη.

7.1.1 Γραμμικότητα

Χρησιμοποιώντας τις πράξεις της πρόσθεσης, του πολλαπλασιασμού με σταθερά και της ολίσθησης, είναι δυνατό να εκφράσουμε κάθε δισδιάστατο σήμα $x(n_1, n_2)$ ως γραμμικό συνδυασμό ολισθημένων δισδιάστατων μοναδιαίων κρουστικών $\delta(n_1, n_2)$:

$$x(n_1, n_2) = \sum_{l_1=-\infty}^{\infty} \sum_{l_2=-\infty}^{\infty} x(l_1, l_2) \delta(n_1 - l_1, n_2 - l_2) \quad (7.3)$$

Στην τελευταία σχέση η $\delta(n_1 - l_1, n_2 - l_2)$ είναι η δισδιάστατη μοναδιαία κρουστική μετατοπισμένη στη θέση $n_1 = l_1$ και $n_2 = l_2$ και οι τιμές $x(l_1, l_2)$ μπορούν να θεωρηθούν ως οι αντίστοιχοι πολλαπλασιαστές.

Όπως ήδη αναπτύξαμε στο Κεφάλαιο 3, για τα γραμμικά συστήματα (linear systems) ισχύει η αρχή της υπέρθεσης. Αυτό σημαίνει, ότι αν θέλουμε να υπολογίσουμε την απόκριση του γραμμικού συστήματος σε ένα σύνθετο σήμα που αποτελείται από το γραμμικό συνδυασμό επιμέρους απλών σημάτων, αρκεί να λάβουμε τον αντίστοιχο γραμμικό συνδυασμό των αποκρίσεων του συστήματος στα απλά αυτά σήματα. Έτσι, αν η συμπεριφορά του συστήματος συμβολιστεί με $L(\cdot)$ και το σήμα εισόδου σε αυτό είναι το σήμα της σχέσης (7.3) τότε η έξοδος είναι:

$$y(n_1, n_2) = L \left[\sum_{l_1=-\infty}^{\infty} \sum_{l_2=-\infty}^{\infty} x(l_1, l_2) \delta(n_1 - l_1, n_2 - l_2) \right] \quad (7.4)$$

Για ένα γραμμικό σύστημα, η (7.4) μπορεί να ξαναγραφεί ως

$$\begin{aligned} y(n_1, n_2) &= \sum_{l_1=-\infty}^{\infty} \sum_{l_2=-\infty}^{\infty} x(l_1, l_2) L[\delta(n_1 - l_1, n_2 - l_2)] \\ &= \sum_{l_1=-\infty}^{\infty} \sum_{l_2=-\infty}^{\infty} x(l_1, l_2) h_{l_1, l_2}(n_1, n_2) \end{aligned} \quad (7.5)$$

όπου h_{l_1, l_2} είναι η απόκριση του συστήματος στη δισδιάστατη κρουστική που βρίσκεται στη θέση (l_1, l_2) .

7.1.2 Αμεταβλητότητα στην ολίσθηση

Ένα φίλτρο καλείται *αμετάβλητο στην ολίσθηση* (shift invariant) αν συγκεκριμένη μετατόπιση του σήματος εισόδου προκαλεί την ίδια μετατόπιση στο σήμα της εξόδου. Έτσι, αν

$$y(n_1, n_2) = T[x(n_1, n_2)] \quad (7.6)$$

για αμετάβλητο στην ολίσθηση φίλτρο ισχύει

$$T[x(n_1 - l_1, n_2 - l_2)] = y(n_1 - l_1, n_2 - l_2) \quad (7.7)$$

Η αμεταβλητότητα στην ολίσθηση και η γραμμικότητα ενός φίλτρου είναι ιδιότητες ανεξάρτητες μεταξύ τους.

Δεδομένου ότι η αμεταβλητότητα στην ολίσθηση συνεπάγεται

$h_{l_1, l_2}(n_1, n_2) = h(n_1 - l_1, n_2 - l_2)$, για ένα γραμμικό φίλτρο που έχει και τις δύο ιδιότητες η σχέση (7.5) γράφεται ως

$$y(n_1, n_2) = \sum_{l_1=-\infty}^{\infty} \sum_{l_2=-\infty}^{\infty} x(l_1, l_2) h(n_1 - l_1, n_2 - l_2) \quad (7.8)$$

Η σχέση (7.8) περιγράφει τη δισδιάστατη συνέλιξη (two-dimensional convolution). Καθορίζει ότι το σήμα εξόδου $y(n_1, n_2)$ μπορεί να βρεθεί, αν η κρουστική απόκριση $h(n_1, n_2)$ του δισδιάστατου ψηφιακού φίλτρου μετατοπιστεί κατάλληλα $h(n_1 - l_1, n_2 - l_2)$ και πολλαπλασιαστεί κάθε φορά με το αντίστοιχο δείγμα του σήματος $x(l_1, l_2)$. Στο τέλος αθροίζουμε όλα τα γινόμενα για να λάβουμε το $y(n_1, n_2)$. Η σχέση (7.8) μπορεί ισοδύναμα να γραφεί και ως

$$y(n_1, n_2) = \sum_{l_1=-\infty}^{\infty} \sum_{l_2=-\infty}^{\infty} h(l_1, l_2) x(n_1 - l_1, n_2 - l_2) \quad (7.9)$$

Η δισδιάστατη συνέλιξη θα συμβολίζεται στο εξής με διπλό *:

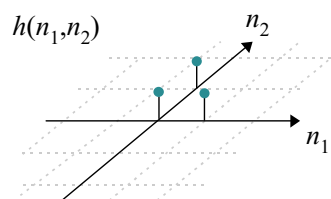
$$y = x ** h \quad \text{ή} \quad y = h ** x \quad (7.10)$$

Παράδειγμα 7.1

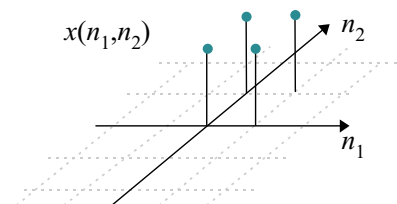
Η κρουστική απόκριση $h(n_1, n_2)$ δισδιάστατου ψηφιακού φίλτρου έχει όπως φαίνεται στο Σχήμα 7.2α, με όλους τους συντελεστές ίσους με $1/3$. Να βρεθεί το σήμα εξόδου $y(n_1, n_2)$ αν για το σήμα εισόδου $x(n_1, n_2)$ έχουμε $x(0,0) = x(1,0) = x(0,1) = x(1,1) = 1$ και 0 σε όλες τις άλλες θέσεις (n_1, n_2) (Σχήμα 7.2β).

Σχήμα 7.2

α) Η κρουστική απόκριση $h(n_1, n_2)$ του φίλτρου του Παραδείγματος 7.1 και β) το σήμα εισόδου $x(n_1, n_2)$.



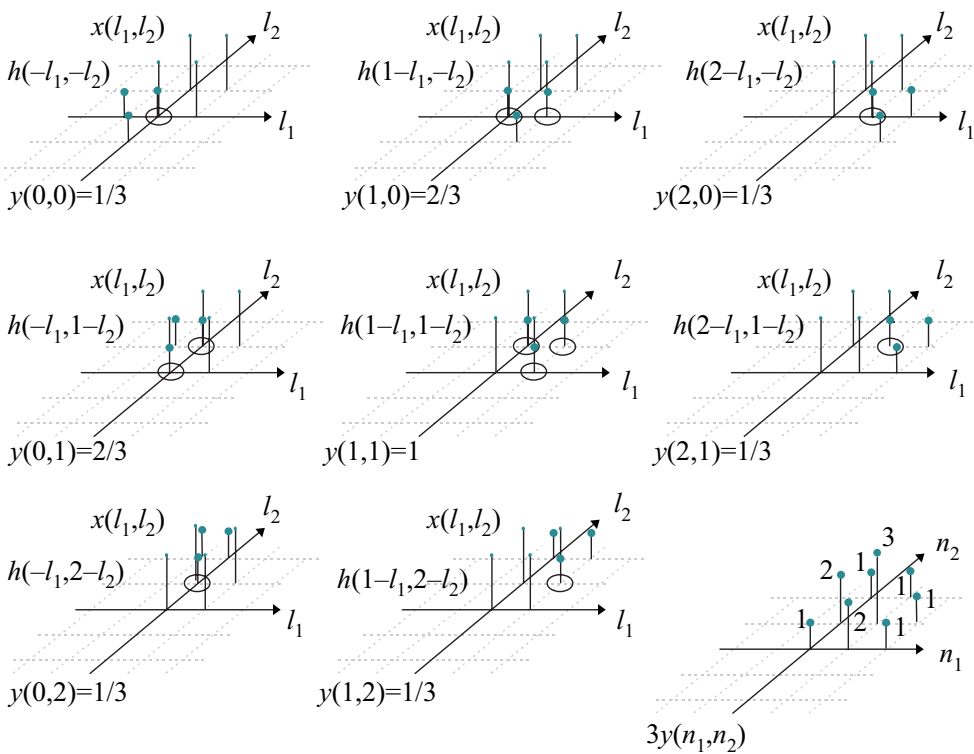
(α)



(β)

Λύση

Για να υπολογίσουμε το σήμα εξόδου θα χρησιμοποιήσουμε τη σχέση (7.8). Σύμφωνα με αυτή, η έξοδος μπορεί να σχηματιστεί αν λάβουμε αντίγραφα της $h(l_1, l_2)$ μετατοπισμένα κατάλληλα σε θέσεις (n_1, n_2) όπου υπάρχουν κοινές θέσεις με τους όρους του σήματος εισόδου $x(l_1, l_2)$. Το άθροισμα των γινομένων των όρων με τις κοινές θέσεις θα δώσει τους όρους της σήματος εξόδου $y(n_1, n_2)$. Η διαδικασία παρουσιάζεται στο Σχήμα 7.3.



Σχήμα 7.3

Υπολογισμός της εξόδου $y(n_1, n_2)$ με υλοποίηση της συνέλιξης που περιγράφεται από τη σχέση (7.8). Μόνο το γινόμενο των όρων $x(l_1, l_2)$ και $h(l_1, l_2)$ που συμπίπτουν και είναι σημειωμένοι με κύκλο, συνεισφέρουν στον τελικό σχηματισμό της $y(n_1, n_2)$.

Παράδειγμα 7.2

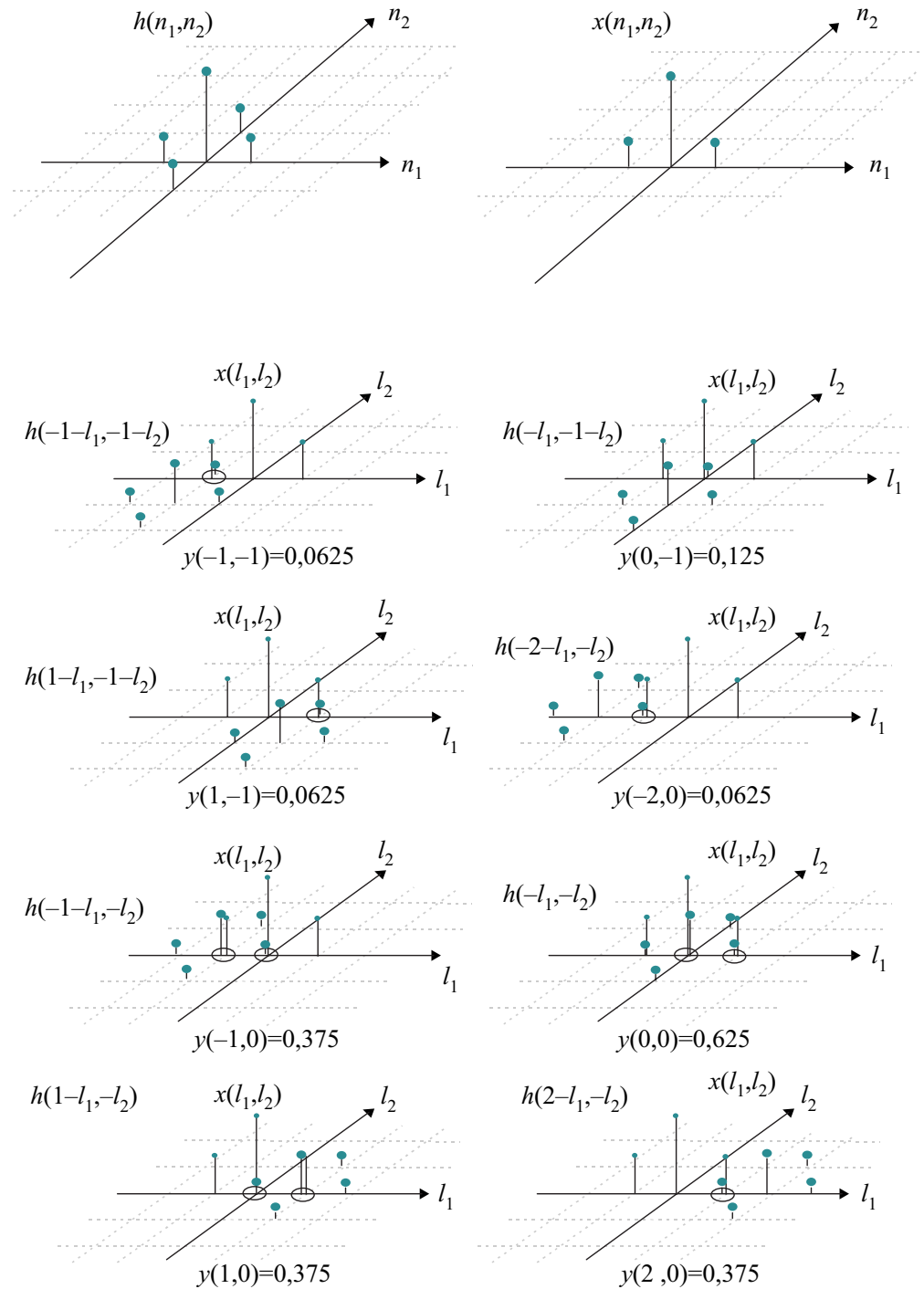
Ψηφιακό φίλτρο δύο διαστάσεων έχει κρουστική απόκριση $h(n_1, n_2)$ τέτοια, ώστε $h(0,0) = 0,5$, $h(1,0) = h(-1,0) = h(0,1) = h(0,-1) = 0,125$ και 0 οπουδήποτε αλλού. Να σχεδιαστεί η κρουστική απόκριση και να βρεθεί το σήμα εξόδου $y(n_1, n_2)$ αν το σήμα εισόδου $x(n_1, n_2)$ είναι $x(0,0) = 1$, $x(-1,0) = x(1,0) = 0,5$.

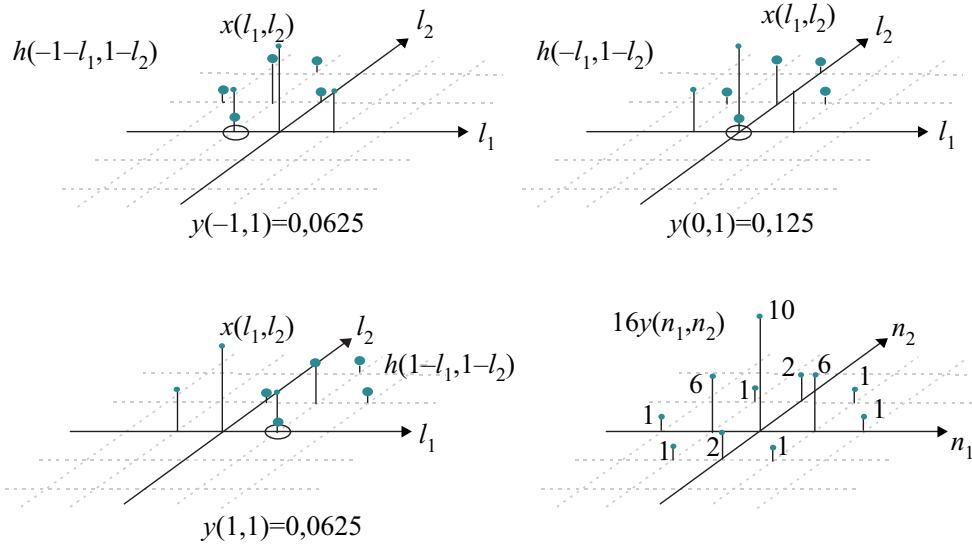
Λύση

Η κρουστική απόκριση $h(n_1, n_2)$ του φίλτρου αλλά και το σήμα εισόδου $x(n_1, n_2)$ δίνονται στο Σχήμα 7.4. Για να υπολογίσουμε το σήμα εξόδου $y(n_1, n_2)$ θα χρησιμοποιήσουμε τη σχέση (7.8). Σύμφωνα με αυτή η έξοδος $y(n_1, n_2)$ μπορεί να σχηματιστεί αν

λάβουμε αντίγραφα της $h(n_1, n_2)$ σε θέσεις (n_1, n_2) όπου το σήμα εισόδου $x(n_1, n_2)$ είναι μη μηδενικό. Τα αντίγραφα αυτά θα πολλαπλασιαστούν με την αντίστοιχη τιμή του σήματος και τα αποτελέσματα θα προστεθούν. Η διαδικασία παρουσιάζεται στο Σχήμα 7.5.

Σχήμα 7.4
 Η κρουστική απόκριση $h(n_1, n_2)$ του παραδείγματος 7.2 και το σήμα εισόδου $x(n_1, n_2)$



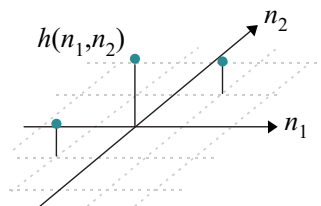


Σχήμα 7.5
Υπολογισμός της εξόδου $y(n_1, n_2)$ με υλοποίηση της συνέλιξης που περιγράφεται από τη σχέση (7.8).

Δραστηριότητα 7.1

Να σχεδιάσετε την κρουστική απόκριση δισδιάστατου ψηφιακού φίλτρου όταν αυτή έχει τρεις όρους και είναι συμμετρική γύρω από τη θέση $(n_1, n_2) = (0, 0)$.

Η κρουστική απόκριση θα είναι συμμετρική με τον έναν όρο να βρίσκεται στη θέση $(n_1, n_2) = (0, 0)$ και τους άλλους δύο εκατέρωθεν αυτού και ίσους μεταξύ τους. Με τρεις μόνο όρους η συμμετρική κρουστική απόκριση μπορεί να έχει διάφορες μορφές (συνολικά 4). Μία από αυτές δίνεται στο Σχήμα 7.6.

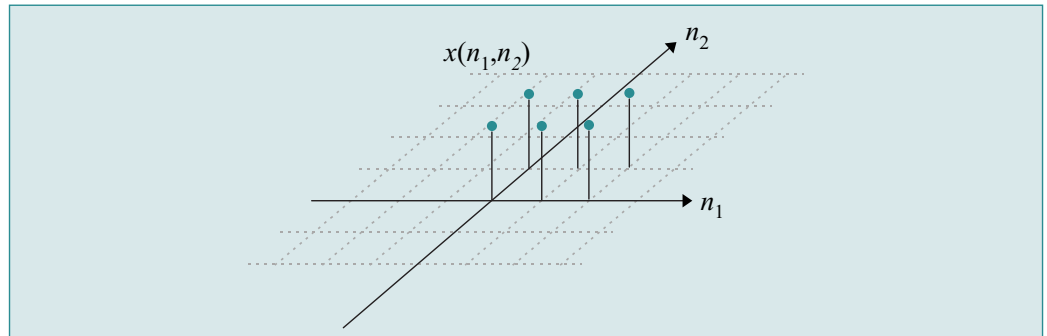


Σχήμα 7.6
Συμμετρική δισδιάστατη κρουστική απόκριση με τρεις όρους.

Άσκηση Αυτοαξιολόγησης 7.1

Θέλουμε να υπολογίσουμε την έξοδο $y(n_1, n_2)$ δισδιάστατου ψηφιακού φίλτρου με κρουστική απόκριση $h(n_1, n_2)$ τέτοια, ώστε $h(0, 0) = 0,5$, $h(0, 1) = h(0, -1) = 0,25$ και 0 οπουδήποτε αλλού. Σύμφωνα με τη σχέση (7.9) το σήμα εισόδου που φαίνεται στο Σχήμα 7.7 πρέπει να υποστεί 5, 10, 15 ή 20 μετατοπίσεις πριν από την τελική άθροιση των γινομένων;

Σχήμα 7.7
Το σήμα εισόδου $x(n_1, n_2)$ για την Άσκηση Αυτοαξιολόγησης 7.1.



Σύνοψη ενότητας

Στην ενότητα αυτή γνωρίσαμε τα γραμμικά, αμετάβλητα στην ολίσθηση δισδιάστατα ψηφιακά φίλτρα. Ο τρόπος δράσης των φίλτρων αυτών καθορίζεται από τη δισδιάστατη κρουστική τους απόκριση. Πιο συγκεκριμένα, το σήμα εξόδου υπολογίζεται ως η δισδιάστατη συνέλιξη της κρουστικής απόκρισης και του σήματος εισόδου. Η δισδιάστατη συνέλιξη είναι αποτέλεσμα της γραμμικότητας και της αμεταβλητότητας των φίλτρων στην ολίσθηση.

7.2 Χαρακτηριστικά των δισδιάστατων ψηφιακών συστημάτων

7.2.1 Τρόποι σύνδεσης των δισδιάστατων ψηφιακών φίλτρων

Η πράξη της δισδιάστατης συνέλιξης έχει σημαντικές ιδιότητες οι οποίες είναι χρήσιμες στην ανάλυση των δισδιάστατων, γραμμικών αμετάβλητων στην ολίσθηση συστημάτων. Έτσι, αν δύο συστήματα με κρουστικές αποκρίσεις $h(n_1, n_2)$ και $g(n_1, n_2)$ είναι συνδεδεμένα σε σειρά (cascade), όπως φαίνεται στο Σχήμα 7.8, τότε η τελική έξοδος $y(n_1, n_2)$ μπορεί να υπολογιστεί χρησιμοποιώντας το σήμα εισόδου $x(n_1, n_2)$, το ενδιάμεσο σήμα $w(n_1, n_2)$ και το γεγονός ότι κάθε ένα φίλτρο υλοποιεί δισδιάστατη συνέλιξη:

$$\begin{aligned} w &= x ** h \\ y &= w ** g = (x ** h) ** g \end{aligned} \quad (7.11)$$

Δεδομένου ότι για τη δισδιάστατη συνέλιξη ισχύει η προσεταιριστική ιδιότητα, η σχέση (7.11) μπορεί να γραφεί

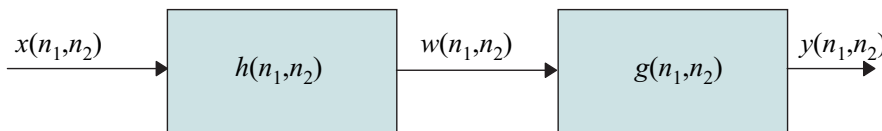
$$y = x ** (h ** g) = x ** h_{\text{σοδ}} \quad (7.12)$$

Από τη σχέση 7.12 προκύπτει ότι τα δύο σε σειρά φίλτρα ισοδυναμούν με ένα τρίτο που έχει κρουστική απόκριση αυτή που προκύπτει από τη συνέλιξη των επιμέρους

κρουστικών αποκρίσεων. Επίσης, δεδομένου ότι για τη δισδιάστατη συνέλιξη ισχύει η αντιμεταθετική ιδιότητα

$$h ** g = g ** h,$$

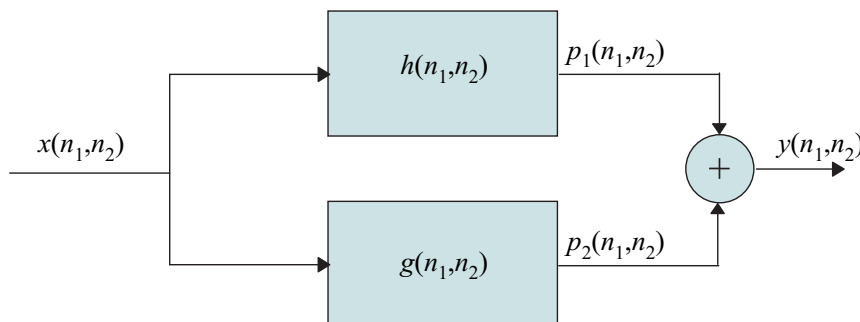
η συνολική συμπεριφορά του δισδιάστατου φίλτρου δεν αλλάζει αν εφαρμοστεί στο σήμα $x(n_1, n_2)$ πρώτα το φίλτρο g και μετά το h .



Στην περίπτωση που έχουμε παράλληλη συνδεσμολογία δύο δισδιάστατων, γραμμικών αμετάβλητων στην ολίσθηση ψηφιακών φίλτρων, όπως στο Σχήμα 7.9, η έξοδος $y(n_1, n_2)$ δίνεται ως

$$\begin{aligned} y(n_1, n_2) &= p_1(n_1, n_2) + p_2(n_1, n_2) = \\ &= x ** h + x ** g = \\ &= x ** (h + g) = \\ &= x ** h_{\text{ισοδ}} \end{aligned} \quad (7.13)$$

Η σχέση (7.13) δηλώνει ότι, αν δύο δισδιάστατα, γραμμικά και αμετάβλητα στην ολίσθηση ψηφιακά συστήματα συνδεθούν παράλληλα, ισοδυναμούν με ένα τρίτο που έχει κρουστική απόκριση το άθροισμα των επιμέρους κρουστικών αποκρίσεων.



Παράδειγμα 7.3

Για δύο δισδιάστατα, γραμμικά, αμετάβλητα στην ολίσθηση φίλτρα οι κρουστικές αποκρίσεις $h(n_1, n_2)$ και $g(n_1, n_2)$ έχουν όπως παρουσιάζονται στο Σχήμα 7.10. Να σχεδιαστεί η κρουστική απόκριση του φίλτρου που προκύπτει αν τα φίλτρα αυτά συνδεθούν α) σε σειρά και β) παράλληλα.

Σχήμα 7.8

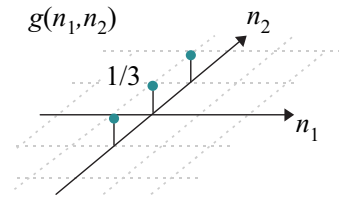
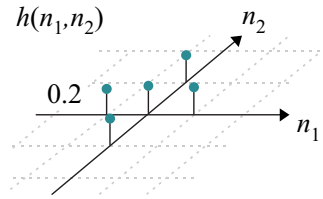
Δισδιάστατα γραμμικά φίλτρα συνδεδεμένα σε σειρά. Η κρουστική απόκριση του συνδυασμού προκύπτει από τη δισδιάστατη συνέλιξη $h ** g$. Η αλλαγή της σειράς των δύο φίλτρων δίνει θεωρητικά το ίδιο πάντα αποτέλεσμα.

Σχήμα 7.9

Το παράλληλο σύστημα, δισδιάστατων, γραμμικών και αμετάβλητων στην ολίσθηση ψηφιακών φίλτρων ισοδυναμεί με φίλτρο που έχει κρουστική απόκριση $h_{\text{ισοδ}}(n_1, n_2) = h(n_1, n_2) + g(n_1, n_2)$.

Σχήμα 7.10

Κρουστικές αποκρίσεις δισδιάστατων, γραμμικών, αμετάβλητων στην ολίσθηση ψηφιακών φίλτρων.



Λύση

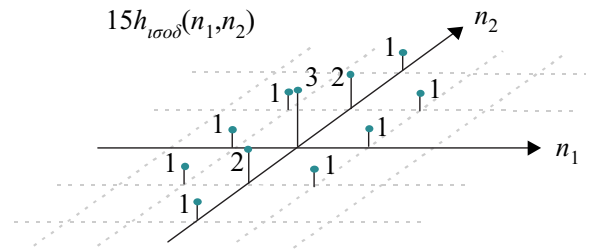
Στην περίπτωση σύνδεσης σε σειρά η κρουστική απόκριση του συνολικού φίλτρου δίνεται από τη συνέλιξη των κρουστικών αποκρίσεων του Σχήματος 7.10:

$$h_{ισοδ} = h^{**}g$$

Για να κάνουμε αυτόν τον υπολογισμό θα μετατοπίσουμε την $h(n_1, n_2)$ σε όλες εκείνες τις θέσεις που έχει με την $g(n_1, n_2)$ σύμπτωση κάποιων όρων. Η διαδικασία είναι παρόμοια με τη συνέλιξη που περιγράφεται στο Σχήμα 7.5. Το τελικό αποτέλεσμα φαίνεται στο Σχήμα 7.11.

Σχήμα 7.11

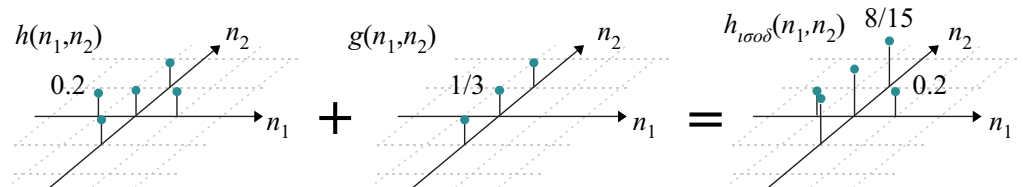
Η κρουστική απόκριση του εν σειρά συνδυασμού των φίλτρων h και g προκύπτει από τη συνέλιξη των δύο κρουστικών αποκρίσεων.



Σχήμα 7.12

Η κρουστική απόκριση του παράλληλου συνδυασμού των φίλτρων h και g προκύπτει από την απλή άθροιση των δύο κρουστικών αποκρίσεων.

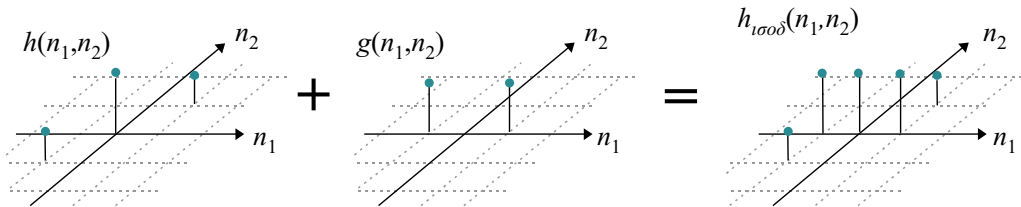
Στην περίπτωση της σύνδεσης των δύο φίλτρων παράλληλα, η ισοδύναμη κρουστική απόκριση προκύπτει από την απλή άθροιση των δύο επιμέρους κρουστικών αποκρίσεων σύμφωνα με τη σχέση (7.13) και φαίνεται στο Σχήμα 7.12.



Δραστηριότητα 7.2

Ποια είναι η κρουστική απόκριση $h_{ισοδ}$ δισδιάστατου γραμμικού φίλτρου που προκύπτει από τον παράλληλο συνδυασμό του δισδιάστατου φίλτρου της Δραστηριότητας 7.1, με δισδιάστατο φίλτρο που έχει κρουστική απόκριση $g(n_1, n_2) = 0,5[\delta(n_1-1, n_2) + \delta(n_1 + 1, n_2)]$;

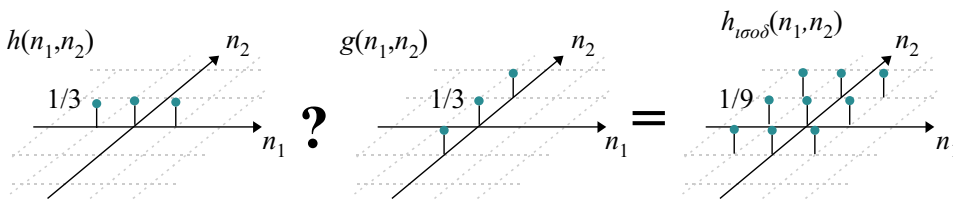
Η ισοδύναμη κρουστική απόκριση θα προκύψει από την πρόσθεση των δύο κρουστικών αποκρίσεων όπως φαίνεται στο Σχήμα 7.13. Στο σχήμα αυτό δίνεται η συμμετρική $h(n_1, n_2)$ της Δραστηριότητας 7.1, η $g(n_1, n_2)$ και η $h_{ισοδ}(n_1, n_2)$.

**Σχήμα 7.13**

Η κρουστική απόκριση του παράλληλου συνδυασμού των φίλτρων h και g προκύπτει από την απλή άθροιση των δύο κρουστικών αποκρίσεων.

Άσκηση Αυτοαξιολόγησης 7.2

Στο Σχήμα 7.14 δίνονται οι κρουστικές αποκρίσεις $h(n_1, n_2)$ και $g(n_1, n_2)$ δύο φίλτρων. Η $h_{ισοδ}(n_1, n_2)$, που δίνεται δίπλα προκύπτει από τον παράλληλο ή τον εν σειρά συνδυασμό τους;

**Σχήμα 7.14**

Κρουστικές αποκρίσεις δισδιάστατων γραμμικών αμετάβλητων στην ολίσθηση ψηφιακών φίλτρων.

7.2.2 Διαχωρίσιμα δισδιάστατα ψηφιακά συστήματα

Ένα διαχωρίσιμο (separable) δισδιάστατο σύστημα έχει κρουστική απόκριση $h(n_1, n_2)$ τέτοια ώστε να μπορεί να γραφεί ως γινόμενο δύο μονοδιάστατων ακολουθιών:

$$h(n_1, n_2) = h_1(n_1) \cdot h_2(n_2) \quad (7.14)$$

Στην περίπτωση αυτή η έξοδος $y(n_1, n_2)$ του συστήματος για είσοδο $x(n_1, n_2)$ μπορεί να γραφεί:

$$\begin{aligned} y(n_1, n_2) &= \sum_{l_1=-\infty}^{\infty} \sum_{l_2=-\infty}^{\infty} x(n_1 - l_1, n_2 - l_2) h_1(l_1) h_2(l_2) \\ &= \sum_{l_1=-\infty}^{\infty} h_1(l_1) \sum_{l_2=-\infty}^{\infty} x(n_1 - l_1, n_2 - l_2) h_2(l_2) \end{aligned} \quad (7.15)$$

Το γινόμενο στο δεύτερο άθροισμα της τελευταίας σχέσης αντιπροσωπεύει ένα δισδιάστατο πίνακα αριθμών. Εάν θέσουμε

$$f(n_1, n_2) = \sum_{l_2=-\infty}^{\infty} x(n_1, n_2 - l_2) h_2(l_2) \quad (7.16)$$

η σχέση (7.15) γίνεται

$$y(n_1, n_2) = \sum_{l_1=-\infty}^{\infty} h_1(l_1) f(n_1 - l_1, n_2) \quad (7.17)$$

Η διαδικασία αυτή δηλώνει ότι μπορούμε να υπολογίσουμε πρώτα την ακολουθία $f(n_1, n_2)$ πραγματοποιώντας μονοδιάστατη συνέλιξη στις στήλες του πίνακα $x(n_1, n_2)$ ($n_1 = \text{σταθερό}$). Στη συνέχεια υπολογίζουμε τις μονοδιάστατες συνέλιξεις στις γραμμές της $f(n_1, n_2)$ ($n_2 = \text{σταθερό}$) σύμφωνα με τη σχέση (7.17).

Παράδειγμα 7.4

Να αποδειχθεί ότι η κρουστική απόκριση που περιγράφεται από τη σχέση

$$h(n_1, n_2) = \begin{cases} \exp\left(-\frac{n_1^2 + n_2^2}{5}\right) & |n_1| \leq 5, |n_2| \leq 5 \\ 0 & \text{οπουδήποτε αλλού} \end{cases} \quad (7.18)$$

αντιστοιχεί σε διαχωρίσιμο δισδιάστατο σύστημα.

Λύση

Η δισδιάστατη κρουστική απόκριση $h(n_1, n_2)$ μπορεί να γραφεί ως

$$h(n_1, n_2) = \exp\left(-\frac{n_1^2 + n_2^2}{5}\right) = \exp\left(-\frac{n_1^2}{5}\right) \exp\left(-\frac{n_2^2}{5}\right) = h_1(n_1) h_2(n_2) \quad (7.19)$$

δηλαδή, ως γινόμενο δύο μονοδιάστατων κρουστικών αποκρίσεων. Επομένως, είναι διαχωρίσιμη.

7.2.3 Ευσταθή δισδιάστατα συστήματα

Όπως και στην περίπτωση των μονοδιάστατων φίλτρων, τα δισδιάστατα συστήματα που μπορούν να χρησιμοποιηθούν είναι μόνο τα ευσταθή. Ένα τέτοιο φίλτρο δίνει πεπερασμένη έξοδο για κάθε πεπερασμένη σε πλάτος είσοδο. Η αναγκαία και ικανή συνθήκη για να είναι ευσταθές ένα δισδιάστατο, γραμμικό, αμετάβλητο στην ολίσθηση ψηφιακό φίλτρο είναι η ακόλουθη:

$$\sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} |h(n_1, n_2)| = \alpha < \infty \quad (7.20)$$

όπου η $h(n_1, n_2)$ είναι η κρουστική απόκριση του φίλτρου.

Άσκηση Αυτοαξιολόγησης 7.3

Η κρουστική απόκριση που δίνεται στο Παράδειγμα 7.4 αντιστοιχεί σε ευσταθές ή ασταθές δισδιάστατο φίλτρο και γιατί;

7.3 Απόκριση στη συχνότητα των δισδιάστατων ψηφιακών συστημάτων

Στις προηγούμενες παραγράφους εξηγήσαμε τον τρόπο με τον οποίο υπολογίζεται η έξοδος $y(n_1, n_2)$ ενός δισδιάστατου ψηφιακού συστήματος. Προς τούτο είναι απαραίτητο να πραγματοποιηθεί η πράξη της δισδιάστατης συνέλιξης μεταξύ του σήματος εισόδου $x(n_1, n_2)$ και της κρουστικής απόκρισης $h(n_1, n_2)$ του συστήματος. Στην ουσία η κρουστική απόκριση $h(n_1, n_2)$ καθορίζει πλήρως την απόκριση του συστήματος για κάθε (n_1, n_2) . Στη συνέχεια θα δείξουμε τον τρόπο που η συνάρτηση $h(n_1, n_2)$ καθορίζει τη συμπεριφορά του φίλτρου στις χωρικές συχνότητες ω_1 και ω_2 , επομένως θα επηρεάζει και το φάσμα του σήματος εισόδου (όπως αυτό το γνωρίσαμε στην Ενότητα 6.4).

Για να προχωρήσουμε στην ανάλυση της συμπεριφοράς του δισδιάστατου συστήματος στη συχνότητα θα χρησιμοποιήσουμε το σήμα

$$x(n_1, n_2) = e^{j\omega_1 n_1 + j\omega_2 n_2}$$

Το σήμα αυτό αντιστοιχεί σε ένα μιγαδικό δισδιάστατο ημιτονοειδές με ω_1 και ω_2 πραγματικούς αριθμούς που αντιστοιχούν στις χωρικές συχνότητες ω_1 και ω_2 αντίστοιχα. Η έξοδος $y(n_1, n_2)$ σύμφωνα με τη σχέση της συνέλιξης (7.9) έχει ως εξής:

$$\begin{aligned} y(n_1, n_2) &= \sum_{r_1=-\infty}^{\infty} \sum_{r_2=-\infty}^{\infty} e^{j\omega_1(n_1-r_1) + j\omega_2(n_2-r_2)} h(r_1, r_2) \\ &= e^{j\omega_1 n_1 + j\omega_2 n_2} \sum_{r_1=-\infty}^{\infty} \sum_{r_2=-\infty}^{\infty} e^{-j\omega_1 r_1 - j\omega_2 r_2} h(r_1, r_2) \\ &= e^{j\omega_1 n_1 + j\omega_2 n_2} H(\omega_1, \omega_2) \end{aligned} \quad (7.21)$$

Η τελευταία σχέση δείχνει ότι το σήμα εξόδου $y(n_1, n_2)$ αποτελείται από τις ίδιες

συχνότητες με το σήμα εισόδου $x(n_1, n_2)$, αλλά το πλάτος τους και η φάση τους έχουν υποστεί την επίδραση της συνάρτησης $H(\omega_1, \omega_2)$. Η συνάρτηση αυτή, όπως και στην περίπτωση των μονοδιάστατων σημάτων του Κεφαλαίου 3, καλείται «απόκριση συχνότητας του φίλτρου» και προφανώς ισούται με:

$$H(\omega_1, \omega_2) = \sum_{r_1=-\infty}^{\infty} \sum_{r_2=-\infty}^{\infty} e^{-j\omega_1 r_1 - j\omega_2 r_2} h(r_1, r_2) \quad (7.22)$$

Το μέτρο $|H(\omega_1, \omega_2)|$ της συνάρτησης καθορίζει το πώς θα επηρεαστούν οι συχνότητες του σήματος από το δισδιάστατο σύστημα. Για τις συχνότητες (ω_1, ω_2) όπου $|H(\omega_1, \omega_2)| \rightarrow 0$, οι αντίστοιχες συχνότητες του σήματος εισόδου αποκόπτονται. Ουσιαστικά η σχέση (7.22) δηλώνει ότι η απόκριση συχνότητας του συστήματος είναι ο δισδιάστατος διακριτός μετασχηματισμός Fourier της κρουστικής απόκρισης του συστήματος.

Η απόκριση συχνότητας $H(\omega_1, \omega_2)$ είναι περιοδική συνάρτηση των συχνοτήτων ω_1 και ω_2 με περίοδο 2π :

$$\begin{aligned} H(\omega_1 + 2\pi, \omega_2) &= H(\omega_1, \omega_2) \\ H(\omega_1, \omega_2 + 2\pi) &= H(\omega_1, \omega_2) \end{aligned} \quad (7.23)$$

Η περιοδικότητα της απόκρισης των ψηφιακών συστημάτων είναι αναμενόμενη λόγω της ψηφιακής φύσης αυτών. Ίδια συμπεριφορά έχει και το φάσμα των δισδιάστατων ψηφιακών σημάτων (Ενότητα 6.4). Ανάλογη δε συμπεριφορά παρουσιάζουν και τα μονοδιάστατα ψηφιακά σήματα και συστήματα (Κεφάλαια 1 έως 3).

Παράδειγμα 7.5

Να υπολογιστεί το μέτρο της απόκρισης στη συχνότητα του δισδιάστατου ψηφιακού συστήματος με κρουστική απόκριση

$$h(n_1, n_2) = \delta(n_1 + 1, n_2) + \delta(n_1 - 1, n_2) + \delta(n_1, n_2)$$

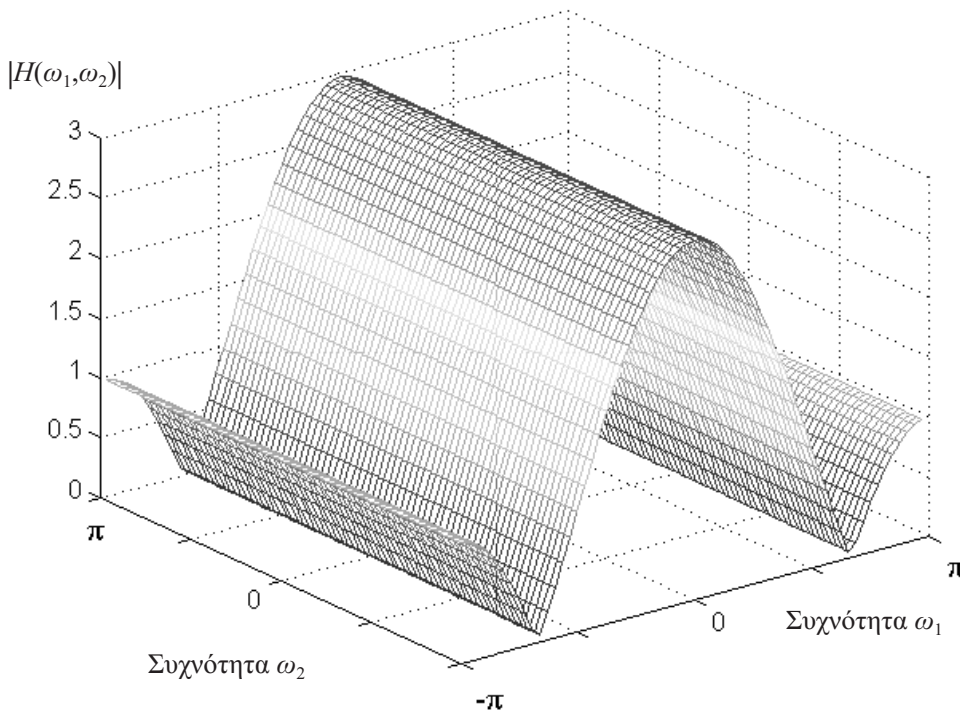
Λύση

Δεδομένου ότι η κρουστική απόκριση έχει μόνο τρεις όρους ίσους με τη μονάδα στις θέσεις (n_1, n_2) : $(-1, 0)$, $(1, 0)$ και $(0, 0)$, η σχέση (7.22) γράφεται

$$\begin{aligned} H(\omega_1, \omega_2) &= e^0 + e^{+j\omega_1} + e^{-j\omega_1} = \\ &= 1 + 2\cos(\omega_1) \end{aligned} \quad (7.24)$$

Η απόκριση συχνότητας που περιγράφεται από τη σχέση (7.24) είναι πραγματική,

επομένως δεν αλλάζει τη φάση των αρμονικών συνιστωσών του σήματος, μεταβάλλεται δε μόνο στην κατεύθυνση της ω_1 . Χρησιμοποιώντας κατάλληλο λογισμικό σχεδίασης (λ.χ. MATLAB), μπορούμε να δούμε ότι η συνάρτηση $|H(\omega_1, \omega_2)|$ έχει τη μορφή που δείχνεται στο Σχήμα 7.15.



Σχήμα 7.15

Η μορφή της συνάρτησης $|H(\omega_1, \omega_2)| = |1 + 2\cos(\omega_1)|$, όπως έχει ληφθεί με το MATLAB.

Δραστηριότητα 7.3

Να υπολογιστεί το μέτρο της απόκρισης στη συχνότητα δισδιάστατου ψηφιακού φίλτρου μέσης τιμής της επιλογής σας.

Ένα φίλτρο μέσης τιμής υπολογίζει τη μέση τιμή ορισμένων δειγμάτων του σήματος εισόδου $x(n_1, n_2)$ που βρίσκονται γύρω από τη θέση (n_1, n_2) , και δίνει το αποτέλεσμα στο δείγμα του σήματος εξόδου $y(n_1, n_2)$ που βρίσκεται στην αντίστοιχη θέση. Επειδή η διαδικασία αυτή επαναλαμβάνεται, καθώς μετακινούμαστε σε όλες τις θέσεις (n_1, n_2) , το φίλτρο καλείται και *φίλτρο του κινούμενου μέσου*. Πρόκειται για ένα βαθυ-

περατό φίλτρο του οποίου οι συχνότητες αποκοπής εξαρτώνται από το πλήθος των δειγμάτων που θα ληφθούν από το σήμα εισόδου. Ως βαθυπερατό φίλτρο, αποκόπτει τις υψηλές συχνότητες, «λειαίνει» (smooths) το θόρυβο, αλλά προκαλεί θόλωση (blurring) στις ακμές των εικόνων.

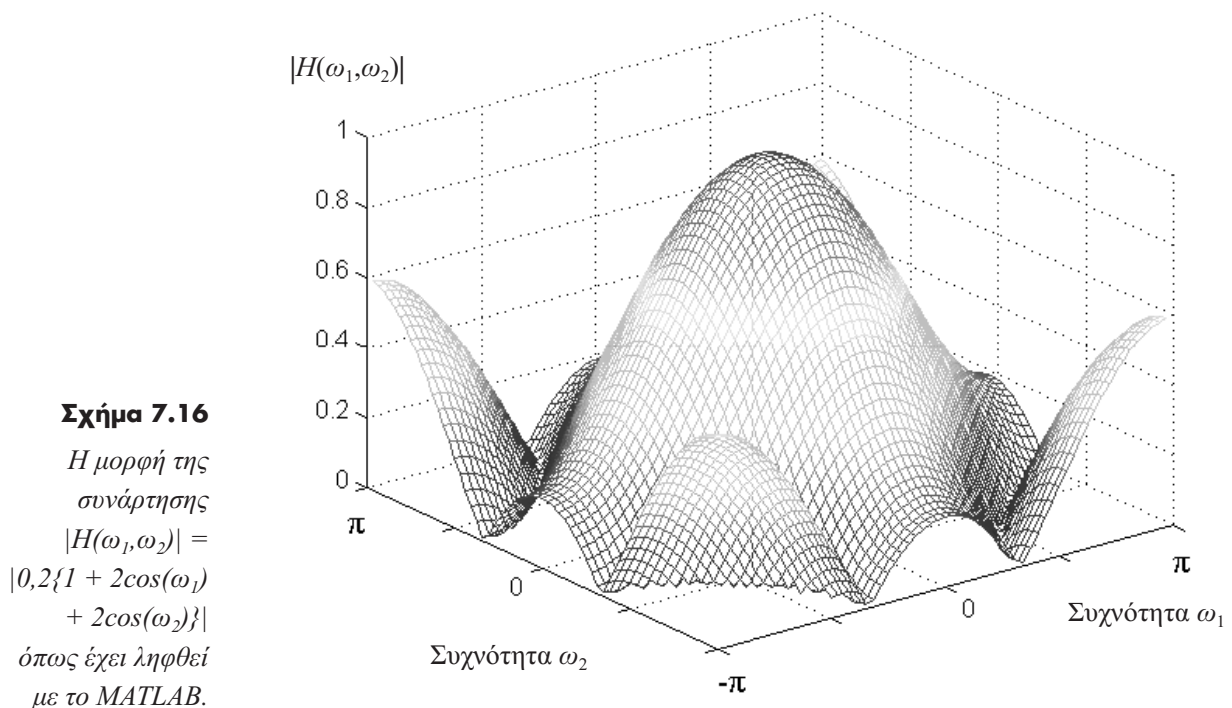
Ένα απλό φίλτρο μέσης τιμής είναι αυτό που περιγράφηκε στο Παράδειγμα 7.5. Σ' αυτό κάθε τιμή του σήματος εξόδου $y(n_1, n_2)$ υπολογίζεται από τρεις τιμές του σήματος εισόδου, αυτές για τις οποίες η κρουστική απόκριση του φίλτρου $h(n_1, n_2)$ είναι μη μηδενική. Ένα άλλο φίλτρο μέσης τιμής είναι αυτό του οποίου η κρουστική απόκριση έχει ως εξής:

$$h(n_1, n_2) = \frac{1}{5} [\delta(n_1, n_2) + \delta(n_1 - 1, n_2) + \delta(n_1 + 1, n_2) + \delta(n_1, n_2 - 1) + \delta(n_1, n_2 + 1)]$$

Κάθε δείγμα του σήματος εξόδου υπολογίζεται από τη μέση τιμή των πέντε δειγμάτων του σήματος εισόδου $x(n_1, n_2)$, $x(n_1 - 1, n_2)$, $x(n_1 + 1, n_2)$, $x(n_1, n_2 - 1)$ και $x(n_1, n_2 + 1)$. Σύμφωνα με την εξίσωση (7.22) το μέτρο της απόκρισης συχνότητας $H(\omega_1, \omega_2)$ έχει ως εξής:

$$\begin{aligned} H(\omega_1, \omega_2) &= \frac{1}{5} [e^0 + e^{j\omega_1} + e^{-j\omega_1} + e^{j\omega_2} + e^{-j\omega_2}] = \\ &= \frac{1}{5} [1 + 2 \cos \omega_1 + 2 \cos \omega_2] \end{aligned} \quad (7.25)$$

Δεδομένου ότι η τελευταία ποσότητα είναι πραγματική το μέτρο της είναι η απόλυτη τιμή της, η οποία γραφικά έχει όπως φαίνεται στο Σχήμα 7.16.



Άσκηση Αυτοαξιολόγησης 7.4

Είδαμε στα προηγούμενα ότι είναι δυνατό η απόκριση συχνότητας του δισδιάστατου φίλτρου να είναι πραγματική. Στην περίπτωση αυτή το φίλτρο δεν αλλοιώνει τη φάση των αρμονικών του σήματος (φίλτρο μηδενικής φάσης). Ποιο από τα επόμενα χαρακτηριστικά πρέπει, κατά τη γνώμη σας, να έχει η κρουστική απόκριση του φίλτρου για να είναι αυτό μηδενικής φάσης;

- α. Οι όροι της κρουστικής απόκρισης να είναι όλοι ίσοι.
- β. Η κρουστική απόκριση του φίλτρου να είναι συμμετρική.
- γ. Η κρουστική απόκριση να είναι διαχωρίσιμη.

Σύνοψη ενότητας

Στην ενότητα αυτή είδαμε ότι η κρουστική απόκριση ενός δισδιάστατου ψηφιακού φίλτρου καθορίζει και την απόκρισή του στη συχνότητα. Όταν η κρουστική απόκριση είναι συμμετρική, η απόκριση στη συχνότητα είναι πραγματική και δεν επηρεάζει τη φάση των συνιστωσών του σήματος. Στην περίπτωση αυτή τα φίλτρα καλούνται «φίλτρα μηδενικής φάσης» (zero phase filters).

7.4 Σχεδίαση δισδιάστατων ψηφιακών φίλτρων

Η σχεδίαση των δισδιάστατων ψηφιακών φίλτρων είναι δυσκολότερη από αυτή των μονοδιάστατων. Επίσης, η υλοποίησή τους είναι πολύπλοκη γιατί στις περισσότερες περιπτώσεις οι δισδιάστατες συναρτήσεις δεν μπορούν να παραγοντοποιηθούν, έτσι ώστε να οδηγήσουν σε απλοποιημένες δομές.

Στην παρούσα ενότητα θα προσδιορίσουμε αρχικά την κρουστική απόκριση ενός δισδιάστατου ψηφιακού φίλτρου από τις επιθυμητές προδιαγραφές της απόκρισής του $H(\omega_1, \omega_2)$ στις συχνότητες ω_1 και ω_2 . Στη συνέχεια θα γνωρίσουμε μία από τις μεθόδους σχεδίασης των δισδιάστατων FIR φίλτρων, αυτή των παραθύρων. Θα κλείσουμε την ενότητα με αναφορά σε εφαρμογές των δισδιάστατων ψηφιακών φίλτρων στην επεξεργασία της εικόνας.

7.4.1 Υπολογισμός της κρουστικής απόκρισης από την απόκριση συχνότητας

Σύμφωνα με τη σχέση (7.22) η απόκριση συχνότητας $H(\omega_1, \omega_2)$ ενός δισδιάστατου ψηφιακού φίλτρου αποτελεί το δισδιάστατο μετασχηματισμό Fourier της κρουστι-

κής του απόκρισης $h(n_1, n_2)$. Με άλλα λόγια, η $h(n_1, n_2)$ είναι ο δισδιάστατος αντίστροφος μετασχηματισμός Fourier της $H(\omega_1, \omega_2)$:

$$h(n_1, n_2) = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} H(\omega_1, \omega_2) e^{j\omega_1 n_1 + j\omega_2 n_2} d\omega_1 d\omega_2 \quad (7.26)$$

Η ολοκλήρωση υλοποιείται σε μία περίοδο της συνάρτησης $H(\omega_1, \omega_2)$ στις κατευθύνσεις ω_1 και ω_2 .

Παράδειγμα 7.6

Να υπολογιστεί η κρουστική απόκριση ιδανικού βαθυπερατού φίλτρου του οποίου η απόκριση συχνότητας $H(\omega_1, \omega_2)$ έχει ως εξής:

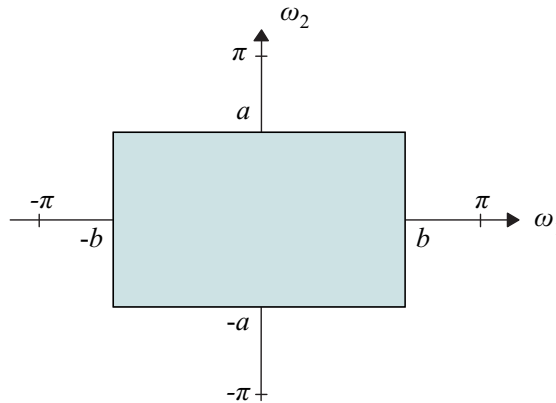
$$H(\omega_1, \omega_2) = \begin{cases} 1 & |\omega_1| \leq a < \pi, \quad |\omega_2| \leq b < \pi \\ 0 & \text{οπουδήποτε αλλού} \end{cases} \quad (7.27)$$

Σχήμα 7.17

Απόκριση συχνότητας ιδανικού βαθυπερατού φίλτρου τετραγωνικής μορφής. Στη σκιασμένη περιοχή η $H(\omega_1, \omega_2)$ είναι 1 και έξω από αυτήν 0. Η απόκριση είναι περιοδική (περίοδος 2π) και επαναλαμβάνεται στο επίπεδο (ω_1, ω_2) προς όλες τις κατευθύνσεις.

Λύση

Η απόκριση στη συχνότητα του φίλτρου θα έχει τη μορφή που φαίνεται στο Σχήμα 7.17.



Με βάση τη σχέση (7.26) η κρουστική απόκριση $h(n_1, n_2)$ θα ισούται με:

$$\begin{aligned} h(n_1, n_2) &= \frac{1}{4\pi^2} \int_{-a}^a \int_{-b}^b e^{j\omega_1 n_1 + j\omega_2 n_2} d\omega_1 d\omega_2 \\ &= \frac{1}{2\pi} \int_{-a}^a e^{j\omega_1 n_1} d\omega_1 \cdot \frac{1}{2\pi} \int_{-b}^b e^{j\omega_2 n_2} d\omega_2 \\ &= \frac{\sin(an_1)}{\pi n_1} \frac{\sin(bn_2)}{\pi n_2} \end{aligned} \quad (7.28)$$

Πρέπει να σημειώσουμε εδώ ότι η κρουστική απόκριση $h(n_1, n_2)$, που υπολογίστηκε, είναι διαχωρίσιμη, δεδομένου ότι μπορεί να γραφεί ως γινόμενο $h_1(n_1)h_2(n_2)$. Η μονοδιάστατη κρουστική απόκριση $h(n)$ είναι της ίδιας μορφής με αυτή που συναντήσαμε στη σχεδίαση των μονοδιάστατων φίλτρων {σχέση (4.7)} και προκύπτει με τον ίδιο ακριβώς τρόπο που γνωρίσαμε στην Ενότητα 4.3. Η μορφή της απόκρισης συχνότητας $H(\omega_1, \omega_2)$ μπορεί να ποικίλει ανάλογα με τις απαιτήσεις του σχεδιαστή. Όταν θέλουμε το φίλτρο να παρουσιάζει κυκλική συμμετρία, δηλαδή να συμπεριφέρεται προς όλες τις κατευθύνσεις με τον ίδιο τρόπο, τότε η σκιασμένη περιοχή του Σχήματος 7.17 είναι κύκλος και η λύση για την $h(n_1, n_2)$ είναι περισσότερο πολύπλοκη.

7.4.2 Σχεδίαση δισδιάστατων FIR φίλτρων με τη μέθοδο των παραθύρων

Τα δισδιάστατα FIR φίλτρα, σε αντίθεση με τα IIR, έχουν πεπερασμένη κρουστική απόκριση. Αυτό έχει ως αποτέλεσμα τα FIR φίλτρα να είναι πάντα ευσταθή, αφού για αυτά ισχύει η σχέση (7.20). Ένα άλλο επίσης σημαντικό πλεονέκτημα των δισδιάστατων FIR φίλτρων είναι ότι μπορούν να σχεδιαστούν ώστε η απόκριση συχνότητάς τους να είναι πραγματική, με αποτέλεσμα να μην προκαλούν μεταβολή στη φάση των διαφόρων συχνοτήτων του σήματος. Για το λόγο αυτό καλούνται «φίλτρα μηδενικής φάσης» και χρειάζεται προς τούτο να έχουν συμμετρική κρουστική απόκριση:

$$h(n_1, n_2) = h(-n_1, -n_2) \quad (7.29)$$

Τις πεπερασμένες αυτές κρουστικές αποκρίσεις τις πετυχαίνουμε αν αποκόψουμε την απεριόριστη σε μήκος δισδιάστατη κρουστική απόκριση που λαμβάνουμε με τον υπολογισμό της σχέσης (7.28). Η αποκοπή της αρχικής κρουστικής απόκρισης $h_D(n_1, n_2)$ που αντιστοιχεί στις ιδανικές ή επιθυμητές προδιαγραφές μπορεί να γίνει αν πολλαπλασιάσουμε τους όρους της $h_D(n_1, n_2)$, έναν προς έναν, με τους όρους της δισδιάστατης διακριτής συνάρτησης $w(n_1, n_2)$ που καλείται *παράθυρο* και η οποία έχει πεπερασμένο μήκος

$$h(n_1, n_2) = h_D(n_1, n_2) \cdot w(n_1, n_2) \quad (7.30)$$

Η δισδιάστατη διακριτή συνάρτηση $w(n_1, n_2)$ έχει πεπερασμένο μήκος και πρέπει να είναι συμμετρική, σύμφωνα με τη σχέση

$$w(n_1, n_2) = w(-n_1, -n_2), \quad (7.31)$$

για να διατηρηθεί η ιδιότητα της μηδενικής φάσης στο τελικό φίλτρο. Η τελική απόκριση συχνότητας $H(\omega_1, \omega_2)$ προκύπτει από τη δισδιάστατη συνέλιξη στη συχνότη-

τα της αρχικής επιθυμητής απόκρισης $H_D(\omega_1, \omega_2)$ με το φάσμα $W(\omega_1, \omega_2)$ του παραθύρου $w(n_1, n_2)$, παρόμοια με τη μονοδιάστατη περίπτωση (Ενότητα 3/ Κεφάλαιο 4).

Η επιλογή της συνάρτησης του παραθύρου είναι καθοριστική για την τελική μορφή της απόκρισης του δισδιάστατου φίλτρου. Όπως και στη μονοδιάστατη περίπτωση, το παράθυρο θα πρέπει να έχει κατάλληλη μορφή για να αποφεύγεται το φαινόμενο Gibbs και ο βαθμός της εξασθένησης στη ζώνη απόρριψης του φίλτρου να είναι ικανοποιητικός. Για το λόγο αυτό δεν χρησιμοποιείται το τετραγωνικό παράθυρο, το οποίο εξάλλου αντιστοιχεί στην απλή αποκοπή της αρχικής κρουστικής απόκρισης $h_D(n_1, n_2)$, και επομένως, συνδέεται άμεσα με τη δημιουργία του φαινομένου Gibbs. Μπορούν όμως να χρησιμοποιηθούν όλες οι συναρτήσεις παραθύρων που υπάρχουν στη βιβλιογραφία, ανάλογα κάθε φορά με τις τελικές απαιτήσεις για το φίλτρο. Ο πιο εύκολος τρόπος για να δημιουργηθεί η δισδιάστατη ψηφιακή συνάρτηση $w(n_1, n_2)$ είναι από το γινόμενο δύο μονοδιάστατων παραθύρων:

$$w(n_1, n_2) = w(n_1) \cdot w(n_2) \quad (7.32)$$

Παράδειγμα 7.7

Να σχεδιαστεί δισδιάστατο FIR φίλτρο με κρουστική απόκριση που να έχει συνολικά 11×11 όρους. Η επιθυμητή απόκριση στη συχνότητα $H_D(\omega_1, \omega_2)$ θα πρέπει να είναι:

$$H_D(\omega_1, \omega_2) = \begin{cases} 1 & |\omega_1| \leq \pi/2, |\omega_2| \leq \pi/2 \\ 0 & \text{οπουδήποτε αλλού} \end{cases} \quad (7.33)$$

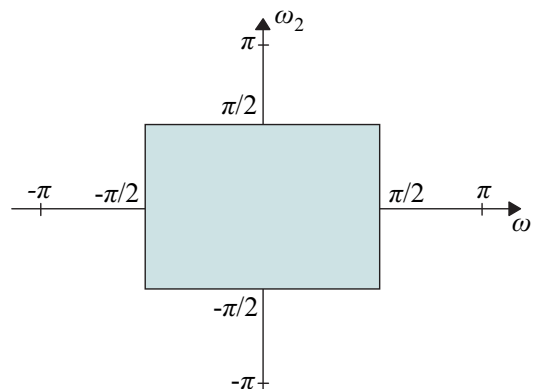
Να γίνει χρήση του μονοδιάστατου παραθύρου του Hamming.

Σχήμα 7.18

Απόκριση συχνότητας ιδανικού βαθυπερατού φίλτρου τετραγωνικής μορφής. Στη σκιασμένη περιοχή η $H(\omega_1, \omega_2)$ είναι 1 και έξω από αυτή 0. Η απόκριση είναι περιοδική (περίοδος 2π) και επαναλαμβάνεται στο επίπεδο (ω_1, ω_2) προς όλες τις κατευθύνσεις.

Λύση

Η επιθυμητή απόκριση του φίλτρου στις συχνότητες ω_1 και ω_2 , η οποία περιγράφεται από τη σχέση (7.33) έχει γραφικά όπως φαίνεται στο Σχήμα 7.18.



Η ιδανική κρουστική απόκριση θα υπολογιστεί με βάση τη σχέση (7.26) ως ο αντίστροφος μετασχηματισμός Fourier της απόκρισης συχνότητας $H_D(\omega_1, \omega_2)$:

$$\begin{aligned} h_D(n_1, n_2) &= \frac{1}{4\pi^2} \int_{-\pi/2}^{\pi/2} \int_{-\pi/2}^{\pi/2} e^{+j\omega_1 n_1 + j\omega_2 n_2} d\omega_1 d\omega_2 \\ &= \frac{1}{2\pi} \int_{-\pi/2}^{\pi/2} e^{+j\omega_1 n_1} d\omega_1 \frac{1}{2\pi} \int_{-\pi/2}^{\pi/2} e^{+j\omega_2 n_2} d\omega_2 \\ &= \frac{1}{4} \frac{\sin(\frac{\pi n_1}{2})}{\frac{\pi n_1}{2}} \frac{\sin(\frac{\pi n_2}{2})}{\frac{\pi n_2}{2}} \end{aligned} \quad (7.34)$$

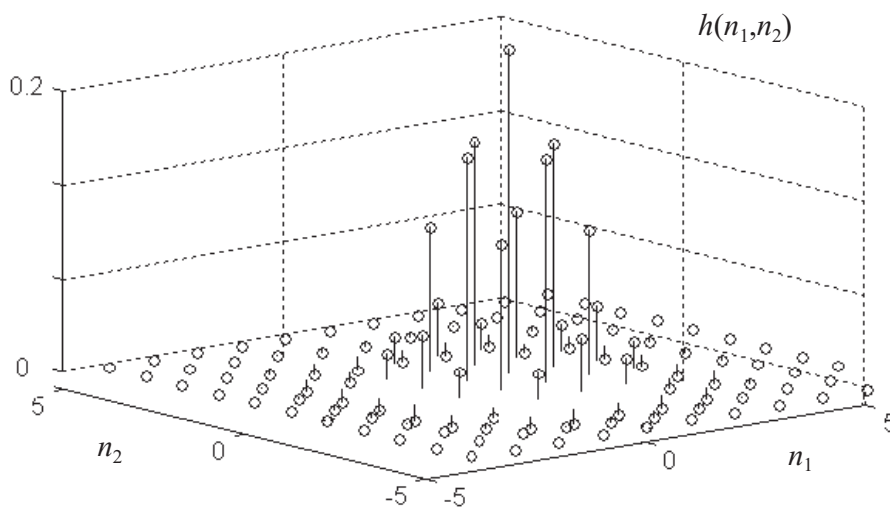
Το μονοδιάστατο παράθυρο Hamming δίνεται από τη σχέση

$$w(n) = 0,54 + 0,46 \cos\left(\frac{\pi n}{N}\right) \quad -N \leq n \leq N \quad (7.35)$$

Για $N = 5$ το μονοδιάστατο παράθυρο $w(n)$ έχει συνολικά 11 όρους. Έτσι, το δισδιάστατο $w(n_1, n_2)$ που θα προκύψει από τη σχέση (7.32), θα έχει συνολικά $11 \times 11 = 121$ όρους. Η τελική κρουστική απόκριση, που παρουσιάζεται στο Σχήμα 7.19, υπολογίζεται ως

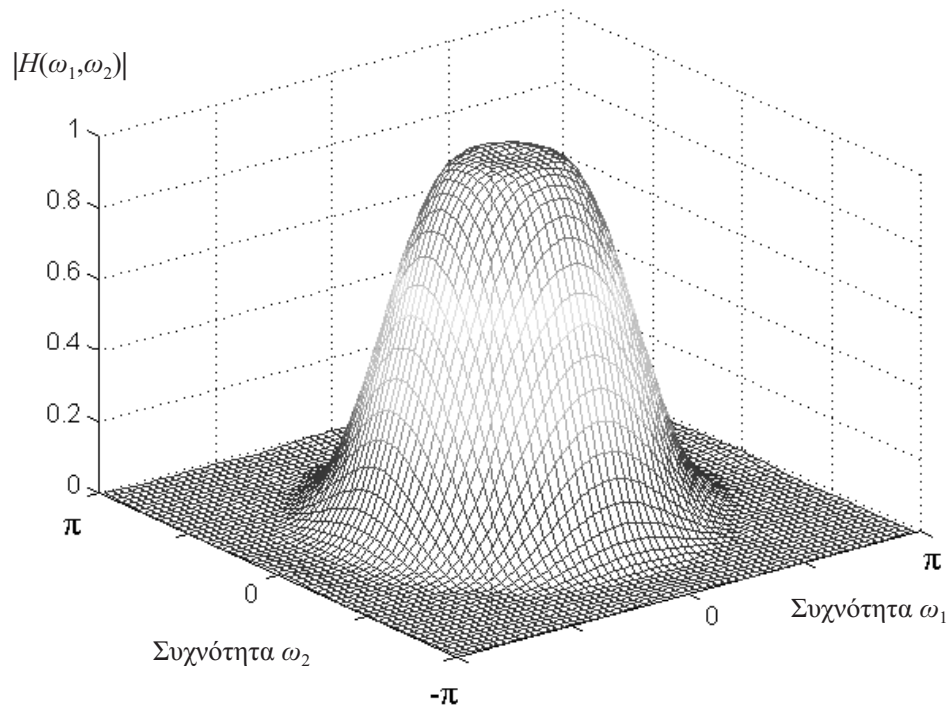
$$h(n_1, n_2) = h_D(n_1, n_2) \cdot w(n_1, n_2)$$

Αφού η κρουστική απόκριση είναι γνωστή, η απόκριση συχνότητας $H(\omega_1, \omega_2)$ του τελικού φίλτρου θα υπολογιστεί όπως ακριβώς στο Παράδειγμα 7.5. Η απόκριση αυτή δίνεται στο Σχήμα 7.20.



Σχήμα 7.19

Η τελική κρουστική απόκριση του δισδιάστατου FIR φίλτρου του Παραδείγματος 7.7, με $11 \times 11 = 121$ όρους.



Σχήμα 7.20

Το μέτρο της απόκρισης συχνότητας $|H(\omega_1, \omega_2)|$ του δισδιάστατου FIR φίλτρου του παραδείγματος 7.7.

7.4.3 Άλλες μέθοδοι σχεδίασης των δισδιάστατων FIR φίλτρων

Η μέθοδος σχεδίασης των δισδιάστατων FIR φίλτρων με τα παράθυρα που παρουσιάστηκε προηγούμενα, είναι η απλούστερη διαδικασία σχεδίασης που μπορεί να βρει ο αναγνώστης στη βιβλιογραφία. Με τη μέθοδο αυτή είναι δυνατό να σχεδιαστούν φίλτρα κυκλικά συμμετρικά στις συχνότητες, και όχι μόνο με ορθογώνια $H(\omega_1, \omega_2)$ όπως αυτές που μελετήθηκαν στα Παραδείγματα 7.6 και 7.7. Άλλες μέθοδοι σχεδίασης των FIR φίλτρων είναι οι ακόλουθες:

- *Μέθοδος σχεδίασης βέλτιστων δισδιάστατων FIR φίλτρων.* Η μέθοδος προτάθηκε από τον McClellan και αποδίδει δισδιάστατα FIR φίλτρα με τις καλύτερες προδιαγραφές για την $H(\omega_1, \omega_2)$, όταν το πλήθος των όρων της κρουστικής απόκρισης είναι συγκεκριμένο. Πρόκειται ουσιαστικά, για αλγόριθμο βελτιστοποίησης που αποτελεί τη δισδιάστατη έκδοση της αντίστοιχης μεθόδου για τα μονοδιάστατα FIR φίλτρα. Η μέθοδος απαιτεί υψηλή υπολογιστική ισχύ για φίλτρα μεσαίων προδιαγραφών (λ.χ., για τον προσδιορισμό κρουστικής απόκρισης με $33 \times 33 = 1089$ όρους).
- *Σχεδίαση των δισδιάστατων FIR φίλτρων χρησιμοποιώντας μετασχηματισμούς.* Με τη μέθοδο αυτή, την οποία επίσης πρότείνει ο McClellan, επιτυγχάνεται ο σχεδιασμός δισδιάστατων FIR φίλτρων με απόκριση συχνότητας κυκλικά συμμε-

τρική (κυρίως στις χαμηλές συχνότητες). Τα δισδιάστατα FIR φίλτρα προκύπτουν από κατάλληλο μετασχηματισμό μονοδιάστατων FIR φίλτρων. Αν τα μονοδιάστατα FIR φίλτρα είναι βέλτιστα, τότε και τα δισδιάστατα που προκύπτουν είναι κατά προσέγγιση βέλτιστα.

Ολοκληρώνοντας τη συζήτηση για τη σχεδίαση των δισδιάστατων φίλτρων θα πρέπει να αναφερθούμε για λίγο στα δισδιάστατα IIR φίλτρα. Όπως είναι γνωστό από τα κεφάλαια 4 και 5 τα μονοδιάστατα IIR φίλτρα σε σχέση με τα μονοδιάστατα FIR παρουσιάζουν τα ακόλουθα δύο μειονεκτήματα:

- Δεν είναι πάντα ευσταθή και γι' αυτό το λόγο χρειάζεται προσοχή στην τοποθέτηση των πόλων του φίλτρου.
- Δεν έχουν γραμμική απόκριση φάσης.

Τα ίδια ακριβώς μειονεκτήματα παρουσιάζουν και τα δισδιάστατα IIR φίλτρα. Επιπρόσθετα, οι δισδιάστατες συναρτήσεις μπορούν να ελεγχθούν ως προς την ευστάθειά τους αρκετά δύσκολα (είναι γενικά δύσκολη η παραγοντοποίηση και ο καθορισμός των πόλων), ενώ η έλλειψη γραμμικής φάσης είναι καταστροφική για την εικόνα, όπως αναφέραμε στην Ενότητα 6.4. Για τους δύο αυτούς λόγους αποφεύγεται η χρησιμοποίηση των δισδιάστατων IIR φίλτρων στην ψηφιακή επεξεργασία εικόνας, παρόλο που μπορούν να πετύχουν την ίδια απόκριση στη συχνότητα με λιγότερους συντελεστές.

7.4.4 Εφαρμογές των δισδιάστατων FIR φίλτρων

Οι εφαρμογές των δισδιάστατων FIR φίλτρων, κυρίως στην επεξεργασία της ψηφιακής εικόνας, είναι ποικίλες δεδομένης της ιδιότητας των φίλτρων αυτών να μην αλλάζουν τη φασική σχέση των αρμονικών του σήματος. Μεταξύ των εφαρμογών αυτών μπορούμε να αναφέρουμε επιλεκτικά τις ακόλουθες:

- Εξαγωγή του θορύβου Gauss με τη βοήθεια βαθυπερατού FIR φίλτρου (βλ. Σχήμα 7.21).
- Ανίχνευση των ακμών μιας εικόνας (edge detection) με τη βοήθεια υψηπερατού FIR φίλτρου (βλ. Σχήμα 7.22).
- Χρήση των FIR φίλτρων στην αποκατάσταση της εικόνας από συστηματικά σφάλματα (image restoration) και γενικότερα στην περίπτωση επίλυσης αντίστροφων προβλημάτων.
- Χρήση των FIR φίλτρων στη διαδικασία της σύνθεσης τομών αντικειμένου από προβολικές εικόνες (αξονική τομογραφία).

Πρέπει να σημειωθεί ότι τα FIR φίλτρα ως γραμμικά συστήματα έχουν και μειονεκτήματα. Το κυριότερο από αυτά είναι η θόλωση που προκαλεί στις ακμές της εικόνας, όταν προσπαθούμε να εξάγουμε θόρυβο (βλέπε Σχήμα 7.21γ).



α



β

Σχήμα 7.21

α) Η αρχική εικόνα, β) η εικόνα με θόρυβο Gauss. γ) Η μείωση του θορύβου με τη βοήθεια δισδιάστατου FIR φίλτρου προκαλεί ανεπιθύμητη θόλωση των ακμών της εικόνας.



γ



Σχήμα 7.22
Η χρήση υπερπαρατού FIR φίλτρου βοηθά στην απομόνωση και ενίσχυση των ακμών της εικόνας.



Σύνοψη ενότητας

Στην παράγραφο αυτή αναφερθήκαμε σε τρόπους σχεδίασης των δισδιάστατων FIR φίλτρων. Αναλυτικά, γνωρίσαμε τη μέθοδο σχεδίασης με τη χρήση των παραθύρων. Τα FIR φίλτρα είναι τα γραμμικά φίλτρα που χρησιμοποιούνται περισσότερο. Το γεγονός αυτό οφείλεται στην ιδιότητα που έχουν να μην αλλάζουν τη φασική σχέση των αρμονικών του σήματος (φίλτρα μηδενικής φάσης), αν σχεδιαστούν έτσι ώστε να παρουσιάζουν συμμετρική κρουστική απόκριση.

Σύνοψη κεφαλαίου

Στο κεφάλαιο αυτό γνωρίσαμε τον τρόπο που επιδρούν τα γραμμικά, αμετάβλητα στην ολίσθηση δισδιάστατα ψηφιακά φίλτρα πάνω στα δισδιάστατα σήματα. Στην ουσία οι δύο αυτές ιδιότητες, γραμμικότητα και αμεταβλητότητα στην ολίσθηση, οδηγούν στην πράξη της δισδιάστατης γραμμικής συνέλιξης. Τα φίλτρα αυτά χαρακτηρίζονται από την κρουστική τους απόκριση, που είναι μια δισδιάστατη διακριτή συνάρτηση και καθορίζει την απόκριση του φίλτρου στις χωρικές συχνότητες ω_1 και ω_2 . Τέλος, γνωρίσαμε μία από τις μεθόδους σχεδίασης των δισδιάστατων FIR φίλτρων. Τα FIR φίλτρα είναι κατάλληλα στην επεξεργασία της ψηφιακής εικόνας διότι δεν αλλάζουν τις φασικές σχέσεις των αρμονικών του φάσματος της εικόνας (φίλτρα μηδενικής φάσης), επομένως δεν καταστρέφουν τη μορφή της εικόνας.

Βιβλιογραφία κεφαλαίου

ΕΛΛΗΝΟΓΛΩΣΣΗ

I. Πήτας, *Ψηφιακή επεξεργασία Εικόνας*, Θεσσαλονίκη 1996.

ΞΕΝΟΓΛΩΣΣΗ

K.R. Castleman, *Digital Image Processing*, Prentice Hall, 1996

D.E. Dudgeon and R.M. Mersereau, *Multidimensional Digital Signal Processing*, Prentice–Hall, 1984.

A.K. Jain, *Fundamentals of Digital Image Processing*, Prentice Hall, 1989.

R. Klette and P. Zamperoni, *Handbook of Image Processing Operators*, John Wiley and Sons, Chichester, 1996.

J.S. Lim, *Two–Dimensional Signal and Image Processing*, Prentice Hall, 1990.

MATLAB, Image Processing Toolbox, MathWorks.

ΟΔΗΓΟΣ ΠΕΡΑΙΤΕΡΩ ΜΕΛΕΤΗΣ

I. Πήτας, *Ψηφιακή επεξεργασία Εικόνας*, Θεσσαλονίκη 1996.

Έμφαση μπορεί να δοθεί στα κεφάλαια 3 και 5.

Απόδοση αγγλικών όρων στην ελληνική

Cascade	σε σειρά
Convolution	Συνέλιξη
Edge detection	Ανίχνευση ακμών
Image restoration	Αποκατάσταση της εικόνας
Linear systems	Γραμμικά συστήματα
Separable	Διαχωρίσιμος
Shift–Invariant	Αμετάβλητο στην ολίσθηση
Two–dimensional	Δισδιάστατος
Zero phase response	Απόκριση μηδενικής φάσης

Ο μετασχηματισμός συχνότητας στον αναλογικό χώρο από βαθυπερατό σε ζωνοδιαβατό επιτυγχάνεται με βάση τη σχέση (5.19) και μας δίνει:

$$H_{BP}(s) = H_p(s) \Bigg|_{s = \Omega_c \frac{s^2 + \Omega_l' \Omega_u'}{s(\Omega_u' - \Omega_l')}} = \frac{s^2 + 0,32 \cdot 0,51}{s(0,51 - 0,32)} = \frac{s^2 + 0,17}{s \cdot 0,19}$$

$$= \frac{1}{s+1} \Bigg|_{s = \frac{s^2 + 0,17}{0,19s}} = \frac{1}{\left(\frac{s^2 + 0,17}{0,19s}\right) + 1} = \frac{0,19s}{s^2 + 0,19s + 0,17}$$

3. Εφαρμόζουμε τον διγραμμικό μετασχηματισμό στη συνάρτηση μεταφοράς του ζωνοδιαβατού αναλογικού φίλτρου, που μόλις υπολογίσαμε, αντικαθιστώντας στην τελευταία σχέση όπου s το $(z-1)/(z+1)$

$$H(z) = H_{BP}(s) \Bigg|_{s = \frac{z-1}{z+1}} = \frac{0,19s}{s^2 + 0,19s + 0,17} \Bigg|_{s = \frac{z-1}{z+1}} = \frac{0,19 \left(\frac{z-1}{z+1}\right)}{\left(\frac{z-1}{z+1}\right)^2 + 0,19 \left(\frac{z-1}{z+1}\right) + 0,17} =$$

$$= \frac{0,19(z^2 - 1)}{1,36z^2 - 1,66z + 0,98} = \frac{0,14(1 - z^{-2})}{1 - 1,22z^{-1} + 0,72z^{-2}}$$

Η αντίστοιχη εξίσωση διαφορών του IIR ψηφιακού φίλτρου είναι:

$$y(n) = 0,14[x(n) - x(n-2)] + 1,22y(n-1) - 0,72y(n-2).$$

Σας αξίζουν συγχαρητήρια αν απαντήσατε στην άσκηση αυτή. Πρόκειται για σχετικά δύσκολη άσκηση, αφού η διαδικασία απαιτεί τη μετατροπή από αναλογικό βαθυπερατό φίλτρο πρώτης τάξης σε αναλογικό ζωνοδιαβατό φίλτρο δεύτερης τάξης.

Στην περίπτωση που δεν τα καταφέρατε, μην απογοητευθείτε. Καταρχήν διαβάστε και πάλι το Παράδειγμα 5.2. Στη συνέχεια, επικεντρώστε τη προσοχή σας στις σχέσεις (5.18) και (5.19) και στα σχόλια που τις συνοδεύουν. Θα διαπιστώσετε ότι σε δεύτερης τάξης ζωνοδιαβατό φίλτρο καταλήγουμε από πρώτης τάξης βαθυπερατό φίλτρο. Τα υπόλοιπα βήματα είναι τα ίδια με εκείνα που είδαμε στα προηγούμενα παραδείγματα. Ξαναπροσπαθήστε...

6.1

Η ασπρόμαυρη εικόνα χρειάζεται συνολικά $128 \times 128 = 16.384$ bytes μνήμης. Αντίθετα, η έγχρωμη απαιτεί 3 bytes/εικονοστοιχείο, δηλαδή συνολικά $512 \times 512 \times 3 =$

1.537.536 bytes. Το τελευταίο μέγεθος είναι 48 φορές μεγαλύτερο από το πρώτο. Επομένως, η σωστή απάντηση είναι η τρίτη.

Στην περίπτωση που έχετε υπολογίσει το σωστό αποτέλεσμα σημαίνει ότι έχετε κατανοήσει την ισοδυναμία εικονοστοιχείου και byte, καθώς επίσης το γεγονός ότι το έγχρωμο εικονοστοιχείο χρειάζεται τρία bytes για αποθήκευση. Αυτό θα σας βοηθήσει να προχωρήσετε με ευχέρεια στις πράξεις πάνω στην εικόνα.

Αν δεν υπολογίσατε το σωστό αποτέλεσμα, δώστε περισσότερη προσοχή στον τρόπο υπολογισμού του πλήθους των εικονοστοιχείων και της αντιστοιχίας τους με τον αριθμό των bytes που χρειάζονται στην ασπρόμαυρη και την έγχρωμη εικόνα. Ίσως εκεί είναι το λάθος σας.

6.2

Το σύνολο των πράξεων ισούται με το πλήθος των εικονοστοιχείων (31×31) επί το πλήθος των συντελεστών του παραθύρου (5×5). Επομένως η σωστή απάντηση είναι η δεύτερη.

Αν έχετε υπολογίσει τη σωστή απάντηση, βρίσκεστε σε πολύ καλό δρόμο, δεδομένου ότι θα κατανοήσετε χωρίς πρόβλημα τον τρόπο υλοποίησης της διδιάστατης συνέλιξης που ακολουθεί στο Κεφάλαιο 7. Πρόκειται για τοπική πράξη με καθοριστική σημασία στην επεξεργασία εικόνας.

Αν δε βρήκατε την σωστή απάντηση προσπαθήστε πάλι. Δώστε προσοχή στο ότι για τον υπολογισμό κάθε εικονοστοιχείου στην εικόνα εξόδου χρειάζονται τόσες πράξεις, όσο είναι το μέγεθος του παραθύρου (ή γειτονιάς) που χρησιμοποιούμε (στην περίπτωσή μας 5×5). Ίσως το λάθος σας να είναι εκεί.

6.3

Σε μία ανεστραμμένη εικόνα η συντεταγμένη x παραμένει αμετάβλητη, αφού αφορά τη θέση του εικονοστοιχείου στον οριζόντιο άξονα. Αντίθετα, η θέση του εικονοστοιχείου στον άξονα n_2 αλλάζει και γίνεται $N + 1 - n_2$. Επομένως, ο σωστός μετασχηματισμός είναι ο τρίτος.

Αν δώσατε τη σωστή απάντηση αυτό σημαίνει ότι έχετε κατανοήσει τον τρόπο τοποθέτησης των καρτεσιανών συντεταγμένων πάνω στην εικόνα. Αυτό θα σας βοηθήσει σε δύσκολα προβλήματα μετασχηματισμών.

Αν δεν βρήκατε το σωστό αποτέλεσμα, δώστε περισσότερη προσοχή στον τρόπο που μεταβάλλεται η συντεταγμένη n_2 ή y .

7.1

Καταρχήν γράφουμε πάλι τη σχέση (7.9)

$$y(n_1, n_2) = \sum_{l_1=-\infty}^{\infty} \sum_{l_2=-\infty}^{\infty} h(l_1, l_2) x(n_1 - l_1, n_2 - l_2) \quad (7.9)$$

Το σήμα $x(l_1, l_2)$ θα μετατοπιστεί μόνο στις θέσεις (n_1, n_2) για τις οποίες οι όροι της $h(l_1, l_2)$ έχουν κοινές θέσεις με αυτό. Οι θέσεις αυτές είναι συνολικά 10.

Εάν η απάντησή σας είναι σωστή σημαίνει ότι έχετε κατανοήσει τη βασική διαδικασία της συνέλιξης και κυρίως την πράξη της ολίσθησης της μιας συνάρτησης πάνω στην άλλη.

Μην απογοητευτείτε αν το αποτέλεσμα που δώσατε είναι λάθος. Προσπαθήστε πάλι να μετατοπίσετε τη μία συνάρτηση πάνω στην άλλη και να βρείτε τις περιπτώσεις όπου έχουν κάποια επικάλυψη.

7.2

Η $h_{ισοδ}(n_1, n_2)$ του Σχήματος 7.14 μπορεί να προκύψει μόνο από τη συνέλιξη των $h(n_1, n_2)$ και $g(n_1, n_2)$. Επομένως, υλοποιείται από τον εν σειρά συνδυασμό των δύο φίλτρων.

Εάν έχετε απαντήσει σωστά τότε αποδεικνύεται ότι έχετε κατανοήσει τόσο τον τρόπο συνδυασμού των φίλτρων, όσο και την πράξη της συνέλιξης. Σε αντίθετη περίπτωση ξαναδιαβάστε τον τρόπο που υπολογίζεται η κρουστική απόκριση συνδυασμού φίλτρων και προσπαθήστε πάλι δοκιμάζοντας τόσο τον παράλληλο, όσο και τον σε σειρά συνδυασμό των δύο φίλτρων.

7.3

Η κρουστική απόκριση αντιστοιχεί σε ευσταθές φίλτρο δεδομένου ότι έχει πεπερασμένο πλήθος όρων ($n_1 = n_2 = 5$). Συνεπώς ικανοποιείται η σχέση (7.20), αφού το άθροισμα των όρων αυτών θα είναι ένας πεπερασμένος αριθμός ($< \infty$).

Αν η απάντησή σας δεν είναι σωστή μην απογοητεύεστε. Πιθανόν δεν προσέξατε την παράγραφο 7.2.3.

7.4

Για να προκύψει πραγματική η $H(\omega_1, \omega_2)$ θα πρέπει οι όροι της κρουστικής απόκρισης του φίλτρου να εμφανίζονται σε ζευγάρια συμμετρικά, έτσι ώστε να απαλείφε-

ται ο φανταστικός όρος των εκθετών της σχέσης (7.22). Αυτό ακριβώς έγινε στο Παράδειγμα 7.5 και τη Δραστηριότητα 7.3. Επομένως η σωστή απάντηση είναι η β.

Σωστή απάντησή σας σημαίνει ότι έχετε κατανοήσει τον τρόπο δημιουργίας φίλτρων μηδενικής φάσης. Αυτό θα σας είναι πολύ χρήσιμο δεδομένου ότι στην ψηφιακή επεξεργασία εικόνας χρησιμοποιούμε κυρίως φίλτρα γραμμικής φάσης.

Αν κάνατε λάθος, προσπαθήστε πάλι διαβάζοντας την προηγούμενη δραστηριότητα και την υποενότητα 4.1.1.

Απαντήσεις Δραστηριοτήτων

4.4

Θα υπολογίσουμε τους συντελεστές του φίλτρου στην περίπτωση που $N = 3$ (third-band φίλτρο). Στην περίπτωση αυτή ισχύει $\omega_c = \pi/3$ και από την εξίσωση 4.7 έχουμε:

$$h_D(n) = \frac{\omega_c \sin(n\omega_c)}{\pi n\omega_c} = \frac{\omega_c = \pi/3}{\pi} \rightarrow \frac{1 \sin(n\pi/3)}{3 (n\pi/3)}$$

Η κρουστική απόκριση h_D είναι συμμετρική γύρω από το $n = 0$, έχει άπειρους όρους και κάθε τρίτος όρος είναι μηδενικός (n πολλαπλάσιο του 3).